5 | Phase Transitions

> The lobster demands to be cut while alive.
>
> —Brillat-Savarin

5.1 Liquid–Gas Coexistence and Critical Point

I borrow from M. E. Fisher this vivid description of carbon dioxide's critical point [Fish83, p. 6]:

The first critical point to be discovered was in carbon dioxide. Suppose one examines a sealed tube containing CO_2 at an overall density of about 0.5 gm/cc, and hence a pressure of about 72 atm. At a temperature of about 29° C one sees a sharp meniscus separating liquid (below) from vapor (above). One can follow the behavior of the liquid and vapor densities if one has a few spheres of slightly different densities close to 0.48 gm/cc floating in the system. When the tube is heated up to about 30°C one finds a large change in the two densities since the lighter sphere floats up to the very top of the tube, i.e. up into the vapor, while the heaviest one sinks down to the bottom of the liquid. However, a sphere of about "neutral" density (in fact "critical density") remains floating "on" the meniscus. There is, indeed, still a sharp interface between the two fluids, but they have approached one another closely to density. Further slight heating to about 31°C brings on the striking phenomenon of critical opalescence. If the carbon dioxide, which is quite transparent in the visible region of the spectrum, is illuminated from the side, one observes a strong intensity of scattered light. This has a bluish tinge when viewed normal to the direction of illumination, but has a brownish-orange streaky appearance, like a sunset on a smoggy day, when viewed from the forward direction (i.e., with the opalescent fluid illuminated from behind). Finally, when the temperature is raised a further few tenths of a degree, the opalescence disappears and the fluid becomes completely clear again. Moreover, the meniscus separating "liquid" from "vapor" has completely vanished: no trace of it remains! *All* differences between the two phases have gone: indeed only one, quite homogeneous, "fluid" phase remains above the critical temperature ($T_c = 31.04$°C).
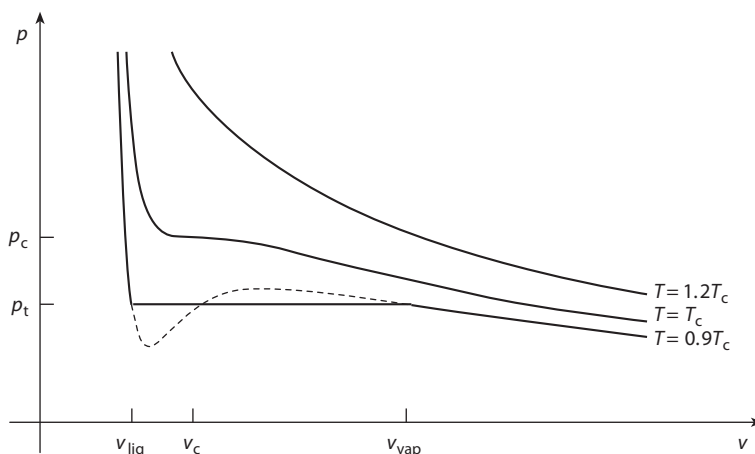
FIGURE 5.1. Pressure-volume isotherm of a fluid according to the van der Waals equation. The dashed line represents the nonphysical part of the isotherm in the coexistence region.

We can describe this behavior in a pressure–volume diagram like the one shown in figure 5.1. This diagram is based on the (approximate) van der Waals equation of state, which we will discuss in the next section. Above a characteristic temperature value T_c, called the **critical temperature**, the pressure p is a monotonically decreasing function of the specific volume per particle $v = V/N$. For large values of v, p is inversely proportional to v and therefore follows the equation of state for perfect gases. When p becomes very large, v cannot decrease beyond a certain limit, which is the molecule's *intrinsic volume*. As temperature diminishes, the isotherm tends to become flatter in the intermediate region, until for $T = T_c$, it shows a horizontal inflection point corresponding to the value $v = v_c = 1/\rho_c$. This value of density is called the **critical density**. The value p_c for pressure corresponds to this value of specific volume. Below the critical temperature, the isotherm displays a flat region, between $v = v_{liq} = 1/\rho_{liq}$ and $v = v_{vap} = 1/\rho_{vap}$, where $v_{liq,vap}$ are the densities of the liquid and vapor, respectively. Only one pressure value corresponds to each value of density included between these two values; it is called the **transition pressure** (or **vapor pressure**), and we will denote it by $p_t(T)$. If, as mentioned earlier, we consider a sample of critical density and we place it at a temperature below T_c, we observe the coexistence of two phases at different densities. By increasing the temperature, the difference in density decreases until it disappears at T_c.

By plotting this behavior in a pressure–temperature diagram, we obtain the **phase diagram** shown in figure 5.2. Let us note that in this case, the diagram's axes both correspond to *intensive* quantities. The curve $p = p_t(T)$ in the plane (p, T) defines the line in which there is a coexistence of the two phases: vapor on the left and liquid on the right. This curve terminates in correspondence with the point (p_c, T_c), without exhibiting any singularities.

Other curves in the same plane describe the coexistence of fluid and solid phases—they meet in the **triple point** (p_{tr}, T_{tr}), where all three phases coexist. One should note that
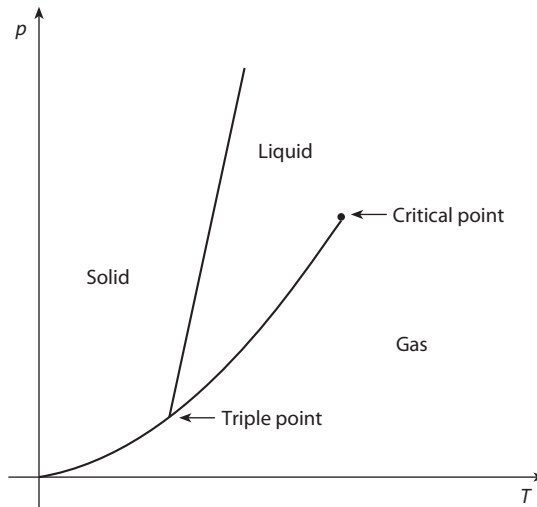
FIGURE 5.2. Phase diagram (schematic, not to scale) of a fluid on the plane (p, T). Note the liquid–gas coexistence curve, which terminates at the critical point, and the fusion (solid–liquid coexistence) and sublimation (solid–gas coexistence) curves, which meet the preceding curve at the triple point.

neither the solid–liquid nor the solid–vapor coexistence exhibit a critical point, unlike the liquid–vapor coexistence.

5.2 Van der Waals Equation

A first interpretation of this curious thermodynamic behavior can be found in van der Waals' doctoral thesis [Waal73]. James C. Maxwell was so enthusiastic about this work (to which he made a substantial contribution [Maxw74] as we will see) that he said it was worthwhile to learn Dutch even just to be able to read it. (It has since been translated.) Van der Waals' idea is to start from the equation of state for ideal gases (we denote by v the volume per particle V/N):

$$pv = k_{\mathrm{B}} T. \tag{5.1}$$

We will take two effects into account:

1. The *intrinsic volume* effect of the gas particles: the volume actually available to each particle is not equal to v, but is smaller, and equal to $(v - b)$, where b (the **covolume**) is a parameter related to the particles' intrinsic volume.

2. The effect of the particles' reciprocal attraction: the pressure exerted by the particles on the walls is lower than the actual pressure acting on them within the sample, because they tend to be attracted to the other particles. In order to evaluate this effect, let us denote by u the potential of a particle within the fluid, due to the attraction by the other particles, and by u_{w}

the corresponding value of the potential for a particle close to the wall. One then expects $u_w = u/2$, since only the particles within a half-space attract a particle close to the wall:

$$u_w = \frac{u}{2} = -\frac{N}{2V}u_0, \qquad (5.2)$$

where u_0 is the integral of the pair potential $u(r)$:

$$u_0 = \int d^3 r\, u(r). \qquad (5.3)$$

This allows us to evaluate the local density near the wall:

$$
\begin{aligned}
\left.\frac{N}{V}\right|_w &= \frac{N}{V}\exp\left[-\frac{1}{k_B T}(u_w - u)\right] \\
&= \frac{N}{V}e^{-Nu_0/2Vk_B T}.
\end{aligned}
\qquad (5.4)
$$

Supposing that the pressure exerted by the molecules near the wall is given at least approximately by the ideal gas law, we obtain

$$
\begin{aligned}
p &= k_B T \left.\frac{N}{V}\right|_w = k_B T \frac{N}{V}e^{-Nu_0/2Vk_B T} \simeq k_B T \frac{N}{V}\left(1 - \frac{Nu_0}{2Vk_B T}\right) \\
&= k_B T \frac{N}{V} - \frac{1}{2}\left(\frac{N}{V}\right)^2 u_0.
\end{aligned}
\qquad (5.5)
$$

One thus obtains the well-known **van der Waals equation**:

$$\left(p + \frac{a}{v^2}\right)(v - b) = k_B T, \qquad (5.6)$$

where $a = u_0/2$ and b are two parameters characteristic of the material.

Maxwell realized that the isotherms' behavior, below a temperature $T_c = 8a/27b$, was no longer monotonically decreasing, and that this contradicted thermodynamic stability, according to which the isothermal compressibility:

$$K_T = -\frac{\partial v}{\partial p}\bigg|_T \qquad (5.7)$$

must be a positive quantity. He therefore suggested that below T_c, a portion of the isotherm should be replaced with a horizontal segment, which represented phase coexistence. One could determine the position of this straight line by making use of the construction (which was later given his name) whereby the total work performed by first going along the horizontal line and then along the isothermal curve would vanish. This construction is equivalent to imposing the equality of Helmholtz free energy per molecule for the two coexisting phases:

$$\mu_{liq} = \mu_{vap}. \qquad (5.8)$$

The van der Waals equation of state is just a clever interpolation formula, which does however possess a significant heuristic value. It is, in effect, possible to estimate the two parameters, a and b that appear in the equation, from measures on gases at low densities, and then substitute the results in the expression of the critical parameters:

$$k_B T_c = \frac{8a}{27b}, \qquad v_c = 3b, \qquad p_c = \frac{a}{27b^2}. \tag{5.9}$$

The results one obtains are reasonable. Moreover, one can choose to measure the temperature, pressure, and volume in terms of the corresponding critical values, introducing the ratios:

$$\mathcal{T} \equiv \frac{T}{T_c}, \qquad \mathcal{P} \equiv \frac{p}{p_c}, \qquad \mathcal{V} \equiv \frac{v}{v_c}. \tag{5.10}$$

One thus obtains a "universal" equation, independent of the nature of the fluid being considered:

$$\left(\mathcal{P} + \frac{3}{\mathcal{V}^2} \right) \left(\mathcal{V} - \frac{1}{3} \right) = \frac{8}{3} \mathcal{T}. \tag{5.11}$$

This equation is called the **law of corresponding states**. It yields good results if one is not too close to the critical temperature.

More particularly, it follows from the van der Waals equation that for T close enough to T_c, the difference between the gas's and the liquid's specific volume vanishes as a power of $|T - T_c|$:

$$v_{vap} - v_{liq} \propto |T - T_c|^\beta. \tag{5.12}$$

It follows from the van der Waals equation (as one can easily see) that $\beta = 1/2$. Experimentally, one can observe that this quantity actually behaves as a power of $|T - T_c|$, but with an exponent β closer to 1/3 than to 1/2. It is also interesting to remark that the value of this exponent is apparently the same (within the errors) for several different liquids, but also that the equation of state has an apparently universal form, as can be seen in figure 5.3.

5.3 Other Singularities

As a consequence of the van der Waals equation, one can show that isothermal compressibility exhibits a singularity at the critical point, which can be verified experimentally:

$$\chi = -\frac{\partial v}{\partial p}\bigg)_T \sim |T - T_c|^{-\gamma}. \tag{5.13}$$

The van der Waals equation, however, implies that $\gamma = 1$, while experimentally, γ is closer to 1.25.

As the system approaches the critical temperature, also the specific heat C exhibits a singularity that is well described by a power law of the form

$$C \sim |T - T_c|^{-\alpha}, \tag{5.14}$$

where the exponent α has a value included between 1/8 and 1/9 and is the same (within the errors) for all fluids. As we will see, this behavior contrasts with the simpler theories of phase transitions.
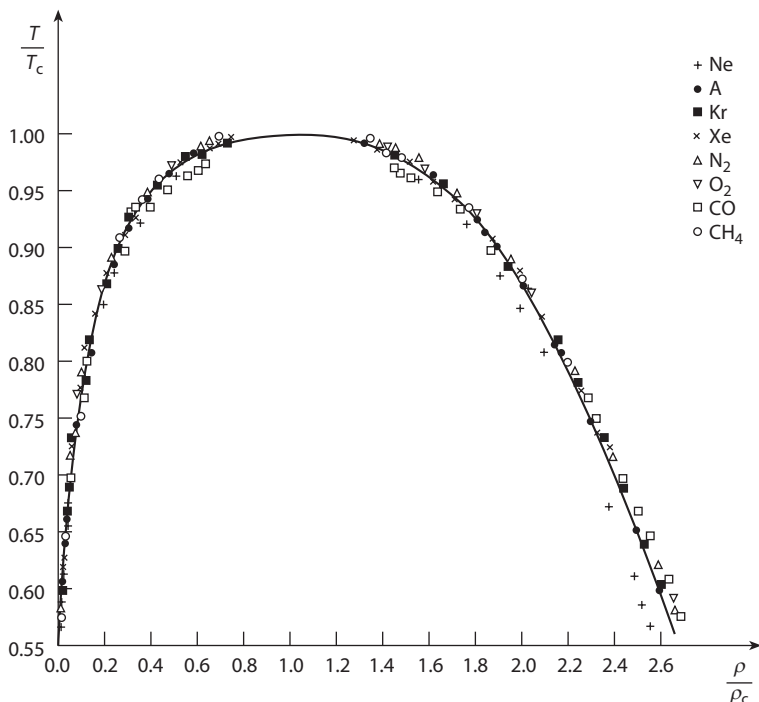
FIGURE 5.3. Coexistence curve ρ/ρ_c as a function of T/T_c for eight fluids in the proximity of the critical temperature. The continuous curve corresponds to $\rho_{\text{liq,vap}} - \rho_c \propto |T - T_c|^\beta$, with $\beta = 1/3$. From [Gugg45], with permission.

5.4 Binary Mixtures

Another system that exhibits an analogous behavior is a fluid mixture composed of two chemical species. One usually uses organic compounds whose names are difficult to pronounce—one of the most convenient (also from a phonetic point of view) is aniline–cyclohexane, because its critical point is close to ambient temperature at atmospheric pressure. We will refer to one of the components as A and to the other as B. By fixing the pressure, it is possible to observe the coexistence of *three* phases within a certain temperature interval, and they are arrayed from the bottom up in a decreasing order of density: an α phase rich in A, a β phase rich in B, and a vapor phase. As the temperature varies, the composition (and therefore also the density) of the α and β phases move closer to each other until, for a particular temperature value T_c (as long as we have selected the correct composition of the sample) the two liquid phases are replaced by a homogeneous phase. The point at which this occurs is called the **consolution point**.

The behavior of this system is fairly similar to that of the liquid–gas system. More particularly, it also exhibits critical opalescence in proximity of the consolution point. Let us note some differences:

1. The quantity that distinguishes the two phases is the relative concentration of one of the components—A, for example. We refer to A's molar fraction as x_A and its value in phase α as x_A^α—then, $x_A^\alpha - x_A^\beta \to 0$ when $T - T_c$ along the coexistence curve.

2. Fairly frequently, the consolution point corresponds to a temperature *lower* than the coexistence temperature. In other words, a homogeneous phase is replaced by a phase coexistence when the sample is heated. This does not occur with simple fluids.

3. One can observe, however, that the exponent β that describes the behavior of this system is also close to 1/3 and indistinguishable from the value it assumes in simple fluids.

5.5 Lattice Gas

In order to obtain a more microscopic description of phase coexistence, let us build a simple statistical model of a "fluid." We consider a system of N particles, described by the Hamiltonian:

$$H = K + U, \tag{5.15}$$

where K is the kinetic energy

$$K = \sum_{\alpha=1}^{N} \frac{p_\alpha^2}{2m}, \tag{5.16}$$

and U is the potential energy, which depends only on the particles' positions. When we calculate the partition function:

$$Z = \frac{1}{N!} \int \prod_{\alpha=1}^{N} \frac{d\mathbf{p}_\alpha d\mathbf{r}_\alpha}{h^3} \exp\left(-\frac{H}{k_B T}\right), \tag{5.17}$$

the integral over momenta $\mathbf{p}_\alpha$ factorizes and results in λ^{-3N}, where λ is the thermal de Broglie wavelength:

$$\lambda^2 = \frac{h^2}{2\pi m k_B T}. \tag{5.18}$$

In the two previous equations, h is Planck's constant, which appears to account for the volume of phase space occupied by the quantum states. In this fashion, we obtain the partition function expressed as a function of an integral over the system's configurations:

$$Z = \frac{\lambda^{-3N}}{N!} \int \prod_{\alpha=1}^{N} d\mathbf{r}_\alpha \exp\left(-\frac{U}{k_B T}\right). \tag{5.19}$$

The interaction between particles is usually described by a potential that exhibits a hard core—it becomes large and positive when the particles get closer than a certain distance a_0, where $a_0/2$ is the particle radius—and an attractive tail, i.e., it becomes negative at intermediate distances, eventually vanishing when the particles are far apart.

We can provide a drastically simplified description of this state of affairs with the following hypotheses:

1. The particles can occupy only the sites on a simple cubic lattice whose lattice constant is a_0.

2. At every site, there can be one particle at most.

3. For each pair of particles that occupy nearest-neighbor sites, the system's potential energy decreases by a quantity ϵ.

Let us denote the state of occupation of site i by a variable τ_i, which is equal to 0 if the position is empty and 1 if it is occupied. Then the partition function assumes the form

$$Z_N = \lambda^{-3N} \sum_{\{\tau_i\}} \delta_{N,\Sigma_i \tau_i} \exp\left(\sum_{\langle ij \rangle} \frac{\epsilon}{k_B T} \tau_i \tau_j\right). \tag{5.20}$$

In this equation, the Kronecker delta expresses the condition that the total number of particles in the system is exactly N:

$$\sum_i \tau_i = N. \tag{5.21}$$

The expression $\sum_{\langle ij \rangle}$ indicates that the sum runs over all pairs of nearest-neighbor sites on the lattice. The combinatorial factor $N!$ has disappeared because this expression is invariant with respect to the permutation of particle labels. The model defined in this manner is called the **lattice gas**.

In order to make this expression more manageable, it is useful to eliminate the Kronecker delta—one can do this by shifting to the grand canonical ensemble. In this fashion, we define the grand partition function Z_{GC}:

$$Z_{GC} = \sum_{N=0}^{\infty} \exp\left(\frac{\mu N}{k_B T}\right) Z_N. \tag{5.22}$$

The grand partition function Z_{GC} is related to the pressure p by the equation

$$\ln Z_{GC} = pV/k_B T, \tag{5.23}$$

where V is the system's volume. It is easy to see that Z_{GC} can be expressed as follows:

$$Z_{GC} = \sum_{\{\tau_i\}} \exp\left(-\frac{H}{k_B T}\right), \tag{5.24}$$

where the Hamiltonian H is defined by

$$-\sum_{\langle ij \rangle} \epsilon \tau_i \tau_j - \sum_i \tilde{\mu} \tau_i. \tag{5.25}$$

In this equation, we have introduced the expression

$$\frac{\tilde{\mu}}{k_B T} = \frac{\mu}{k_B T} - 3 \ln \lambda. \tag{5.26}$$

The particle density per site is given by the quantity

$$x = \langle \tau \rangle = \frac{1}{N}\sum_i \langle \tau_i \rangle. \tag{5.27}$$

Along the coexistence curve $\tilde{\mu} = \tilde{\mu}*(T)$, for $T < T_c$, this quantity exhibits a discontinuity:

$$x^+ = \lim_{\tilde{\mu} \to \tilde{\mu}*(T)^+} x(\tilde{\mu}, T) > x^- = x^+ = \lim_{\tilde{\mu} \to \tilde{\mu}*(T)^-} x(\tilde{\mu}, T). \tag{5.28}$$

The discontinuity tends to zero when the temperature gets close to the critical temperature. If the model correctly describes the experimental data, we can expect the discontinuity to tend to zero as a power of $|T < T_c|$:

$$x^+ - x^- \propto |T - T_c|^\beta. \tag{5.29}$$

5.6 Symmetry

The model we just described cannot be solved exactly. It is possible, however, to manipulate it to highlight certain properties. In place of the variable τ_i, which assumes the values 0 and 1, let us consider the variable σ_i, defined by

$$\sigma_i = 2\tau_i - 1. \tag{5.30}$$

This variable assumes the values -1 (when $\tau_i = 0$) and $+1$ (when $\tau_i = 1$). Expressing the τ_i as a function of the σ_i, and substituting in the Hamiltonian (5.25), we obtain the expression (up to an irrelevant additive constant):

$$H = -\sum_{\langle ij \rangle} J\sigma_i\sigma_j - \sum_i h\sigma_i, \tag{5.31}$$

We have defined

$$J = \epsilon/4, \tag{5.32}$$

$$h = (\tilde{\mu} + \epsilon\zeta)/2, \tag{5.33}$$

where ζ is the lattice's **coordination number**—in other words, the number of each site's nearest neighbors. In the case of the simple cubic lattice, $\zeta = 6$ (and it is equal to $2d$ in the case of a cubic lattice in d dimensions). This Hamiltonian can also describe binary mixtures, if one stipulates that $\sigma_i = +1$ corresponds to the presence of a molecule of type A in i, $\sigma_i = -1$ to a molecule of type B, and $2h = \mu_B - \mu_A$.

This expression makes explicit the fact that if $h = 0$, the value that H assumes for a certain configuration $\sigma = \{\sigma_i\}$ remains invariant with respect to the transformation I, where

$$\mathsf{I}\sigma_i = -\sigma_i, \quad \forall i. \tag{5.34}$$

We can express this fact by saying that when $h = 0$, the Hamiltonian H is **symmetrical** with respect to the transformation group **G** composed by I and the identity **E**. Since $\mathsf{I}^2 = \mathsf{E}$, it is not necessary in this case to consider other transformations.

The model we obtained is the most famous model of statistical mechanics. It was invented by Lenz, who proposed it as a thesis to one of his students, and for this reason, it is given the student's name, *Ising*. It was actually proposed to describe the behavior of ferromagnetic systems. In this case, σ_i is interpreted as the component z of the magnetic moment of an electron (apart from a factor), and h is proportional to the applied magnetic field. The factor J represents the tendency of electrons belonging to contiguous atoms to align themselves parallel because of the exchange interaction—it is therefore known as the exchange integral. In spite of the quantum origin of this interaction the Ising model is a classical model (in the sense that all the quantities commute).

5.7 Symmetry Breaking

We can now look at phase coexistence from another point of view. In the case of the ferromagnet, the two coexisting phases are equivalent from a physical standpoint and differ only because of the prevailing direction of their spin alignment. In other words, although the Hamiltonian is invariant with respect to the transformation group **G**, the thermodynamic state is not. This situation is called **spontaneous symmetry breaking**. It can be made manifest by identifying some thermodynamic quantities that should vanish in a thermodynamic state that is symmetrical with respect to the transformation group **G** and that instead do not vanish.

Let us consider, for instance, the **magnetization** M in the Ising model:

$$\langle M \rangle = \sum_i \langle \sigma_i \rangle = \frac{1}{Z} \sum_\sigma \sum_i \sigma_i \exp\left[-\frac{H(\sigma)}{k_B T} \right], \tag{5.35}$$

where the first sum runs over all of the system's 2^N configurations $\sigma = (\sigma_i)$, and $H(\sigma_i)$ is Ising's Hamiltonian (5.31).

Let us assume that $h = 0$, and let us consider the contributions to $\langle \sigma_i \rangle$. For each configuration σ in which $\sigma_i = +1$, there will be a configuration $I\sigma_i = (-\sigma_i)$ in which $\sigma_i = -1$. Since $H(I\sigma) = H(\sigma)$, this configuration will, in principle, have the same Boltzmann weight as the preceding one. We can therefore expect $\langle \sigma_i \rangle$ to always vanish. But then how can phase coexistence occur? The point is that in certain circumstances, a small nonsymmetrical perturbation is sufficient to cause a macroscopic violation of symmetry. It is clear that states in which the spin is prevalently oriented up or down have a much lower energy than those in which the spins are arranged in a disordered fashion. The latter are much more numerous, and therefore prevail at high temperatures. But at low temperatures, they are not accessible (they cost too much energy). The system therefore prefers that the spins be aligned prevalently in one direction. Which direction is indifferent from an energetic point of view—a small perturbation (due to boundary conditions that slightly favor the up direction, or to a weak magnetic field, for instance) is therefore sufficient to make the system prefer one direction over another.

It is possible in this manner to obtain a thermodynamic state in which magnetization is not equal to zero, even though the external macroscopic field h vanishes:

$$\langle M \rangle = \sum_i \langle \sigma_i \rangle = Nm \neq 0. \tag{5.36}$$

The Hamiltonian's symmetry implies that if $M_0 = \langle M \rangle \neq 0$ defines a state of thermodynamic equilibrium, there exists another state of thermodynamic equilibrium (which corresponds to the same values of the intensive variables, but perhaps to different boundary conditions) in which the magnetization is equal to $-M_0$. Therefore, if $\langle M \rangle \neq 0$ for $h = 0$, several states characterized by different values of the magnetization can coexist.

The magnetization's value can be fixed by means of an artifice called **quasi-average**. One applies an external field h—a positive one, for example—which explicitly breaks symmetry, and one evaluates the magnetization in the limit in which this field goes to zero:

$$Nm_0(T) = \lim_{h \to 0^+} \langle M \rangle. \tag{5.37}$$

Because of the symmetry, one has

$$\lim_{h \to 0^-} \langle M \rangle = -Nm_0(T). \tag{5.38}$$

Therefore, $m_0(T)$ is a measure of the degree of spontaneous symmetry breaking. This quantity is known as the **spontaneous magnetization** per particle and vanishes at the critical point.

Let us note that the formal equivalence between the Ising model and the lattice gas allows us to identify the coexistence curve for the latter in the plane (μ, T):

$$\tilde{\mu}_c(T) = -\epsilon \zeta. \tag{5.39}$$

5.8 The Order Parameter

Let us now consider the same problem from a slightly more abstract point of view. Let us suppose that we have a system described by the Hamiltonian H, which is invariant with respect to a transformation group **G** applied to the microscopic states σ:

$$H(g\sigma) = H(\sigma), \quad \forall g \in \mathbf{G}. \tag{5.40}$$

If the thermodynamic state is not invariant with respect to all the group's transformations, one says that spontaneous symmetry breaking is present. In order to define the degree of symmetry breaking, one considers an observable M that has the following characteristics:

1. M is extensive: M's value for a system composed of several subsystems in contact is equal to the sum of each subsystem's values.

2. M is transformed according to a nontrivial representation of the group **G**:

$$M(g\sigma) = \mathsf{T}(g)M(\sigma), \quad \forall g \in \mathbf{G}, \tag{5.41}$$

   where $\mathsf{T}(g)$ is a linear operator belonging to some representation of **G**.

The symmetry breaking then manifests itself by a means of a nonvanishing value of $\langle M \rangle$. With the quasi-average method, it is possible to obtain various values of $\langle M \rangle$ that correspond

to the same value of temperature (and of other possible intensive parameters). Because of symmetry, they can be obtained from one another by applying a $\mathsf{T}(g)$ transformation.

A quantity endowed with these characteristics is called an **order parameter**. The introduction of this concept into phase transition theory is due to Lev D. Landau (see [Land80], ch. XIV).

Let us now consider a certain number of systems, the transitions that characterize them, and the corresponding order parameters:

1. **Ising antiferromagnet**: Let us assume that the exchange integral J in the Ising Hamiltonian is negative. In the lowest energy states, the neighboring spins tend to be aligned in *antiparallel* fashion. In this case, reflection symmetry ($\sigma_i \rightarrow -\sigma_i$) and translational symmetry $\sigma_i \rightarrow \sigma_{i+\delta}$ are broken at the same time.

   One can then use the **staggered magnetization** as the order parameter. In a simple cubic lattice in d dimensions, one can identify two intertwined sublattices, such that each point of each sublattice is surrounded by nearest-neighbor points belonging to the other sublattice. Let us now assign a factor ϵ_i, whose value is $+1$ in one sublattice and -1 in the other, to each point i. The staggered magnetization is then defined by

$$N \equiv \sum_i \epsilon_i \sigma_i. \tag{5.42}$$

   The symmetry group $\mathbf{G}$ is still the group with two elements (E, I) that correspond to spin identity and spin inversion.

2. **Vector order parameter**: The Ising model describes a ferromagnet in which the spins are constrained to align in parallel or antiparallel fashion in a fixed direction. One often finds situations in which the spins have more degrees of freedom and can point in arbitrary directions in a plane or in space. We will disregard here the quantum nature of the spins. If we denote the spin's components by $\sigma_i^\alpha (\alpha = 1, 2, 3)$, we can define the order parameter M as follows:

$$M^\alpha = \sum_i e^\alpha \sigma_i^\alpha, \tag{5.43}$$

   where e^α is the versor of the α axis. More generally, one can consider vectorial order parameters with n components, which are transformed by the rotational symmetry in n dimensions, represented by the group $O(n)$. The generalization of Ising's model for $n = 2$ is often called the **XY model**, while that for $n = 3$ is called the **Heisenberg model**.

3. **Einstein condensation**: Sometimes the order parameter is not directly accessible to the experiment. In the Einstein condensation, for example, the order parameter is the condensate's wave function, and the broken symmetry is the gauge symmetry of the first kind, which expresses the invariance with respect to the multiplication of the wave functions by a phase:

$$\Psi \rightarrow e^{i\alpha}\Psi. \tag{5.44}$$

   Since the order parameter is a complex number, and the gauge transformation is isomorphic to a rotation in the plane, it can be assimilated to a vector order parameter with $n = 2$.

There are more complex situations (for example, in liquid crystals) that require the introduction of less intuitive order parameters. What needs to be emphasized is that the identification of the order parameter can be a fairly difficult problem by itself, because there is no method that allows one to identify it a priori. One needs to be guided by physical intuition.

5.9 Peierls Argument

The existence of spontaneous symmetry breaking in the two-dimensional Ising model can be proven by means of an argument made by Peierls [Peie34]. This argument can easily be generalized to more than two dimensions.

I will try to describe this argument without dwelling excessively on rigor—the argument expounded in this fashion can, however, easily be made rigorous. This section will therefore also serve as an introduction to the problems and methods of a rigorous statistical mechanics.

Let us then consider an Ising model defined over a square lattice with $L \times L$ spin. A spin variable $\sigma_i = \pm 1$ is associated with each lattice site i. The system's microstate σ is defined by the collection of the values (σ_i) of these variables for each site i. The system's Hamiltonian is the usual Ising Hamiltonian:

$$\mathcal{H} = -\sum_{\langle ij \rangle} J\sigma_i\sigma_j, \tag{5.45}$$

where the sum over $\langle ij \rangle$ runs over all the pairs of nearest-neighbor sites. This Hamiltonian is invariant with respect to spin inversions defined by

$$\sigma' = \mathsf{I}\sigma, \qquad \sigma'_i = -\sigma_i, \qquad \forall i. \tag{5.46}$$

Spontaneous symmetry breaking occurs when, in the thermodynamic limit, one can have different states of equilibrium in correspondence with the same temperature value T.

In order to clarify our ideas, let us assume that we impose on the system the boundary conditions +, in which the spins external to the system are directed up. In this situation, we expect that there will be a slight prevalence of up spins in the entire system at any temperature. On the other hand, if we impose the boundary conditions −, in which the spins external to the system are directed down, we will have a prevalence of down spins. The problem is whether this prevalence due to the boundary conditions will remain in the thermodynamic limit.

Let us consider the spin placed at the origin, and let us denote by p the probability that it is equal to $+1$, with the boundary conditions +. Symmetry breaking occurs if p remains strictly larger than $1/2$ in the thermodynamic limit. The symmetry of the Hamiltonian guarantees in fact that p is also equal to the probability that the spin at the origin is equal to -1 with the boundary conditions −. Therefore, if $p > 1/2$, there will be two different equilibrium states at the same temperature T, which are selected by the boundary conditions.
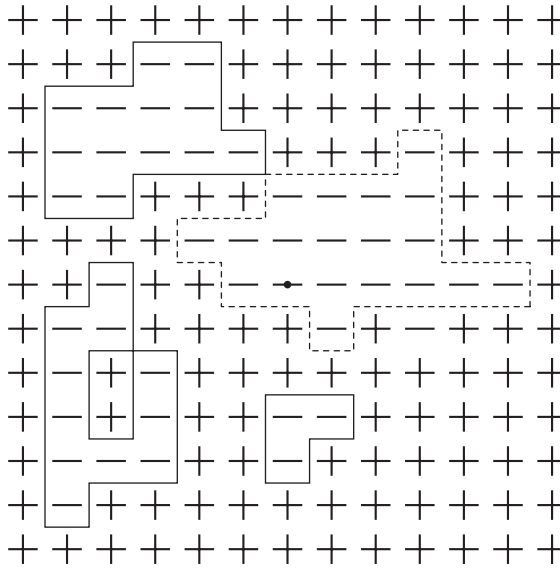
FIGURE 5.4. Configuration of boundaries associated with a spin configuration. The spin at the origin is marked by a black dot.

Let us now evaluate p with the boundary conditions $+$ and show that, if the temperature is low enough, one has $p > 1/2$. Let us consider an arbitrary configuration σ of the spins of the system. We can associate a configuration Γ of boundaries between up and down spins with each such configuration. When the boundary conditions are fixed, this correspondence is such that to each σ corresponds one Γ and vice versa.

Let us focus our attention on the spin at the origin (denoted by a black dot in figure 5.4), and distinguish the configurations into two sets: the set $\sum^+$ contains all σ configurations in which this spin is equal to $+1$, and the set $\sum^-$ contains all the others in which the spin at the origin is equal to -1. It is possible to associate a precise configuration $\sigma^+ \in \sum^+$ with each configuration $\sigma^- \in \sum^-$, thus flipping the direction of all the spins inside the smallest boundary that also contains the origin within it. In the case shown in figure 5.4, it is the discontinuous boundary. The same configuration σ^+ can obviously be obtained starting from several configurations $\sigma^- \in \sum^-$.

One obviously has

$$p = \frac{1}{Z} \sum_{\sigma \in \Sigma^+} e^{-\mathcal{H}(\sigma)/k_B T}, \tag{5.47}$$

where Z is the partition function, and the sum runs over all the configurations belonging to $\sum^+$. Analogously,

$$q = 1 - p = \frac{1}{Z} \sum_{\sigma \in \Sigma^-} e^{-\mathcal{H}(\sigma)/k_B T}. \tag{5.48}$$

It is clear that the energy of a configuration σ is equal to the energy E_0 of the ground state ($\sigma_i = +1$, $\forall i$) plus the contribution of the boundaries. Since each pair of antiparallel nearest-neighbor spins contributes $+2J$ to energy, one has

$$H(\sigma) = E_0 + 2J \,|\,\Gamma\,|, \tag{5.49}$$

where $|\Gamma|$ is the total length of the boundaries that appear in σ. We will denote the set of all configurations $\sigma' \in \sum^-$ that are applied in σ, by removing the smallest boundary that contains the origin, with $\Psi(\sigma)$, where $\sigma \in \sum^+$, as shown earlier. We will denote this boundary with $\gamma_0(\sigma')$ and its length with $|\gamma_0(\sigma')|$. We will then have

$$p - q = \frac{1}{Z} \sum_{\sigma \in \Sigma^+} e^{-\mathcal{H}(\sigma)/k_B T} \left[1 - \sum_{\sigma' \in \Psi(\sigma)} e^{-2J|\gamma_0(\sigma')|/k_B T} \right]. \tag{5.50}$$

We now want to show that the sum $\sum$ over σ' that appears in the right-hand side is certainly smaller than 1 in the thermodynamic limit, if the temperature is low enough. We collect the configurations σ' according to the length $\ell = |\gamma_0(\sigma')|$ of the smallest boundary that contains the origin. Obviously, $\ell \geq 4$. Let us now evaluate the number $g(\ell)$ of configurations σ' that correspond to a certain value of ℓ. Let us assume that the boundary starts from a point i—the first segment can go in four directions, because the boundary cannot retrace its steps. If we neglect the constraint that the path closes, the total number of paths of ℓ length that leave from and return to i is certainly smaller than $4 \times 3^{\ell-1} = 4/3\ 3^\ell$. The starting point i, however, can be arbitrarily chosen among the ℓ boundary points, and the boundary itself can be traveled in two directions. The number of boundaries is therefore certainly smaller than $(4/3)(1/2\ell)3^\ell = 2/(3\ell)3^\ell$. Since the origin could be in any point inside the boundary, one needs to multiply by the area included in the boundary, which is certainly smaller than or equal to the area of a square whose side is $\ell/4$. We thus obtain

$$g(\ell) < \frac{4}{3\ell}3^\ell \left(\frac{\ell}{4}\right)^2 = \frac{\ell}{24}3^\ell. \tag{5.51}$$

One should note that in following this line of reasoning, we omitted the constraints imposed by the presence of other boundaries determined by the $\sigma \in \sum^+$ configuration, over which σ' is applied—we are therefore (significantly) overestimating the number of configurations σ'.

To conclude,

$$\sum = \sum_{\sigma' \in \Psi(\sigma)} e^{-2J|\gamma_0(\sigma')|/k_B T} < \sum_{\ell=4}^{\infty} \frac{\ell}{24} w^\ell, \tag{5.52}$$

where

$$w = 3e^{-2J/k_B T}. \tag{5.53}$$

Therefore,

$$\sum < \frac{1}{24} \sum_{n=2}^{\infty} (2n)\, w^{2n} = \frac{w^4 (2 - w^2)}{12 (1 - w^2)^2}. \tag{5.54}$$

When $w < w_0 = 0.869756\ldots$, the expression in the right-hand side is certainly smaller than 1. This occurs for temperatures below

$$T_0 = 2J/[k_B \ln(3/w_0)] \simeq 1.61531\, J/k_B. \tag{5.55}$$

We therefore see that $p - q > 0$ for temperatures below T_0, and that therefore spontaneous symmetry breaking occurs, and the corresponding critical temperature is higher than T_0. It is obviously possible to obtain a better bound on T_c with little effort. The exact result, for the two-dimensional Ising model, is $T_c = 2.269\, J/k_B$.

Exercise 5.1 Generalize the preceding result to $d = 3$ dimensions.

5.10 The One-Dimensional Ising Model

The preceding argument cannot be applied in one dimension. It is possible in effect to show that the Ising model in one dimension does not exhibit spontaneous symmetry breaking. The most convincing line of reasoning entails the exact solution of the model, but it is possible to show, by means of an argument due to Landau, that there cannot be discrete symmetry breaking in a one-dimensional system with short-range interactions. This also holds for continuous symmetry. In this section, we arrive at the exact solution of the one-dimensional Ising model, and we describe a (nonrigorous) version of Landau's argument.

5.10.1 Solution of the One-Dimensional Ising Model

The model is defined by the Hamiltonian:

$$\mathcal{H} = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1}, \tag{5.56}$$

where we have imposed periodic boundary conditions:

$$\sigma_{i+N} = \sigma_i, \quad \forall i. \tag{5.57}$$

The partition function is expressed by

$$Z = \sum_{\sigma} e^{-\mathcal{H}(\sigma)/k_B T} = \sum_{\{\sigma_1,\ldots,\sigma_N\}} \prod_{i=1}^{N} e^{K\sigma_i\sigma_{i+1}}, \tag{5.58}$$

where $K = J/k_B T$. By introducing the **transfer matrix** $\mathsf{T} = (T_{\sigma\sigma'})$, defined by

$$T_{\sigma\sigma'} = e^{K\sigma\sigma'}, \tag{5.59}$$

we see that the preceding expression assumes the form of a product of the matrices T, which, taking into account the boundary conditions that impose $\sigma_{n+1} = \sigma_1$, can be written

$$Z = \mathrm{Tr}\,\mathsf{T}^N. \tag{5.60}$$

The transfer matrix T has two eigenvalues, $t_+ = 2\cosh K$ and $t_- = 2\sinh K$, with $t_+ > t_-$. Therefore for $N \gg 1$,

$$\mathrm{Tr}\, \mathsf{T}^N = t_+^N + t_-^N \simeq t_+^N. \tag{5.61}$$

In the thermodynamic limit, therefore,

$$f = \lim_{N\to\infty} \frac{\ln Z}{N} = \ln 2\cosh K. \tag{5.62}$$

This expression does not exhibit singularities for $0 \le K < \infty$, and the model therefore does not exhibit phase transitions at finite temperature.

Exercise 5.2 Consider the spin correlation function

$$C_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle,$$

in the one-dimensional Ising model with a vanishing field. Show that one has

$$\lim_{N\to\infty} C_{ij} = (\tanh K)^{|i-j|}.$$

This expression decays as $e^{-|r_j - r_i|/\xi}$, where ξ is called the coherence length and is equal to

$$\xi = a_0 / |\ln \tanh K|,$$

where a_0 is the lattice step.

Exercise 5.3 Show that, in the presence of a nonvanishing magnetic field h, in the one-dimensional Ising model, one has

$$f = \lim_{N\to\infty} \frac{\ln Z}{N} = \ln\left[e^K \cosh \lambda + \sqrt{e^{2K} \sinh^2 \lambda + e^{-2K}} \right],$$

where $\lambda = h/k_{\mathrm{B}} T$.

Exercise 5.4 Obtain the same result with the following method, introduced by Lifson [Lifs64]. Let us consider the system with $\sigma_0 = +1$, and with σ_N free. Then (analogously with what was done with the Peierls argument), the spin configurations σ are univocally defined by the configurations Γ of the boundaries—in other words, of the pairs of nearest-neighbor spins with opposite values. The system is subdivided into n intervals, of length $\ell_i \ge 1$, such that sites belonging to the same interval have the same value as σ_i of the spin.

1. Write the canonical partition function as a function of the ℓ_i's.

2. Move to the corresponding grand canonical partition function by introducing the spin fugacity z.

3. Evaluate the grand canonical partition function as a function of z.

4. Evaluate $\langle N \rangle$ and prove that, in order to obtain the thermodynamic limit $N \to \infty$, it is necessary that $z \to z^*$, where z^* is the value of z closest to the origin for which Z admits a singularity. Evaluate z^*.

5. Show that $f = \lim_{N \to \infty} \ln Z/N = -\ln z^*$, and compare the result with what was already obtained.

5.10.2 The Landau Argument

At the temperature $T = 0$, a one-dimensional Ising system will be in one of the two states of minimal energy—for example, in the one in which all spins are aligned in the $+$ direction. The problem is whether this state can persist in the presence of thermal fluctuations at $T > 0$. Let us assume that we switch the direction of a certain number of consecutive spins—we will obtain an excited state, with an energy that is $2J$ above the minimum. The number of states that have this energy is proportional to $\ln N$, because we can arrange each boundary between $+$ and $-$ spins in N different fashions. Therefore, as soon as $k_B T \ln N > 2J$, reversing a macroscopic number of spins leads to an advantage in free energy. But for $N \to \infty$, this will be true for any positive value of T. This argument can immediately be generalized to all cases of short-range interaction.

Exercise 5.5 Let us suppose that the interaction is long range, so that the coupling constant J_{ij} between the i spin and the j spin decays as $|i-j|^{-a}$, where a is a positive constant. Show that Landau's argument is not valid if $a < 2$.

5.11 Duality

The critical temperature of the two-dimensional Ising model was found in 1941 by Kramiers and Wannier [Kram41] before the exact solution to the model, which is due to Onsager [Onsa44]. Kramers and Wannier's argument introduces a transformation—called **duality**—which transforms a high-temperature Ising model into an analogous low-temperature model. This transformation is of great importance, and it can be generalized to other statistical models, even though, in the large majority of cases, the transformed model is not identical to the initial one. We will deal with it only in the case of the Ising model in $d = 2$, by means of an argument that is simpler than the one initially formulated by Kramers and Wannier. This argument will also allow us to introduce the high- and low-temperature expansions, which allow us to obtain information about the behavior of models that cannot be solved exactly.

Let us therefore consider the expression of the partition function of a two-dimensional Ising model at temperature T in a vanishing external field:

$$Z(K) = \sum_{\sigma} \exp\left\{\sum_{\langle ij \rangle} K\sigma_i \sigma_j\right\}, \tag{5.63}$$

where the sum over σ runs over all the 2^N spin configurations and that over $\langle ij \rangle$ runs over the $2N$ pairs of nearest-neighbor sites. Since

$$e^{K\sigma\sigma'} = \cosh K + \sigma\sigma' \sinh K = \cosh K(1 + \sigma\sigma' \tanh K), \tag{5.64}$$

Z can also be written

$$Z(K) = (\cosh K)^{2N} \sum_\sigma \prod_{\langle ij \rangle} (1 + \sigma_i \sigma_j \tanh K). \tag{5.65}$$

We can now expand the expression on the right-hand-side into a power series of $t = \tanh K$. The terms of this expansion can be represented by diagrams—we associate a factor $t\sigma_i\sigma_j$ to each segment of the diagram that connects site i with site j (where i and j are nearest neighbors). For each pair $\langle ij \rangle$ of nearest-neighbor sites, there can be at most one $t\sigma_i\sigma_j$ factor. We thus obtain

$$Z(K) = (\cosh K)^{2N} \sum_\sigma \sum_{\mathcal{G}} t^{|\mathcal{G}|} \prod_i \sigma_i^{n_i}. \tag{5.66}$$

where the sum runs over all the $\mathcal{G}$ diagrams that can be drawn on the lattice, $|\mathcal{G}|$ is the number of bonds that appear in diagram $\mathcal{G}$, and for each site i, n_i is the number of bonds that contain the i site in diagram $\mathcal{G}$. When we sum all the spins' configurations, all the diagrams $\mathcal{G}$ in which n_i is odd for one or more i sites will result in a zero contribution; all the other diagrams will result in a contribution equal to $(\cosh K)^{2N} t^{|\mathcal{G}|}$. Therefore,

$$Z(K) = 2^N (\cosh K)^{2N} \sum_{\mathcal{G}} {}' t^{|\mathcal{G}|}. \tag{5.67}$$

where the sum runs over all the diagrams in which the number of bonds that contain each i site is even (in other words, $n_i = 0, 2, 4, \ldots, \forall i$).

These diagrams can be interpreted as the configuration Γ of boundaries between opposite spins in an Ising model defined on a **dual lattice** that results from the original one, and in which a spin variable is associated with each four-spin **plaquette** in the original lattice.

In figure 5.5, the original model is defined in the vertices of the lattice identified by dashed lines. A diagram $\mathcal{G}$ that appears in the expansion of the partition function of this model is drawn with a continuous line. This diagram can be interpreted as a configuration Γ of boundaries in a spin model defined over the dual lattice—the corresponding spin configuration associates an Ising variable with each plaquette.

We thus obtain

$$\sum_{\mathcal{G}} {}' = \sum_\Gamma t^{|\Gamma|}, \tag{5.68}$$

where $|\Gamma|$ is the total length of the boundaries that appear in Γ.

We now consider the expression of the partition function of the Ising model over the dual lattice, and we denote its temperature with T'. We impose boundary conditions $+$. One has
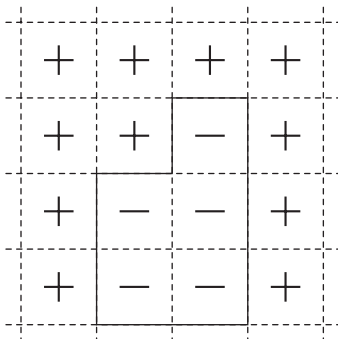

FIGURE 5.5. Diagram $\mathcal{G}$ in the high-temperature expansion of the Ising model interpreted as a configuration Γ of boundaries in an Ising model defined over the dual lattice.

$$Z(K') = e^{2NK'} \sum_{\Gamma} e^{-2K'|\Gamma|}, \tag{5.69}$$

where $K' = J/k_B T'$. By comparing this with the expression of $Z(K)$, we obtain the relation

$$t = \tanh K = e^{-2K'}. \tag{5.70}$$

This is the duality relation for temperatures. It is a duality relation in the sense that if we define $K'(K)$ so that it is satisfied, one gets $K'(K'(K)) = K$—in other words, the $K \longrightarrow K'$ transformation is equal to its inverse. Let us note that if $K \to 0$, one has $K' \to \infty$, and that duality therefore maps high-temperature systems in low-temperature systems, and vice versa. By making use of the relation (5.70), we obtain

$$Z(K') = e^{2NK'} \frac{1}{2^N} \frac{1}{(\cosh K)^N} Z(K) = \frac{1}{(\sinh 2K)^N} Z(K). \tag{5.71}$$

If the Ising model admits a single transition at $h = 0$, it must occur at a point in which $K' = K$. We thus obtain the critical value K_c:

$$\tanh K_c = e^{-2K_c}, \tag{5.72}$$

from which one obtains

$$K_c = \frac{1}{2} \ln\left(1 + \sqrt{2}\right) = 0.44069\ldots, \tag{5.73}$$

which corresponds to $k_B T_c \simeq 2.2692 J$.

Exercise 5.6 (Duality in the Three-Dimensional Ising Model) By considering the expansion of the partition function of the three-dimensional Ising model as a function of the Γ distribution of boundaries, show that the dual of this model is defined by the Hamiltonian

$$\mathcal{H} = -J \sum_{\mathcal{P}} \prod_{i \in \mathcal{P}} \sigma_i,$$

where the sum runs over all the plaquettes $\mathcal{P}$ composed of the dual lattice's four nearest-neighbor pairs of spin. Find the relation between the temperatures of the Ising model and the corresponding dual model.

Note that, since the dual of the Ising model is not identical to the initial model, this argument does not provide us with information about the critical temperature for $d = 3$.

5.12 Mean-Field Theory

In order to qualitatively calculate the phase diagram for these systems, one resorts to the **mean-field theory**. This approximated theory is extremely useful, and it is the first instrument one resorts to when exploring new models. Obviously, we will present it in the simplest case—in other words, in the Ising model.

Let us suppose that we mentally isolate a spin (which we will denote by a 0) from its environment, in the Ising model. If the spin's state changes, passing for instance from

$\sigma_0 = +1$ to $\sigma_0' = -1$, while all the rest of the environment remains constant, the system's energy varies by

$$\Delta H = 2h + 2 \sum_{i \in \text{p.v.}(0))} J\sigma_i = -\left(h + \sum_{i \in \text{p.v.}(0)} J\sigma_i\right)\Delta\sigma_0, \tag{5.74}$$

where $\Delta\sigma_0 = \sigma_0' - \sigma_0$, and the sum is extended to the nearest-neighbor sites of 0.

This is the same variation that one would obtain for a paramagnet immersed in a field whose value is equal to

$$h_{\text{eff}} = h + \sum_{i \in \text{p.v.}(0)} J\sigma_i. \tag{5.75}$$

The spins close to the one being considered fluctuate, however. As a first approximation, we can assume to describe their action on the spin 0 by this field's mean value—hence, the name **mean-field theory**. We thus obtain

$$h_{\text{eff}} = h + \sum_{i \in \text{p.v.}(0)} J\langle\sigma_i\rangle. \tag{5.76}$$

Let us calculate the mean value of σ for an isolated spin, subjected to an arbitrary external field h. One has

$$\langle\sigma\rangle = \frac{1}{z_0}\sum_\sigma \sigma e^{-h\sigma/k_B T}, \qquad z_0 = \sum_\sigma e^{-h\sigma/k_B T}. \tag{5.77}$$

A simple calculation shows that

$$\langle\sigma\rangle = \tanh\left(\frac{h}{k_B T}\right). \tag{5.78}$$

By exploiting this result, and substituting h with h_{eff}, we obtain

$$\langle\sigma_0\rangle = \tanh\left(\frac{h_{\text{eff}}}{k_B T}\right) = \tanh\left[\frac{1}{k_B T}\left(h + \sum_{i \in \text{p.v.}(0)} J\langle\sigma_i\rangle\right)\right]. \tag{5.79}$$

We can now observe that in the ordered phase, we expect that the mean value of $\langle\sigma_i\rangle$ be the same for all spins. This observation is not as trivial as it seems—if in fact we were to change J's sign, all the observations we made so far would be just as valid, except for this one, because we would instead expect that $\langle\sigma_i\rangle$ assume opposite signs on each of the intertwined sublattices. More generally, the most difficult point of the mean-field theory is that one needs to make a good guess about the type of order one expects. In our case, we posit

$$\langle\sigma_i\rangle = m, \qquad \forall i, \tag{5.80}$$

from which we obtain an equation for m:

$$m = \tanh\left(\frac{h + zJm}{k_B T}\right), \tag{5.81}$$

where ζ is the lattice's coordination number—in other words, the number of nearest-neighbors of a given site.

It is easier to discuss this equation in the form

$$\tanh^{-1} m = \frac{h + zJm}{k_B T}.$$ (5.82)

This equation can be solved graphically. Let us consider figure 5.6, in which the curves defined by

$$y = \tanh^{-1} x,$$ (5.83)

$$y = \left(\frac{h}{k_B T}\right) + \left(\frac{zJ}{k_B T}\right) x.$$ (5.84)

have been traced.

The solution corresponds to the intersection of these two curves. The derivative of $\tanh^{-1} m$ assumes its minimum value in the origin, where it is equal to 1. Therefore, if $k_B T > \zeta J$, the curve and the straight line have only one intersection, and there cannot be finite magnetization (in other words, coexistence). Instead, if $k_B T < \zeta J$, and the external field h is small enough, there will be *three* solutions.
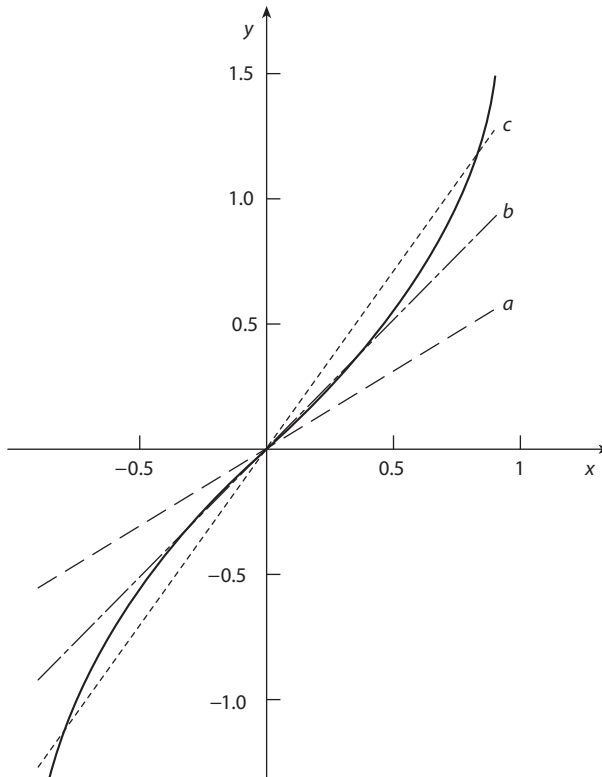

FIGURE 5.6. Graphical solution of the mean-field equation. Continuous line: equation $y = \tanh^{-1} x$. Dashed lines: $y = \zeta J/k_B T$. (a) $T > T_c = \zeta J/k_B$; (b) $T = T_c$; and (c) $T < T_c$.
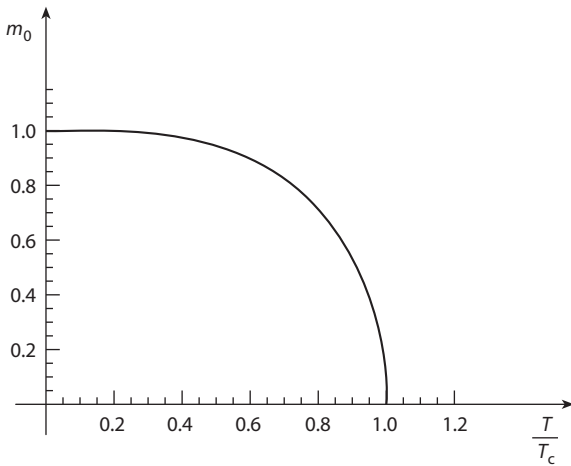
FIGURE 5.7. Spontaneous magnetization for spin m_0 in an Ising model, as a function of $T/T_c = k_B T/\zeta J$, in the mean-field approximation.

Which one is the "physical" solution? Thermodynamics tells us that we must choose the one in which the free energy $E - TS - hM$ is minimal. In our case, we are consistent if we estimate the entropy S from the corresponding valid expression for independent spins:

$$S = -Nk_B \left[\left(\frac{1-m}{2} \right) \ln \left(\frac{1-m}{2} \right) + \left(\frac{1+m}{2} \right) \ln \left(\frac{1+m}{2} \right) \right]. \tag{5.85}$$

By making use of this expression, it is easy to see that the "physical" solution is the one in which m assumes the maximum value (in modulus) of the same sign as the external magnetic field h. We thus obtain the spontaneous magnetization as a function of T, represented by the curve shown in figure 5.7.

In this curve, we have reproduced m as a function of T/T_c, where $k_B T = \zeta J$. We can also calculate the specific heat with a vanishing external field:

$$
\begin{aligned}
C &= T \frac{\partial S}{\partial T}\bigg)_h = T \frac{dS}{dm} \frac{\partial m}{\partial T}\bigg)_h = \\
&= \begin{cases} Nk_B \left(zJm/k_B T \right)^2 \left[1/(1-m)^2 - zJ/k_B T \right]^{-1}, & \text{for } T < T_c, \\ 0, & \text{for } T > T_c. \end{cases}
\end{aligned}
\tag{5.86}
$$

One should note the specific heat's discontinuity at the critical temperature—in our approximation, the specific heat vanishes above the critical temperature. The phase diagram is particularly simple in the (h, T) plane. The coexistence curve is simply the segment $(0, 0) - (0, T_c)$. The magnetization m is subject to a discontinuity along this segment, going from positive h's to negative h's.

5.13 Variational Principle

The reasoning we reproduced in the preceding section someow lacks consistency. Once we had introduced the concept of mean field, we saw that there can be several thermodynamic

states (defined by different values of magnetization by spin m) that are solutions of the corresponding equations. To resolve this ambiguity, we recalled the variational principle of free energy, approximating entropy with its expression as obtained for the paramagnet.

In this section, we prove that mean-field theory can be obtained from a variational principle, which allows us to derive both the self-consistency equations we described and the variational principle of free energy in the form we utilized it. This method also allows us to clarify the hypotheses that provide the foundation of the method we used.

The starting point is an inequality valid for each real function $f(x)$ and for each probability distribution:

$$\langle \exp f(x) \rangle \geq \exp \langle f(x) \rangle. \tag{5.87}$$

This inequality is a consequence of the concavity of the exponential function. One can prove this relation by setting

$$\langle e^{f} \rangle = e^{\langle f \rangle} \langle e^{f - \langle f \rangle} \rangle,$$

and using the well-known inequality [equivalent to $x \geq \ln(1 + x)$]

$$e^{x} \geq 1 + x.$$

One thus obtains

$$\langle e^{f} \rangle \geq e^{\langle f \rangle} \langle 1 + f - \langle f \rangle \rangle = e^{\langle f \rangle}.$$

Let us consider, for instance, a variable x that can assume only two values: x_1 with probability p and x_2 with probability $(1 - p)$. As p varies between 0 and 1, $\langle x \rangle$ can assume any value between x_1 and x_2, and $e^{\langle x \rangle}$ will be the value of the corresponding ordinate in the plot of $\exp(x)$ (see figure 5.8). On the other hand, the value of $\langle e^{x} \rangle$ will be the value of the corresponding ordinate in the plot of the linear function $pe^{x_1} + (1 - p)e^{x_2}$. Since this exponential function is concave, the plot of this function always remains above the plot of $\exp(x)$ in the interval (x_1, x_2).

This inequality suggests that we can use a variational principle when calculating the free energy. Let us introduce a "trial Hamiltonian" H_0 (arbitrary for the time being). We can then write

$$
\begin{aligned}
Z &= \sum_{\sigma} \exp\left[-\frac{H(\sigma)}{k_B T} \right] \\
&= Z_0 \left\langle \exp\left[-\frac{H(\sigma) - H_0(\sigma)}{k_B T} \right] \right\rangle_0 \\
&\geq Z_0 \exp\left[-\left\langle \frac{H(\sigma) - H_0(\sigma)}{k_B T} \right\rangle_0 \right],
\end{aligned}
\tag{5.88}
$$

where we have defined

$$Z_0 = \sum_{\sigma} \exp\left[-\frac{H_0(\sigma)}{k_B T} \right] = \exp\left[-\frac{(\langle H \rangle_0 - TS_0)}{k_B T} \right], \tag{5.89}$$

$$\langle A(\sigma) \rangle_0 = \frac{1}{Z_0} \sum_{\sigma} A(\sigma) \exp\left[-\frac{H_0(\sigma)}{k_B T} \right]. \tag{5.90}$$
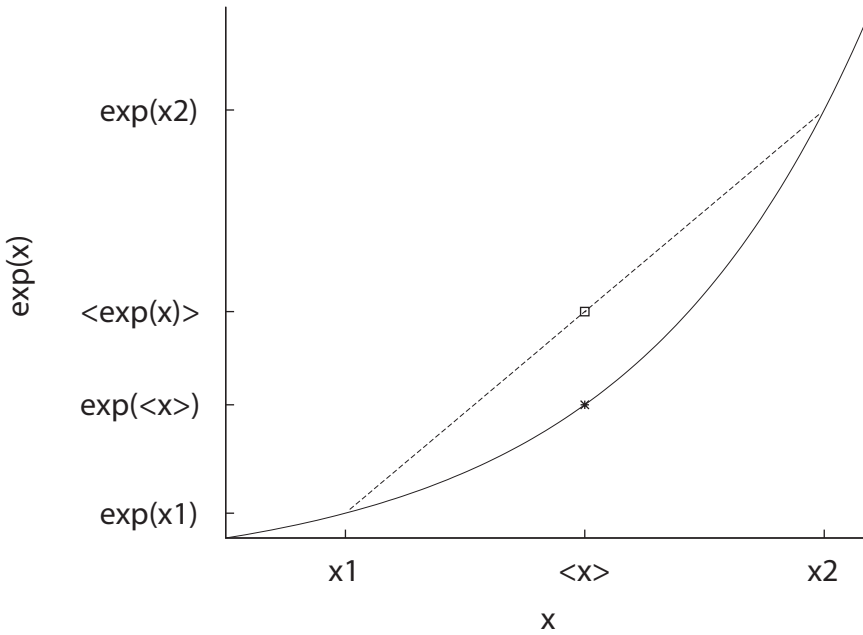
FIGURE 5.8. Concavity of the exponential function. The dashed line represents $\langle \exp x \rangle$ and lies always above the continous curve representing $\exp\langle x \rangle$, when x can take one of two values.

The idea is therefore to look for the maximum of the expression

$$\mathcal{Z} = Z_0 \exp\left[-\left\langle \frac{H - H_0}{k_B T} \right\rangle_0\right], \tag{5.91}$$

with respect to the possible choices of the trial Hamiltonian H_0. The easiest choice is that of a Hamiltonian with independent spins:

$$H_0 = -\sum_i \lambda \sigma_i. \tag{5.92}$$

Let us now shift to the logarithms, introducing the true free energy $F = -k_B T \ln Z$ and the trial free energy $\mathcal{F} = -k_B T \ln \mathcal{Z}$. The variational principle then takes the form

$$F \leq \mathcal{F}(T, h; \lambda) \equiv \langle H_0 \rangle - T S_0, \tag{5.93}$$

where the internal energy $\langle H \rangle_0$ and the entropy S_0 are evaluated by means of the trial Hamiltonian. Since there is a one-to-one correspondence between the parameter λ and the magnetization (in this simple case!), we can choose m rather than λ as variational parameter. We thus obtain

$$H_0 = -\frac{NzJ}{2} m^2 - Nhm; \tag{5.94}$$

$$S_0 = -Nk_B\left[\left(\frac{1-m}{2}\right)\ln\left(\frac{1-m}{2}\right)\right.$$
$$\left. + \left(\frac{1+m}{2}\right)\ln\left(\frac{1+m}{2}\right)\right]. \tag{5.95}$$

The variational equations are

$$zJm + h = k_\mathrm{B}T \tanh^{-1}m, \tag{5.96}$$

which is what we had obtained in the preceding section. This time, however, we additionally obtain the criterion by which to choose this equation's solutions and, even more importantly, the way to generalize it—it will be sufficient to choose some trial Hamiltonians H_0 that are always obtained as sums of Hamiltonians with a few variables independent from the others (so as to allow the calculation of the Z_0's and of the averages $\langle A \rangle_0$'s), but that contain a part of the spin interactions. These generalizations of mean-field theory are called **cluster approximations**.

5.14 Correlation Functions

In the Ising model, the spin variables σ_i relative to different sites are not independent. A measure of their dependence is given by the **correlation function**:

$$C_{ij} \equiv \langle \sigma_i \sigma_j \rangle_c \equiv \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle. \tag{5.97}$$

This quantity vanishes if σ_i and σ_j are independent.

In this section, we will calculate the C_{ij} in the context of mean-field theory, using a method due to Ornstein and Zernike [Orns14]. In order to do so, let us suppose that we rewrite the partition function, expressing it as a function of an interaction matrix K_{ij}, which is a priori arbitrary (but symmetrical), and of an external field λ_i, which can vary from point to point:

$$Z = \exp\left[\sum_{ij} K_{ij}\sigma_i\sigma_j + \sum_i \lambda_i\sigma_i\right]. \tag{5.98}$$

The physical case corresponds to

$$K_{ij} = \begin{cases} J/k_\mathrm{B}T, & \text{if } i \text{ and } j \text{ are nearest neighbors,} \\ 0, & \text{otherwise} \end{cases} \tag{5.99}$$

and $\lambda_i = h/k_\mathrm{B}T$.

We observe that

$$\left.\frac{\partial \ln Z}{\partial \lambda_i}\right|_{\mathrm{phys}} = \langle \sigma_i \rangle = m_i, \tag{5.100}$$

$$\left.\frac{\partial^2 \ln Z}{\partial \lambda_i \partial \lambda_j}\right|_{\mathrm{phys}} = \langle \sigma_i \sigma_j \rangle_c. \tag{5.101}$$

The *phys* notation indicates that the derivatives must be calculated with the physical values of K and λ reproduced in equation (5.99) and the line following it. Let us now try to calculate these quantities, using the mean-field expression of the partition function that we derived earlier. We have

$$\ln Z = -S_0/k_B + \sum_{ij} K_{ij} m_i m_j + \sum_i \lambda_i m i, \tag{5.102}$$

where

$$S_0 = -k_B \sum_i \left[\left(\frac{1-m_i}{2} \right) \ln \left(\frac{1-m_i}{2} \right) + \left(\frac{1+m_i}{2} \right) \ln \left(\frac{1+m_i}{2} \right) \right]. \tag{5.103}$$

Taking the derivative of this expression with respect to λ_i, we obtain

$$\frac{\partial \ln Z}{\partial \lambda_i} = \sum_j \frac{\partial \ln Z}{\partial m_j} \bigg|_{\text{phys}} \frac{\partial m_j}{\partial \lambda_i} + m_i, \tag{5.104}$$

where the local magnetization m_i is given by the mean-field equation:

$$0 = \frac{\partial \ln Z}{\partial m_i} \bigg|_{\text{phys}} = -\tanh^{-1} m_i + \sum_{j \in \text{p.v.}(i)} K_{ij} m_j + \lambda_i \bigg|_{\text{phys}}. \tag{5.105}$$

By taking the derivative of this equation with respect to λ_j, we obtain an equation for the correlation function $C_{ij} = \partial m_i / \partial \lambda_j$:

$$\frac{1}{1 - m_i^2} C_{ij} = \sum_\ell K_{i\ell} C_{\ell j} + \delta_{ij}. \tag{5.106}$$

In order to solve this equation, it is useful to transform it according to Fourier, taking into account the fact that both K_{ij} and C_{ij} depend only on the vector $r_{ij} = r_j - r_i$, the distance between sites i and j.

We define the Fourier transform according to

$$C(k) = \sum_j \exp(i k \cdot r_{ij}) C_{ij} \tag{5.107}$$

and analogously for K. Equation (5.106) becomes

$$C(k) = (1 - m^2)[K(k) C(k) + 1], \tag{5.108}$$

and one therefore has

$$C(k) = \frac{1 - m^2}{1 - (1 - m^2) K(k)}. \tag{5.109}$$

We can make this expression explicit by keeping in mind that for a simple cubic lattice in d dimensions, one has

$$K(k) = \sum_{j \in \text{p.v.}(i)} \frac{J}{k_B T} \exp(i k \cdot r_{ij}) = \sum_{\alpha=1}^d \frac{J}{k_B T} 2 \cos(k_\alpha a_0), \tag{5.110}$$

where a_0 is the lattice step. For small values of $|k|$, one has

$$K(k) \simeq \frac{J}{k_B T} (\zeta - k^2 a_0^2), \tag{5.111}$$

where $\zeta = 2d$ is the lattice's coordination number. We thus obtain

$$
\begin{aligned}
C(k) &= (1-m^2)\left[1-(1-m^2)\,(J/k_{\mathrm{B}}T)\sum_\alpha 2\cos{(k_\alpha a_0)}\right]^{-1} \\
&\simeq (1-m^2)\left[1-(1-m^2)\frac{T_c}{T}\left(1-k^2 a_0^2\right)\right]^{-1},
\end{aligned}
\tag{5.112}
$$

in which we have accounted for the fact that the transition temperature is given by $T_c = \zeta J/k_B$.

Let us remark that the Fourier transform of the correlation function, evaluated at $k = 0$, is proportional to the magnetic susceptibility per spin. In fact, one has

$$
C(k=0) = \sum_j \frac{\partial\langle\sigma_i\rangle}{\partial\lambda_j} = k_{\mathrm{B}}T\,\frac{\partial\langle\sigma_i\rangle}{\partial h} = k_{\mathrm{B}}T\chi.
\tag{5.113}
$$

Calculating this expression for $T > T_c$ (and therefore for $m = 0$), we obtain

$$
\chi = \frac{1}{k_{\mathrm{B}}(T-T_c)},
\tag{5.114}
$$

which diverges when $T\to T_c$. This law is called the **Curie-Weiss law**. On the other hand, it is possible to introduce a length ξ that measures the range of the correlations. Let us set

$$
\xi^2 = \frac{1}{2d}\frac{1}{C(k=0)}\sum_j r_{ij}^2 C_{ij}.
\tag{5.115}
$$

where the $1/2d$ factor has been introduced for convenience's sake. The quantity ξ is called the **coherence** or **correlation length**. It is easy to see that

$$
\xi^2 = \sum_\alpha \left.\frac{\partial\ln C(k)}{\partial k_\alpha^2}\right|_{k=0} = \frac{Ta_0^2}{2(T-T_c)} \propto (T-T_c)^{-1}.
\tag{5.116}
$$

We thus obtain the important result that the correlations' range diverges when $T\to T_c$. This result is not unexpected, given that the divergence of the per spin susceptibility implies that the number of terms different from zero in the expression of the $C(k=0) = \sum C_{ij}$ must diverge, since each term is limited.

It is useful to give an approximate expression of the correlation function in the r space that corresponds to the expressions we derived in the k space. Let us therefore assume that the correlation function in k space is given by the Ornstein-Zernike formula:

$$
C(k) \propto \frac{1}{k^2 + \xi^{-2}}.
\tag{5.117}
$$

The Fourier antitransform of this expression is given by

$$
C(r) \propto \int d^d k\,\frac{e^{-i k\cdot r}}{k^2 + \xi^{-2}}.
\tag{5.118}
$$

In order to evaluate this expression, let us introduce the identity

$$
\frac{1}{x} = \int_0^\infty du\,e^{-ux}, \qquad x > 0.
\tag{5.119}
$$

We obtain

$$C(r) \propto \int d^d k \int_0^\infty du\, e^{-u(k^2 + \xi^{-2}) - ik\cdot r}.$$ (5.120)

We can now evaluate the Gaussian integral over the k's and obtain

$$C(r) \propto \int_0^\infty du\, u^{-d/2} e^{-u\xi^{-2} - r^2/(4u)}.$$ (5.121)

Let us change integration variable, setting

$$w = ur^{-2}.$$ (5.122)

We obtain

$$C(r) \propto r^{2-d} \int_0^\infty dw\, w^{-d/2} \exp\left[-w\left(\frac{r}{\xi}\right)^2 - \frac{1}{4w}\right].$$ (5.123)

If $r \ll \xi$, we can disregard r's dependence in this integral (at least if $d > 2$; otherwise, the integral diverges for large w's). In this manner, we obtain $C(r) \sim r^{2-d}$. If instead $r \gg \xi$, we can evaluate the integral with the saddle-point method, which one has for $w = w_c = \xi/2r$. We thus obtain

$$C(r) \sim e^{-r/\xi}.$$ (5.124)

These two behaviors can be summed up by the approximate formula

$$C(r) \sim \frac{e^{r/\xi}}{r^{d-2}},$$ (5.125)

which is exact for $d = 3$.

5.15 The Landau Theory

The mean-field theory allows us to obtain an approximate expression for the phase diagram of the system being considered. If, however, we want to aim for the more limited goal of describing the system's behavior in proximity to the critical point, we can resort to a theory that is formally simpler and more easily generalizable—Landau theory.

The mean-field theory requires that we look for the minimum of an effective free energy, defined by

$$\mathcal{F}(h, T; m) = N[f(T; m) - hm],$$ (5.126)

expressed as a function of the magnetization per spin m. The trial free energy per spin, expressed as a function of m, has the form

$$\begin{aligned} f(T; m) &= -\frac{zJ}{2}m^2 - Ts(m) \\ &= -\frac{zJ}{2}m^2 + k_B T\left[\left(\frac{1-m}{2}\right)\ln\left(\frac{1-m}{2}\right)\right. \\ &\quad \left. +\left(\frac{1+m}{2}\right)\ln\left(\frac{1+m}{2}\right)\right]. \end{aligned}$$ (5.127)

In proximity of the critical point, the equilibrium value of the order parameter m is small, and we can therefore expand this function into a power series:

$$f(T; m) = f(T; 0) = \frac{1}{2} a(T) m^2 + \frac{1}{4!} b(T) m^4 + \ldots \tag{5.128}$$

Only the even powers of m appear in this expansion, because the Hamiltonian is symmetrical with respect to the transformation $m \to -m$. It is obviously possible to calculate the coefficients $a(T)$, $b(T)$, and so on explicitly. We have

$$f(T; 0) = k_B T \ln 2, \tag{5.129}$$

$$a(T) = k_B T - zJ \propto (T - T_c), \tag{5.130}$$

$$b(T) = 10 k_B T. \tag{5.131}$$

We see that the coefficient a changes sign at the critical temperature, while the coefficient b remains positive. In effect, when we look for $\mathcal{F}$'s minimum with respect to m, we expect to find a vanishing solution for $T > T_c$ and two opposite minima, different from zero, for $T < T_c$. The behavior of $f(T; m)$ for small values of m, and for temperatures close to T_c, is shown in figure 5.9.

Therefore, continuing to follow our goal of studying T_c's neighborhood, we can set:

$$a(T) \simeq a'(T - T_c), \qquad a' > 0, \tag{5.132}$$

$$b(T) \simeq b(T_c) = b > 0. \tag{5.133}$$

The equations for $\mathcal{F}$'s minimum take the form

$$a'(T - T_c) m + \frac{b}{3!} m^3 = h. \tag{5.134}$$

For $T > T_c$ we have a real solution:

$$m \simeq \frac{1}{a'(T - T_c)} h = \chi h, \tag{5.135}$$
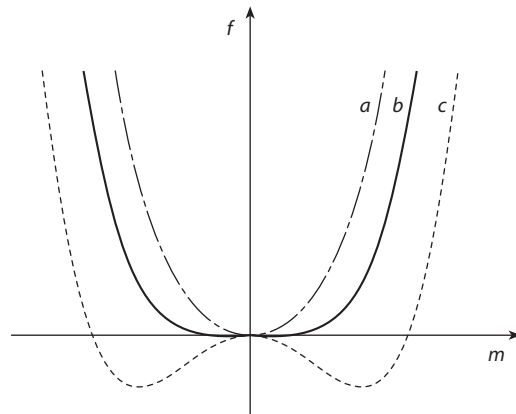
which confirms the result we had already arrived at:

FIGURE 5.9. Landau free energy $f = f(T; m)$ as a function of magnetization m near the critical temperature. (a) $T > T_c$; (b) $T = T_c$; and (c) $T < T_c$.

$$\chi \propto (T - T_c)^{-1}. \tag{5.136}$$

On the other hand, for $T < T_c$, we have *three* solutions—

$$m \simeq \frac{1}{a'(T - T_c)} h, \tag{5.137}$$

but there is one which corresponds to a *maximum* of $\mathcal{F}$ (as is also obvious from figure 5.9). The other two solutions correspond to minima and are given approximately by

$$\frac{b}{3!} m^2 = a' |T - T_c|, \tag{5.138}$$

with additional contributions that vanish for $h \to 0$. We thus obtain the behavior for the spontaneous magnetization close to T_c:

$$m_0(T) = \lim_{h \to 0^+} m(T, h) = \left(\frac{3! a' |T - T_c|}{b} \right)^{1/2} \propto |T - T_c|^{1/2}. \tag{5.139}$$

It is also interesting to calculate m's behavior as a function of h exactly at the critical temperature—we have

$$\frac{b}{3!} m^3 = h, \tag{5.140}$$

and as result, the magnetization is proportional to $h^{1/3}$.

One can also calculate the specific heat. In fact, since

$$S = -\frac{\partial F}{\partial T}, \tag{5.141}$$

one obtains

$$C_h = -\frac{\partial^2 F}{\partial T^2}\bigg|_h = \begin{cases} 0, & \text{for } T > T_c, \\ \left(\frac{3}{4} a'\right)^2 & \text{for } T < T_c. \end{cases} \tag{5.142}$$

The specific heat is therefore discontinuous at the transition.

All the results for the critical behavior of thermodynamic quantities derive from the hypothesis that the thermodynamic value of the order parameter can be obtained by means of a variational principle, in which a function $f(T; m)$ appears that is (1) analytic and (2) symmetrical with respect to a certain transformation group of m. In our case, this group is composed of the identity and the inversion $m \to -m$.

One can consider more general cases, in which the symmetry group is more complicated. In this case, in the expansion of $f(T; m)$, only the invariants that can be built starting from m can appear. Let us suppose for example that the order parameter is a vector $\mathbf{m} = (m_\alpha)$ in n dimensions and that it transforms according to the rotation group in n dimensions, $O(n)$. The only invariant we can construct is $m^2 = \sum_\alpha m_\alpha^2$. Therefore, $f(T; m)$ is in fact only a function of m^2, and its expansion into a Taylor series has the same form as in the Ising model. The results we obtained are therefore still essentially valid.[1]

---

[1] There are, however, some effects related to the vectorial nature of the order parameter, such as the existence of a transverse susceptibility, that diverges for $h \to 0$ below T_c.

On the other hand, in more complicated situations, there can be cubic invariants, or invariants proportional to higher odd powers of m. When this happens, the Landau theory predicts that the transition becomes discontinuous. In effect, exactly at the critical temperature, the f has to admit a minimum in $m = 0$, but this is impossible if the term proportional to m^3 does not vanish. Cubic invariants of this type are always present at the fluid–solid transition—therefore, the order parameter of the fluid–solid transition cannot vanish continuously (at least within the context of mean-field theory).

The analyticity of f is due to the very nature of mean-field approximation—the effect of the rest of the sample on a degree of freedom (or on a small number of degrees of freedom) is represented by effective interaction parameters, such as the "effective" magnetic field λ. The trial free energy is then obtained by calculating the partition sum over these few degrees of freedom—it is obvious that this procedure cannot produce nonanalytic dependencies in the order parameter.

It is possible, in principle, to improve the mean-field approximation by considering a larger, but finite, number of degrees of freedom—this improvement, however, cannot modify those results of the mean-field theory that depend only on the analyticity of the trial free energy.

5.16 Critical Exponents

We have thus obtained the behaviors of the various thermodynamic quantities in T_c's neighborhood. These behaviors are described by power-laws, whose exponents are called **critical exponents**. The results we obtained are summarized in table 5.1. These values of the critical exponents follow from the analyticity hypothesis and are called the **classical values**. Given the generality of the hypotheses underlying Landau theory, it was a surprise to see that the experimentally observed values of the critical exponents are not the same as the classical ones—physical systems can be grouped into broad classes, called **universality classes**, within which the exponents' values are constant, and it is actually possible to map the critical behavior of one system into that of another by means of some simple transformations.

Table 5.1   Classical Critical Exponents

Quantity	Behavior	Region	Exponent
C_h	$|T - T_c|^{-\alpha}$	$h = 0$	$\alpha = 0$ (discont.)
χ_T	$|T - T_c|^{-\gamma}$	$h = 0$	$\gamma = 1$
M	$|T - T_c|^{-\beta}$	$h = 0$	$\beta = 1/2$
	$h^{1/\delta}$	$T = T_c$	$\delta = 3$
ξ	$|T - T_c|^{-\nu}$	$h = 0$	$\nu = 1/2$

Two challenges remain:

1. Identify the reason for the Landau theory's lack of success.

2. Identify the reason for the existence of the universality classes, and if possible, provide a method for calculating the critical behavior within each class.

Let us observe that the hypothesis that the order parameter has a well-specified value, obtained from free energy's minimum condition, corresponds to the hypothesis that it is possible to disregard its fluctuations. We will see further on that this allows us to define a criterion (called the **Ginzburg criterion**) to evaluate the region of validity for mean-field theory. In order to make this criterion explicit, however, it is necessary to describe the behavior of the critical fluctuations.

5.17 The Einstein Theory of Fluctuations

In order to understand the critical behavior, it is necessary to treat the fluctuations of thermodynamic quantities in small, but macroscopic, systems. This theory of fluctuations is essentially due to Einstein.

In order to clarify our ideas, let us consider a system described by the Ising model, at a given value of the temperature T and magnetic field h. Within this system, let us focus our attention on a small but macroscopic region S. We want to evaluate the probability that the energy and the magnetizations of the spins contained in this region have respectively the values E_S and M_S.

To evaluate this probability, let us consider those microstates ν_S of the subsystem S such that the internal energy of the subsystem is equal to E_S and its magnetization is equal to M_S. The probability that the subsystem is in any microstate of this kind is proportional to the Boltzmann factor $\exp[-(E_S - hM_S)/k_B T]$, with the given values of T and h. Thus, the probability that the subsystem S has these values of internal energy and magnetization is obtained by summing this probability over all the microstates satisfying the condition just mentioned:

$$
\begin{aligned}
p(E_S, M_S) \;=\; & \frac{1}{Z}\sum_{\nu_S} \delta[E_S - E(\nu_S)]\delta[M_S - M(\nu_S)] \\
& \times \exp\left(-\frac{E_S - hM_S}{k_B T}\right).
\end{aligned}
\tag{5.143}
$$

In this expression, Z is the partition function, which imposes the normalization condition on the probability distribution of the microstates of S, and the delta functions impose the condition on the values of E_S and M_S.

The sum over all microscopic states gives the exponential of the entropy, expressed as a function of the internal energy and the magnetization:

$$
\sum_{\nu_S} \delta[E_S - E(\nu_S)]\delta[M_S - M(\nu_S)] = \exp\left[\frac{S_S(E_S, M_S)}{k_B}\right].
\tag{5.144}
$$

Therefore,

$$
\begin{aligned}
p(E_S, M_S) &= \frac{1}{Z} \exp\left[\frac{E_S - TS_S(E_S, M_S) - hM_S}{k_B T}\right] \\
&= \exp\left[-\frac{\Delta\mathcal{F}(E_S, M_S; T, h)}{k_B T}\right],
\end{aligned}
\tag{5.145}
$$

where

$$
\Delta\mathcal{F}(E_S, M_S; T, h) = E_S - TS_S(E_S, M_S) - hM_S - F(T, h),
\tag{5.146}
$$

is the **availability**. A normalization constant is understood in this expression, and $F(T, h) = k_B T \ln Z$ is the Helmholtz free energy, expressed as a function of the temperature and magnetic field. Note that the availability vanishes for the equilibrium values of energy and magnetization by virtue of the variational principle of the Legendre transform. Availability is often called *free energy difference*—let us remember, however, that unlike actual free energy, it is a function both of intensive variables (in our case, the magnetic field and the temperature) and of fluctuating extensive variables.

Let us observe that we have obtained an apparently nonnormalized probability distribution. Indeed, the quantities that appear explicitly in equation (5.146) are extensive, and are defined up to subextensive corrections. One can easily check that, since the normalization factor behaves like a power of the subsystem's size, imposing the normalization also leads to subextensive corrections, which are therefore negligible.

If, as usually occurs, entropy is a regular function of its arguments around the equilibrium values, then it is possible to deduce a certain number of important relations for the fluctuations.

Let us first consider a system enclosed by nonadiabatic walls, in which the internal energy can fluctuate. We obtain

$$
\begin{aligned}
\langle \Delta E^2 \rangle &= -k_B \left.\frac{\partial^2 S}{\partial E^2}\right|_V^{-1} = -k_B \left.\frac{\partial(1/T)}{\partial E}\right|_V^{-1} \\
&= k_B T^2 \left.\frac{\partial E}{\partial T}\right|_V = k_B T^2 C_V.
\end{aligned}
\tag{5.147}
$$

This result obviously coincides with the one obtained for the fluctuations in the canonical ensemble.

Let us now focus our attention on a certain collection (X_i), $i = 1, 2, \ldots, r$ of *extensive* quantities. Let us denote the relative equilibrium values with X_i^0. The first nonvanishing term of the Taylor expansion of availability will be the second one:

$$
\frac{\Delta\mathcal{F}}{k_B T} \simeq \frac{1}{2k_B} \sum_{ij} \frac{\partial^2 S}{\partial X_i \partial X_j} \Delta X_i \Delta X_j,
\tag{5.148}
$$

where $\Delta X_i = X_i - X_i^0$. The linear terms of the expansion vanish due to the variational principle of the Legendre transform. Because of entropy's convexity, this quadratic form must be positive-semidefinite. More particularly, if it is positive-definite, and if it is possible to disregard the further terms of the Taylor expansion, we deduce from equation (5.148) that

the distribution of fluctuations of the (X_i)'s is a Gaussian. The correlation matrix is the inverse of the one that appears in this formula:

$$C_{ij}^{-1} = -\frac{1}{k_B}\frac{\partial^2 S}{\partial X_i \partial X_j^0} = \frac{1}{k_B}\frac{\partial(f_j/T)}{\partial X_i^0}\bigg)_X. \tag{5.149}$$

We have made explicit the fact that the derivative is taken by keeping the average values of the other extensive variables constant. This result coincides with what can be obtained by considering the corresponding generalized ensemble:

$$C_{ij} = \langle \Delta X_i \Delta X_j \rangle = \frac{\partial^2 \ln Z}{\partial(f_i/k_B T)\,\partial(f_j/k_B T)} = k_B \frac{\partial X_i^0}{\partial(f_j/T)}\bigg)_{f/T}. \tag{5.150}$$

In this case, the derivative is taken by keeping the values of the ratios f/T constant. Let me remind you that the force conjugated with internal energy is $1/T$. These expressions are equivalent, since in effect one has

$$\sum_\ell \frac{\partial X_i^0}{\partial(f_\ell/T)}\bigg)_{f/T}\frac{\partial(f_\ell/T)}{\partial X_j^0}\bigg)_X = \delta_{ij}. \tag{5.151}$$

This approach to the statistics of fluctuations is not limited to just the usual thermodynamic quantities. More specifically, it can be generalized to the study of the spatial distribution of magnetization density.

We can, for instance, evaluate the probability that a certain Fourier component of magnetization density has a given value ϕ_k. The component of the magnetization density is an observable, because it is expressed by

$$\phi_k = \sum_j e^{-ik\cdot r_i}\sigma_j. \tag{5.152}$$

This probability is proportional to $\exp(-\Delta\mathcal{F}(\phi_k)/k_B T)$, where, for small fluctuations, one has

$$\frac{\Delta\mathcal{F}(\phi_k)}{k_B T} \simeq \frac{1}{2N}\frac{|\phi_k|^2}{C(k)}. \tag{5.153}$$

In this expression, $C(k)$ is the Fourier transform of the correlation function:

$$C(k) = \frac{1}{N}\langle|\phi_k|^2\rangle = \sum_j e^{ik\cdot r_j}\langle\sigma_j\sigma_j\rangle_c. \tag{5.154}$$

For small values of $|k|$, one can approximate $C(k)$ by means of its Ornstein-Zernike expression:

$$C(k) = \frac{k_B T\chi}{1 + k^2\xi^2}, \tag{5.155}$$

where ξ is the coherence length, and χ the susceptibility per spin. The expressions of χ and ξ are not necessarily given correctly by mean-field theory. In order for this expression to be valid, it is sufficient that $C^{-1}(k)$ can be expanded into a Taylor series as a function of k^2—in other words, basically that the correlations be of finite range.

We have so far considered the fluctuations of extensive quantities in a finite region of the system, which are in contact with a much larger system, and therefore characterized by well-defined values of the corresponding intensive quantities (generalized forces and temperature). If we consider an isolated system, we can ask ourselves whether it is possible to observe analogous fluctuations in the intensive quantities. This problem is still the subject of discussion. One can formally associate corresponding fluctuations of the intensive variables with the fluctuations of the extensive variables by means of the equations of state. This procedure is fairly arbitrary however and, above all, does not allow one to obtain further physical insights.

5.18 Ginzburg Criterion

Let us now suppose that we are considering the fluctuations of the magnetization as we approach the critical temperature. We can estimate the availability by using the results of mean-field theory (or, equivalently, of the Landau theory):

$$\Delta\mathcal{F} \simeq \frac{N}{2} a'(T - T_c) m^2 + \frac{Nb}{4!} m^4. \tag{5.156}$$

As long as $T > T_c$, the first term dominates the second one when $N \to \infty$. The magnetization fluctuations are therefore Gaussian and $O(N^{-1/2})$. The variance diverges as we get closer to the critical temperature, and exactly at T_c, the fluctuations are no longer Gaussian:

$$p(m) \propto \exp\left(-\frac{Nb}{4! k_B T_c} m^4\right). \tag{5.157}$$

Close to the critical temperature, the fluctuations become very intense and are spatially correlated over very large distances—we have in fact seen that the coherence length diverges for $T \to T_c$. It is therefore essential to understand under which conditions it is possible to neglect the consequences of the fluctuations, and therefore the mean-field's theory's predictions can be considered valid, and when they become unreliable.

In order to estimate the fluctuations' relative importance, we employ a criterion that is due to Ginzburg [Ginz60]. The idea is to compare the mean value of the order parameter as predicted by mean-field theory with the fluctuations predicted by Ornstein-Zernicke's theory.

Let us consider the expression (5.155) of the Fourier transform of the correlation function, which we will rewrite (up to a constant factor) in the form

$$C(k) = \frac{1}{k^2 + t}, \tag{5.158}$$

where $t \propto |T - T_c|$.

By evaluating the Fourier antitransform of this expression, we obtain an estimate of the local spin fluctuations:

$$\langle \Delta\sigma_i^2 \rangle = \langle \sigma_i^2 \rangle - \langle \sigma_i \rangle^2 \propto \int d^d k\, C(k) \propto t^{(d-2)/2} + \text{regular terms}. \tag{5.159}$$

The "regular" terms come from larger values of $|\boldsymbol{k}|$, for which the expression (5.155) is no longer applicable. They do not have consequences for the critical behavior (at least if $d > 2$), and we can forget them as far as our argument is concerned.

The order of magnitude of the spin's critical fluctuations is therefore given by

$$\Delta\sigma \propto \sqrt{\langle \Delta\sigma_i^2 \rangle} \propto t^{(d-2)/4}. \tag{5.160}$$

We want to compare this fluctuation with the mean value of the order parameter, and we therefore have

$$\langle \sigma \rangle \propto t^{1/2}. \tag{5.161}$$

As long as $\Delta\sigma \ll \langle\sigma\rangle$ the mean-field theory's predictions remain valid. This also occurs for $t \to 0$, as long as $d > 4$. We can therefore expect some deviations from mean-field behavior when $d \leq 4$.

This argument can be made more quantitative by reinstating the coefficients we gave as understood, and by evaluating them sufficiently far from the critical point. It can also be formulated in a more rigorous manner, by introducing an approximation scheme that reproduces mean-field theory at the lowest order and provides corrections to it as a systematic power series expansion of a fictitious parameter. The Ginzburg criterion then follows from the comparison of the second and first terms of this expansion [Amit84, p. 105].

In conclusion, we can expect that in two and three dimensions, the asymptotic behavior in proximity of the critical point will not be well described by mean-field theory.

5.19 Universality and Scaling

The values of the critical exponents we calculated on the rare occasions in which it is possible to do so exactly, or that were obtained from real or numerical experiments, are in effect not the same as those of the classical exponents. In table 5.2, we report the values of the critical exponents for a certain number of physical systems. The exponent η characterizes the behavior of the correlation function exactly at the critical temperature—in Fourier space, one has

$$C(\boldsymbol{k}, T = T_c) \propto |\boldsymbol{k}|^{-2+\eta}. \tag{5.162}$$

These values are obviously not equal to the classical values. One can notice, however, that the exponents of the first three systems are mutually compatible, and so are the exponents of the last two—the exponents seem to depend only on the dimensionality of the order parameter. This conjecture is corroborated by considering the exponents obtained numerically (with different methods) for some statistical mechanics models on a lattice, which are shown in table 5.3. Analogously, the measurable exponents of the λ transition of helium correspond to those of the planar model, while the ferromagnetic or antiferromagnetic transitions characterized by a vectorial order parameter in three dimensions are described by the same exponents. These observations have led to the formulation of the **universality hypothesis**—phase transitions can be sorted into a small number of

Table 5.2 Critical Exponents of Different Physical Systems

	Xe	Bin. mixt.	β-brass	^{4}He	Fe	Ni
n	1	1	1	2	3	3
α	< 0.2	$0.113 \pm .005$	$0.05 \pm .06$	-0.014 ± 0.16	$-0.03 \pm .12$	$0.04 \pm .12$
β	$0.35 \pm .15$	$0.322 \pm .002$	$0.305 \pm .005$		$0.37 \pm .01$	$0.358 \pm .003$
γ	$1.3 \pm_{.2}^{.1}$	$1.239 \pm .002$	$1.25 \pm .02$		$1.33 \pm .015$	$1.33 \pm .02$
δ	$4.2 \pm_{.3}^{.6}$				4.3 ± 1	$4.29 \pm .05$
η	$0.1 \pm .1$	$.0017 \pm .015$	$0.08 \pm .07$		$0.07 \pm .04$	$0.041 \pm .01$
ν	≈ 0.57	$0.615 \pm .006$	$0.65 \pm .02$	$0.672 \pm .001$	$0.69 \pm .02$	$0.64 \pm .1$

Note: The first three columns are systems with a scalar order parameter ($n = 1$): Xe at its critical point, different binary mixtures at their consolution point, and β-brass, a form of copper and tin alloy. The transition λ(He$^{\rm I}$-He$^{\rm II}$) corresponds to a planar order parameter ($n = 2$), while Fe and Ni are Heisenberg ferromagnets ($n = 3$) with weak anisotropy. I have included only exponents that have been actually measured. More specifically, in the case of the λ transition of helium, the order parameter is not accessible, and therefore the exponents β, γ, δ, and η cannot be measured.

Table 5.3 Critical Exponents for Different Models as a Function of the Dimensionality n of the Order Parameter and the Dimensionality d of Space.

	Mean-field	Ising $d = 2$	Ising $d = 3$	Heisenberg	Spherical
(n, d)		(1,2)	(1.3)	(3,3)	(∞, 3)
α	0 (disc.)	0 (log)	$0.119 \pm .006$	$-0.08 \pm .04$	-1
β	$1/2$	$1/8$	$0.326 \pm .004$	$0.38 \pm .03$	$1/2$
γ	1	$7/4$	$1.239 \pm .003$	$1.38 \pm .02$	2
δ	3	(15)	$4.80 \pm .05$	$4.65 \pm .29$	5
η	0	$1/4$	$0.024 \pm .007$	$0.07 \pm .06$	0
ν	$1/2$	1	$0.627 \pm .002$	$0.715 \pm .02$	1

Note: The exponent δ of the two-dimensional Ising model is conjectured; the others are obtained from the exact solution. Exponents obtained from an exact solution of the spherical model, which corresponds to $n \to \infty$, are also shown.

universality classes, characterized by the dimensions of the system and of the order parameter, and the critical exponents are equal each class.

In fact, one can observe a stronger relation. Let us consider the equation of state for the order parameter, expressed as a function of h and of

$$t \equiv \frac{T - T_{\rm c}}{T_{\rm c}} \tag{5.163}$$

for two different systems:

$$m^{(1)} = m^{(1)}(t, h), \tag{5.164}$$

$$m^{(2)} = m^{(2)}(t, h). \tag{5.165}$$

It is then possible to map the two equations of state onto each other by rescaling m, h, and t by an appropriate positive factor:

$$m^{(1)}(t, h) = \lambda_m m^{(2)}(\lambda_t t, \lambda_h h). \tag{5.166}$$

A relation of this kind, on the other hand, must also be valid if one is discussing the *same system*—by arbitrarily choosing one of the scale factors, we can choose the other two so that

$$m(t, h) = \lambda_m m(\lambda_t t, \lambda_h h). \tag{5.167}$$

Let us consider what occurs at $h = 0$, $t < 0$. On the one hand, we obtain

$$m \propto |t|^{\beta}, \tag{5.168}$$

and on the other,

$$m(t, h = 0) = \lambda_m m(\lambda_t t, h = 0). \tag{5.169}$$

By choosing $\lambda_t = |t|^{-1}$, we obtain

$$\lambda_m = \lambda_t^{-\beta}. \tag{5.170}$$

Following the same reasoning at $t = 0$, we obtain

$$\lambda_m = \lambda_h^{-1/\delta}. \tag{5.171}$$

The magnetization is therefore a general homogeneous function of its arguments:

$$m(\lambda t, \lambda^{\beta\delta} h) = \lambda^{\beta} m(t, h). \tag{5.172}$$

Since the magnetization is the derivative of the free energy with respect to h, an analogous relation also exists for free energy density (or, more exactly, for its singular part):

$$f(\lambda t, \lambda^{\beta\delta} h) = \lambda^a f(t, h). \tag{5.173}$$

We can determine the exponent a from the behavior of specific heat. We know that

$$C_h \propto \frac{\partial^2 f}{\partial t^2} \tag{5.174}$$

One therefore has

$$C_h(\lambda t, 0) = \lambda^{a-2} C_h(t, 0), \tag{5.175}$$

from which we obtain

$$a - 2 = -\alpha. \tag{5.176}$$

The free energy density therefore satisfies the general homogeneity relation:

$$f(\lambda t, \lambda^{\beta\delta} h) = \lambda^{2-\alpha} f(t, h). \tag{5.177}$$

This expression vanishes at the critical point—obviously, this does not mean that *all* the free energy vanishes at the critical point, but only its singular part.

By taking the derivative of this relation, one can express the exponents of the thermodynamic quantities as a function of only two exponents, α and $\beta\delta$, which appear in the expression of the free energy. By taking its derivative with respect to h, for example, one obtains the expression of the magnetization:

$$\lambda^{\beta\delta} m\left(\lambda t, \lambda^{\beta\delta} h\right) = \lambda^{2-\alpha} m\left(t, h\right). \tag{5.178}$$

One therefore has

$$\beta = 2 - \alpha - \beta\delta. \tag{5.179}$$

By taking once more the derivative, one obtains the relation for susceptibility:

$$\lambda^{2\beta\delta} \chi\left(\lambda t, \lambda^{\beta\delta} h\right) = \lambda^{2-\alpha} \chi\left(t, h\right), \tag{5.180}$$

from which we obtain the exponent γ:

$$-\gamma = 2 - \alpha - \beta\delta. \tag{5.181}$$

These two expressions imply a relation between the first three exponents:

$$\alpha + 2\beta + \gamma = 2. \tag{5.182}$$

Relations of this type (known as **scaling laws**) can be written between any three exponents, and follow from the homogeneity of the free energy and from the fact that only two independent exponents appear in it. They are well satisfied by the experimentally measured exponents, and currently there are few doubts that they are valid for the asymptotic exponents.

Another relation of this type connects β, γ, and δ:

$$\gamma = \beta\left(\delta - 1\right). \tag{5.183}$$

It is enlightening to consider the implication of these relations for the coherence length and for the correlation function. The lines of reasoning we have followed show that the coherence length must itself also satisfy a general homogeneity relation:

$$\xi(\lambda t, \lambda^{\beta\delta} h) = \lambda^{-\nu} \xi(t, h). \tag{5.184}$$

An analogous relation must also hold for the correlation function, where one will, however, also have to rescale the wave vector:

$$C(\lambda^b k, \lambda t, \lambda^{\beta\delta} h) = \lambda^c C(k, t, h). \tag{5.185}$$

In order to determine the exponents b and c, let us observe that for $k = 0$, the correlation function is proportional to the susceptibility. We thus obtain

$$c = -\gamma. \tag{5.186}$$

On the other hand, at $t = h = 0$, we must have

$$C(k) \propto k^{-2+\eta}. \tag{5.187}$$

Therefore,

$$\frac{\gamma}{b} = 2 - \eta. \tag{5.188}$$

Then, recalling that the coherence length is defined by

$$\xi^2 \propto \frac{1}{C(0)} \left. \frac{\partial C(k)}{\partial k^2} \right|_{k=0}, \tag{5.189}$$

and comparing with the relation for the ξ, we obtain

$$b = \nu. \tag{5.190}$$

We have thus obtained another scaling law:

$$\frac{\gamma}{\nu} = 2 - \eta. \tag{5.191}$$

And last, we want to connect the exponent ν to the other thermodynamic variables. To do this, we observe that due to Einstein's theory of fluctuations, we expect that the free energy of a fluctuation (which, as we know, is correlated over distances on the order of the coherence length) should be of the order of magnitude of $k_B T_c$—in other words, a constant. When the coherence length is multiplied by a factor $\ell = \lambda^{-\nu}$, the contribution of the fluctuations to the free energy density is multiplied by a factor $\ell^{-d} = \lambda^{d\nu}$, where d is the system's dimensionality. We thus obtain the scaling law:

$$2 - \alpha = d\nu. \tag{5.192}$$

5.20 Partition Function of the Two-Dimensional Ising Model

In this section, I report the solution of the two-dimensional Ising model in zero magnetic field, as reported by Vdovichenko [Vdov64 and Vdov65]. The model was first solved by Lars Onsager in 1944 [Onsa44] by a mathematical *tour de force*. Simpler derivations were found later by Kac and Ward [Kac52]; Schulz, Mattis, and Lieb [Schu64]; and others. The present solution was inspired by Kac and Ward's solution, but is simpler. It became widely known due to Landau and Lifshitz's treatise on theoretical physics [Land80, p. 4798ff] .

We consider a system of $N = L^2$ Ising spins placed on a square lattice. Thus the spin placed at the (k, ℓ) lattice point is denoted by $\sigma_{k\ell}$ and one has $\sigma_{k\ell} = \pm 1$, $k, \ell \in \{1, \dots, L\}$. The Hamiltonian $H(\{\sigma\})$ is given by

$$H(\{\sigma\}) = -\sum_{k\ell} \left[J(\sigma_{k\ell}\sigma_{k,\ell+1} + \sigma_{k\ell}\sigma_{k+1,\ell}) + h\sigma_{k\ell} \right], \tag{5.193}$$

where we have assumed periodic boundary conditions:

$$\sigma_{k+L,\ell} = \sigma_{k,\ell+L} = \sigma_{k\ell}, \qquad \forall k, \ell. \tag{5.194}$$

We set $h = 0$ from now on. Then, we have seen in section 5.11 that the partition function can we written

$$Z(K) = \left(\frac{2}{1-t^2}\right)^N \sum_{\mathcal{G}} {}' t^{|G|}, \tag{5.195}$$

where

$$t = \tanh \frac{J}{k_{\mathrm{B}}T}, \tag{5.196}$$

and the sum runs over all diagrams $\mathcal{G}$ that can be drawn on the lattice, such that (1) each bond appears at most once, and (2) at each vertex, only an even number of bonds (zero, two, or four) can meet. In this expression, $|\mathcal{G}|$ is the number of bonds that appear in the diagram $\mathcal{G}$. Then, this expression can be written in the form

$$S = \sum_{\mathcal{G}} {}' t^{|\mathcal{G}|} = \sum_{r} t^r g_r, \tag{5.197}$$

where g_r is the total number of diagrams satisfying the two preceding rules and containing exactly r bonds.

We will now evaluate this expression by transforming it into a sum over loops. The resulting expression will then be evaluated by reducing it to a random-walk problem.

A generic diagram $\mathcal{G}$ can be considered as a collection of **loops**. A loop is the trajectory of a walk that starts and ends on the same site. However, the decomposition of a diagram into loops is ambiguous if there are self-intersections, that is, if there are vertices where four bonds meet. Let us consider, for example, the diagram in figure 5.10. It can be considered as the collection of two loops that meet at one vertex (a), or as a single loop whose path does intersect itself (c) or does not (b). In order to obtain a nonambiguous sum, we assign to each diagram a factor $(-1)^n$, where n is the number of intersections. In this situation, the contribution of case (c) will be opposite to that of case (b), and they cancel out, leaving only the contribution of case (a). One can easily realize, then, that with this convention, the contribution of diagrams in which three bonds meet at a vertex identically vanishes, as can be seen in figure 5.11. In this way, the sum over all diagrams $\mathcal{G}$ is reduced to a sum over all loops, in which each loop appears with a weight proportional to $(-1)^n$, where n is the number of self-intersections. Notice that we do not allow vertices connected to only one bond and, therefore, the possibility that a walker gets back in its steps.
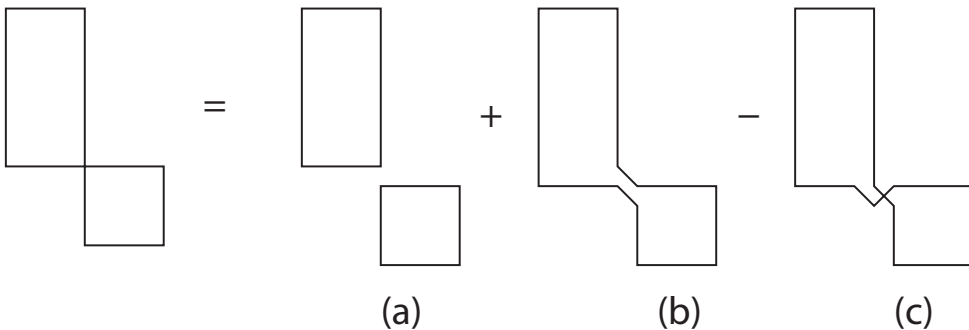

FIGURE 5.10. A diagram with self-intersections can be decomposed in several different ways into loops.
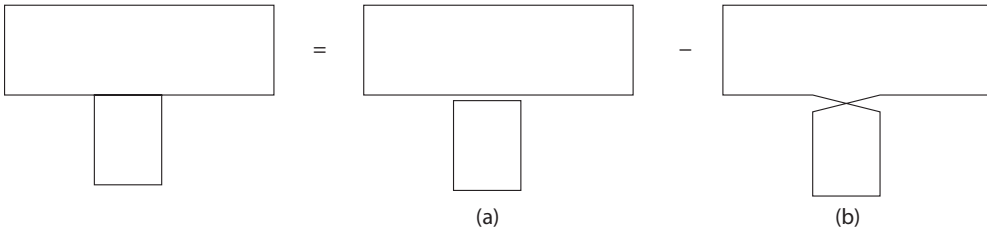
FIGURE 5.11. A diagram with a three-bond vertex can be obtained as the sum of two diagrams with numbers of intersections that differ by one. Their contributions cancel out.

Now, we can express the number of self-intersections of a loop by means of the following trick. It is well known that the total angle through which the tangent angle to the trajectory of a walker performing a loop turns around is given by $2\pi(\ell + 1)$, where the parity of ℓ is equal to the parity of the number of intersections n. Thus, if we assign a factor $e^{i\phi/2}$ to each lattice point with turning angle ϕ, then we will have at the end of the loop a factor $(-1)^{\ell+1} = (-1)^{n+1}$, where n is the number of intersections. With this counting, each diagram made up of s loops will give a contribution proportional to $(-1)^{s+n}$. Thus we have to multiply this contribution by $(-1)^s$ in order to have the required sign in equation (5.197).

In order to evaluate the angle ϕ, it is convenient to deal with directed loops. Let us denote by f_r the sum over all undirected loops consisting of r bonds (taking into account the factors t^r and $e^{i\phi n/2}$). Then, the sum over all double loops of ℓ bonds will be given by

$$\frac{1}{2} \sum_{r_1+r_2=\ell} f_{r_1} f_{r_2},$$

taking into account the possible permutations of the loops. Thus we have in general

$$S = \sum_{s=1}^{\infty} (-1)^s \frac{1}{s!} \sum_{r_1, r_2, \ldots, r_s=1} f_{r_1} f_{r_2} \cdots f_{r_s} = \exp\left\{-\sum_{r=1}^{\infty} f_r\right\}. \tag{5.198}$$

In going from undirected to directed loops, each loop is encountered twice, and thus if we denote by v_r the sum of the contributions of directed loops with r bonds, we have

$$S = \exp\left\{-\frac{1}{2} \sum_{r=1}^{\infty} v_r\right\}. \tag{5.199}$$

We will now evaluate v_r. Given a lattice point (k, ℓ), let denote the possible directions as follows:

N: $(k, \ell) \longrightarrow (k, \ell + 1),$
E: $(k, \ell) \longrightarrow (k + 1, \ell),$
S: $(k, \ell) \longrightarrow (k, \ell - 1),$
W: $(k, \ell) \longrightarrow (k - 1, \ell).$

Let us denote by $W_r(k\ell v | k_0 \ell_0 v_0)$ the sum of all contributions of r-bond diagrams starting from lattice point $k_0 \ell_0$ in the direction $v_0 \in \{N, E, S, W\}$ and ending in lattice point (k, ℓ) in

the direction ν. Each bond occurs with a factor $te^{i\phi/2}$, where ϕ is the change of direction in going to the next bond.

Then, it is possible to write a linear recursion relation for W_r:

$$W_{r+1}(k\ell\nu|k_0\ell_0\nu_0) = \sum_{k'\ell'\nu'} T_{k\ell\nu,k'\ell'\nu'} W(k'\ell'\nu'|k_0\ell_0\nu_0). \tag{5.200}$$

The transition matrix $\mathsf{T} = (T_{k\ell\nu,k'\ell'\nu'})$ has the expression

$$T_{k\ell\nu,k'\ell'\nu'} = tA_{\nu\nu'}\delta_{k',k+\alpha(\nu)}\delta_{\ell',\ell+\beta(\nu')}, \tag{5.201}$$

where

$$\alpha(N) = 0, \quad \beta(N) = -1,$$

$$\alpha(E) = -1, \quad \beta(E) = 0,$$

$$\alpha(S) = 0, \quad \beta(S) = +1,$$

$$\alpha(W) = +1, \quad \beta(N) = 0.$$

The matrix $\mathsf{A} = (A_{\nu\nu'})$ (where $\nu, \nu' = N, E, S, W$) is given by

$$\mathsf{A} = \begin{pmatrix} 1, \omega, 0, \omega* \\ \omega*, 1, \omega, 0 \\ 0, \omega*, 1, \omega \\ \omega, 0, \omega*, 1 \end{pmatrix}, \tag{5.202}$$

where

$$\omega = e^{i\pi/4}, \tag{5.203}$$

and $\omega*$ is the complex conjugate of ω.

The connection between the weights W and the loop contributions v_r is given by

$$v_r = \frac{b_r}{r}, \tag{5.204}$$

where

$$b_r = \mathrm{Tr}\,W_r = \sum_{k\ell\nu} W(k\ell\nu|k\ell\nu). \tag{5.205}$$

The factor $1/r$ comes from the fact that a single diagram with r bonds can be obtained from r different walks, with different the starting points. Now, from equation (5.200), we have

$$\mathrm{Tr}\,W_r = \mathrm{Tr}\,\mathsf{T}^r = \sum_i \lambda_i^r, \tag{5.206}$$

where λ_i are the eigenvalues of the matrix T. From this equation, taking into account equations (5.199) and (5.204), we obtain

$$S = \exp\left\{-\frac{1}{2}\sum_{r,i}\frac{1}{r}\lambda_i^r\right\} = \exp\left\{\frac{1}{2}\sum_i \ln(1-\lambda_i)\right\} = \prod_i \sqrt{1-\lambda_i}. \tag{5.207}$$

Thus the problem boils down to the diagonalization of the matrix T. One can see from equation (5.201) that T depends only on the differences in the indices k, ℓ. It can thus be diagonalized by a Fourier transformation. We set

$$T_{\nu\nu'}(m, n) = \sum_{k\ell} e^{-2\pi i(mk+n\ell)} T_{k\ell\nu, 00\nu'}. \tag{5.208}$$

We then find that

$$T_{\nu\nu'}(m, n) = t \begin{pmatrix} \gamma^*(n), & \omega\gamma^*(m), & 0, & \omega^*\gamma(m) \\ \omega^*\gamma^*(n), & \gamma^*(m), & \omega\gamma(n), & 0 \\ 0, & \omega^*\gamma^*(m), & \gamma(n), & \omega\gamma(m) \\ \omega\gamma^*(n), & 0, & \omega^*\gamma(n), & \gamma(m) \end{pmatrix}, \tag{5.209}$$

where

$$\gamma(m) = e^{2\pi i m/L}. \tag{5.210}$$

Thus, for given values of (m, n), we have

$$\prod_{i=1}^{4} \{1 - \lambda_i(m, n)\} = \det[I - T(m, n)]$$
$$= (1 + t^2)^2 - 2t(1 - t^2)\left(\cos\frac{2\pi m}{L} + \cos\frac{2\pi n}{L}\right). \tag{5.211}$$

Thus we obtain

$$Z = 2^N (1 - t^2)^{-N} \prod_{mn}\left[(1 + t^2)^2 - 2t(1 - t^2)\left(\cos\frac{2\pi m}{L} + \cos\frac{2\pi n}{L}\right)\right]^{1/2}, \tag{5.212}$$

where the product runs over L consecutive values of m and of n.

Setting $p = 2\pi m/L$ and $q = 2\pi n/L$, the Helmholtz free energy is given by

$$F(T) = -Nk_B T\left\{\ln 2 - \ln(1 - t^2)\right.$$
$$\left. + \frac{1}{2}\int_{-\pi}^{+\pi} \frac{dpdq}{(2\pi)^2} \ln\left[(1 + t^2)^2 - 2t(1 - t^2)(\cos p + \cos q)\right]\right\}. \tag{5.213}$$

Let us consider the contribution of the integral. The minimum value of the integrand is reached for $p = q = 0$ and is given by

$$\ln\left[(1 + t^2)^2 - 4t(1 - t^2)\right] = \ln[t^2 + 2t - 1]^2.$$

The argument of the logarithm vanishes for

$$t = t_c = \sqrt{2} - 1,$$

which corresponds to the transition temperature T_c given by equation (5.72):

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

In order to understand the behavior of F in the neighborhood of this temperature, let us introduce $\tau = t - t_c$ and expand the integrand for small values of τ and of p, q. One has

$$F(T) = \frac{1}{2} \int_{-\pi}^{+\pi} \frac{dp\,dq}{(2\pi)^2} \ln\left[c_1 \tau^2 + c_2 (p^2 + q^2)\right] + \text{regular terms},$$

where c_1 and c_2 are constants. Integrating, one obtains

$$F(T) = -a\tau^2 \ln|\tau| + \text{regular terms},$$

where $a > 0$ is a constant. The specific heat C is proportional to $-\mathrm{d}^2 F/\mathrm{d}\tau^2$. Thus we have

$$C \simeq a \ln|\tau| + \text{regular terms}, \tag{5.214}$$

indicating that the specific heat exhibits a logarithmic divergence at the critical temperature.

The evaluation of the spontaneous magnetization $m_0 = \langle \sigma \rangle$ proceeds in a similar way [Vdov65], starting, e.g., from the relation

$$m_0^2 = \lim_{k \to \infty} \langle \sigma_{1\ell} \sigma_{1,\ell+k} \rangle.$$

One obtains $m_0 = 0$ for $t < t_c$ given earlier, and

$$m_0 = \left\{ 1 - \left(\frac{t^{-1} - t}{2} \right)^4 \right\}^{1/8} \tag{5.215}$$

for $t > t_c$, i.e., below the transition temperature. Thus for small positive values of $\tau = t - t_c$, one has

$$m_0 \propto \tau^\beta, \tag{5.216}$$

where the exponent β is given by

$$\beta = \frac{1}{8}. \tag{5.217}$$

The connection between the Ising model and the statistics of loops can be interpreted more deeply as the equivalence between the model and a system of noninteracting fermions. This correspondence is exploited by Schulz, Mattis and Lieb [Schu64] in their solution of the Ising model.

Recommended Reading

The phenomenology of phase transitions is explained in H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Oxford, UK: Oxford University Press, 1971. The books devoted to phase transitions are now legion: it is sufficient to quote the series of volumes edited by C. Domb and M. Green (and later by C. Domb and J. L. Lebowitz), *Phase Transitions and Critical Phenomena*, London: Academic Press, 1971–present. The mean-field and Landau theories are discussed in detail in L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed., part 1, Oxford, UK: Pergamon Press, 1980. A simple presentation of the different ordered phases in condensed matter can be found in the excellent

volume by P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge, UK: Cambridge University Press, 1995. One book is entirely devoted to Landau phase transition theory in complex systems: J.-C. Tolédano and P. Tolédano, *The Landau Theory of Phase Transitions: Application to Structural, Incommensurate, Magnetic, and Liquid Crystal Systems*, Singapore: World Scientific, 1987.