Contents

Chapter 1

Introduction to Phase Transitions

Take a large piece of material and measure some of its macroscopic properties, for example its density, compressibility or magnetisation. Now divide it into two roughly equal halves, keeping the external variables (like pressure and temperature) the same. The macroscopic properties of each piece will then be the same as those of the whole. The same holds true if the process is repeated. But after many iterations, something different must happen, because we know that matter is made up of atoms whose individual properties are quiet different from those of the matter which they constitute. The length scale at which the overall properties of atoms begin to differ markedly from those of the original gives a measure of what is termed the *correlation length* of the material. It is the distance over which the fluctuations of the microscopic degrees of freedom are significantly correlated with each other.

Usually this length is of the order of a few interatomic spacings, so we may consider really a quite small collections of atoms to get a very good idea of the macroscopic behaviour of the material. However, the actual value of the correlation length, depends on the external conditions of the system, such as temperature and pressure. It is well known that systems may abruptly change their macroscopic behaviour as these quantities are smoothly varied. The points at which this happens are called critical points, and they usually mark a phase transition from one state of matter to another. There are basically two possible ways in which a transition may occur.

In the first scenario, the two (or more) states coextist exactly at the critical point. However, they are distinct from each other, in that they have different macroscopic properties. Away from the critical point there is generally a unique phase whose properties are continuously connected to one of the co-existent phases at the critical point. In this case, we expect to find discontinuous behaviour in various thermodynamic quantities as we pass through the critical point. These transitions are termed *discontinuous* or *first-order* phase transitions. Examples of this are the melting of a 3D solid, or the condensation of a gas into a liquid. These transitions often exhibit hysteresis, or memory effect, since the continuation of a given state into the opposite phase may be metastable, and the system may take a macroscopically long time to readjust. The correlation length in a first-order transition is generally finite.

In a *continuous* transition the situation is quite different, and the correlation length becomes effectively infinite. The fluctuations are correlated over all distance scales, and the system is in a unique, critical phase. At a continuous transition, the two phases on either side of the critical point become identical as it is approached. Not only does the correlation length diverge in a continuous fashion as such a critical point is approached, but the differences in the various thermodynamic

quantities between the competing phases go to zero smoothly.
Simple examples of continuous transitions occur at the liquid-gas critical point in a fluid, and at the Curie temperature in a ferromagnet.

1.1 Simple fluids

A simple overview of phase transitions can be done analysing the phase diagram of a fluid, focusing on the liquid-gas transition. The phase diagram of a generic substance in the temperature-pressure plane usually has the form shown in Figure 1.1. Let us focus on the part of the diagram close to the liquid-gas critical point (T_c, p_c). As we will see, it looks quite similar to that for a ferromagnet in the (T, H)-plane (where H is an external magnetic field). We know from thermodynamics that across the liquid-gas phase boundary, the density ρ is discontinuous. We are in presence of a first-order transition. The jump in the density $\rho_{liquid} - \rho_{gas}$ approaches zero at the critical end point (T_c, p_c) (Figure 1.2 shows the isotherms curves of p versus ρ at constant T). This is an example of a *critical end point*, at which the first-order transition becomes continuous.
It is known that most of the quantities of interest exhibit power law behaviour close to the critical point. For a fluid near the critical point, for example, we have:

- specific heat, $C_V \propto |t|^{-\alpha}$ at $\rho = \rho_c$;
- $\rho_{liquid} - \rho_{gas} \propto (-t)^{\beta}$ gives the shape of the coexistence curve near the critical point;
- isothermal compressibility $\chi_T \propto |t|^{-\gamma}$;
- $p - p_c \propto |\rho_{liquid} - \rho_{gas}|^{\delta}$ gives the shape of the critical isotherm near the critical point.

Where we defined the *reduced temperature*, $t \equiv (T - T_c)/T_c$. We will give more precise definitions of these exponents in describing ferromagnetic systems. One of the most remarkable results of universality is that the critical exponents of a simple fluid are identical with those of uni-axial ferromagnets, showing that, in phase transitions, the behaviours of the relevant physical quantities weakly depends on the specific system.

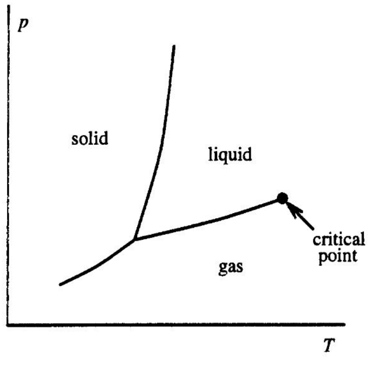


Figure 1.1. Phase Diagram of a typical substance.

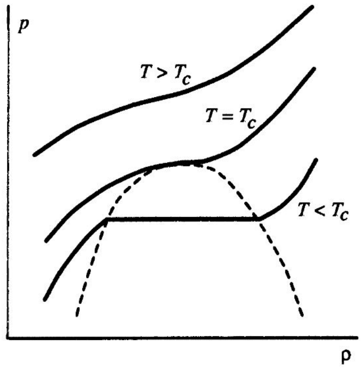


Figure 1.2. Isotherms near the liquid-gas critical point. The dashed line is the coexistence curve.

1.2 Magnetic Systems

1.2.1 Uniaxial ferromagnets

We will now discuss another example of phase transition: a ferromagnetic material. In a ferromagnet there are two interesting external parameters which may be varied: the temperature T and the applied magnetic field h (in a fluid the two variables were the temperature and the external pressure). In the most straightforward case, the local magnetisation is constrained to lie parallel or anti-parallel to a particular axis. The phase diagram is shown in Figure 1.3. All the thermodynamic quantities are smooth analytic functions of T and h except on the line $h = 0$, $T = T_c$. Across the line $T < T_c$, the magnetisation M, as a function of h, is discontinuous, having the form illustrated in Figure 1.4, a). This is thus a first-order transition, with a finite correlation length. As T approaches the Curie point T_c, from below, however, the discontinuity approaches to zero, and the correlation length diverges (Figure 1.4, b) and c)). The point $h = 0$, $T = T_c$ is another example of a critical end point.

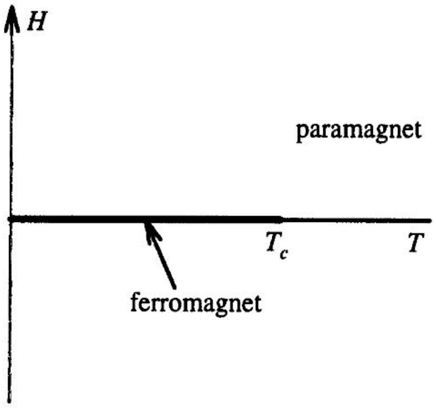
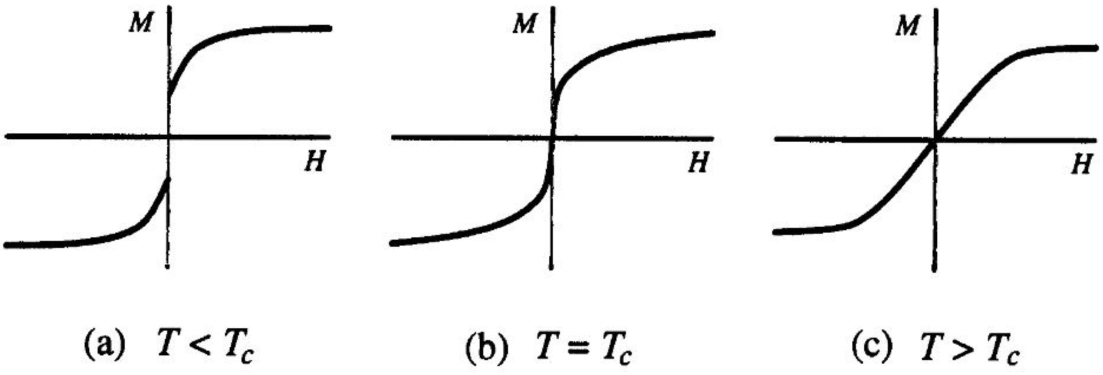


Figure 1.3. Phase Diagram of a uniaxial ferromagnet.



(a) $T < T_c$ (b) $T = T_c$ (c) $T > T_c$

Figure 1.4. Magnetisation versus applied field.

When $T < T_c$, the two limits $h \to 0^+$ and $h \to 0^-$ give different possible values $\pm M_0$ for magnetisation. Which one the system chooses depends on its previous history. This is an example of *spontaneous symmetry breaking*: the Hamiltonian is invariant under simultaneous reversal of all the local magnetic degrees of freedom, but this symmetry is not respected by the equilibrium state. This is a common (but not universal) characteristic of continuous critical point. The magnetisation M measures the amount of magnetic order in the material, and it is called *order parameter* for this transition (in most examples of critical behaviour it is possible to find an order parameter). In general, the behaviour of the local fluctuations of the order parameter provides a useful way of characterising the nature of the transition.

As mentioned in the previous paragraph, most of the quantities involved in the transition exhibit power law. We now give the definition of the principal critical exponents. We define again the reduced temperature $t \equiv (T - T_c)/T_c$ and the reduced external magnetic field $\tilde{h} = h/k_B T_c$.

α : The specific heat in zero field $C \sim A|t|^{-\alpha}$, apart from terms regular in t. In principle, one should consider the possibility of different exponents α and α' and amplitudes A and A' for $t < 0$ and $t > 0$ respectively. However it is an immediate consequence of the renormalization group that $\alpha = \alpha'$. α can be positive or negative, corresponding to either a divergent spike or a cusp in the specific heat when plotted against T. Although the exponent is universal, the amplitude A is not. Meanwhile, renormalization group shows that the ratio A'/A is universal.

β : The spontaneous magnetisation $\lim_{h\to 0^+} M \propto (-t)^\beta$.

γ : The zero field susceptibility $\chi \equiv (\partial M)/\partial h) \mid_{h=0} \propto |t|^{-\gamma}$. Again one should define different exponents for different values of t, but theory indicates that they should be equal.

δ : At $T = T_c$, the magnetisation varies with $\tilde{h}$ according to $M \propto |\tilde{h}|^{1/\delta}$.

ν : The correlation length diverges as $t \to 0$, with $h = 0$, according to $\xi \propto |t|^{-\nu}$.

η : The correlation function of the fluctuations in the local magnetization, $G(r)$, does not decay exponentially at the critical point, but rather according to $G(r) \propto 1/r^{d-2+\eta}$.

z : This exponent is related to the relaxation time of the system τ near the critical point. It diverges like $\tau \propto \xi^z$.

1.2.2 General properties of magnetic systems

Let us analyse a generic magnetic system. For each configuration of the system we can define the corresponding magnetization $\vec{S}(x)$ for each point x of the physical space. Microscopically $\vec{S}(x)$ receives the contributions both from the motion of the electrons and from their spins. In many cases the electrons that mainly contribute to $\vec{S}(x)$ are localized near the atoms lattice, so it is convenient to define $\vec{S}(x)$ only on the points of the lattice. It is interesting to study the response of the material to the action of an external magnetic field $h(x)$. The Hamiltonian of the system can be written as:

$$\mathscr{H}(C) = \mathscr{H}_0(C) - \int_V \mathrm{d}^3x\, \vec{h}(x) \cdot \vec{S}(x) \tag{1.1}$$

with $\vec{h}(x)\cdot\vec{S}(x) = \sum_{\nu=1}^{3} h_\nu(x)S_\nu(x)$. $\mathscr{H}_0(C)$ is the Hamiltonian in absence of magnetic field, V is the volume occupied by the system. C denotes a generic configuration of the system. From Statistical Mechanics we know that most of the information on the properties of the system are contained in the free energy F. In this case F is a functional of the external field $h(x)$, $F = F[h]$. If the magnetic field does not depend on x, we can write the free energy as a function of h, $F(h) \equiv F[h] \mid_{\vec{h}(x)=\vec{h}}$, and define a free energy density $f(h) = F(h)/V$.

The expectation values of the local and total magnetization ($m_n u(x)$ and M_v) are given as

$$m_\nu(x) = \langle S_\nu(x)\rangle; \qquad M_\nu = \langle S_\nu^T\rangle : \quad S_\nu^T \equiv \int_V \mathrm{d}^3x\; S_n u(x) \tag{1.2}$$

From the form of the Hamiltonian 1.1, it is easy to see that the local magnetization can be computed as

$$m_\mu(x) = -\frac{\delta F[h]}{\delta h_\mu(x)}, \qquad m_\mu = -\frac{\partial f(h)}{\partial h_\mu} \tag{1.3}$$

If the Hamiltonian is spatially homogeneous (i.e. invariant under continuous or discrete translations), the magnetization will be x independent, $m_\nu(x) = m_\nu$. The free energy is given by:

$$\begin{aligned} F[h] &= -\frac{1}{\beta}\ln(Z[h]), \\ Z[h] &= \int \mathrm{d}C\; \exp[-\beta\mathscr{H}(C)] \end{aligned} \tag{1.4}$$

A very important quantity in this kind of systems is the magnetic susceptibility, defined by

$$\chi_{\mu\nu} = \frac{\partial m_\mu}{\partial h_\nu} \tag{1.5}$$

And one can also consider a point-dependent magnetic susceptibility:

$$\chi_{\mu\nu}(x,y) = \frac{\delta m_\mu(x)}{\delta h_\nu(y)} \tag{1.6}$$

If the system is invariant under translations, the dependence of the susceptibility on x and y reduces to a dependence on the difference $x - y$:

$$\chi_{\mu\nu}(x,y) = \chi_{\mu\nu}(x-y) \qquad \text{and} \qquad \chi_{\mu\nu} = \frac{1}{V}\int_V \mathrm{d}^3x\; \chi_{\mu\nu}(x) \tag{1.7}$$

We see from 1.3 that

$$\chi_{\mu\nu} = -\frac{\partial^2 f}{\partial h_\mu \partial h_\nu}, \qquad \chi_{\mu\nu}(x,y) = -\frac{\delta^2 f}{\delta h_\mu(x)\delta h_\nu(y)} \tag{1.8}$$

From these equations we get the symmetry properties of the magnetic susceptibility:

$$\chi_{\mu\nu} = \chi_{\nu\mu} \qquad \chi_{\mu\nu}(x,y) = \chi_{\nu\mu}(y,x) \tag{1.9}$$

We will now perform the explicit derivatives from the partition function in order to find a very important result.

Before doing this, it is useful to define the connected average of two functions A and B:

$$\langle AB\rangle_C \equiv \langle AB\rangle - \langle A\rangle\langle B\rangle \tag{1.10}$$

Now we start from the statistical definition of the local magnetization m_l as mean value of the magnetization S_l (eq. 1.2):

$$m_l = \langle S_l\rangle \equiv \frac{1}{Z}\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_l \tag{1.11}$$

Where Z is the partition function (eq. 1.4), and the integral is extended to all possible configurations of the system. We want to compute the magnetic susceptibility (the Hamiltonian is the 1.1):

$$\begin{aligned}
\chi_{lk} = \frac{\partial m_l}{\partial h_k} &= \frac{1}{Z^2}\left\{ Z\frac{\partial}{\partial h_k}\Big(\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_l\Big) - \Big(\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_l\Big)\frac{\partial Z}{\partial h_k}\right\} = \\
&= \frac{1}{Z}\Big(\int \mathrm{d}C\; S_l\, S_k\; e^{-\beta\mathscr{H}}\Big) - \frac{1}{Z^2}\left[\Big(\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_l\Big)\frac{\partial Z}{\partial h_k}\right] = \\
&= \frac{1}{Z}\Big(\int \mathrm{d}C\; S_l\, S_k\; e^{-\beta\mathscr{H}}\Big) - \frac{1}{Z}\Big(\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_l\Big)\; \frac{1}{Z}\Big(\int \mathrm{d}C\; e^{-\beta\mathscr{H}} S_k\Big) = \\
&= \langle S_l\, S_k\rangle - \langle S_l\rangle\langle S_k\rangle \equiv \langle S_l\, S_k\rangle_C
\end{aligned}$$

$$\boxed{\chi_{\mu\nu} = \langle S_l\, S_k\rangle_C} \tag{1.12}$$

This equation is a particular case of a more general theorem, called the linear response theorem, which states that the linear response of a system is proportional to the random thermodynamical fluctuations.

1.2.3 Spontaneous symmetry breaking

As we've seen, ferromagnetic system have spontaneous magnetization below the Curie temperature. If a system in zero external field has a magnetization in a given direction, this magnetization will not change with time. However if we apply an external magnetic field and then remove it, we can have the same system with a time-indipendent spontaneous magnetization oriented in a different direction. In other words the system may stay in two (or more) states characterized by the direction of the magnetization; however in the presence of an external magnetic field the magnetization will be oriented in the direction of the external field.

From the microscopic point of view it is clear that for each configuration C there is another configuration C' such that C and C' have the same energy but opposite magnetization, i.e. $S(C) = -S(C')$. The configuration C' con be obtained from C by inverting the sign of all the velocities and the spins of all the particles of the system.

When spontaneous magnetization is present, the Hamiltonian of the system is invariant under a transformation that changes the sign of the magnetization (the two configurations have the same energy); however, the state of the system, i.e. its probability distribution, is not. When this happens we say that the symmetry of the system is spontaneously broken (in the presence of an external magnetic field it would be explicitly broken).

At first sight spontaneous symmetry breaking seems to be in conflict with the fundamental principles of statistical mechanics: if $\mathscr{H}$ is invariant, P must be invariant. This argument is correct for a finite volume, but it fails for an infinite system in which Z is infinite. To explain it, we consider the magnetization density and the free-energy density as functions of h in the infinite-volume limit. The symmetry argument tells us that $f(h) = f(-h)$, but $m(h) = -\partial f/\partial h = -m(-h)$, i.e. f is an even function of h while m is an odd function. We have now two possibilities:

a) $m(h)$ is continuous at $h = 0$, so $m(0) = 0$. In this case we have $f(h) = f(0) + o(h^\alpha), \qquad \alpha > 1$

b) $m(h)$ is discontinuous, then $m(0)$ is not defined. So we have $f(h) = f(0) - m_s|h| + o(h^\alpha)$.

In the case a), $f(h)$ is differentiable with respect to h and $m(h)$ is continuous. In this case the magnetization at $h = 0$ is really zero.
In the case b), $f(h)$ is not differentiable at $h = 0$ and $m(h)$ is a discontinuous function:

$$m^+ \equiv \lim_{h\to 0^+} m(h) = m_s \neq m^- \equiv \lim_{h\to 0^-} m(h) = -m_s \tag{1.13}$$

The symmetry relation implies only $m^+ = -m^-$. In this second case, an infinitesimal magnetic field is enough to produce a nonzero magnetization. If this happens, we say that the system has a spontaneous magnetization. In the limit $h \to 0$ the Hamiltonian becomes invariant under the transformation $S \to -S$, but the statistical expectation values are not invariant under the same transformation, so we say that the symmetry is spontaneously broken. For most ferromagnets, this happens at low temperature at least in three dimensions.
It is important to underline that $m_\Omega(h = 0) = 0$ is indeed correct for a *finite* system (Ω is the sample region). It fails in the thermodynamic limit, however, because $f(h)$ can develop a discontinuity in its first derivative $\partial f/\partial h$ (as in the case b). We know that $f(h)$ is a convex up function, i.e.

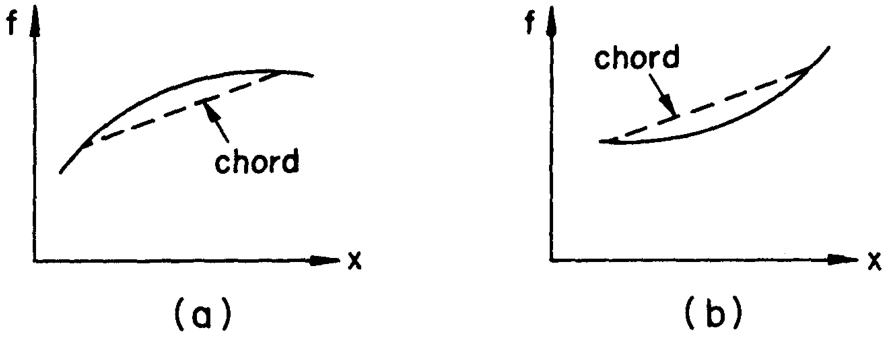


Figure 1.5. (a) A convex up function. (b) A convex down function.

$$f(\alpha\beta_1 + (1-\alpha)\beta_2) \geq \alpha f(\beta_1) + (1-\alpha) f(\beta_2)\,. \tag{1.14}$$

The proof of this is rather simple: in a finite volume, where $f(\beta)$ is an analytic function, its second derivative with respect to any parameter is always negative because it can be written as the expectation value of minus a perfect square, e.g., $\frac{\mathrm{d}^2 f}{\mathrm{d}\beta^2} = -\frac{1}{\beta V}\langle HH\rangle_c$. It is well known that a function whose second derivative is never positive is convex. The condition $f(h) = f(-h)$ does *not* imply $m(h = 0) = 0$ unless we make the additional assumption that $f(h)$ is smooth at $h = 0$ and the left and right derivatives are equal.

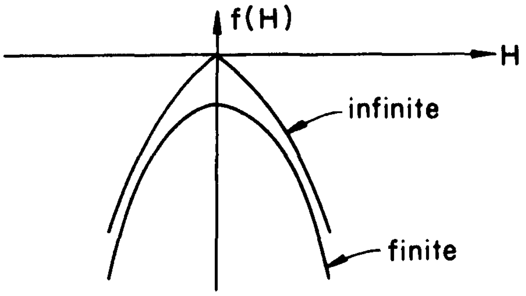


Figure 1.6. The free energy density as a function of magnetic field h for finite and infinite systems, for $T < T_c$.

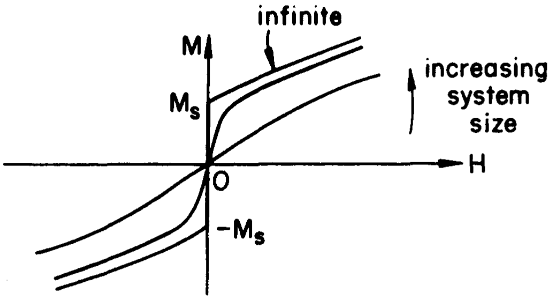


Figure 1.7. Magnetisation m plotted against external field h.

Notice also that the limits

a) $N(\Omega) \to \infty$

b) $h \to 0$

do *not* commute:

$$\lim_{N(\Omega)\to\infty} \lim_{h\to 0} \frac{1}{N(\Omega)} \frac{\partial F_\Omega(h)}{\partial h} = 0 \tag{1.15}$$

whereas

$$\lim_{h\to 0} \lim_{N(\Omega)\to\infty} \frac{1}{N(\Omega)} \frac{\partial F_\Omega(h)}{\partial h} \neq 0 \tag{1.16}$$

Notice also that the value of the spontaneous magnetisation m_s, is a function of temperature: at zero temperature, we already argued that m_s should be unity, because the spins will be either all up or all down. As the temperature rises towards T_c, the value of the spontaneous magnetisation is reduced, as an increasingly greater fraction of spins are flipped by thermal fluctuations. At T_c, the spontaneous magnetisation has fallen to zero, as depicted in Fig.(1.8).

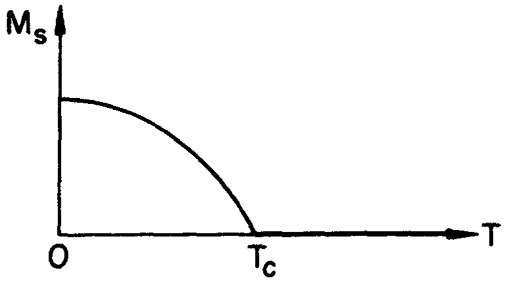


Figure 1.8. Spontaneous magnetisation as a function of temperature.

1.2.4 First-order phase transitions

The nondifferentiability of the free energy with respect to a parameter and the existence of two or more equilibrium states is the distinctive feature of a first-order transition. According to Ehrenfest classification, if the free energy is differentiable $(k-1)$ times, but not k times, the singularity point is called a transition of order k (note that the free energy is always continuous)[1]. However, it is customary to call any transition of order greater than the first a second-order transition.
If a first-order transition is induced by changing the temperature, we are in the presence of a latent heat:

$$u_+ \equiv \lim_{T\to T_c^+}\left(\frac{\partial(\beta f)}{\partial\beta}\right) \neq u_- \equiv \lim_{T\to T_c^-}\left(\frac{\partial(\beta f)}{\partial\beta}\right) \tag{1.17}$$

It is crucial to understand that phase transitions (and by consequence spontaneous symmetry breaking) are possible only if the volume of the system is infinite.
Indeed, assume to have a finite-volume system, with a non-singular Hamiltonian. If the configuration space is compact, it is easy to see that Z is an entire function:

$$|Z(\beta)| = \left|\int \mathrm{d}C\ \exp(-\beta\mathscr{H}(C))\right| \leq V_c \exp\Big(|\beta|\max(|\mathscr{H}(C)|)\Big) \tag{1.18}$$

Where $V_c = \int \mathrm{d}C$ is the volume of the configuration space. This is valid even for complex β.
If the configuration space is not compact, but the Hamiltonian is well defined so that $Z(\beta)$ is finite (i.e. the integral is convergent) for any β real and positive, the same integral must exist for any complex β such that $\mathrm{Re}(\beta)$, because $|Z(a+ib)| \leq Z(a)$.
We have proved that the partition function of a reasonable Hamiltonian Z (such that $Z(\beta)$ is defined for positive β) is an analytic function in the positive half of the complex β plane. A similar result may be obtained for Z as a function of the magnetic field.

What happens to the free energy, $f = -(1/\beta V)\ln(Z(\beta))$? If Z is analytic, its logarithm may be singular only at the points where $Z = 0$. But for real values of the parameters, Z is the sum of positive terms, so it cannot have zeros, consequently the free energy is analytic near the positive real axis for a finite-volume system. However, in the infinite-volume limit a nonanalytic free energy may be produced if the complex zeros of the partition function pinch the real parameter axis. To show how this can work, just consider a toy model, given by a system of volume V, which can stay in only two states of equal energy at zero magnetic field, having total magnetization $\pm V$ respectively. We can easily compute the partition function and then the free energy as functions of the magnetic field h:

$$Z(h) = 2\cosh(\beta h V), \qquad f(h) = -\frac{1}{\beta V}\ln[2\cosh(\beta h V)] \Rightarrow m(h) = \tanh(\beta h V) \tag{1.19}$$

And the partition function is zero for

$$Z(h) = 0 \quad \Rightarrow \quad h = \frac{(2n+1)\pi i}{2\beta V} \tag{1.20}$$

---

[1] The free energy is a convex function. As we have just said, the limit of differentiable functions may not be differentiable, but the limit of convex functions must be convex (convexity is written under the form of an inequality). It is also well known that convexity implies continuity.

When V goes to infinity, the zeros of Z pinch the real h axis at $h = 0$, and the expectation value of the magnetization becomes $m = \text{sign}(h)$.

In a real macroscopic system of finite extent the internal energy will change a finite amount when the temperature changes by 10^{-23} K, and for all practical purposes we are in the presence of a discontinuity.

Chapter 2

The Low-Temperature and High-Temperature Expansion

2.1 Peierl's Argument

In 1936 R. Peierls published an article with the title On the Model of Ising for the Ferromagnetism in which he proved that the Ising model in two or higher dimensions has a low-temperature region in which the spontaneous magnetization is different from zero. Since at high temperature the system is disordered, it follows that there must exist a critical value of the temperature at which a phase transition takes place. Peierls's argument starts with the initial observation that to each configuration of spins there corresponds a set of closed lines that separate the regions in which the spins assume values +1 from those in which they assume values -1, as shown in Fig. 4.1. If it is possible to prove that at sufficiently low temperatures the mean value of the regions enclosed by the closed lines is only a small fraction of the total volume of the system, one has proved that the majority of the spins is prevalently in the state in which there is a spontaneous magnetization. There are several versions of the original argument given by Peierls. The simplest generalizes the argument already used in the one-dimensional case and concerns the stability of the state with a spontaneous magnetization. Let's consider the two-dimensional Ising model at low temperatures and suppose that it is in the state of minimal energy in which all the spins have values +1. The thermal fluctuations create domains in which there are spin flips, such as the domain in Fig. 4.1. The creation of such domains clearly destabilizes the original ordered state. There is an energetic cost to the creation of the domain shown in Fig. 4.1, given by

2.2 The low-temperature expansion

Chapter 3

The Landau-Ginsburg Model

3.1 Motivations

The statistical models we have analyzed so far are defined on a lattice and they have a microscopic length-scale given by the lattice spacing a. In all these models there is, however, another length-scale provided by the correlation length: this is a function of the coupling constants and can be varied by varying the external parameters of the systems. When the system is sufficiently close to its critical point, the correlation length is much larger than the microscopic scale, $\xi \gg a$. It is then natural to assume that the configurations of the system are sufficiently smooth on many lattice spacings and to adopt a formalism based on continous quantities like a field $\varphi(x)$.

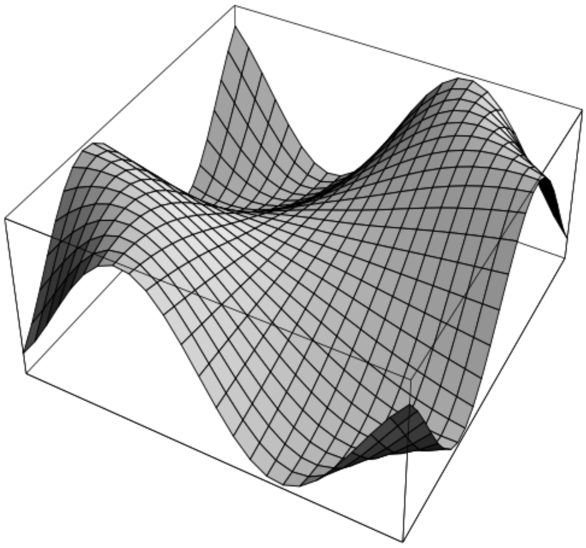


Figure 3.1. Continous formulation in terms of a field theory.

At the transition point the singular terms in the thermodynamic functions are produced by fluctuations on large scales (i.e., at small momenta in Fourier space); the fine details of the probability distribution of the magnetization on a small scale seen not to be very important for deriving general properties, such as the value of the critical exponents. We introduce as a relevant variable the field $\varphi(x)$,[1] which is the average of S_i around the point x (e.g., $\varphi(x) \propto \sum_i \exp[-(x-i)^2/2\delta^2]\, S_i$). The radius δ over which the microscopic magnetization S_i is averaged must be large enough, that $\varphi(x)$ is a sufficiently smooth function of x, but it should not be larger than the correlation lenght, otherwise trivial results are obtained; in particular, if we want to keep a simple form of the effective action, δ should not diverge at the transition point; δ equal to a few lattice spacings is likely to be a good choice. As we will show in the sequel, the quantum field theory formulation of statistical models has the important advantage of greatly simplifying the study of critical phenomena: it helps us to select the most important aspects of phase transitions – those related to the symmetries and the dimensionality of the system – and to reach results of great generality. The reason why QFT plays a central role both in the context of elementary particles and critical phenomena is due, in a nutshell, to the principle of universality. This is a primary aspect of all local interactions and it is noteworthy that it naturally emerges from the analysis of the renormalization group.

3.2 Construction of the Effective Hamiltonian

Let's start our discussion with the functional formalism of the euclidean QFT that is at the root of the continuous formulation of statistical models. This formalism relies on the possibility to substitute the sum over the classical discrete variables $\{s_i\}$ in terms of a functional integral on the continuous variables $\varphi(x)$, also classical. This happens near a phase transition point, when the correlation length ξ is much larger than the lattice spacing a:

$$Z = \sum_{\{s_i\}} e^{-H(\{s_i\})} \simeq \int \mathscr{D}\varphi(x)\, e^{-S(\{\varphi\})}, \qquad \xi \gg a. \tag{3.1}$$

Let's comment on this expression. The first problem that arises in the functional approach is the identification of the order parameter of the statistical system. As already discussed previously, to solve this problem one has to rely on the symmetry of the hamiltonian and on some physical intuition. For instance, in the presence of a Z_2 symmetry, the role of the order parameter can be played by a scalar quantity $\varphi(x)$ that takes values on all of the real axis, odd under the Z_2 transformation, $\varphi(x) \to -\varphi(x)$. For a system that is instead invariant under $O(n)$ symmetry, just to make another example, one can take as order parameter a field with n components $\Phi(x) = [\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)]$ that transforms as a vector under the $O(n)$ transformations.

3.2.1 The Boltzmann weight

Once the order parameter is identified, one needs next to introduce the Boltzmann weight associated to its different configurations. Only in this way, in fact, can one further proceed to compute statistical averages, correlation functions, and all the other thermodynamic quantities. In analogy with what was done for the statistical systems defined on a lattice, and remembering the path integral expression of the partition function, the probability of the field configuration can be assumed to be proportional to

$$W(\varphi, \{g\}) = e^{-S(\{\varphi, \{g\})} = \exp\left\{-\int \mathrm{d}^D x\, \mathscr{L}(x)\right\}, \tag{3.2}$$

---

[1] The field φ is a *continuous random variable* and can assume all the values $-\infty < \varphi < +\infty$.

where S is the *euclidean action* of the theory, given by an integral on the euclidean lagrangian density $\mathscr{L}(x)$. The latter is a local quantity, generically expressed in terms of a polynomial of the fields and their derivatives. To simplify the notation, in the following we focus our attention on a QFT of a scalar field $\varphi(x)$, odd under the Z_2 symmetry. In this case, restricting attention to those terms that are at most of degree 2 in the derivatives,[2] the most general expression of the action is given by

$$S = \int \mathrm{d}^D x \, \left[\frac{1}{2} (\partial_j \varphi)^2 + g_1 \varphi + \frac{g_2}{2!} \varphi^2(x) + \cdots + \frac{g_n}{n!} \varphi^n(x) + \dots \right] \tag{3.3}$$

In D-dimensional euclidean space, the definition of the derivative term is meant to be a sum over the repeated indices

$$(\partial_j \varphi)^2 \equiv (\partial_j \varphi)(\partial_j \varphi) = \sum_{j=1}^{D} \left(\frac{\partial \varphi}{\partial x_j} \right)^2 = |\nabla \varphi(x)|^2 . \tag{3.4}$$

It is also convenient to introduce the concept of the *manifold of the coupling constants*, defined as the space spanned by the set of all couplings $\{g\} = (g_1, \dots, g_n, \dots)$. Once the lagrangian is given, the partition function of the system is obtained by summing up all possible configurations of the order parameter

$$Z = \int \mathscr{D}\varphi(x) \, \exp\left[-S(\varphi, \{g\})\right] \tag{3.5}$$

In writing this expression we have emphasized that the partition function depends both on the coupling constants g_i and the *microscopic cut-off* a provided by the lattice spacing of the original theory. Even if we have adopted a continuous formalism to describe a statistical model, it is in fact necessary to take into account the microscopic scales of the systems, and we will see later several effects of such a dependence. Notice that an obvious reason to introduce the microscopic scale a is related to the definition of the measure $\mathcal{D}\varphi$: with this notation we mean a measure on all possible values of the field $\varphi(x)$. Since φ is a continuous quantity defined on each point of the space, $\mathcal{D}\varphi$ is not a priori well-defined. In order to make sense of it, one can proceed in two equivalent ways.

3.2.2 The Measure

The first approach to define a measure consists of considering the field as a collection of discrete quantities φ_i, defined only on N sites of a lattice with spacing length a, so that $\mathcal{D}\varphi$ can be expressed as a product of the differentials of all these variables, whose number can be enormously large but in any case finite:

$$\mathscr{D}\varphi = \prod_i^N \mathrm{d}\varphi_i \, . \tag{3.6}$$

The second equivalent approach makes use of the translation invariance of the system. This invariance allows us to decompose the field into its Fourier components

$$\varphi(x) = \frac{1}{\sqrt{N}} \sum_k \varphi(k) e^{ikx} \, . \tag{3.7}$$

When N is finite, the frequencies are discrete. Futhermore, in the presence of a microscopic scale a, they satisfy the condition

$$|k| \leq \Lambda \simeq \frac{1}{a} \, . \tag{3.8}$$

---

[2] This can be justified by demanding the causality of the theory.

The lattice space a acts then as an ultraviolet cut-off. This turns out to be a very useful quantity, since it permits us also to regularize the divergent terms coming from the perturbative formulation of the theory. In the second approach the measure $\mathscr{D}\varphi$ is also given by the differential of a finite number of variables:

$$\mathscr{D}\varphi = \prod_{0 \leq |k| \leq 1/a}^{N} \mathrm{d}\varphi(k) \,. \tag{3.9}$$

Notice that in both cases the problem to control the behavior of $\mathcal{D}\varphi$ when $N \to \infty$, or, equivalently, $a \to 0$ still remains open. This is a problem not only of the measure but of the entire quantum field theory.

3.2.3 Engineering dimensions

As a matter of fact, the ultraviolet cut-off a also enters other key aspects. Consider, for instance, the engineering dimensions of the coupling constants in the action. To determine such quantities, it is necessary to fix initially the dimension of the scalar field φ. Since the action is a dimensionless quantity, each term of the lagrangian should have dimension a^{-D}. Consider then the kinetic term $(\partial_j \varphi)^2$: imposing the dimension of the field equal to $[\varphi] = a^{x_\varphi}$, we have the condition $a^{-2} a^{2x_\varphi} = a^{-D}$ and therefore

$$[\varphi] = a^{1-D/2} \tag{3.10}$$

Once the dimension of $\varphi(x)$ is known, it is easy to obtain the dimensions of the various coupling constants

$$[g_m] = a^{mD/2-m-D} \equiv a^{\delta_m} \,. \tag{3.11}$$

It is interesting to observe that each coupling constant has a particular dimension $D_s^{(m)}$ (the so-called *upper critical dimension*) in which it is dimensionless. For instance g_3 is dimensionless for $D = 6$, g_4 for $D = 4$, and so on. Notice that the quantity δ_m is positive when

$$D \geq D_s^{(m)} = \frac{2m}{m-2} \,. \tag{3.12}$$

3.2.4 Critical behavior

On the basis of the information above, we can already formulate some educated guesses on the critical behavior of the theory – guesses that need however to be refined by further analysis. For a lagrangian with higher coupling constant given by gn, the corresponding statistical theory is expected to present two different regimes by varying D:

- for $D \geq D_s^{(m)}$, the critical behavior is expected to be described by the mean field theory, with a classical value for the critical exponents;
- for $D \leq D_s^{(m)}$ the system is instead expected to present strong fluctuations with a corresponding significant change of its thermodynamic singularities.

The simplest way to understand these two different critical behaviors is to study the sign of the exponent δ_m: when $\delta_m > 0$ (i.e. $D \geq D_s^{(m)}$), sending to zero the lattice space a, the corresponding coupling constant becomes smaller, while when $\delta_m < 0$ ($D \geq D_s^{(m)}$) the coupling constant becomes larger. Consequently, for what concerns the critical behavior, in the first case the microscopic fluctuations are expected to be irrelevant while in the second case to be relevant. Anticipating the results and the terminology of the renormalization group that will be discussed in the next chapter, the coupling constants gn with $\delta_m > 0$ are called *irrelevant*, those with $\delta_m < 0$ are called *relevant*, and, finally, those with $\delta_m = 0$, *marginal.* The previous analysis was carried out for a theory invariant under a Z_2 symmetry but the same scenario holds for other theories with different internal symmetry. Namely, each theory has a lower critical dimension D_i, below which there is no longer a phase transition, and an upper critical dimension D_s, beyond which the critical exponents take classical values. The strong fluctuation regime of the order parameters is expected to occur in between, i.e. in the range of dimensions D satisfying

$$D_i \leq D \leq D_s \tag{3.13}$$

For systems with short-range interactions and a discrete symmetry, such as the Ising or the Potts models, the lower critical dimension is always $D_i = 1$, whereas for those with a continuous symmetry, such as the $O(n)$ model, $D_i = 2$. In the range (7.2.11) the critical exponents assume values that are different from their mean field solution and their determination requires more sophisticated theoretical tools.

3.3 The Landau-Ginsburg Model

It is much more convenient to take into account only a little part of the more general action previously written

$$\boxed{H_{ef}[\varphi] = S_{LG} = \int \mathrm{d}^D x \, \left[\frac{1}{2}\left(\partial_j \varphi\right)^2 + \frac{\mu}{2}\varphi^2(x) + \frac{g}{4!}\varphi^4(x) + h(x)\,\varphi(x) + \frac{1}{2\Lambda^2}\left(\Delta\varphi(x)\right)^2\right]}\,, \tag{3.14}$$

where we have set $\mu = g_2 = m^2$ the mass parameter,[3] and $g = g_4$. Note that the parameter g must be positive if the integral in (3.1) is to converge, since $\lim_{\varphi \to \pm\infty} H_{ef}[\varphi] = +\infty$; but the sign of μ is arbitrary. In this expression we have added the coupling $h(x)\,\varphi(x)$ to an external applied $h(x)$-field. This choice of the action defines the *Landau-Ginsburg model.* The form of $H_{ef}[\varphi]$ is not chosen at random. If $\mu > 0$, in fact, the minimum of the effective Hamiltonian is just at $\varphi = 0$; on the contrary, for $\mu < 0$, $H_{ef}[\varphi]$ has *two* minima, suggesting the possibility of a broken simmetry. Expliciting the calculus with $h(x) = 0$ we have

$$\begin{cases} \mu\varphi + \dfrac{g}{6}\varphi^3 = 0 \\ \mu + \dfrac{g}{2}\varphi^2 > 0 \end{cases} \implies \varphi = \begin{cases} 0, & \mu > 0 \\ \pm\sqrt{-\dfrac{6\mu}{g}}, & \mu < 0\,. \end{cases} \tag{3.15}$$

If we approximate the free energy with the minimum of $H_{ef}[\varphi]$,[4] , i.e.

$$F = \min_{\{\varphi\}} H_{ef}[\varphi] \tag{3.16}$$

---

[3] In the canonical quantization of the theory, m can indeed be identified with the mass of the particle created by the field $\varphi(x)$.

[4] This approximation is good when the entropy of the system is small.

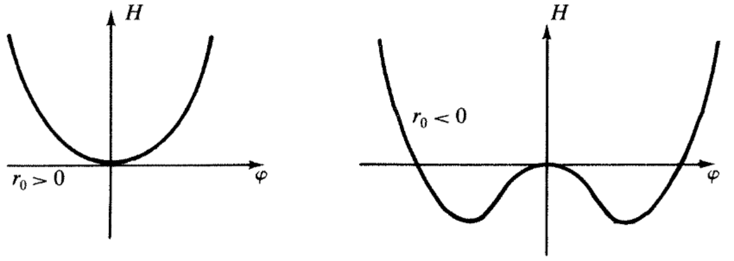


Figure 3.2. $r_0 = \mu$.

we find that spontaneous magnetization is present as soon as $\mu < 0$. In this approximation the critical temperature T_c is at $\mu = 0$; we therefore have

$$\mu = \alpha\,(T - T_c) + o(T - T_c)^2\ , \tag{3.17}$$

where α is a constant. How it has been already said, $g > 0$, otherwise the functional integral would diverge. It must be nonzero or we would recover the Gaussian model, which is not definite for $T < T_c$. The term $(\partial_j\varphi)^2$ describes a ferromagnetic interaction, so its coefficient must be positive, since the effective energy must increase (and consequently the probability decrease) if φ changes with x (i.e., if spins are less aligned). The term $\frac{1}{2\Lambda^2}\,(\Delta\varphi(x))^2$ is very large when the function φ is not smooth. It is needed in two or higher dimensions to suppress those configurations for which φ changes too fast. Of coure, the term $(\partial\varphi)^2$ also suppresses these configurations, but it is not so efficient as the term proportional to $1/(2\Lambda^2)$. As we shall see, at $\Lambda = \infty$, $\langle\varphi^2\rangle$ is divergent as soon as $D \geq 2$.

It should be clear that if the field φ is the average value of the magnetization, the true effective Hamiltonian would be much more complex. The main advantage of the Landau-Ginsburg model is its extreme simplicity: no term may be removed from it without changing the physics of the problem in a deep way. We shall see (Parisi 11.3) that the insertion of other terms in the effective Hamiltonian does not change the results for the critical exponents: From our point of view these other terms are irrelevant. It is normally assumed that the parameters g, Λ, and μ are smooth functions of the temperature near the critical point; since the change in sign of μ drives the transition in the approximation (3.16), the variation of μ is the most important. In computing the singular terms of the thermodynamic functions, it is therefore sufficient to consider g, Λ, and β as constants and to concentrate all the temperature dependence in μ. In approximation (3.16), the transition point corresponds to $\mu = 0$; if this approximation is removed, the critical value of μ (μ_c) will change. Then the previous equation is replaced by

$$\mu - \mu_c = \alpha\,(T - T_c)\ , \tag{3.18}$$

where we have neglected the higher orders in $T - T_c$.

3.3.1 Discrete Lattice

We can also write the effective Hamiltonian defining the field φ only on the points of a lattice (which may or may not coincide with the original lattice). The effective moldel Hamiltonian then becomes

$$H_{ef}[\varphi] = a^D \left[\sum_{i,k} \frac{1}{2a^2} J_{ik} \left(\varphi_i - \varphi_k\right)^2 + \sum_i \left(\frac{\mu}{2}\,\varphi_i^2 + \frac{g}{4!}\,\varphi_i^4 + h_i \varphi_i \right) \right] , \tag{3.19}$$

where a is the lattice spacing and where we have introduced the matrix J_{ik}; this matrix is equal to J if i and k are nearest neighbors and it is zero elsewhere and it is zero elsewhere. The introduction of a lattice clarifies the mathematical structure by reducing the problem to a finite number of degrees of freedom. The continuum limit can be considered as the limit $a \to 0$ of the lattice theory.

3.4 The Gaussian Model Again

3.4.1 Diagonalization of a matrix invariant under translations

Consider a (real) matrix A_{st}, $0 \leq s, t \leq N-1$, such that A_{st} depends only on the difference $(s-t)$. Assume periodic boundary conditions: $p+N \equiv p$ for al integer p. The matrix A_{st} is diagonalized by a Fourier trasformation on a lattice, corresponding to the unitarity trasformation $U_{sq} = \frac{1}{\sqrt{N}} e^{iqx_s}$, with

$$x_s = sa; \quad q = \frac{2\pi p}{Na} , \tag{3.20}$$

where s and p are integers between 0 and $N-1$, and a is the lattice spacing. This result is well known, for instance in the theory of vibrational normal modes. We rederive it briefly:

$$A_{qq'} = \frac{1}{N} \sum_{s,t} e^{iq'x_s} A_{st} e^{-iqx_t} = \frac{1}{N} \sum_s e^{-i(q-q')x_s} \sum_t A(s-t)\, e^{-iq(x_t - x_s)} . \tag{3.21}$$

The sum over t is independent of s, by virtue of the condition $q = 2\pi p/Na$, and of the periodic boundary conditions. Call this sum $\widetilde{A}(q)$; we obtain

$$A_{qq'} = \delta_{qq'}\, \widetilde{A}(q) . \tag{3.22}$$

In general $A(s-t)$ depends only on $|s-t|$, and $\widetilde{A}(q)$ is real. Because the trace is invariant under a similarity trasformation, we have

$$\mathrm{Tr}\, A_{st} = \sum_q \widetilde{A}(q) . \tag{3.23}$$

To take the continuum limit in proves convenient to define $\widetilde{A}(q)$ by

$$\widetilde{A}(q) = a \sum_t A(s-t)\, e^{-iq(x_t - x_s)} \to \textstyle\int \mathrm{d}x\ A(x)\, e^{iqx} , \tag{3.24}$$

where we have gone from a Riemann sum to an integral. Similarly, the sum over q can be replaced by an integral on applying

$$\sum_q \to \frac{Na}{2\pi} \int \mathrm{d}q , \tag{3.25}$$

appropriate because successive values of q are spaced by $2\pi/Na$. On the other hand, instead of the limits $0 \leq q \leq 2\pi/a$, it is generally more convenient to choose

$$-\frac{\pi}{a} \leq q \leq \frac{\pi}{a} . \tag{3.26}$$

Under these conditions the inverse Fourier trasform is given by

$$A(x) = \frac{1}{Na} \sum_q e^{-iqx} \widetilde{A}(q) \;\rightarrow\; \int_{-\pi/a}^{\pi/a} \frac{\mathrm{d}q}{2\pi} \widetilde{A}(q)\, e^{-iqx} \,, \tag{3.27}$$

and the trace relation becomes

$$\mathrm{Tr}\, A_{st} = Na \int \frac{\mathrm{d}q}{2\pi} \widetilde{A}(q) \,. \tag{3.28}$$

Note that because we are in one dimension $Na = V$, where V is the volume of the system. In D dimensions these equations generalize to

$$\begin{aligned} \widetilde{A}(q) &= \int \mathrm{d}^D x\; e^{iq\cdot x} A(x) \\ A(x) &= \int \frac{\mathrm{d}^D q}{(2\pi)^D}\; e^{-iq\cdot x} \widetilde{A}(q) \\ \mathrm{Tr}\, A(x) &= V \int \frac{\mathrm{d}^D q}{(2\pi)^D}\; \widetilde{A}(q) \,. \end{aligned} \tag{3.29}$$

3.4.2 The free energy

If we set $g = 0$ we recover the Gaussian model; The partition function can be rewritten with several integration by parts in this way

$$\begin{aligned} Z &= \int \mathscr{D}\varphi(x)\, \exp\left(-\int \mathrm{d}^D x \left[\frac{1}{2} (\nabla\varphi)^2 + \frac{1}{2} m^2 \varphi^2 + \frac{1}{2\Lambda^2} (\Delta\varphi(x))^2 + h(x)\varphi(x) \right] \right) \\ &= \int \mathscr{D}\varphi(x)\, \exp\left(-\frac{1}{2} \int \mathrm{d}^D x\, \varphi(x) \left[-\partial^2 + m^2 + \frac{1}{\Lambda^2} \partial^4 \right] \varphi(x) - \int \mathrm{d}^D x\, h(x)\varphi(x) \right) \\ &= \int \mathscr{D}\varphi(x)\, \exp\left(-\frac{1}{2} \int \mathrm{d}^D x\, \varphi(x)\, G_\Lambda^{-1}\, \varphi(x) - \int \mathrm{d}^D x\, h(x)\varphi(x) \right) \,, \end{aligned} \tag{3.30}$$

where $\partial^2 = \partial_j \partial_j = \Delta$ and $\partial^4 = \partial_j^2\, \partial_j^2 = \Delta^2$ which simbolically indicates the sum of 4-order partial derivatives. G_Λ^{-1} now plays the role of A_{ij}^{-1}. The functional integral over $\varphi(x)$ is now a Gaussian integral

$$Z = \frac{1}{(\det G_\Lambda)^{1/2}}\, \exp\left\{ \frac{1}{2} \int \mathrm{d}^D x\, \mathrm{d}^D y\; h(y)\, G_\Lambda(x-y)\, h(x) \right\} \,, \tag{3.31}$$

where $G_\Lambda(x-y)$ is the correlation function of the Gaussian model (with the factor Λ in the probability distribution). To compute its expression note that $G_\Lambda(x-y)$ is, formally, the inverse of $\left[-\partial_x^2 + m^2 + \frac{\partial_x^4}{\Lambda^2} \right]$, that means that

$$\left[-\partial_x^2 + m^2 + \frac{\partial_x^4}{\Lambda^2} \right] G_\Lambda(x-y) = \delta^D(x-y) \,; \tag{3.32}$$

using the Fourier representation of the correlation function and the delta we obtain

$$\begin{aligned} \left[-\partial_x^2 + m^2 + \frac{\partial_x^4}{\Lambda^2} \right] \int \frac{\mathrm{d}^D p}{(2\pi)^D}\, \widetilde{G}_\Lambda(p)\, e^{-ip\cdot(x-y)} &= \int \frac{\mathrm{d}^D p}{(2\pi)^D} \left[p^2 + m^2 + \frac{p^4}{\Lambda^2} \right] \widetilde{G}_\Lambda(p)\, e^{-ip\cdot(x-y)} \\ &= \int \frac{\mathrm{d}^D p}{(2\pi)^D}\, e^{-ip\cdot(x-y)} \,, \end{aligned} \tag{3.33}$$

da cui segue che

$$\boxed{\widetilde{G}_\Lambda(p) = \frac{1}{p^2 + m^2 + \frac{p^4}{\Lambda^2}} \quad \Longrightarrow \quad G_\Lambda(x-y) = \int \frac{\mathrm{d}^D p}{(2\pi)^D} \, \frac{e^{-ip\cdot(x-y)}}{p^2 + m^2 + \frac{p^4}{\Lambda^2}}} \,. \tag{3.34}$$

Now the free energy assumes the form

$$\begin{aligned} F = -\ln Z &= \frac{1}{2}\ln(\det G_\Lambda) - \frac{1}{2}\int \mathrm{d}^D x\, \mathrm{d}^D y \; h(x)\, G_\Lambda(x-y)\, h(y) \\ &= \frac{1}{2}\,\mathrm{Tr}\,(\ln G_\Lambda) - \frac{1}{2}\int \mathrm{d}^D x\, \mathrm{d}^D y \; h(x)\, G_\Lambda(x-y)\, h(y) \end{aligned} \tag{3.35}$$

where we used the identity $\ln(\det G_\Lambda) = \mathrm{Tr}\,(\ln G_\Lambda)$. Now using the third of the (3.29) we obtain

$$\boxed{F = \frac{V}{2(2\pi)^D}\int \mathrm{d}^D p \; \ln \widetilde{G}_\Lambda(p) - \frac{1}{2}\int \mathrm{d}^D x\, \mathrm{d}^D y \; h(x)\, G_\Lambda(x-y)\, h(y)} \,. \tag{3.36}$$

A similar formula holds on the lattice

$$\begin{aligned} &F_L = \frac{V}{2(2\pi)^D}\int_B \mathrm{d}p \; \ln \widetilde{G}_L(p) - \frac{1}{2}a^2 \sum_l \sum_k h_l h_k G_L(l-k) \\ &G_L(k) = \frac{1}{(2\pi)^D}\int_B \mathrm{d}p \; \widetilde{G}_L(p)\, e^{iap\cdot k} \\ &\widetilde{G}_L(p) = \frac{1}{\mu + \frac{1}{a^2}\sum_\nu \left(2 - 2\cos(p_\nu a)\right)} \qquad \int_B \mathrm{d}p = \int_{-\pi/a}^{\pi/a} \prod_{\nu=1}^{D} \mathrm{d}p_\nu \,. \end{aligned} \tag{3.37}$$

3.4.3 Large distance behavior of the free propagator

We return to

$$G(r) = \int \frac{\mathrm{d}^D p}{(2\pi)^D} \, \frac{e^{-ip\cdot r}}{p^2 + m^2} \tag{3.38}$$

to find its behaviour at large distance. If we write

$$\frac{1}{p^2 + m^2} = \int_0^\infty \mathrm{d}\alpha \; e^{-\alpha\left(p^2 + m^2\right)} \,, \tag{3.39}$$

the integrations over p are Gaussian and we find

$$G(r) = \int \frac{\mathrm{d}^D p}{(2\pi)^D} \int_0^\infty \mathrm{d}\alpha \; e^{-ip\cdot r - \alpha\left(p^2+m^2\right)} = \frac{1}{(4\pi)^{D/2}}\int_0^\infty \frac{\mathrm{d}\alpha}{\alpha^{D/2}} \; \exp\left\{-\left(\frac{r^2}{4\alpha} + m^2\alpha\right)\right\} \,. \tag{3.40}$$

If we rescale $\alpha \to \frac{r}{2m}\alpha$, we obtain

$$G(r) = \frac{1}{(4\pi)^{D/2}}\left(\frac{2m}{r}\right)^{D/2-1}\int_0^\infty \frac{\mathrm{d}\alpha}{\alpha^{D/2}} \; \exp\left\{-\frac{mr}{2}\left(\alpha + \frac{1}{\alpha}\right)\right\} \,. \tag{3.41}$$

For mr large the integration over α is governed by the saddle-point α_0 solution of $\frac{\mathrm{d}}{\mathrm{d}\alpha}(\alpha+1/\alpha)=0$, i.e by the point $\alpha_0 = 1$ on the integration contour. If we write $\alpha = 1 + x$, then $\alpha + 1/\alpha \simeq 1+x+1-x+x^2+O(x^3) = 2+x^2+O(x^3)$ and the Gaussian integration[5] over x gives the result

$$G(r) \overset{r\to\infty}{\longrightarrow} \frac{1}{2(4\pi)^{D/2}} \left(\frac{2m}{r}\right)^{D/2-1} e^{-mr} \int_{-\infty}^{+\infty} \mathrm{d}x \, \exp\left\{-\frac{mr}{2}x^2\right\} \sim \frac{e^{-mr}}{r^{(D-1)/2}} \,. \tag{3.42}$$

Altrimenti funzioni di bessel Parisi pag 100.

---

[5] Note that in the limit of $r \to \infty$ the integration over the x can be extended from $\int_{-1}^{\infty}$ to $\int_{-\infty}^{+\infty}$ and dividing this last integral by 2. For large r the integral is governed by the saddle point $\alpha = \pm 1$, but only one of this values contribute to the integral: this is the reason of the 2 factor. This method is a sort of *steepest descendent method.*

Chapter 4

The Perturbation Expansion and Feynman Diagrams

4.1 Wick's Theorem and the generating functional

4.1.1 The generating function with a single variable

Let $P(\varphi)$ be a probability distribution over a random variable; thus $P(\varphi) \geq 0$, but we do not assume that it is normalized, whence $\int P(\varphi)\mathrm{d}\varphi$ need not equal 1. The *generating function* $Z(j)$ is defined by

$$Z(j) = \int \mathrm{d}\varphi \; P(\varphi) e^{j\varphi} \, ; \tag{4.1}$$

we assume tacitly that $P(\varphi)$ falls fast enough at infinity for the integral to converge. One is interested in $Z(j)$ because differentiatig it one can obtain the moments $\langle \varphi^n \rangle$ of the probability distribution $P(\varphi)$:

$$\langle \varphi^n \rangle = \frac{\int \mathrm{d}\varphi \, \varphi^n P(\varphi)}{\int \mathrm{d}\varphi \, P(\varphi)} = \frac{1}{Z(0)} \left. \frac{\partial^n Z}{\partial j^n} \right|_{j=0} . \tag{4.2}$$

Conversely, $Z(j)/Z(0)$ is given as a function of the $\langle \varphi^n \rangle$ by

$$\frac{Z(j)}{Z(0)} = \sum_{n=0}^{\infty} \frac{j^n}{n!} \langle \varphi^n \rangle \, . \tag{4.3}$$

For a Gaussian probability distribution (Gaussian model)

$$P(\varphi) = \exp\left(-\frac{1}{2} \varphi \frac{1}{A} \varphi \right) , \tag{4.4}$$

performing a Gaussian integral, we obtain

$$Z(j) = Z(0) \exp\left(\frac{1}{2} j A j \right) . \tag{4.5}$$

Further (with the index 0 indicating averages over a Gaussian distribution) one has

$$\langle \varphi^{2n} \rangle_0 = \frac{\partial^{2n}}{\partial j^{2n}} \exp\left(\frac{1}{2} j A j \right) \bigg|_{j=0} = \frac{\partial^{2n}}{\partial j^{2n}} \frac{1}{n!} \frac{1}{2^n} (jAj)^n = \frac{(2n)!}{2^n \, n!} A^n = (2n-1)!! \, A^n \, . \tag{4.6}$$

The secondo equality follows on expanding the exponential and noting that only one term of the expansion can contribute, in view of the $2n$ derivatives and the condition $j = 0$; remember also that for a integer n is valid the relation $2^n\, n! = (2n)!!$. All odd moments of course vanish: $\langle \varphi^{2n+1} \rangle_0 = 0$. The latter equation can be rewritten in terms of the second moment $\langle \varphi^2 \rangle_0 = A$

$$\boxed{\langle \varphi^{2n} \rangle_0 = (2n-1)!!\ \langle \varphi^2 \rangle_0^n} . \tag{4.7}$$

If $P(\varphi)$ is not a Gaussian, but assumes say the form

$$P(\varphi) = \exp\left(-\frac{1}{2}\varphi\frac{1}{A}\varphi + f(\varphi)\right) , \tag{4.8}$$

then

$$Z(j) = \int \mathrm{d}\varphi\, \exp\left(-\frac{1}{2A}\varphi^2 + f(\varphi) + j\varphi\right) . \tag{4.9}$$

The generating function can be written as

$$Z(j) = \exp\left(f\left(\frac{\partial}{\partial j}\right)\right) \int \mathrm{d}\varphi\ \exp\left(-\frac{1}{2A}\varphi^2 + j\varphi\right) . \tag{4.10}$$

The equality follows from

$$f\left(\frac{\partial}{\partial j}\right) e^{j\varphi} = f(\varphi)e^{j\varphi} , \tag{4.11}$$

which in turn can be proved through a Taylor expansion of $f\left(\frac{\partial}{\partial j}\right)$ near the origin.

4.1.2 Wick's theorem

By generalizing the argument of the preceding section to N variables we can prove a fundamental theorem for Gaussian integrals which is well known to probabilists. Start from a Gaussian probability distribution for N variables,

$$P(\varphi_1, \ldots, \varphi_N) = \exp\left(\frac{1}{2}\varphi^T A^{-1}\varphi\right) \tag{4.12}$$

where

$$\varphi^T A^{-1}\varphi = \sum_{i,j} \varphi_i A^{-1}_{ij} \varphi_j , \tag{4.13}$$

and define a generating function $Z(j)$,

$$Z(j_1, \ldots, j_N) = \int \prod_{i=1}^{N} \mathrm{d}\varphi_i\ \exp\left(-\frac{1}{2}\varphi^T A^{-1}\varphi + j^T\varphi\right) . \tag{4.14}$$

Performing this multidimensional Gaussian integral we have

$$Z(j) = Z(0)\exp\left(\frac{1}{2}j^T A j\right) , \qquad \text{with } Z(0) = \frac{(2\pi)^{N/2}}{(\det A)^{1/2}} \tag{4.15}$$

which in turn yields the moments of $P(\varphi)$ on differentiations; thus

$$\langle \varphi_{i_1} \dots \varphi_{i_{2n}} \rangle_0 = \frac{\partial^{2n}}{\partial j_{i_1} \dots \partial j_{i_{2n}}} \frac{1}{n!} \frac{1}{2^n} \left(j^T A j \right)^n . \tag{4.16}$$

From this equation we can prove the generalization of (4.6): *all the moments of a Gaussian distribution can be expressed as functions of the second moments alone.* In quantum field theory this result is known as *Wick's theorem.* The proof is straightforward. Start with the second moment

$$\langle \varphi_{i_1} \varphi_{i_2} \rangle_0 = \frac{\partial^2}{\partial j_{i_1} \partial j_{i_2}} \left(\frac{1}{2} \sum_{k,l} j_k A_{kl}\, j_l \right) = A_{i_1, i_2} = \overbracket{\varphi_{i_1} \varphi_{i_2}} . \tag{4.17}$$

The quantity $\overbracket{\varphi_{i_1} \varphi_{i_2}}$ is called the *contraction* of φ_{i_1} and φ_{i_2}. The differentiation of (4.16) yield $(2n)!$ terms; we must however divide by $2^n\, n!$, whence the total number of terms in the right of (4.16) is $(2n-1)!!$. But this is simply the number of ways of choosing the pairs $\overbracket{\varphi_{i_1} \varphi_{i_2}}\, \overbracket{\varphi_{i_3} \varphi_{i_4}} \dots \overbracket{\varphi_{i_{2n-1}} \varphi_{i_{2n}}}$. To see this, note that there are $(2n-1)$ ways of forming the first pair $\overbracket{\varphi_{i_1} \varphi_{i_2}}$, $(2n-3)$ ways of forming the second pair, and so on. Thus one finds

$$\begin{aligned} \langle \varphi_{i_1} \dots \varphi_{i_{2n}} \rangle_0 &= \overbracket{\varphi_{i_1} \varphi_{i_2}}\, \overbracket{\varphi_{i_3} \varphi_{i_4}} \dots \overbracket{\varphi_{i_{2n-1}} \varphi_{i_{2n}}} \\ &\qquad + \text{permutations}, \\ \text{with} \qquad \overbracket{\varphi_{i_1} \varphi_{i_2}} &= \langle \varphi_{i_1} \varphi_{i_2} \rangle_0 = A_{i_1, i_2} . \end{aligned} \tag{4.18}$$

It is important to realize that the total number of terms does not change even if some of the indices are the same. In fact, if all the indices are the same one recovers (4.6).

4.1.3 The generating functional

The results above can be applied to a continuum theory considered as the limit of a theory on a lattice. Take the probability density as given by a Hamiltonian of the Ginzburg-Landau type

$$\boxed{P[\varphi] = \exp\left(-H_{ef}[\varphi]\right) = \exp\left(- \int \mathrm{d}^D x \left(\frac{1}{2} \left(\nabla \varphi\right)^2 + \frac{1}{2} m^2 \varphi^2 + \frac{g}{4!} \varphi^4 \right) \right)} ; \tag{4.19}$$

then the *generating functional* of the correlation functions, $Z[j]$, is defined by

$$Z[j] = \int \mathscr{D}\varphi(x)\, \exp\left(- \int \mathrm{d}^D x \left(\frac{1}{2} \left(\nabla \varphi(x)\right)^2 + \frac{1}{2} m^2 \varphi^2(x) + \frac{g}{4!} \varphi^4(x) + j(x) \varphi(x) \right) \right) . \tag{4.20}$$

The function $j(x)$ is called the source of the field φ; it plays the same role as did the field $h(x)$. The moments of order $2n$, which are nothing but the $2n$-point correlation functions, are found by differentiating the functional $Z[j]$; thus

$$G^{(2n)}\left(x_1, \dots, x_{2n}\right) = \frac{1}{Z[0]} \left. \frac{\partial^{2n} Z[j]}{\partial j(x_1) \dots \partial j(x_{2n})} \right|_{j=0} . \tag{4.21}$$

It is convenient to subdivide the Hamiltonian $H_{ef}[\varphi]$ into a Gaussian part H_0 (quadratic in φ)

$$H_0 = \frac{1}{2} \int \mathrm{d}^D x \left[(\nabla \varphi)^2 + m^2 \varphi^2 \right] = \frac{1}{2} \int \mathrm{d}^D x\, \varphi(x) \left[-\partial^2 + m^2 \right] \varphi(x) , \tag{4.22}$$

plus the so called *interaction term* H_I

$$H_I = \frac{g}{4!}\int \mathrm{d}^D x\, \varphi^4(x)\,. \tag{4.23}$$

More generally one could replace $(g/4!)\varphi^4$ by a polynomial in φ, which must be even if one wishes to respect the symmetry $\varphi \to -\varphi$; further, one could introduce so-called *derivative interactions* like $\varphi^2\,(\nabla\varphi)^2$. However, we shall ignore these interactions for the time being in order to keep the discussion simple. Then, for an interaction

$$H_I = \int \mathrm{d}^D x\, \mathscr{H}_I(\varphi) = \int \mathrm{d}^D x\, \left(\frac{g}{4!}\varphi^4 + \frac{g_6}{6!}\varphi^6 + \dots\right)\,, \tag{4.24}$$

the expression (4.10) generalized to

$$Z[j] = \exp\left(-\int \mathrm{d}^D x\, \mathscr{H}_I\left(\frac{\partial}{\partial j(x)}\right)\right) \int \mathscr{D}\varphi(x)\, \exp\left(-H_0 + \int \mathrm{d}^D x\, j(x)\varphi(x)\right)\,. \tag{4.25}$$

The integral over $\varphi(x)$ is a Gaussian; integrating by parts we obtain

$$\begin{aligned}&\int \mathscr{D}\varphi(x)\, \exp\left(-\int \mathrm{d}^D x\, \left[\frac{1}{2}\,(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2 - j(x)\varphi(x)\right]\right)\\ &= Z_0[j=0]\exp\left(\frac{1}{2}\int \mathrm{d}^D x\, \mathrm{d}^D y\; j(y)\, G_0(x-y)\, j(x)\right)\end{aligned} \tag{4.26}$$

where $G_0(x-y)$ is the two-point correlation function of the Gaussian model ($G_0(x,y)$ now plays the role of A_{ij}). Its expression can be computed noting that $G_0(x-y)$ is, formally, the inverse of $\left[-\partial_x^2 + m^2\right]$, that means that

$$\left[-\partial_x^2 + m^2\right] G_0(x-y) = \delta^D(x-y)\,; \tag{4.27}$$

using the Fourier rapresentation of the correlation function and the delta we obtain

$$G_0(x-y) = \int \frac{\mathrm{d}^D p}{(2\pi)^D}\, \widetilde{G}_0(p)\, e^{-ip\cdot(x-y)} = \int \frac{\mathrm{d}^D p}{(2\pi)^D}\, \frac{e^{-ip\cdot(x-y)}}{p^2+m^2} = \langle \varphi(x)\,\varphi(y)\rangle_0 = \wick{\c{\varphi}(x)\c{\varphi}(y)}\,. \tag{4.28}$$

From a physical point of view, it describes the propagation of a fluctuation of the field $\varphi(x)$ from position x to y. Accordingly, up to a multiplicative constant the generating functional $Z[j]$ may be written as

$$Z[j] = Z_0[0]\exp\left(-\int \mathrm{d}^D x\, \mathscr{H}_I\left(\frac{\partial}{\partial j(x)}\right)\right) \exp\left(\frac{1}{2}\int \mathrm{d}^D x\, \mathrm{d}^D y\; j(y)\, G_0(x-y)\, j(x)\right) \tag{4.29}$$

Throughout the following, $Z_0[0]$ represents a normalization constant for generating functionals. In general such constants do not matter and need not be specified. Wick's theorem in this continuum theory now reads

$$\begin{aligned}\langle \varphi(x_1)\varphi(x_2)\dots\varphi(x_N)\rangle_0 &= \wick{\c{\varphi}(x_1)\c{\varphi}(x_2)}\,\wick{\c{\varphi}(x_3)\c{\varphi}(x_4)}\dots\wick{\c{\varphi}(x_{N-1})\c{\varphi}(x_N)}\\ &\quad + \text{permutations},\\ \text{with}\qquad \wick{\c{\varphi}(x_1)\c{\varphi}(x_2)} &= \langle\varphi(x_1)\varphi(x_2)\rangle_0 = G_0(x_1-x_2)\,.\end{aligned} \tag{4.30}$$

The direct application of this equations will give rise to $(N-1)!!$ different terms; this would be very tedious at high orders. It is convenient to use the equivalent relation

$$\begin{aligned}\langle\varphi(x_1)\,\varphi(x_2)\ldots\varphi(x_N)\rangle_0 &= G_0(x_1-x_2)\,\langle\varphi(x_3)\ldots\varphi(x_N)\rangle_0\\ &+G_0(x_1-x_3)\,\langle\varphi(x_2)\,\varphi(x_4)\ldots\varphi(x_N)\rangle_0+\ldots\\ &+G_0(x_1-x_N)\,\langle\varphi(x_2)\ldots,\varphi(x_{N-1})\rangle_0\\ &=\sum_{i=2}^{N}G_0(x_1-x_i)\left\langle\prod\nolimits_{k\neq i}\,{}_{2}^{N}\,\varphi(x_k)\right\rangle_0 .\end{aligned} \tag{4.31}$$

We will see that this expression enable us to collect equal terms at each step.
This form of $Z[j]$, however compact, cannot be evaluated directly. We shall have to expand it in a perturbation series.

4.2 The perturbation expansion of $G^{(2)}$

We want to show how to compute the correlation functions for the Landau-Ginsburg Hamiltonian as an expansion in series of powers of g. For the time being it is convenient to suppose that $D<4$ and Λ is finite, so no divergences are present in the correlation functions. In the next subsection we shall study the limit $a\to 0$ or $\Lambda\to\infty$. We shall study together the cases of the lattice and of the continuum. In order to avoid duplications, we use only the continuum notation. Let us study the correlation function of two φ fields; we get

$$\begin{aligned}G^{(2)}(x-y)=\langle\varphi(x)\,\varphi(y)\rangle &\equiv \frac{\displaystyle\int\mathscr{D}\varphi\;e^{-H_{ef}[\varphi]}\,\varphi(x)\,\varphi(y)}{\displaystyle\int\mathscr{D}\varphi\;e^{-H_{ef}[\varphi]}}=\frac{\left\langle\varphi(x)\,\varphi(y)\,e^{-H_{ef}[\varphi]}\right\rangle_0}{\left\langle e^{-H_{ef}[\varphi]}\right\rangle_0}\\ &=\lim_{V\to\infty}\frac{\displaystyle\sum_n\left(-\frac{g}{4!}\right)^n\frac{1}{n!}\left\langle\left[\int_V \mathrm{d}z\;\varphi^4(z)\right]^n\varphi(x)\,\varphi(y)\right\rangle_0}{\displaystyle\sum_n\left(-\frac{g}{4!}\right)^n\frac{1}{n!}\left\langle\left[\int_V \mathrm{d}z\;\varphi^4(z)\right]^n\right\rangle_0},\end{aligned} \tag{4.32}$$

where $\langle\;\rangle_0$ denote, how usual, the $g=0$ expectation values, i.e. those of the Gaussian model.

4.2.1 The calculation of $G^{(2)}$ to order g

At the first order in g we have

$$G^{(2)}(x-y)=\frac{\langle\varphi(x)\,\varphi(y)\rangle_0-\dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\;\langle\varphi(x)\,\varphi(y)\,\varphi^4(z)\rangle_0}{1-\dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\;\langle\varphi^4(z)\rangle_0}\,. \tag{4.33}$$

Now we use Wick's theorem. It proves convenient to represent these contractions on a diagram, by drawing two *external points* x and y ('external' means that they refer to the arguments of the correlation function), marked by crosses; and an *internal point* or *vertex* z, which stems from the expansion of $\exp(-H_I)$, and over which we shall integrate. Because $\varphi(z)$ enters through its fourth power, at first z is drawn as four separate points. Every contraction is represented by a line joining arguments of φ, as for instance in

$$\overbracket{\varphi(x)\varphi(y)} \;\rightarrow\; \underset{x}{\times}\!\!\text{———}\!\!\underset{y}{\times}\,.$$

Two types of terms are possible for the numerator at the order g (Fig. 4.1):

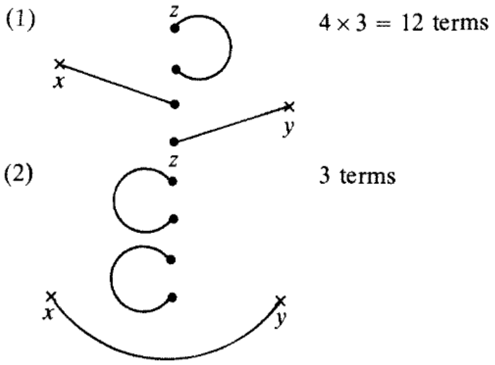


Figure 4.1. Possible terms deriving from the numerator at the first order in g.

$$\begin{aligned} G^{(2)}(x-y) &= \frac{G_0(x-y) - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\, \Big[G_0(x-y)\,\langle\varphi^4(z)\rangle_0 + 4G_0(x-z)\,\langle\varphi(y)\,\varphi^3(z)\rangle_0\Big]}{1 - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\, \langle\varphi^4(z)\rangle_0} \\ &= \frac{G_0(x-y) - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\, \Big[3G_0(x-y)\,G_0(0) + 12G_0(x-z)\,G_0(y-z)\,G_0(0)\Big]}{1 - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\, [3G_0(0)]}\,. \end{aligned} \tag{4.34}$$

In the following figures are represented the diagrams coming from the numerator and the denominator at the first order in g. They can be derived also by merging the four points z into a single point. These diagrams are called *Feynman diagrams* (or *graphs*).

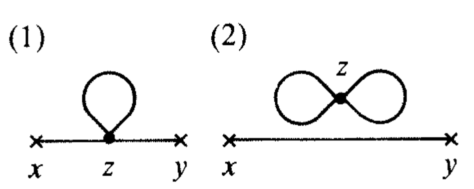


Figure 4.2. The two diagrams of order g comparing in the numerator. The second diagram (2) is a vacuum fluctuation.

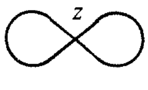


Figure 4.3. The vacuum-fluctuation diagram comparing in the denominator at order g.

At the first order in g we obtain

$$\begin{aligned} G^{(2)}(x-y) &= \frac{G_0(x-y) - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z \left[3G_0(x-y)\,G_0(0) + 12G_0(x-z)\,G_0(y-z)\,G_0(0)\right]}{1 - \dfrac{g}{4!}\displaystyle\int_V \mathrm{d}z\ [3G_0(0)]} \\ &\simeq G_0(x-y) - \frac{g}{2}\int_V \mathrm{d}z\, G_0(x-z)\,G_0(y-z)\,G_0(0) \end{aligned} \tag{4.35}$$

Using the diagrammatical representation introduced before, equation (4.35) can be visualized in Fig. (4.4).

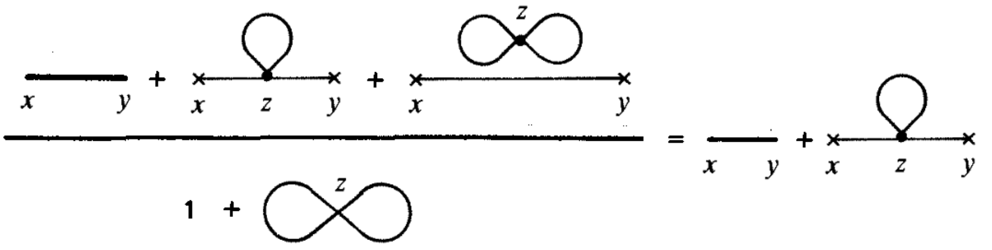


Figure 4.4. The two point correlation at order g: after the cancellations between the numerator and the denominator only connected diagrams survive.

The graph (2) does not feature in the perturbation expansion of G. Diagrams of this type contain parts called *vacuum fluctuations* or 'vacuum to vacuum' (sub)diagrams (a terminology borrowed from quantum field theory), meaning a *subgraph that is completely disconnected from the 'external' points x and y.* It is thus important to note that in the final result only the connected diagram survives, and the disconnected ones cancel between numerator and denominator. This cancellation of disconnected diagrams is present at all orders; if it were not, there would be a serious disaster: disconnected diagrams are proportional to the volume V, and noncanceling disconnected diagrams would appear as terms proportional to V^n in the connected correlation functions, thus destroying the possibility of a perturbative expansion in powers of g. In the literature, the statement of vacuum graph cancellation is sometimes referred to as the *linked cluster theorem.*
The analytic expression corresponding to a graph is evaluated according to the following rules (these are the *Feynman rules in x-space*):

- To every internal point (i.e. vertex) we assign a factor $-g$.
- To every line joining two points x_i and x_j; we assign a factor $G_0(x_i - x_j)$, often called a *propagator* (again a term borrowed from quantum field theory).
- We integrate over all internal points z_i: $\int \mathrm{d}^D z_i$;.
- To every graph we assign a multiplicative numerical factor, called a *symmetry factor.* In the case considered above, this factor equals 1/2.

On taking its Fourier transform, the last equation in p-space becomes

$$\widetilde{G}^{(2)}(p) = \widetilde{G}_0(p) - \frac{1}{2}\,g\,\widetilde{G}_0(p)\left[\int \frac{\mathrm{d}^D k}{(2\pi)^D}\,\widetilde{G}_0(k)\right]\widetilde{G}_0(p)\,. \tag{4.36}$$

This expression suggests the following *Feynman rules in p-space*:

- To every vertex we assign a factor $-g$.
- To every line we assign a factor $G_0(p)$.
- To every *independent loop* (this notion will be explained below) there corresponds an integration $\int \mathrm{d}^D k/(2\pi)^D$.
- Finally, every graph is multiplied by a *symmetry factor*.

The diagrammatic expansion and the Feynman rules follow immediately from Wick's theorem: every term of the perturbation expansion is a product of factors $G_0(x_i{}^{-}x_j)$. The only factor which is not completely selfevident is the symmetry factor. Moreover, we must still specify the notion of *independent loops*. In order to become familiar with the perturbation expansion and with the Feynman rules, it is worth pushing the calculation of $G^{(2)}$ to order g^2.

4.2.2 The calculation of $G^{(2)}$ to order g^2

At the next order we obtain

$$G^{(2)}(x-y) = \langle\varphi(x)\,\varphi(y)\rangle =$$
$$= \frac{\langle\varphi(x)\,\varphi(y)\rangle_0 - \frac{g}{4!}\int \mathrm{d}z_1\ \langle\varphi(x)\,\varphi(y)\,\varphi^4(z_1)\rangle_0 + \frac{1}{2}\frac{g^2}{(4!)^2}\int \mathrm{d}z_1\,\mathrm{d}z_2\ \langle\varphi(x)\,\varphi(y)\,\varphi^4(z_1)\varphi^4(z_2)\rangle}{1 - \frac{g}{4!}\int \mathrm{d}z_1\ \langle\varphi^4(z_1)\rangle_0 + \frac{1}{2}\frac{g^2}{(4!)^2}\int \mathrm{d}z_1\,\mathrm{d}z_2\ \langle\varphi^4(z_1)\varphi^4(z_2)\rangle}\,.$$

Instead of doing all the algebra, let us guide by the results of the first order. The linked cluster theorem tells us that at every order in g we can neglect disconnected diagrams and retain only connected ones

$$G^{(2)}(x-y) = \langle\varphi(x)\,\varphi(y)\rangle = \sum_n \left(-\frac{g}{4!}\right)^n \frac{1}{n!}\int \mathrm{d}z_1\ldots\mathrm{d}z_n\ \Big\langle\varphi(x)\,\varphi(y)\,\varphi^4(z_1)\ldots\varphi(z_n)\Big\rangle_0^c\,. \quad (4.37)$$

A further simplification can be achieved by noticing that the r.h.s. of (4.37) contains $n!$ equal terms that differ from the *permutations of the internal vertices* z_i; we can consider only one such contributions, $\langle\ \rangle_0^{cp}$ and cancel the $1/n!$. We finally get

$$G^{(2)}(x-y) = \langle\varphi(x)\,\varphi(y)\rangle = \sum_n \left(-\frac{g}{4!}\right)^n \int \mathrm{d}z_1\ldots\mathrm{d}z_n\ \Big\langle\varphi(x)\,\varphi(y)\,\varphi^4(z_1)\ldots\varphi(z_n)\Big\rangle_0^{cp}\,. \quad (4.38)$$

The term of second order in g is

$$\begin{aligned}\Big\langle\varphi(x)\,\varphi(y)\,\varphi^4(z_1)\,\varphi^4(z_2)\Big\rangle_0 &= G_0(x-y)\Big\langle\varphi^4(z_1)\,\varphi^4(z_2)\Big\rangle_0\\ &+ 4G_0(x-z_1)\Big\langle\varphi(y)\,\varphi^3(z_1)\,\varphi^4(z_2)\Big\rangle_0\\ &+ 4G_0(x-z_2)\Big\langle\varphi(y)\,\varphi^4(z_1)\,\varphi^3(z_2)\Big\rangle_0\,.\end{aligned} \quad (4.39)$$

The first term is disconnected and can be neglected. The second term gives

$$
\begin{aligned}
4G_0(x-z_1)\left\langle \varphi(y)\,\varphi^3(z_1)\,\varphi^4(z_2)\right\rangle_0 = {} & 4G_0(x-z_1) \\
& \times \left[3G_0(y-z_1)\left\langle \varphi^2(z_1)\,\varphi^4(z_2)\right\rangle_0\right. \\
& \left. +4G_0(y-z_2)\left\langle \varphi^3(z_1)\,\varphi^3(z_2)\right\rangle_0\right] .
\end{aligned}
\tag{4.40}
$$

The first term arises by connecting y with z_1 (3 ways), the second by connecting y with z_2 (4 ways). We can go on: if in the first term we connect z_1 with z_1 (extracting $G_0(z_1-z_1)=G_0(0)$) we are left with $\langle\varphi^4(z_2)\rangle_0$, which yields a disconnected piece. Therefore we must connect z_1 with z_2 (4 ways). In the second term, either we connect z_1 with z_1 (3 ways) or we connect z_1 with z_2 (3 ways). We finally get

$$
\begin{aligned}
& 4G_0(x-z_1)\left\{3G_0(y-z_1)4G_0(z_1-z_2)\left\langle \varphi(z_1)\,\varphi^3(z_2)\right\rangle_0\right. \\
& \left. +4G_0(y-z_2)\left[3G_0(0)\left\langle \varphi(z_1)\,\varphi^3(z_2)\right\rangle_0 + 3G_0(z_1-z_2)\left\langle \varphi^2(z_1)\,\varphi^2(z_2)\right\rangle_0\right]\right\} \\
& = 4G_0(x-z_1)\left\{12G_0(y-z_1)G_0(z_1-z_2)G_0(z_1-z_2)\left\langle \varphi^2(z_2)\right\rangle_0\right. \\
& + 4G_0(y-z_2)\left[3G_0(0)3G_0(z_1-z_2)\left\langle \varphi^2(z_2)\right\rangle_0\right. \\
& \left.\left. +3G_0(z_1-z_2)\left(G_0(0)\left\langle \varphi^2(z_2)\right\rangle_0 + 2G_0(z_1-z_2)\left\langle \varphi(z_1)\,\varphi(z_2)\right\rangle_0\right)\right]\right\}
\end{aligned}
\tag{4.41}
$$

The general diagrammatical rules for computing the order g^n are as general follows: we draw all connected diagrams with n vertices with 4 lines and two external points; diagrams equivalent by the interchange of the vertices must be identified; with each diagram we associate an integrand which is the product of the G_0's corresponding to the lines of the diagrams; we finally integrate over the coordinates of the vertices. The result must be multiplied by $(-g/24)^n$ and by the multiplicity factor of the diagram (how many times the diagram can be obtained); the multiplicity factors for

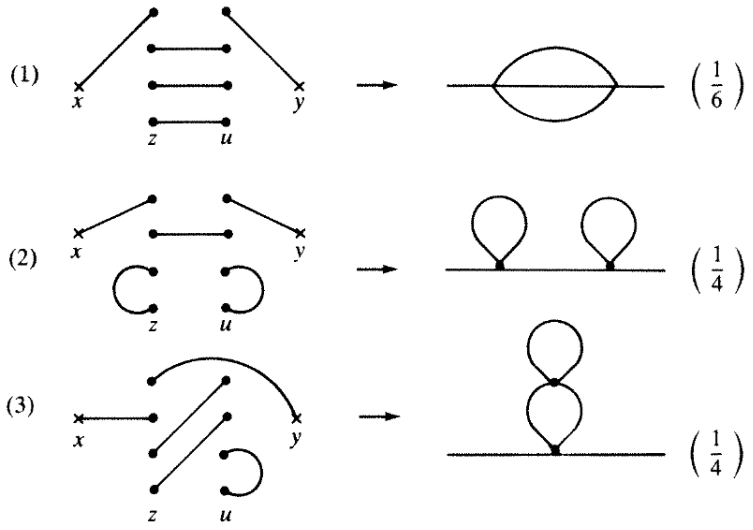


Figure 4.5. Connected diagrams of order g^2.

the diagrams in Fig.(4.5) are, respectively, 96, 144, 144. The computation of the correct multiplicity factor is normally a tedious job. It can be simplified by using the following rule: the multiplicity of a diagram of order n for the φ^4 interaction, divided by $(24)^n$, is equal to one divided by the number of symmetries of the diagram (i.e., the number of permutations of the lines of the diagrams that do not change the diagram). The diagrams of Fig.(4.5) have 6, 4, and 4 symmetries, respectively. We shall settle for examining the contribution $\overline{G}(x-y)$ to the correlation function from graph (1) in Fig.(4.5). Thus

$$\overline{G}(x-y) = \frac{1}{6}\,g^2 \int \mathrm{d}^D z\, \mathrm{d}^D u\; G_0(x-z)\,[G_0(z-u)]^3\, G_0(u-y)\,. \tag{4.42}$$

Ler us write $\overline{G}(x-y)$ as a Fourier trasform, by replacing every factor G_0 by its Fourier representation

$$\begin{aligned}\overline{G}(x-y) = \frac{1}{6}\,g^2 \int \mathrm{d}^D z\, \mathrm{d}^D u\, \frac{\mathrm{d}^D k}{(2\pi)^D}\frac{\mathrm{d}^D k'}{(2\pi)^D}\prod_{l=1}^{3}\left\{\frac{\mathrm{d}^D q_l}{(2\pi)^D}\, e^{-iq_l\cdot(z-u)}\right\}\\ \times e^{-ik\cdot(x-z)}e^{-ik'\cdot(u-y)}\,\widetilde{G}_0(k)\,\widetilde{G}_0(k')\prod_{l=1}^{3}\widetilde{G}_0(q_l)\,.\end{aligned} \tag{4.43}$$

The integration over z and u yield a product of two delta-functions, namely

$$(2\pi)^D\delta^D(k-q_1-q_2-q_3)\times(2\pi)^D\delta^D(k'-q_1-q_2-q_3) \tag{4.44}$$

whence

$$\overline{G}(x-y) = \frac{1}{6}\,g^2 \int \frac{\mathrm{d}^D k}{(2\pi)^D}\, e^{ik\cdot(x-y)}\left[\widetilde{G}_0(k)\right]^2 \int \frac{\mathrm{d}^D q_1}{(2\pi)^D}\frac{\mathrm{d}^D q_2}{(2\pi)^D}\,\widetilde{G}_0(q_1)\,\widetilde{G}_0(q_2)\,\widetilde{G}_0(k-q_1-q_2)\,. \tag{4.45}$$

The last expression shows that $\overline{G}(x-y)$ is the Fourier transform of a function $\overline{G}(k)$,

$$\overline{G}(k) = \frac{1}{6}\,g^2\,\widetilde{G}_0(k)\left[\int \frac{\mathrm{d}^D q_1}{(2\pi)^D}\frac{\mathrm{d}^D q_2}{(2\pi)^D}\,\widetilde{G}_0(q_1)\,\widetilde{G}_0(q_2)\,\widetilde{G}_0(k-q_1-q_2)\right]\widetilde{G}_0(k)\,, \tag{4.46}$$

represented diagrammatically in Fig(4.6). The graph shown features two external propagators $\widetilde{G}_0(k)$, and three internal propagators; because of the two delta-functions $\delta^D(\dots)$, *only two of the three internal lines are independent.* The diagram in Fig.(4.6) allows us to introduce the notion of *independent loops.*

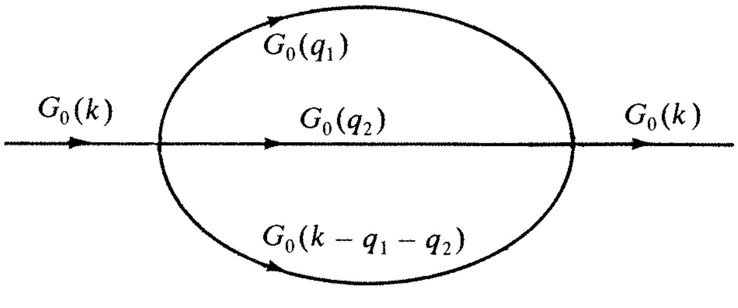


Figure 4.6. Diagrammatic representation of (4.46)

By following the internal propagators one can describe three different closed loops, but on account of the delta-functions δ^D only two of these loops are independent; in other words there are only two integration variables in 4.46 (in general at the n order there will be n integration).
In Fig.(4.7) are represented the diagrams contributing to the perturbation expansion of the correlation function up to the second order in g.

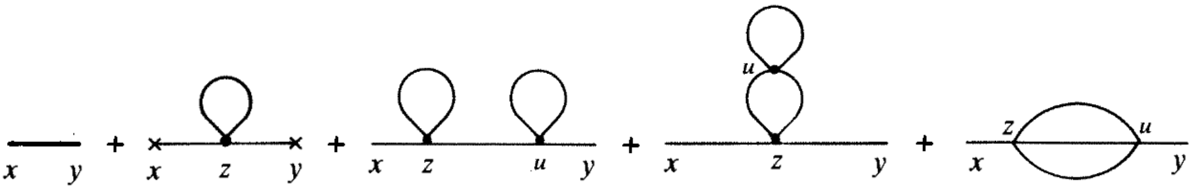


Figure 4.7. Diagrams contributing to the expansion of the correlation function up to the second order in g

4.2.3 The calculation of $G^{(2)}$ to order g^3

Connected diagrams of order g^3 are represented in Fig.(4.8).

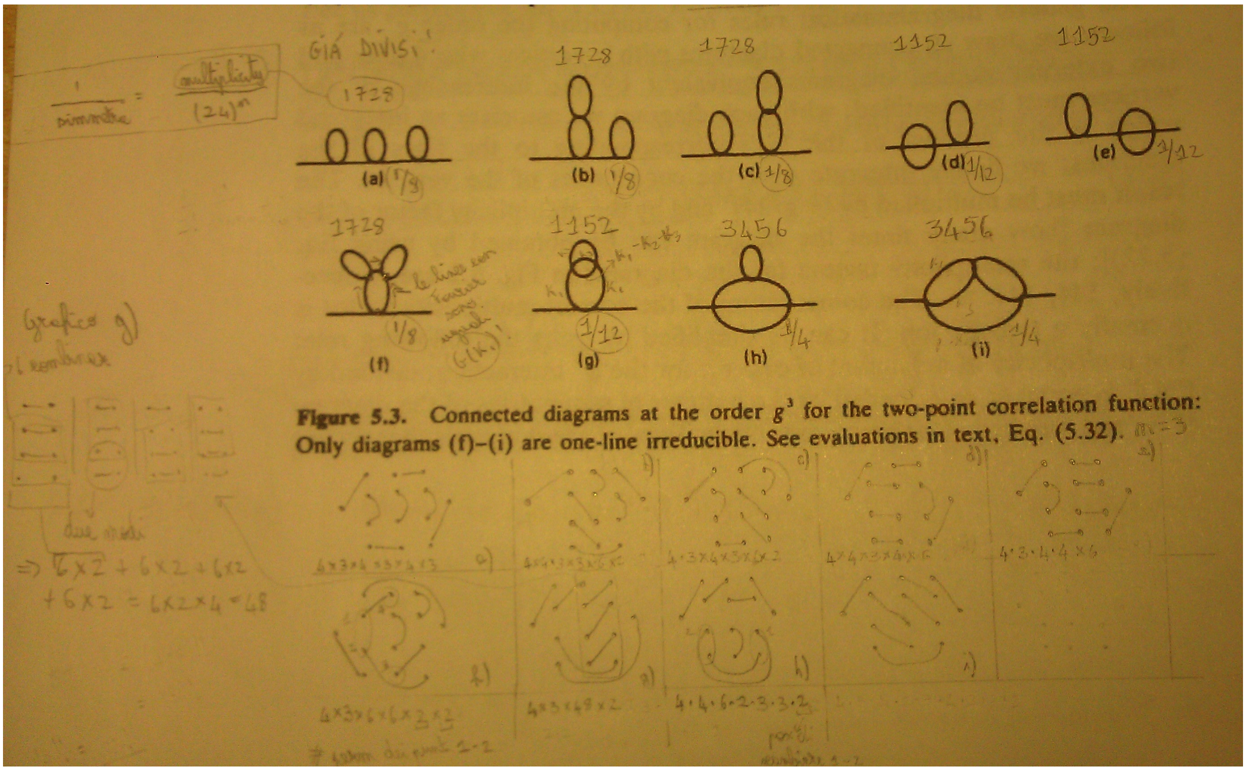


Figure 4.8

We mention again the rules in momentum space (where the evaluation of the integrals is simpler): each external leg of the diagram is associated with a momentum p. Each internal line (α) is associated with a momentum q_α, which runs along the line in a given direction. We must impose the condition that the sum of all incoming minus all outgoing momenta be equal to zero at each vertex. At the order g^n there remain n unfixed momenta (let us call them k's) on which the q_α depend linearly. The contribution of the diagram is given by the integral over the k's of the product of all $\widetilde{G}(q_\alpha)$'s. These rules are clarified by the explicit evaluations of the diagrams of (4.8):

$$a = -\left(\frac{g}{(2\pi)^D}\right)^3 \frac{1}{8} [\tilde{G}_0(p)]^4 \left[\int d^Dk_1\, \tilde{G}_0(k_1)\right]^3$$

$$b = c = -\left(\frac{g}{(2\pi)^D}\right)^3 \frac{1}{8} [\tilde{G}_0(p)]^3 \left[\int d^Dk_1\, \tilde{G}_0(k_1)\right]^2 \int d^Dk_2 (\tilde{G}_0(k_2))^2$$

$$d = e = -\frac{1}{12}\left(\frac{g}{(2\pi)^D}\right)^3 [\tilde{G}_0(p)]^3 \int d^Dk_1\, d^Dk_2\, \tilde{G}_0(k_1)\tilde{G}_0(k_2)$$

$$\times\, \tilde{G}_0(p-k_1-k_2) \int d^Dk_3\, \tilde{G}_0(k_3)$$

$$f = -\frac{1}{8}\left(\frac{g}{(2\pi)^D}\right)^3 [\tilde{G}_0(p)]^2 \int d^Dk_1 [\tilde{G}_0(k_1)]^3 \left[\int d^Dk_2\, \tilde{G}_0(k_2)\right]^2$$

$$g = -\frac{1}{12}\left(\frac{g}{(2\pi)^D}\right)^3 [\tilde{G}_0(p)]^2 \int d^Dk_1\, d^Dk_2\, d^Dk_3 (\tilde{G}_0(k_1))^2$$

$$\times\, \tilde{G}_0(k_2)\tilde{G}_0(k_3)\tilde{G}_0(k_1-k_2-k_3)$$

$$h = -\frac{1}{4}\left(\frac{g}{(2\pi)^D}\right)^3 [\tilde{G}_0(p)]^2 \int d^Dk_1\, d^Dk_2\, d^Dk_3 [\tilde{G}_0(p-k_1-k_2)]^2$$

$$\times\, \tilde{G}_0(k_1)\tilde{G}_0(k_2)\tilde{G}_0(k_3)$$

$$i = -\frac{1}{4}\left(\frac{g}{(2\pi)^D}\right)^3 [\tilde{G}_0(p)]^2 \int d^Dk_1\, d^Dk_2\, d^Dk_3\, \tilde{G}_0(k_1)\tilde{G}_0(k_2)$$

$$\times\, \tilde{G}_0(k_3)\tilde{G}_0(p-k_1-k_3)\tilde{G}_0(p-k_2-k_3)\,.$$

4.3 Proper vertices and self energy

In order to decrease the proliferation of diagrams with order, it is convenient to introduce diagrams that are *one-line irreducible* (or *1-particle irreducible*), i.e., they cannot be divided into two disconnected diagrams by cutting only one internal line (the diagrams in Fig.(4.8) (f)-(i) are one-line irreducible). In k-space, a diagram like that of Fig.(4.9) effectively reads $\hat{G}_1\, G_0(k)\, \hat{G}_2$ and is one-line reducible (or 1-particle reducible). In order to obtain the expression for this diagram it is therefore enough to know how to calculate $\hat{G}_1$ and $\hat{G}_2$ independently.

Finally, a 1-particle irreducible correlation function shorn of its external lines is called a *proper*

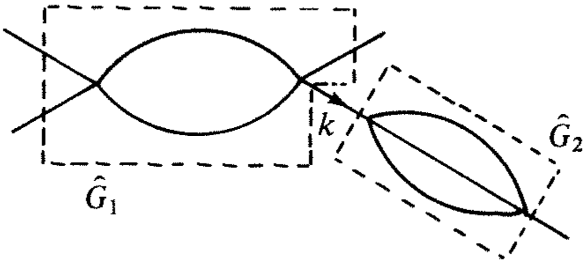


Figure 4.9. A 1-particle reducible diagram

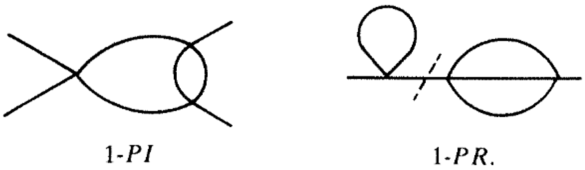


Figure 4.10

vertex (amputated diagrams). For example, the expression of the proper vertex in Fig.(4.11) is

$$-\frac{g^2}{2}\int\frac{\mathrm{d}^Dq}{(2\pi)^D}\,\frac{1}{q^2+m^2}\,\frac{1}{((k-q)^2+m^2)} \tag{4.47}$$

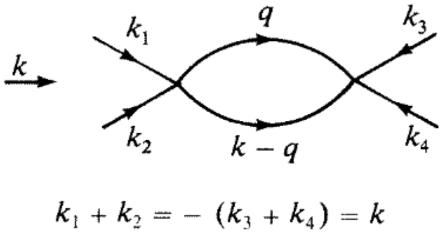


Figure 4.11

No factor $\prod_{i=1}^4 G_0(k_i)$ is associated with the external lines. It proves useful to define the *self-energy*[1] $\Sigma(p)$ as the sum of all two-point 1-particle irreducible diagrams shorn of their external lines (i.e. it is the sum of all proper two-point vertices; see Fig.(4.12)).

Diagrams (f)-(i) in Fig.(4.8) can be written as $\left[\widetilde{G}_0(p)\right]^2\,\Sigma(p)$. In general the correlation function $\widetilde{G}^{(2)}(p)$ can be written in terms of the self energy in this way

$$\widetilde{G}^{(2)}(p)=\widetilde{G}_0(p)+\widetilde{G}_0(p)\,\Sigma(p)\,\widetilde{G}_0(p)+\widetilde{G}_0^3(p)\,\Sigma^2(p)+\cdots=\frac{1}{\widetilde{G}_0^{-1}(p)-\Sigma(p)} \tag{4.48}$$

and then

$$\widetilde{G}^{(2)}(p)=\frac{1}{p^2+m^2-\Sigma(p)}\,. \tag{4.49}$$

---

[1] The name is borrowed from quantum field theory. One must be careful of the fact that $\Sigma(k)$ is given different signs by different writers.

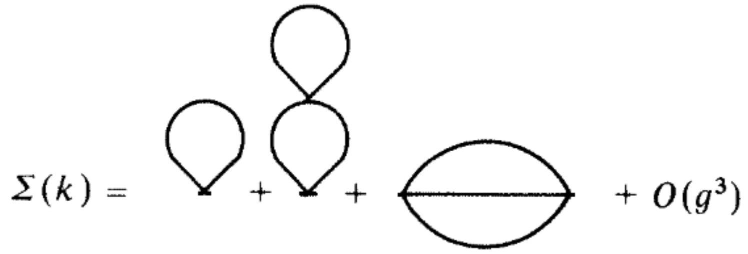


Figure 4.12

This equation is shown diagrammatically in Fig.(4.13).

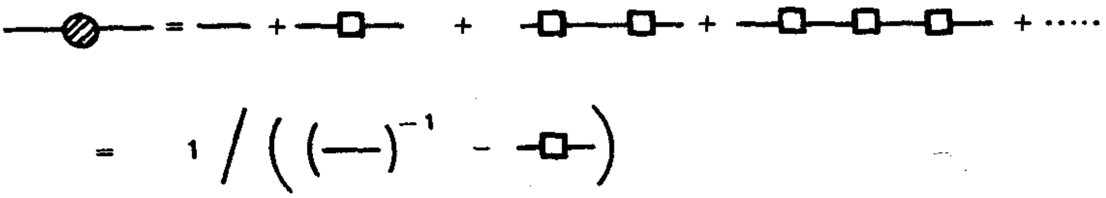


Figure 4.13. The graphical representation of Eq.(4.48) of the text. The square and the blob denote, respectively, the self-energy and the *exact* two-point correlation.

The only delicate point in proving this equation is to check that the weights of the diagrams are correct.

4.4 The four-point correlation function $G^{(4)}$

We can now develop in powers of g the four-field correlation functions

$$G^{(4)}(x_1, x_2, x_3, x_4) = \langle \varphi(x_1)\, \varphi(x_2)\, \varphi(x_3)\, \varphi(x_4) \rangle \ . \tag{4.50}$$

By translation-invariance, this function depends only on three coordinate differences, e.g. on $(x_1 - x_2)$, $(x_2 - x_3)$, $(x_3 - x_4)$. To order g^0, $G^{(4)}$ is the sum of three *disconnected graphs*, see Fig.(4.14)

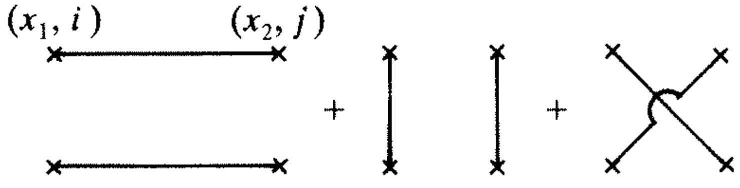


Figure 4.14. $G^{(4)}$ to order g^0

$$G^{(4)}(x_1, x_2, x_3, x_4) = G_0(x_1 - x_2)\, G_0(x_3 - x_4) + G_0(x_1 - x_3)\, G_0(x_2 - x_4) + G_0(x_1 - x_4)\, G_0(x_2 - x_3)\, .$$

To order g there are three types of graphs, see Fig.(4.15)

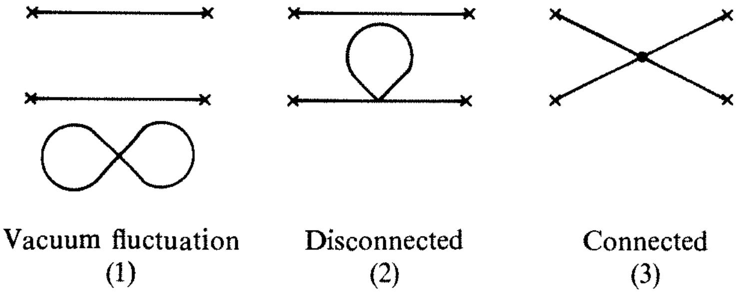


Figure 4.15. $G^{(4)}$ to order g

The graphs of type (1) contain *vacuum-fluctuation* parts, and are cancelled on division by Z(0). The graphs of type (2) can be written as products of two independent factors; these are examples of *disconnected graphs.* Their expressions are known, since they involve only two-point correlation functions, which have been calculated already. In general we have three classes of diagrams shown in Fig.(4.16).

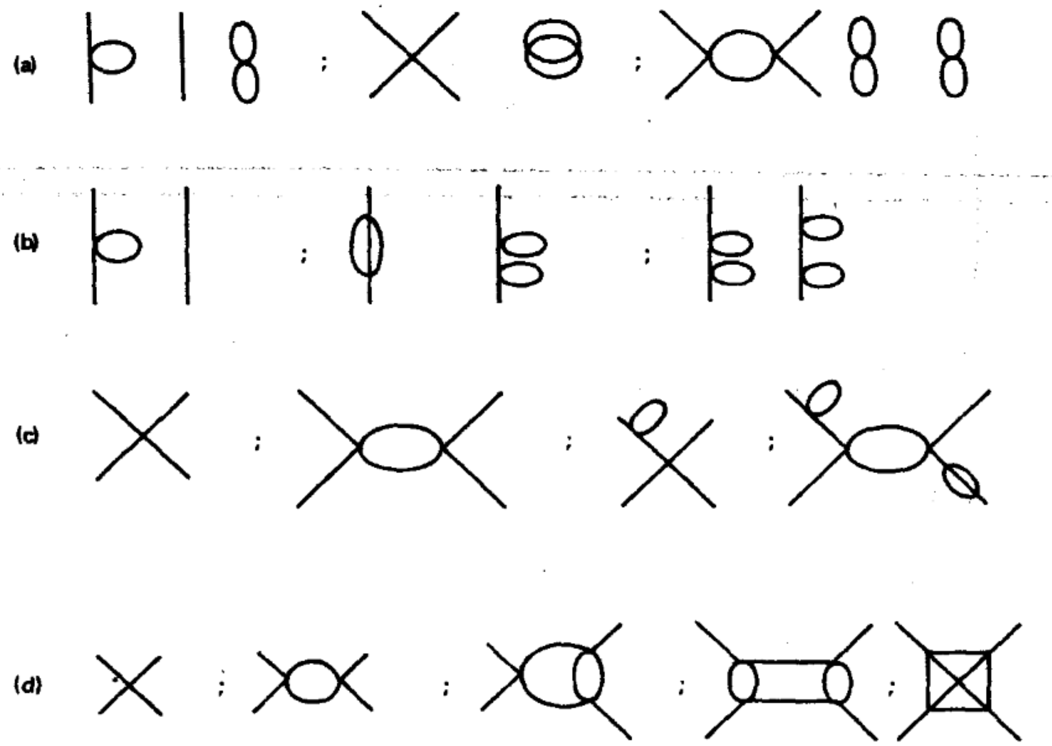


Figure 4.16. Various diagrams contributing to the four-point function: (a) fully disconnected diagrams; (b) partially connected diagrams; (c) connected diagrams; (d) some connected one-line irreducible diagrams contributing to Γ.

- In the first class we find the fully disconnected diagrams, i.e., *there are vertices that are not connected to the external points.* These diagrams cancel with the corresponding ones coming from the expansion of the denominator.

- In the second class all vertices are connected to the external points, but the diagrams can be divided into two *disconnected parts.* These diagrams are the same as those of the two-field correlation functions, and their contributions can be summed to $G(x_1 - x_2)\, G(x_3 - x_4) + G(x_1 - x_3)\, G(x_2 - x_4) + G(x_1 - x_4)\, G(x_2 - x_3)$ (note that the G's do not have the 0: they are correlation function which we have studied in the perturbative approach in the preceding section).
- In the third class we find *fully connected diagrams.* They are the only ones that contribute to the connected correlation functions.

In order to simplify the analysis it is convenient to introduce one-line irreducible amputated diagrams. Their sum, neglecting the external leg, is denoted $\Gamma(p_1, p_2, p_3, p_4)$; at the order g^2 we are left with the diagrams shown in Fig.(4.16), (d). In momentum space we find

$$\begin{aligned} -\Gamma(p_1, p_2, p_3, p_4) &= \delta(p_1 + p_2 + p_3 + p_4)\Big[g - \frac{g^2}{2}\,(I(p_1 + p_2) \\ &\quad + I(p_1 + p_3) + I(p_1 + p_4)) + O(g^3)\Big] \\ I(p) &= \frac{1}{(2\pi)^D}\int d^Dk\, \tilde{G}_0(k)\tilde{G}_0(p - k) \\ G(p_1, p_2, p_3, p_4) &= \tilde{G}(p_1)\tilde{G}(p_3)\delta(p_1 + p_2)\delta(p_3 + p_4) \\ &\quad + \tilde{G}(p_1)\tilde{G}(p_2)\delta(p_1 + p_3)\delta(p_2 + p_4) \\ &\quad + \tilde{G}(p_1)\tilde{G}(p_3)\delta(p_1 + p_4)\delta(p_4 + p_3) \\ &\quad + \tilde{G}(p_1)\tilde{G}(p_2)\tilde{G}(p_3)\tilde{G}(p_4)\Gamma(p_1, p_2, p_3, p_4) \end{aligned}$$

The δ function is the consequence of the translational invariance of the correlation function. It is clear that if we already know the two-field correlation functions, in order to get the four-field correlation functions we need only compute Γ, which receives contributions from only a "few" diagrams. A similar analysis can be performed for the many-field correlation functions.

4.5 Ultraviolet Divergences

Before analyzing these argument in more detail, let us make some general observations. The first-order expansion of G contains a number of factors of $G_0(0)$, the free Green function evaluated at coinciding points. This bears disturbing consequences. To see this, consider $G_0(0)$ evaluated in momentum space:

$$G_0(0) = \int \frac{\mathrm{d}^D p}{(2\pi)^D}\, \frac{1}{p^2 + m^2}\,. \tag{4.51}$$

For dimensions $D > 1$, the integral is divergent at large momenta or short wavelengths; we have met with an *ultraviolet* (UV) *divergence*. Physically, the divergence implies that, already at first order, our expansion runs into a difficulty that is obviously related to the short-distance structure of the system. How can this problem be overcome? One way out is to remember that field theories like the φ^4-model represent effective low-temperature, or long-wavelength, approximations to more microscopic models. The range of applicability of the action must be limited to wavelengths in excess of some microscopic lattice cutoff a (e.g. the lattice spacing), or momenta $k < a^{-1}$. It seems that, once that cutoff has been built in, the convergence problem is solved. However, there is something unsatisfactory in this argument. All our perturbative corrections, and therefore the final result of the analysis, exhibit sensitivity to the microscopic cutoff parameter. But this is not what we expect of a sensible low-energy theory. The UV problem signals that something more interesting is going on than a naive cutoff regularization has the capacity to describe. However, even if we temporarily close our eyes to the UV-phenomenon, there is another problem. For dimensions $d \leq 2$, and in the limit $r \to 0$, $G_0(0)$ also diverges at small momenta, an *infrared* (IR) *divergence*. Being related to structures at large wavelengths, this type of singularity should attract our attention even more than the UV-divergence mentioned above. Indeed, it is intimately related to the accumulation of long-range correlations in the limit $r \to 0$.

4.5.1 The removal of ultraviolet divergences

We want to study what happens when $\Lambda \to \infty$ or $a \to 0$. Some diagrams will be divergent, some will remain finite; we have to take care of the first and to compensate for their divergence by the introduction of appropriate counterterms [Eq. E.14). Here we remark that a diagram whose expression is given by

$$\int \prod_{i=1}^{L} \mathrm{d}^D k_i \prod_{j=1}^{I} \widetilde{G}_0 \left[q_j(k) \right] , \tag{4.52}$$

the q_j being linear functions of the k's, is certainly divergent (for $\Lambda \to \infty$) if the number of powers of k coming from the integrations (DL) is *greater* than or *equal* to the number of powers coming from the denominator $2I$, as can be seen from rescaling all the k's together. If the opposite happens, i.e. ,

$$DL < 2I \iff \mathscr{D} < 0 \, , \tag{4.53}$$

the diagram is said to be *superficially convergent*; $\mathscr{D} = DL - 2I$ is the *degree of divergence* of a diagram. It is possible that a superficially convergent diagram is divergent because a subdiagram (a subintegration) is divergent. The necessary condition for convergence is that the diagram and all the subdiagrams be superficially convergent[2]. We must therefore classify all the superficially divergent diagrams. In fewer than four dimensions this task is quite easy; there are the following relations between:

- the number of external lines E,
- the number of vertices V,
- the number of internal lines I,
- the number L of loops (k integrations):

---

[2] This is essentially the content of the Weinberg theorem

$$\begin{gathered} E+2I=4V \qquad V=L+\frac{E-2}{2} \\ \mathscr{D}=(D-4)\,L+4-E \end{gathered} \tag{4.54}$$

The first relation can be proved noting that if every internal line is cut, then 4 lines arrive at every vertex, while the total number of lines is $2I+E$. The third relation can be obtained putting the first and the second one in the expression of $\mathscr{D}$

$$\mathscr{D}=DL-2I=DL+E-4V=DL+E-4L-2E+4=(D-4)\,L+4-E\,. \tag{4.55}$$

Limiting to the case of the two-field correlation functions which has $E=2$, we have, as we know, $V=L$; this means that the number of k integrations (indipendent loops) are equal to the number of vertex of the graphs, i.e. to the order of the perturbative expansion in g. Putting $E=2$ and $V=L$ we have

$$\mathscr{D}=(D-4)\,V+2\,. \tag{4.56}$$

Let now see how the sign of $\mathscr{D}$ changes varying the dimensions D and the number of vertex V. When $D<4$ we obtain that the graph is divergent if

$$V\leq\frac{2}{4-D}\,. \tag{4.57}$$

Thanks to this relation, we see that if $D<2$ no divergences are present ($V\leq 2/3$ is always satisfied); for $D=2$ only the first order is divergent; for $D=3$ are divergent the first and the second order in g as shown in Fig.(4.17).

Figure 4.17. Divergent diagrams for $D=2,3$

In $D=2$ the value of the only divergent self-energy diagram is (in a general dimension D)

$$\Sigma(p,\mu)=-\frac{g}{(2\pi)^D}\frac{1}{2}\int \mathrm{d}^D k\,\frac{1}{k^2+\mu+k^4/\Lambda^2} \tag{4.58}$$

where conventionally $k^4\equiv(k^2)^2=\left(\sum_{\nu=1}^{D}k_v^2\right)^2$. Let us decide that when we evaluate a diagram we substitute the following expression for $\Sigma(p,\mu)$:

$$\Sigma_R(p,\mu)=\Sigma(p,\mu)-\Sigma(0,\mu_N)\,,\qquad \text{if } D=2\,, \tag{4.59}$$

where μ_N is a given normalization point. This means that these divergences may be removed if we subtract from all the self-energy subdiagrams their value computed at zero external momentum. This substitution corresponds to adding a *counterterm*[3] in the Hamiltonian equal to $\frac{1}{2}\int \mathrm{d}^2x\,\varphi^2(x)$.

---

[3] To add a constant C to Σ is equivalent to changing the two field correlation functions from $G(p)$ to $1/\left(G^{-1}(p)-C\right)$, i.e., to modifying the value of μ; this operation correspond to adding to the Hamiltoniana a new term equal to $C/2\int \mathrm{d}^Dx\,\varphi^2(x)$.

Here Σ_R is finite for any μ (for $\mu = \mu_N$ the contribution of Fig.(4.17) (a) is zero because this diagram produces a momentum-independent $\Sigma(p, \mu)$). In the same way in $D = 3$ the new divergent diagram is given by (in a general dimension D)

$$\begin{aligned} \Sigma(p,\mu) = \frac{1}{6}\left(\frac{g}{(2\pi)^D}\right)^2 \int \mathrm{d}^D k_1\, \mathrm{d}^D k_2\, \frac{1}{k_1^2+\mu+k_1^4/\Lambda^2} \cdot \frac{1}{k_2^2+\mu+k_2^4/\Lambda^2} \\ \cdot \frac{1}{(p-k_1-k_2)^2+\mu+(p-k_1-k_2)^4/\Lambda^2} \end{aligned} \tag{4.60}$$

Chapter 5

The Renormalization Group

We have seen in the preceding chapters that the critical point is characterized by the onset of correlated fluctuations at large distances: their study is particularly difficult, insofar as all the methods we have seen here do not work: the correlation length is zero at $T = 0$ and at $T = \infty$ and at any order in the high-temperature (or low-temperature) expansion we can keep track of the correlated fluctuations of only a finite number of spins. Indeed, far from the critical point thermodynamic quantities may well be estimated by approximating the infinite system with a relatively small finite system. If periodic boundary conditions are used, the finite-volume corrections are proportional to $\exp(-L/\xi)$, where L and ξ are, respectively, the size of the box and the coherence length. In contrast, at the critical point the finite-volume corrections are proportional to a negative power of L and our aim is to study the collective fluctuations that produce these power-law corrections. In other words, at the critical point the number of *relevant* degrees of freedom is actually infinite; the usual reductionistic procedure of considering only a finite number of degrees of freedom fails here, and a new approach is needed.

5.1 Introduction

We have stressed that, at a critical point, all length scales are important. This is an unusual situation: usually physical theories can concentrate on a small range of scales of length. A continuum theory of water waves, ignoring atomic motions, or a theory of the arrangement of nucleons which ignores the atomic environment are essentially exact. So how can we cope with, or even exploit, scale invariance at criticality? The answer lies in a set of theories known as renormalization groups. The ideas behind them can be illustrated using the Monte Carlo simulations in Figs. (5.1) - (5.3). The aim is to change the scale of the system and see how it behaves. This is done by taking each group of nine spins in turn and replacing it by a single spin which takes the same value as the majority of spins in the original cluster. This procedure reduces the scale of the system by a factor $b = 3$. We then keep going to produce the series of snapshots of the spin configuration, essentially seen under different magnifications, shown in the figures. For a starting temperature above the critical temperature, (Fig. 5.1), the scale change soon obliterates any short-range order and the spins on the renormalized lattices become uncorrelated. This corresponds to an infinite temperature: the system has been renormalized by the simple transformation we have defined to $T = \infty$. This will be the case for all temperatures above T_C; the nearer to the critical temperature is the starting point the more steps of the transformation it will take *to lose the short-range order.*

Figure 5.1. A real-space renormalization group transformation for the two-dimensional Ising model on the square lattice. The initial configuration, corresponding to a temperature $T = 1.22T_C$, was generated using a Monte Carlo simulation. A sequence of renormalized configurations is then obtained by replacing successive clusters of nine spins by a single spin which takes the same value as the majority of the spins in the original cluster. Hence the length scale of the lattice is changed by a scale factor $b = 3, 32, 33, and 34$ in (b),(c),(d), and (e) respectively. Note that the correlation length decreases under successive iterations of the renormalization group corresponding to an increase in the temperature. After Wilson, K. G. (1979). *Scientific American*, **241**, 140.

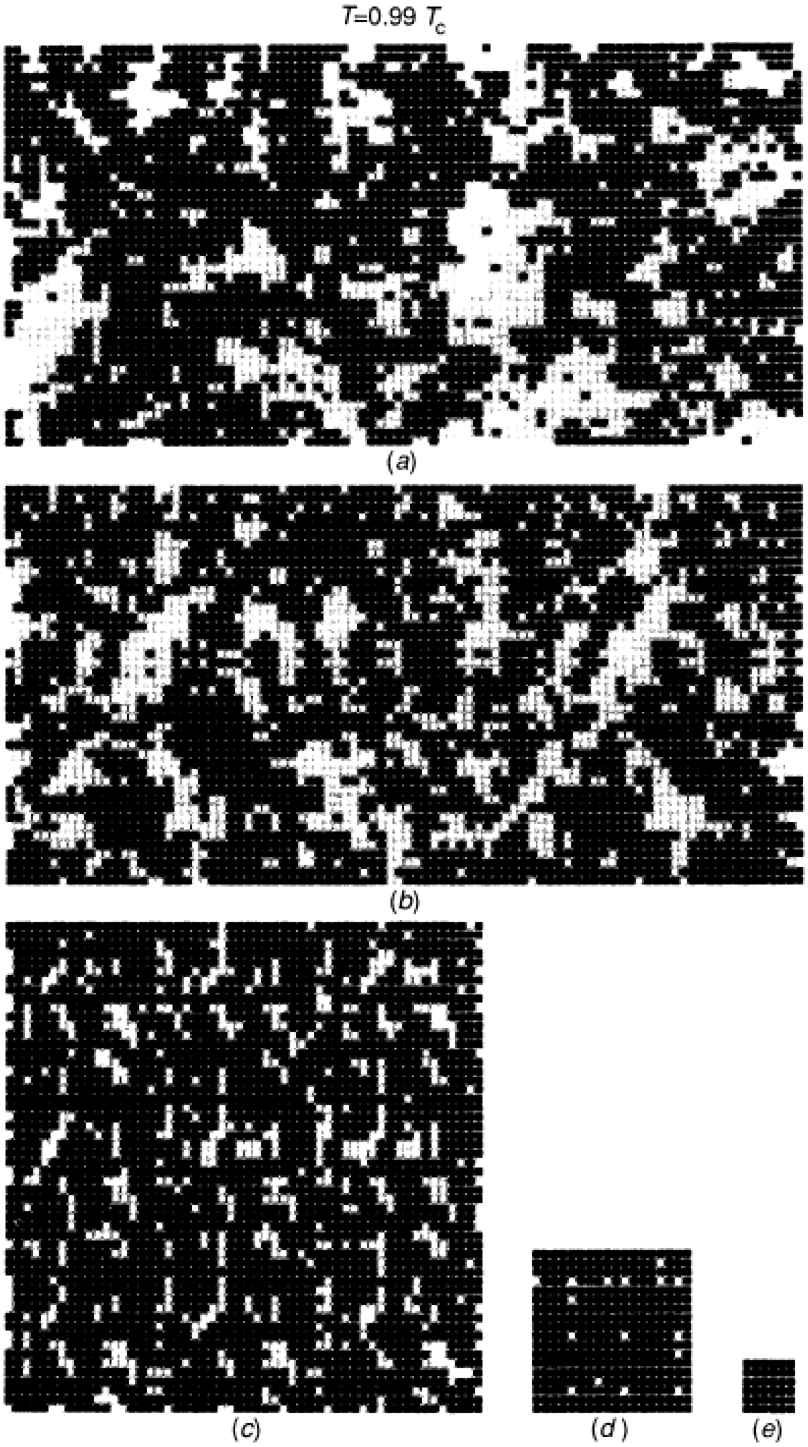


Figure 5.2. As Fig. (5.1) but with a starting temperature $T = 0.99T_C$. Fluctuations relative to the ordered state are suppressed by the change in length scale and the system flows towards zero temperature. After Wilson, K. G. (1979). Scientific American, **241**, 140.

For temperatures below the critical temperature there is an analogous flow as the renormalization

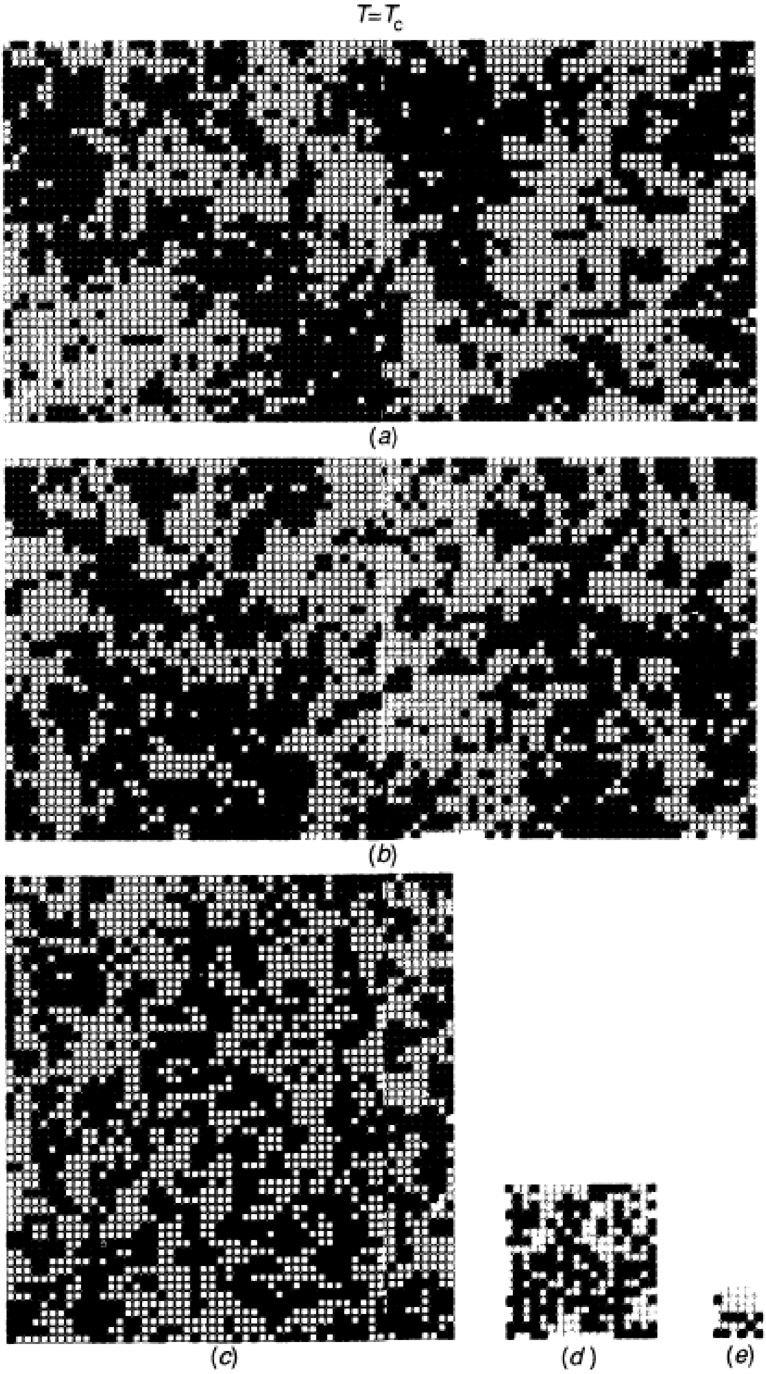


Figure 5.3. As preceding figures but with a starting temperature $T = T_C$. Because the correlation length is initially infinite there is no change in the ordered state under iteration of the renormalization group and the system remains at the critical temperature. After Wilson, K. G. (1979). Scientific American, **241**, 140.

group is iterated. However, now any fluctuations are relative to the ground state and, as these are lost under renormalization, the system flows to a completely ordered state characteristic of zero temperature. This is the case in Fig. (5.2). Only at the critical temperature itself, (Fig. (5.3)), where there are fluctuations on all length scales does the system remain invariant under the renormalization group transformation. This can be exploited to identify the critical point and describe the behaviour of the thermodynamic functions in its vicinity.

5.2 Block Variables

The renormalization group approach is based on the recursive introduction of *block variables.* These can be defined by generalizing the procedure we have used to derive the Landau-Ginsburg model (where the magnetization averaged around a point was the relevant variable on which the effective Hamiltonian depended). This coarse graining procedure can be justified in this way: since spins are correlated on lenghts of order $\xi(T)$, spins on a lenght scale na, with $n > 1$ and where a is the lattice spacing, act in some sense as a *single unit* as long as $a \ll na \ll \xi(T)$ (as we have just said previously, the scale change soon obliterates any short-range order). Let us consider an infinite (or nearly infinite) system and let us divide the system into boxes of side

$$L_n = f(n)\, a \quad \text{for } \; n = 0, 1, 2, ... \tag{5.1}$$

For simplicity we will use $a = 1$ and $f(n) = 2^n$ (in general we have kudkudAHHHHHHHHHHHHHHHHHHHHHH). We label each box by k. In this way[1] we construct a sequence of nested lattices having spacing L_n (for $n = 0$ we recover the original lattice). We define the block variables $\varphi_n(k)$ as

$$\varphi_n(k) = \sum_{i \in k_n} \sigma_i \tag{5.2}$$

where the sum runs over all the spins σ_i in the box k_n. This equation can be written recursively as

$$\begin{aligned} \varphi_0(k) &= \sigma_k \\ \varphi_{n+1}(k) &= \sum_{i_n \in k_{n+1}} \varphi_n(i_n) \end{aligned} \tag{5.3}$$

There are $2^{Dn} = L_n^D$ and 2^D terms in the sums in equation (5.2) and (5.3), respectively. The probability distribution $e^{-\beta H(\sigma)}$ of the σ variables induces naturally a probability distribution on the φ variables. As soon as L_n is definitively larger than the coherence length ξ (i.e., $n \gg n_\xi \simeq \log_2 \xi$) the correlations of block variables belonging to different blocks will be essentially zero[2]. In this case, in the large-n limit, the probability distribution of the $\varphi_n(k)$ is factorized; moreover, each φ_n is the sum of $2^{D(n-n_\xi)}$ practically uncorrelated variables. The *central limit theorem* tells us that these have a Gaussian distribution and their variance can be easily computed, using $\langle \sigma_i\, \sigma_j \rangle = \frac{\chi}{N\beta}$:

$$\langle \varphi_n(k)\, \varphi_n(k) \rangle = \left\langle \left(\sum_{i \in k_n} \sigma_i \right) \left(\sum_{j \in k_n} \sigma_j \right) \right\rangle \simeq \sum_{i \in k_n} \left\langle \sigma_i \left(\sum_j \sigma_j \right) \right\rangle = L_n^D \frac{\chi}{\beta} \equiv L_n^D \widetilde{\chi}\,. \tag{5.4}$$

In deriving this equation we have used the fact that if correlations are exponentially damped, the correlation function of a generic point i of the box (not too close to the boundary) with any point

---

[1] Usually k denotes the coordinates of the box in units of L_n. Sometimes we shall use the notation k_n to stress that we indicate a box of side L_n.

[2] This may happen only if we do not stay at the critical point.

outside the box is completely negligible: the points outside the box may be added without changing the final result. We thus find

$$\mathrm{d}P[\varphi_n] = \prod_{k_n} \frac{1}{(2\pi L_n^D\,\widetilde{\chi})^{1/2}} \exp\left[-\frac{\varphi_n^2(k)}{2L_n^D\,\widetilde{\chi}}\right] \mathrm{d}\varphi_n(k)\,. \tag{5.5}$$

If we introduce rescaled variables $\psi_n = L^{-D/2}\,\varphi_n$, the new variables have finite variance when $n \to \infty$, if we stay at $T \neq T_C$:

$$\mathrm{d}P[\psi_n] = \prod_{k_n} \frac{1}{(2\pi\widetilde{\chi})^{1/2}} \exp\left[-\frac{\psi_n^2(k)}{2\widetilde{\chi}}\right] \mathrm{d}\psi_n(k)\,. \tag{5.6}$$

In the presence of a nonzero magnetic field we find

$$\begin{aligned} &\mathrm{d}P[\psi_n] = \prod_{k_n} \frac{1}{(2\pi L_n^D\,\widetilde{\chi})^{1/2}} \exp\left[-\frac{\left(\varphi_n(k) - L_n^D m\right)^2}{2L_n^D\,\widetilde{\chi}}\right] \mathrm{d}\varphi_n(k) \\ &\langle\sigma\rangle = m\,. \end{aligned} \tag{5.7}$$

In this case we have two options: we can define

$$\begin{cases} \psi_n = \dfrac{\varphi_n}{L_n^D} \\ \widetilde{\psi}_n = \dfrac{\varphi_n - mL_n^D}{L_n^{D/2}} \end{cases} \tag{5.8}$$

The corresponding probability distribution are, for large n:

$$\begin{cases} \mathrm{d}P[\psi_n] = \prod_{k_n} \delta(\psi_n(k) - m)\,\mathrm{d}\psi_n(k) \\ \mathrm{d}P[\widetilde{\psi}_n] = \prod_{k_n} \dfrac{1}{(2\pi\widetilde{\chi})^{1/2}} \exp\left[-\dfrac{\widetilde{\psi}_n^2(k)}{2\widetilde{\chi}}\right] \mathrm{d}\widetilde{\psi}_n(k)\,. \end{cases} \tag{5.9}$$

Below T_C at $h = 0$, if we consider the symmetric nonclustering state $\langle\,\rangle_s = \frac{1}{2}\left[\langle\,\rangle_+ \langle\,\rangle_-\right]$, the precedent equation becomes

$$\mathrm{d}P[\psi_n] = \frac{1}{2} \sum_{\varepsilon=\pm1} \prod_{k_n} \delta(\psi_n(k) - \varepsilon m_s)\,\mathrm{d}\psi_n(k)\,, \tag{5.10}$$

where m_s is the spontaneous magnetization.
Let us consider from now on the case $h = 0$. From the preceding equations we see that, if we write

$$\psi_n(k) = \frac{\varphi_n(k)}{L_n^{D\omega}}\,, \tag{5.11}$$

the value of ω, such that the ψ has finite nonzero variance, jumps from $\frac{1}{2}$ to 1 when we cross the critical temperature[3]. At the critical temperature ω cannot be $\frac{1}{2}$ if $\chi = \infty$, nor can it be 1 if $m = 0$

---

[3] In most cases (possible exceprions may be zero-temperature antiferromagnetic materials), ω must satisfy the bounds $\frac{1}{2} \leq \omega \leq 1$

at $T = T_C$. Let us assume that there is one intermediate value of ω (ω_C), such that the ψ_n are well normalized when $n \to \infty$. The crucial hypothesis, which is at the basis of the renormalization group approach, is that for $\omega = \omega_C$ at the critical point

$$\lim_{n\to\infty} P_n[\psi_n] = P_\infty[\psi_\infty] \,. \tag{5.12}$$

Equation (5.12) states that the well-normalized large-scale block variables have a limiting probability distribution when $n \to \infty$.

The definition of the block variables (5.2) is not the only one possible; we could also write

$$\varphi_n(k) = \sum_i \left(\sigma_i \exp\left[-\frac{d^2(i,k_n)}{L_n^2} \right] \right) \,, \tag{5.13}$$

where $d(i,k_n)$ is the distance of the point i from the center of the block k_n; the advantage of the original definition (5.2) is the possibility of using the recursive relation (5.3). Other recursive definitions are also possible, as we shall see later. It is generally believed that the existence of a limiting probability distribution [Eq.(5.12)] does not depend on the way that the block variables are defined (at least for "reasonable" choices). A pictorial way to interpret Eq.(5.12) is as follows: We consider a two-dimensional system in which up spins are black and down spins are white. Let us take a television picture of the system from larger and larger distances. When the distance is small, we see individual spins; when the distance becomes large, we can only see the average of many spins [the resolution of the TV set acts like Eq.(5.13)]. The minimal size of the region we can resolve increases linearly with the distance. In order to avoid a uniform gray distribution, we must increase the contrast. The scaling hypothesis, Eq.(5.12) tells us that after an appropriate adjustment of the contrast, by looking at a picture taken from a very large distance, we cannot decide from which distance it has actually been taken if the distance is much greater than the lattice spacing. In other words, the probability distribution of the spins is invariant under scale transformations (after rescaling of the spins) in the large-distance region.

5.3 The recursion relation

The definition [equations (5.3 and (5.11))] of the block variables is recursive, i.e., the probability distribution $P_{n+1}[\psi_{n+1}]$ is obtained directly from the probability distribution $P_n[\psi_n]$. We can describe this fact by writing

$$\begin{aligned} &P_{n+1} = R[P_n] \\ &\psi_{n+1}(k) = 2^{-D\omega_n} \sum_{i_n \in k_{n+1}} \psi_n(i_n) \end{aligned} \tag{5.14}$$

where R is a nonlinear operator acting on probability measures. The asymptotic probability P_∞ satisfies the *fixed-point condition*:

$$P_\infty = R[P_\infty] \,. \tag{5.15}$$

The ψ variables satisfy the normalization $\langle \psi_n^2 \rangle = 1$, and the value of ω_n will in general depend on n. The basic idea of the renormalization group approach is to study the transformation R; as we shall see, we can connect in a simple way the singularities of the free energy (critical temperature, critical exponent) to the properties of R, where R remains regular at the transition. We hope that in computing the action of R we can consider correlated fluctuations only on a scale of a few lattice spacings (i.e., of order L_n in the original lattice). The study of R should be possible also in a finite, not very large system; in this way we would bypass the difficulty pointed out at the beginning of this chapter. It is often convenient to introduce the effective Hamiltonians for the block variables[4]

$$P_n[\psi_n] \propto \exp\{-H_n[\psi_n]\}\,. \tag{5.16}$$

Equation (5.15) induce a transformation on the space of all possible Hamiltonians

$$H_{n+1} = R_H[H_n]\,. \tag{5.17}$$

If the original Hamiltonian $H_0 = \beta H$ is short-range, it is believed (see section 7.5 Parisi) that H_n also remains essentially short range, that is, the coefficients of terms of the type $(\psi_k - \psi_{k+l})^2$ are exponentially small when $l \to \infty$, even at the critical point where long- range correlations are present. The space of essentially short-range Hamiltonians is not so huge as the space of all possible probability distributions; it should be possible to use this fact both for a rigorous construction of the operator R_H and for finding simpler approximate transformations, e.g., by projecting equation (5.17) onto the "small" space of finite-range Hamiltonians. It is crucial that the transformation R_H always remain regular, even at the critical point. The present theory of critical phenomena assumes that all the singularities of the free energy and the onset of long-range correlations are generated by a regular R transformation.

5.4 Fixed Points

The crucial ingredient of the RG method is the recognition of the importance and physical significance of fixed points of the RG transformation. In this section, we will develop these ideas in general terms.

5.4.1 Fixed Points and Basin of Attractions

We assume that there is a (Banach?) space of Hamiltonians for which equation (5.17) makes sense. In order to visualize the situation, it is convenient to discuss the case in which we can define a function H_λ such that

$$\begin{aligned} H_\lambda|_{\lambda=n} &= H_n \\ \frac{\mathrm{d}H}{\mathrm{d}\lambda} &= T[H_\lambda] \end{aligned} \tag{5.18}$$

where T is a nonlinear operator (formally, $\exp T = R$). This equation[5] has the graphic advantage of describing a *flow* in a multidimensional space, the parameter λ playing the role of time. If H^* is a fixed point of the transformation (5.17), $R(H^*) = H^*$, we have

$$T[H^*] = 0\,. \tag{5.19}$$

Any H satisfying this equation is said to be a *fixed point of the renormalization group transformation.* Equation (5.12) tells us that, when the "time" λ goes to infinity, H_λ goes to a fixed point. For a

---

[4] It is convenient to consider the temperature as one of the parameters characterizing the probability distribution; for the initial case, it is convenient to write $P_0 \propto \exp[-\beta H] \equiv \exp[-H_0]$.

[5] Note that we have passed from a recursive relation to a differential equation

given fixed point H^*, we can define an *attraction basin* B_f, i.e., the set of all points H_λ of the space such that

$$\lim_{\lambda \to \infty} H_\lambda = H^* \quad \text{if} \quad H_\lambda|_{\lambda=0} = H \tag{5.20}$$

Different points within the same basin have the same large "time" limit. Equations (5.12) and (5.20) tell us that the whole space can be partitioned into attraction basins of different fixed points. A fixed point is said to be *attractive* if all the points nearby belong to its basin of attraction; it is said to be *repulsive* if its basin of attraction contains only one point (the fixed point itself!). There are also intermediate possibilities: the fixed point is attractive when approached from one direction and it is repulsive from the other directions (typical examples are shown in Fig.(5.4)).

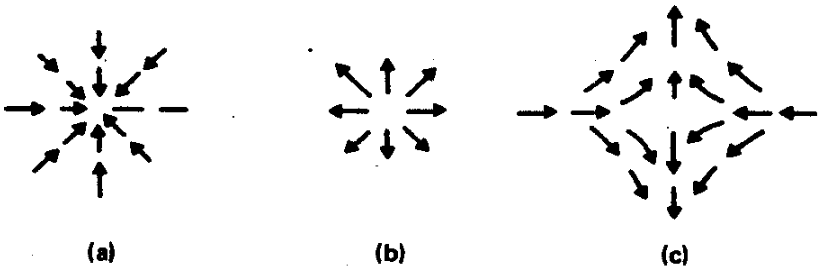


Figure 5.4. Examples of fixed points: (a) attractive; (b) repulsive; and (c) mixed.

5.4.2 Manifold of the coupling constants

To implement successfully the procedure of the RG it is obviously important that the new effective hamiltonian H_{n+1} has the same functional form as H_n, so that the model remains the same at each step of the sequence, beside a change in the value of its coupling constants. As a matter of fact this is impossible if we restrict attention to the hamiltonians with a finite number of couplings, since at each step new couplings are generated: for instance, starting from a hamiltonian with interaction among the next neighbor spins, the new hamiltonian has a new interaction among the spins separated by more than a lattice spacing and, furthermore, interactions that involve more than two spins. For this reason, it is convenient to start from the very beginning with the ensemble of all possible coupling constants that are compatible with the symmetry of the model and the nature of the statistical variables. Let's introduce then the manifold of the coupling constants and denote by $\mathbf{g}^{(n)} = \left(g_1^{(n)}, g_2^{(n)}, \ldots\right)$ the set of all the couplings of the effective hamiltonian H_n. In such a manifold, the application of the recursive relation can be interpreted as a motion of the point $\mathbf{g}$ in a space whose axes are the coupling constants. On iterating the RG transformation, a given system represented by its initial set of coupling constants, traces out a trajectory in coupling constants space. This motion is made in discrete time steps and ruled by

$$\mathbf{g}^{(n+1)} = \mathcal{R}[\mathbf{g}^{(n)}] \tag{5.21}$$

where $\mathcal{R}$ is, in general, a complicated nonlinear transformation. Starting from a point $\mathbf{g}^{(0)}$ and applying (5.21), the point of the system evolves in the sequence $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots$ giving rise in this way to a renormalization group trajectory. It is important to stress that all points of the trajectory describe the same physical situation: they simply correspond to an observation of the system with a different magnifying glass. The set of all such trajectories, generated by different initial sets of coupling constants generates a *renormalization group flow* in coupling constant space. Although it is possible, in principle, for the trajectory of the representative point to trace out *limit cycles*, *strange attractors*, etc., in practice, it is almost always found that the trajectory becomes attracted to fixed points. As we will see, scaling behaviour is invariably associated with the dynamics near a particular sort of fixed point, and the nature of the fixed points and the flows of the representative point in coupling constant space provide important information, allowing the phase diagram of the system to be determined.

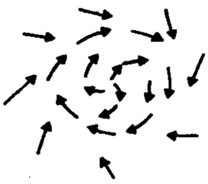


Figure 5.5. A flow producing a limiting cycle.

5.4.3 Physical Significance of Fixed Points

Let us suppose that we know the RG transformation $\mathcal{R}$. Then the fixed point of the RG transformation is a point $\mathbf{g}^*$ in coupling constant space satisfying

$$\mathbf{g}^* = \mathcal{R}[\mathbf{g}^*] \tag{5.22}$$

Now, under the RG transformation Re, length scales are reduced by a factor, say L. For any particular values of the coupling constants, we can compute the correlation length ξ, which transforms under $\mathcal{R}$ according to the rule

$$\xi(\mathbf{g}^{(n+1)}) = \frac{\xi(\mathbf{g}^{(n)})}{L}\,, \tag{5.23}$$

indicating that the system moves further from criticality after a RG transformation has been performed. At a fixed point,

$$\xi(\mathbf{g}^*) = \frac{\xi(\mathbf{g}^*)}{L} \tag{5.24}$$

which implies that $\xi(\mathbf{g}^*)$ can only be zero or infinity.

We will refer to a fixed point with $\xi = \infty$ as a *critical fixed point*, and a fixed point with $\xi = 0$ as a *trivial fixed point.* In general, a RG transformation will have several fixed points. Each fixed point has its own basin of attraction or domain: all points in coupling constant space which lie within the basin of attraction of a given fixed point flow towards and ultimately reach the fixed point after an infinite number of iterations of $\mathcal{R}$. We can easy see that all points in the basin of attraction of a critical fixed point have infinite correlation lenght. Suppose infact that we start with a physical system represented by the point in coupling constant space $\mathbf{g}^{(0)}$. After n iterations of the RG transformation $\mathcal{R}$, the system is now at a representative point denoted by $\mathbf{g}^{(n)}$. Using the recursive relation, we have the sequence of identities

$$\xi(\mathbf{g}^{(0)}) = L\,\xi(\mathbf{g}^{(1)}) = L^2\,\xi(\mathbf{g}^{(2)}) = \cdots = L^N\,\xi(\mathbf{g}^{(N)}) \tag{5.25}$$

fo any N. Taking the limit $N \to \infty$, the right hand side of this equation becomes infinity if $\xi(\mathbf{g}^*) = \infty$; i.e., if $\mathbf{g}^{(0)}$ was in the basin of attraction of a critical fixed point. This set of points

- the basin of attraction of a critical fixed point - is often called the *critical manifold.* The fact that all points on the critical manifold flow towards the same fixed point is the basic mechanism for *universality*, but is by no means the complete explanation. Universality, after all, involves behaviour exhibited by systems close to, but not at the critical point, and we have so far not said anything about this case. To complete our account, we will need to examine the behaviour off the critical manifold. We will shortly see that the critical fixed points describe the singular critical behaviour, whereas the trivial fixed points describe the bulk phases of the system. Knowledge of the location and nature of the fixed points of a RG transformation thus enables the phase diagram to be determined, whilst the behaviour of the RG flows near a critical fixed point determines the critical exponents.
Although we have implicitly assumed that the fixed points are isolated points, this is not necessarily the case. It is possible to have lines and surfaces of fixed points, and later, we will classify fixed points according to their *codimension.*

5.4.4 Local Behavior of RG Flows Near a Fixed Point

The nature of the fixed points can be determined by studying the linear version of equation (5.18). If the g_a are the set of parameters on which the Hamiltonian depends (e.g., $H^{[g]}_\lambda[\psi] = \sum_a g_a(\lambda) w_a[\psi]$, the w's being a "complete set" of functions of the block variables ψ), equation (5.18) can also be written as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\, g_a(\lambda) = \mathcal{T}[g_a]\,. \tag{5.26}$$

(We have obtained the differential equation associated to the recursive relation (5.21)). The fixed-point condition is now

$$\mathcal{T}[g^*_a] = 0\,. \tag{5.27}$$

The number of parameters g_a is infinite in the general case, but it is finite in most of the approximations. In order to study the mathematical properties of equations (5.18) and (5.26) it is convenient to suppose that the number of parameters g_a is finite (let us say M) and that the true transformation T is gentle enough that there are essentially no differences between the finite and the infinite-dimensional cases. Neglecting pathologies, we can now classify the fixed point according to the dimension d_B (or the co-dimension $c_B = M - d_B$) of the basin of attraction (e.g., if the fixed point is attractive $c_B = 0$ and $d_B = M$). Now near the fixed point the transformation (5.18) may be linearized. If $H_\lambda = H^* + \delta H_\lambda$ and $\delta\lambda$ is small, we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\,\delta H_\lambda = T_L\,\delta H_\lambda + O(\delta H^2_\lambda) \tag{5.28}$$

where T_L is now a linear operator. Correspondingly equation (5.26) becomes

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\lambda}\,\delta g_a(\lambda) = \mathcal{T}[g^*_a + \delta g_a] = \sum_{b=1}^{M} \widetilde{\mathcal{K}}_{ab}\,\delta g_b(\lambda) + O(\delta g_a(\lambda)^2) \\ &g_a(\lambda) = g^*_a + \delta g_a(\lambda), \qquad \widetilde{\mathcal{K}}_{ab} = \left.\frac{\partial \mathcal{T}_a}{\partial g_b}\right|_{g_a = g^*_a}\,. \end{aligned} \tag{5.29}$$

Remember always that the operators $\mathcal{R}$, T, $\mathcal{T}$, T_L and $\widetilde{\mathcal{K}}$ depend on the lenght scale L of the transformation.

A well-known theorem states that if all the real parts of the eigenvalues of T_L (or $\widetilde{\mathcal{K}}$) are negative, the fixed point is attractive, if there are n eigenvalues with positive real part, $c_B = n$.[6] In the typical situation that we shall study, there are two attractive fixed points, the *low-* and the *high-temperature* ones: the two basins of attraction are separated by a surface of co-dimension one (the previously cited critical surface). By continuity arguments, if a trajectory starts on this surface, it will remain there forever. Equation (5.12) implies that on this surface there must be at least one fixed point (H_C) which is attractive on the surface, but repulsive from outside, i.e., $\widetilde{\mathcal{K}}$ has one unstable (positive real part) eigenvalue. A two-dimensional example is shown in Fig.(5.6).

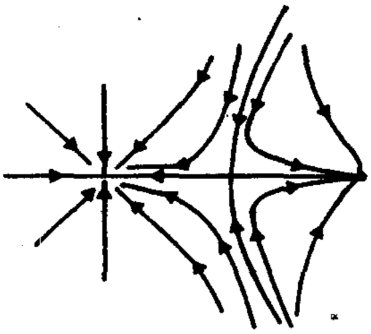


Figure 5.6. A flow with two attractive fixed points separated by a critical surface (dashed line) on which there is a partially attractive fixed point.

5.4.5 Relevant, Irrelevant and Marginal variables

For simplicity we study the linear version of equation (5.21): putting $\mathbf{g} = \mathbf{g}^* + \delta\mathbf{g}$, one has

$$\mathbf{g}^* + \delta\mathbf{g}' = \mathcal{R}[\mathbf{g}^* + \delta\mathbf{g}] \simeq \mathcal{R}[\mathbf{g}^*] + \mathcal{K}\,\delta\mathbf{g} = \mathbf{g}^* + \mathcal{K}\,\delta\mathbf{g} \tag{5.30}$$

namely

$$\delta g'_a = \mathcal{K}_{ab}\,\delta g_b\,, \tag{5.31}$$

where the matrix $\mathcal{K}_{ab}$ is defined as

$$\mathcal{K}_{ab} = \frac{\partial \mathcal{R}_a}{\partial g_b}\,. \tag{5.32}$$

This matrix is not necessarily symmetric and for this reason it is necessary to distin- guish between the right and the left eigenvectors. We denote the eigenvalues and eigenvectors by $\Lambda_L^{(i)}$ and $e_a^{(i)}$ respectively, where i labels the eigenvalues and the subscript a labels the component of the vector $\mathbf{e}$. Using Einstein summation convention, we have

$$\mathcal{K}_{ab}\,e_a^{(i)} = \Lambda_L^{(i)}\,e_b^{(i)}\,. \tag{5.33}$$

In terms of $e_a^{(i)}$ let's now define a linear combination of the displacements

$$u_i = \sum_a e_a^{(i)}\,\delta g_a\,. \tag{5.34}$$

---

[6] We must also consider the nonlinear terms in order to study what happens when the real part of the eigenvalues is exactly equal to zero.

These linear combinations are called *scaling variables*. They have the important quality of transforming in a multiplicative way under the RG transformations

$$u_i' = \sum_a e_a^{(i)}\, \delta g_a' = \sum_{a,b} e_a^{(i)}\, \mathcal{K}_{ab}\, \delta g_b = \Lambda_L^{(i)} \sum_b e_b^{(i)}\, \delta g_b = \Lambda_L^{(i)} u_i\,. \tag{5.35}$$

If L is the rescaling parameter of the block spins, it is common to parameterize[7] $\Lambda_L^{(i)}$ as $\Lambda_L^{(i)} = L^{y_i}$ where the quantities y_i are improperly called the eigenvalues of the renormalization group: we will (maybe) show that they determine the critical exponents of the statistical model. Disregarding the case in which y_i is a complex number, we can have the following cases:

- $y_i > 0$. In this case the corresponding u_i is a *relevant variable*. A repeated application of the transformations moves its value *away* from the critical point.
- $y_i < 0$. In this case u_i is an *irrelevant variable*. Starting sufficiently close to the fixed point, the iteration of the transformation shrinks the initial value to zero.
- $y_i = 0$. In this case u_i is a *marginal variable*. Iterating the transformation, the value of this variable does not change. Their influence cannot be determined by the linearized transformations that we have considered.

The significance of these distinctions is that if we start at $\mathbf{g}$ near $\mathbf{g}^*$, but *not* on the critical manifold then the flows away from $\mathbf{g}^*$ i.e. in directions out of the critical manifold in the vicinity of $\mathbf{g}^*$, are associated with relevant eigenvalues. The irrelevant eigenvalues correspond to directions of flow into the fixed point. The eigenvectors corresponding to the irrelevant eigenvalues *span the critical manifold.* The marginal eigenvalues turn out to be associated with logarithmic corrections to scaling, and are important at the upper and lower critical dimensions. The number of relevant eigenvalues must thus be the codimension c_B of the critical manifold, i.e., the difference between the dimensionalities of the coupling constant space and the critical manifold. It is very important to remember that the terms relevant, irrelevant and marginal are always to be specified with respect to a particular fixed point. A particular term in the Hamiltonian may be relevant at one fixed point, but not at another.

5.4.6 Critical surface

To continue the analysis, let's assume that the dimension of the space of the coupling constants is M and let's consider a fixed point $\mathbf{g}^*$ with N relevant variables and $(M - N)$ irrelevant variables. This means that there exists a $(M-N)$-dimensional surface $\mathcal{C}$ (which has a codimension $c_B = N$), called the *critical surface*, that is the attractive basin for the fixed point $\mathbf{g}^*$. As shown below, on this surface the correlation length is infinite. The coupling constants g_k of the system depend generally on the external parameters of the system, such as temperature, pressure, or magnetic field. Varying these external parameters, the point $\mathbf{g}$ of the coupling constants varies correspondingly. When there are N relevant variables, in order to intercept the critical surface it is necessary to choose appropriately N external control parameters. In all cases of physical interest, the temperature is one of these parameters and its value has to be tuned to its critical value $T = T_C$ to hit the critical surface. This may not be enough: if there are magnetic fields, they must be switched off and it

---

[7] This is due to the *semi-group* property satisfied by the operator $\mathcal{R}$ and $\mathcal{K}$.

may also be necessary to tune appropriately the chemical potential. Once such a fine tuning of the n experimental parameters has been done, the point **g** in on the critical surface. If we now apply the RG transformations, their iterations of the RG move the point toward the critical point g^*, independently of its initial position on $\mathcal{C}$, as shown in Fig.(5.7). This is, in a nutshell, the origin of the universal behavior of the critical phenomena: hamiltonians that differ only for their irrelevant operators give rise to the same critical behavior.

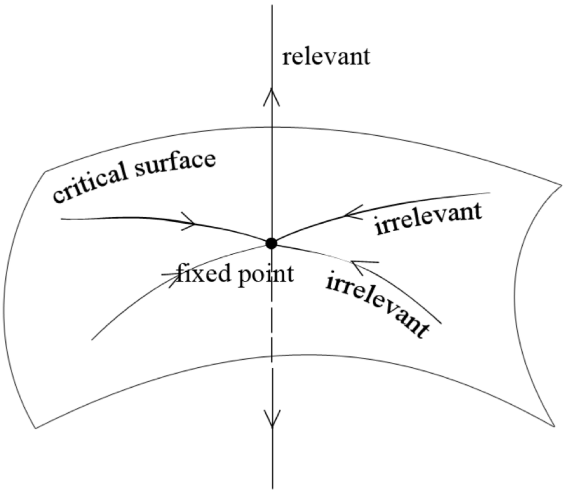
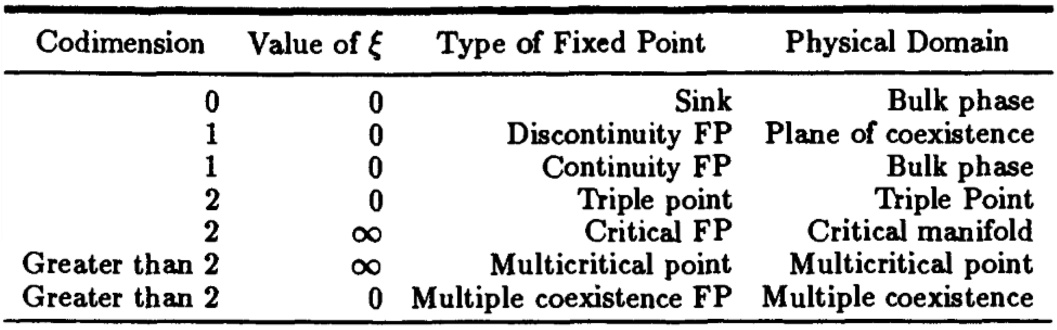


Figure 5.7

Table (9.1) shows a classification of fixed points by their codimension. Those with codimension 0 have no relevant directions, and therefore trajectories only flow into them, giving rise to the name *sink*. The sinks correspond to stable bulk phases, and the nature of the coupling constants at the sink characterise the phase. For example, a simple three dimensional Ising magnet with nearest neighbour ferromagnetic coupling in an external field j has sinks at $h = \pm\infty$, $T = 0$, corresponding to the fact that in a positive (negative) external field, there is a net positive (negative) magnetisation for all temperatures. Starting at any point in the phase diagram (h, T), successive RG iterations will drive the system to the sink corresponding to the appropriate sign of h.

Table 9.1 CLASSIFICATION OF FIXED POINTS



Codimension	Value of ξ	Type of Fixed Point	Physical Domain
0	0	Sink	Bulk phase
1	0	Discontinuity FP	Plane of coexistence
1	0	Continuity FP	Bulk phase
2	0	Triple point	Triple Point
2	∞	Critical FP	Critical manifold
Greater than 2	∞	Multicritical point	Multicritical point
Greater than 2	0	Multiple coexistence FP	Multiple coexistence

Usually the initial Hamiltonian at $\lambda = 0$ depends on β; in this way we have a whole family of trajectories $H(\lambda, \beta)$,

$$\begin{aligned} &H(\lambda, \beta)|_{\lambda=0} = H_0(\beta) \equiv \beta H \,, \\ &\frac{\partial H}{\partial \lambda} = T[H] \,. \end{aligned} \tag{5.36}$$

If $H_0(\beta)$ belongs to two different attraction basins for small and large values of β, respectively, there must be a value of β (β_c) such that $H_0(\beta)$ crosses the critical surface. If ω_R and ω_A are the two largest eigenvalues of T_L at H_c ($\omega_R > 0 > \omega_A$), the corresponding eigenvectors being H_R and H_A, we have for large λ and small $|\beta - \beta_c|$

$$\begin{aligned} &H(\lambda, \beta) \simeq H_c + H_A \, e^{-\lambda|\omega_A|} + (\beta - \beta_c) \, H_R \, e^{-\lambda\omega_R} \qquad \text{if } \; e^{\lambda\omega_R} \, |\beta - \beta_c| \ll 1 \\ &H(\lambda, \beta) \simeq H_\pm \left(\lambda + \frac{\ln\left(|\beta - \beta_c|\right)}{\omega_R} \right) + O(|\beta - \beta_c|^{|\omega_A|/\omega_R}) \qquad \text{if } \; \lambda + \frac{\ln\left(|\beta - \beta_c|\right)}{\omega_R} \gg 0 \end{aligned} \tag{5.37}$$

where H_+ and H_- are the two escape trajectories for $\beta < \beta_c$ and $\beta > \beta_c$, respectively, depending on the nonlinear terms neglected in equation; more precisely, $H_\pm(t) = H_0 \pm e^{t\omega_R} H_R + O\left(e^{2t\omega_R}\right)$. If we do not start exactly from β_c but only close to β_c., the trajectory will stop near H_c for a long "time" proportional to $\ln\left(|\beta - \beta_c|\right)/\omega_R$; for $\lambda \ll -\ln\left(|\beta - \beta_c|\right)/\omega_R$ the probability distribution of the block variables will be very similar to that at the critical point, while for $\lambda \gg -\ln\left(|\beta - \beta_c|\right)/\omega_R \gg 0$ the probability distribution of the block variables will be very close to the low- (or high-) temperature fixed point. In this last region the connected correlation functions of the block variables will be essentially zero. This picture is not essentially changed if there are two unstable eigenvalues of T_L^c: The critical surface now has co-dimension two, and two parameters must be tuned (e.g., temperature and magnetic field, or temperature and pressure) in H_0 in order to reach the critical point.

Chapter 6

Other Models

6.1 The Ising model again

In the preceding chapters we constructed the diagrammatic perturbative expansion for the Ginsburg-Landau model. Here we study the Ising case. We have already seen the relations between the high-temperature expansions of the Ising-Gaussian and random walk models. To avoid painful combinatoric proofs, we proceed in a compact way.

We start from the Ising model Hamiltonian, which we write, as usual (see p. 23),

$$H = -\frac{1}{2} \sum\nolimits_{i,k} J_{ik} \sigma_i \sigma_k - \sum\nolimits_i h_i \sigma_i \,. \tag{11.1}$$

We now have (forgetting factors π):

$$\begin{gathered} \exp[-\beta H] = \det{}^{1/2} \left(\frac{J}{\beta}\right) \int d\varphi_i \exp\left[-\frac{\beta}{2} J^{-1}_{ik} \varphi_i \varphi_k - (\beta\varphi_i + \beta h_i)\sigma_i\right] \\ \sum\nolimits_k J^{-1}_{ik} J_{kl} = \delta_{il} \,. \end{gathered} \tag{11.2}$$

The sum over the configurations of the σ's can be trivially done. We get

$$\begin{aligned} Z &= \det{}^{1/2}\left[\frac{J}{\beta}\right] \int \prod\nolimits_i d\varphi_i \exp\left[-\frac{\beta}{2} \sum_{i,k} J^{-1}_{ik} \varphi_i \varphi_k\right] \sum_{\{\sigma\}} \exp\left[-\beta \sum\nolimits_i \sigma_i (h_i \right. \\ &\qquad \left. + \varphi_i)\right] \\ &= \det{}^{1/2}\left[\frac{J}{\beta}\right] \int \prod\nolimits_i d\varphi_i \exp\left\{-\frac{\beta}{2} \sum\nolimits_{i,k} J^{-1}_{ik} \varphi_i \varphi_k \right. \\ &\qquad \left. + \sum\nolimits_i \ln[\cosh \beta(\varphi_i + h_i)]\right\} . \end{aligned} \tag{11.3}$$

By differentiating Eq. (11.3) with respect to h and using the linear response theory we obtain

$$\begin{aligned} \langle \sigma_i \rangle &= \langle \tanh[\beta(\varphi_i + h_i)] \rangle \\ \langle \sigma_i \sigma_j \rangle_c &= \langle \tanh[\beta(\varphi_i + h_i)] \tanh[\beta(\varphi_j + h_i)] \rangle_c \qquad (11.4) \\ &\quad + \delta_{ij} \left\langle \frac{1}{\cosh^2[\beta(\varphi_i + h_i)]} \right\rangle , \end{aligned}$$

where the expectation values on the r.h.s. of Eq. (11.4) are evaluated with the Hamiltonian

$$\beta H[\varphi] = \frac{\beta}{2} \sum_{i,k} J^{-1}_{i,k} \varphi_i \varphi_k - \ln \cosh[\beta(\varphi_i + h_i)] \, . \qquad (11.5)$$

If we look for the minimum of $H[\varphi]$ in the presence of a constant field ($h_i = h$), we find

$$\textstyle\sum_k J^{-1}_{ik} \varphi_k = \tanh[\beta(\varphi_i + h)] \, , \qquad (11.6)$$

which can also be written as

$$\beta\varphi \equiv \beta\varphi_k = \textstyle\sum_k J_{ik} \tanh[\beta(\varphi_k + h)] = 2D \tanh(\beta\varphi + h) \, . \quad (11.7)$$

If we substitute for the expectation values in Eq. (11.4) their values evaluated at the minimum of $H[\varphi]$, the final result for $\langle \sigma_i \rangle$ can readily be seen to be equivalent to the old mean-field equations.

A better understanding may be obtained by developing $H[\varphi]$ in powers of φ (this may be justified for high temperatures). We obtain at $h = 0$

$$\begin{aligned} \beta H[\varphi] &= \sum_{i,k} \frac{\beta}{2} (J^{-1}_{ik} - \beta\delta_{ik}) \varphi_i \varphi_k + \textstyle\sum_i \left(\frac{\beta^4}{12} \varphi_i^4 - \frac{\beta^6}{45} \varphi_i^6 + O(\varphi_i^8 \beta^8) \right) \\ &= \sum_{i,k} \frac{1}{2} (J^{-1}_{ik} - \beta\delta_{ik}) \tilde{\varphi}_i \tilde{\varphi}_k + \textstyle\sum_i \left[\frac{\beta^2 \tilde{\varphi}_i^4}{12} - \frac{\beta^3 \tilde{\varphi}_i^6}{45} + O(\beta^4 \tilde{\varphi}_i^8) \right] \end{aligned} \qquad (11.8)$$

$$\tilde{\varphi}_i = \beta^{1/2} \varphi_i \, .$$

By comparing Eq. (11.8) with the results of Chapter 4 we see that considering only the quadratic term in $H[\varphi]$ corresponds to the random-walk approximation, while higher orders in φ give interactions of the type we have already studied. If we neglect all powers greater than 4, we get a lattice Landau-Ginsburg-type model, with $g = \frac{1}{2}\beta^2$; the propagator at $g = 0$ given by

$$G(p) = \frac{\sum_{1}^{D}{}_{\nu} 2\cos(p_\nu)}{1 - 2\beta \sum_{1}^{D}{}_{\nu} \cos(p_\nu)} = \frac{1}{J^{-1}(p) - \beta}$$

$$J(p) = \sum_{1}^{D}{}_{\nu} 2\cos p_\nu \tag{11.9}$$

$$\langle \tilde{\varphi}_0 \tilde{\varphi}_k \rangle = \frac{1}{(2\pi)^D} \int_B d^D p \, G(p) \exp(ipk) .$$

In this way we can reconstruct the high-temperature expansion using the rules of the standard perturbative expansion and developing $G(p)$ in powers of β.

The main difference between the Ising and Landau-Ginsburg models is the presence of higher powers of $\tilde{\varphi}$ in the first case. We shall argue later in this chapter that these higher powers do not change the critical exponents and the fixed-point Hamiltonian.

As an exercise we can sum all the diagrams of Fig. 11.1 for the self-energy in the high-temperature phase at $h = 0$. If we write $\ln\cosh \beta^{1/2}\tilde{\varphi} = \Sigma_n c_n \beta^{n/2} \tilde{\varphi}^n$, it is evident that the self-energy is given by

$$\begin{aligned} \Sigma &= -\sum_{4}^{\infty}{}_{n} n(n-1) c_n \beta^{n/2} \langle \tilde{\varphi}_i^{n-2} \rangle = \beta \langle \tanh^2(\beta^{1/2}\tilde{\varphi}) \rangle \\ G(p) &= \frac{1}{J(p)^{-1} - \beta - \Sigma} , \end{aligned} \tag{11.10}$$

where the expectation values are taken with the Hamiltonian (11.8) at $\beta = 0$. The simplest way to evaluate $\langle g(\tilde{\varphi}_i) \rangle_0$, g being an arbitrary function, is to notice that, if the Hamiltonian is quadratic, the probability

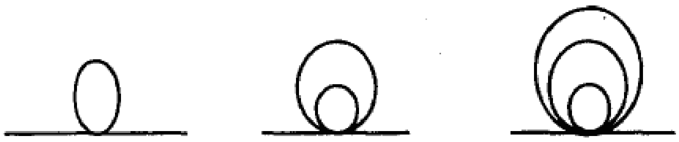


Figure 11.1. Examples of diagrams whose sum is given by Eq. (11.10).

distribution of $\tilde{\varphi}_i$ is a Gaussian with variance $\langle \tilde{\varphi}^2 \rangle_0$. We thus obtain

$$\begin{aligned} \langle g(\tilde{\varphi}_i) \rangle_0 &= \frac{1}{[2\pi \langle \tilde{\varphi}^2 \rangle_0]^{1/2}} \int_{-\infty}^{\infty} dz \exp\left[-\frac{z^2}{2\langle \varphi^2 \rangle_0} \right] g(z) \\ \langle \tilde{\varphi}^2 \rangle_0 &= \frac{1}{(2\pi)^D} \int_B d^D p \, \frac{1}{J(p)^{-1} - \beta} \, . \end{aligned} \tag{11.11}$$

The "gap" equation that corresponds to the sum of all diagrams of the shape shown in Fig. 11.2 is

$$\begin{aligned} \Sigma &= \langle \tanh^2(\beta^{1/2} \tilde{\varphi}^a) \rangle \\ &= \frac{1}{(2\pi \langle \tilde{\varphi}^2 \rangle)^{1/2}} \int_{-\infty}^{\infty} dz \exp\left(-\frac{z^2}{2\langle \tilde{\varphi}^2 \rangle} \right) \tanh^2(\beta^{1/2} z) \\ \langle \tilde{\varphi}^2 \rangle &= \frac{1}{(2\pi)^D} \int_B d^D p \, \frac{1}{J^{-1}(p) - \beta - \Sigma} \, . \end{aligned} \tag{11.12}$$

This equation is a generalization of Eq. (6.6).

The careful reader has certainly already noticed that Eqs. (11.2)–(11.3) do not make sense; the functional integral representation (11.2) is not convergent because the matrix J is not a positive operator [$J(p)$ is not positive]; none of the functional representations we have written are convergent. The remedy to this difficulty is rather simple in momentum space: We must integrate over the φ fields from $-\infty$ to $+\infty$ when $-\pi/2 \le p \le \pi/2$ as before; however, the integration path should be deformed in the complex plane and go from $-i\infty$ to $+i\infty$ when $-\pi \le p \le -\pi/2$ or $\pi/2 \le p \le \pi$.

One can ignore the integration paths if one perturbs around the

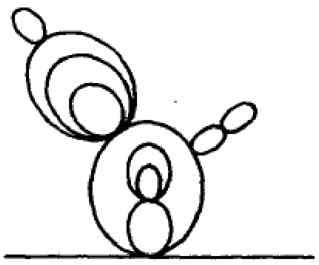


Figure 11.2. An example of a diagram that contributes to Eq. (11.12).

Gaussian integrals and uses the rule

$$\frac{\int dx \exp(-x^2/2\alpha)\, x^2}{\int dx \exp(-x^2/2\alpha)} = \alpha \tag{11.13}$$

even for negative α, as we have implicitly done in the previous equations.

A more subtle problem is the following: the interaction Hamiltonian does not change (apart from a trivial additive factor) if we add to it a term $-\lambda \sum_i \sigma_i^2$, because $\sigma_i^2 = 1$. However, if such a term is included in the Gaussian transformation, the form of the matrix $J(p)$ changes; it becomes

$$\sum_\nu 2 \cos p_\nu + \lambda . \tag{11.14}$$

Obviously if we cut the perturbative expansion at any given order, the results depend on λ; in particular, the critical temperature is shifted.[1]

The motivations for sticking to the case $\lambda = 0$ arise from the need to minimize the diagrams with loops. Indeed, at $\beta \simeq 0$, $\lambda = 0$, $\langle \varphi^2 \rangle = 0$, and the diagrams of Fig. 11.1 give a zero contribution. Moreover, if $\lambda = 0$ (or if λ does not increase with D) when $D \to \infty$, all the diagrams of Fig. 11.2 go to zero and we recover the mean-field result. This can be seen by noticing that if $J(p)$ is given by Eq. (11.9) we have that

$$\begin{gathered} \frac{1}{(2\pi)^D} \int_B d^D p\, J(p) = \frac{1}{(2\pi)^D} \int_B d^D p\, J^3(p) = 0 \\ \frac{1}{(2\pi)^D} \int_B d^D p\, J^2(p) = 2D \qquad \frac{1}{(2\pi)^D} \int_B d^D p\, J^4(p) = 2D(2D+1) . \end{gathered} \tag{11.15}$$

The use of similar identities shows that all the loop diagrams vanish when $D \to \infty$, so that mean-field theory becomes exact in this limit. Of course the infinite-D limit must be taken at fixed $\tilde{\beta} = 2D\beta$ (see p. 36).

The detailed proof of this last statement can be left to the willing reader; the reader could also show how to use the loop expansion to generate the $1/D$ expansion. These considerations are not crucial here. Our aim was to show how the Ising model can be cast in a form reminiscent of the Landau-Ginsburg model.[2]

6.2 The Real Gas (Parisi)

If we have N real classical pointlike particles in a box of volume L^D the Hamiltonian can be written as

$$
\begin{aligned}
H &= \sum_{1}^{N}{}_{i} \frac{1}{2m} p_i^2 + U_N(x) \\
U_N(x) &= \frac{1}{2} \sum\nolimits_{i,j} V(x_i - x_j) ,
\end{aligned}
\tag{11.16}
$$

where we suppose that only two-body interactions are present. The partition function is given by

$$
\begin{aligned}
Z_N &= \int \prod\nolimits_i d^D p_i \, d^D x_i \exp[-\beta H(p, x)] = \left(\frac{2m\pi}{\beta}\right)^{DN/2} \tilde{Z}_N \\
\tilde{Z}_N &= \int \prod\nolimits_i d^D x_i \exp[-\beta U_N(x)] .
\end{aligned}
\tag{11.17}
$$

The thermodynamic limit is obtained when N goes to infinity at fixed density $\rho = N/L^D$.[3] For practical reasons it is more convenient to define the so-called grand partition function

$$
Z_G(\beta, z) = \sum\nolimits_N \frac{z^N}{N!} Z_N(\beta) ,
\tag{11.18}
$$

directly in the very large box, where $z = \exp(\beta\mu)$ is the fugacity, and μ is the chemical potential. The use of the grand partition function, the related physical problems, and the expansion of Z_G in powers of z (the virial expansion) are described in many textbooks of statistical mechanics.[4] Here we want only to rewrite Eq. (11.15) as a functional integral. This can be done by using the following identity:

$$
\begin{aligned}
\exp\left[-\frac{1}{2} \sum_{k,j}^{N} \beta V(x_k - x_j)\right] &= (\det V)^{-1/2} \int d[\varphi] \exp\left[-\int \tfrac{1}{2} V^{-1}(x-y)\right. \\
&\qquad \left. \times \varphi(x)\varphi(y) \, d^D x \, d^D y + i \sum_{1}^{N}{}_{k} \beta^{1/2} \varphi(x_k)\right] \\
\int dy \, V(x-y) V^{-1}(y-z) &= \delta^D(x-z) .
\end{aligned}
\tag{11.19}
$$

Equation (11.19) is the standard Gaussian integral formula. Apart from overall constants we can now write

$$
\begin{aligned}
Z_G(\beta, z) \propto & \int d[\varphi] \\
& \times \Big\{ \textstyle\sum_N \frac{z^N}{N!} \int \prod_{k=1,N} [dx_k \exp(i\beta^{1/2}\varphi_k)] \exp\Big[-\frac{1}{2} \int d^D x \, d^D y \\
& \times \varphi(x)\varphi(y)V^{-1}(x-y) \Big] \\
= & \int d[\varphi] \exp\Big[z \int d^D x \exp(i\beta^{1/2}\varphi(x)) \\
& + \frac{1}{2} \int d^D x \, d^D y \, V^{-1}(x-y)\varphi(x)\varphi(y) \Big] . \qquad (11.20)
\end{aligned}
$$

We again find a functional integral representation where the $\beta = 0$ correlation function is

$$\langle \varphi(x)\varphi(y) \rangle = V(x-y) .$$

We also obtain

$$\rho = \frac{\langle N \rangle}{V} = \frac{1}{V} \frac{\partial}{\partial_z} \ln Z_G(\beta, \mu) = \langle \exp(i\beta^{1/2}\varphi(x)) \rangle . \qquad (11.21)$$

The main differences between this and the Ising case are that all powers (of even and odd) are present and the Hamiltonian depends on two parameters. We expect, therefore, a critical point β_c and μ_c, which is the endpoint of a line of first-order transitions. The model is very similar to the Landau-Ginsburg model in the presence of a magnetic field (if powers of φ greater than 4 are neglected, the correspondence is exact after a shift in φ ($\varphi = \varphi_0 + \tilde{\varphi}^2$, which kills the cubic term).

The first-order transition may be identified with the gas-liquid transition, and the critical point with the endpoint of such a transition. It is clear that after the correct identifications are made, the critical exponents must be the same as for the Ising model, in good agreement with experimental data. It is possible to discuss the details of the phenomenology of the gas-liquid transition and to study the liquid-solid transition in this language. This is beyond the aim of this book.[5] Here we wish only to show that the concept that we have developed to study magnetic systems has a much wider domain of applications.

6.3 The Real Gas (Brezìn)

Consider molecules at thermal equilibrium in a box of volume V, in a range of temperatures and densities, in which one can ignore the quantum effects. If we treat this system in the grand canonical ensemble, the partition function is

$$Z = \sum_N \frac{e^{\beta\mu N}}{h^{3N} N!} \prod_{i=1}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \; \exp\left[-\beta \left(\sum_{i=1}^{N} \frac{p_i^2}{2m} + U(r_1, \ldots, r_N) \right) \right] , \tag{6.1}$$

in which $U(r_1, \ldots, r_N)$ is the interaction energy of N molecules and μ is the chemical potential. Integrating over the momenta, one obtains

$$Z = \sum_N \frac{z^N}{N!} Q_N \, , \tag{6.2}$$

with

$$z = \frac{e^{\beta\mu}}{\lambda^3} = \frac{e^{\beta\mu}}{h^3} \left(2\pi m k T \right)^{3/2} \, , \tag{6.3}$$

and

$$Q_N = \int \mathrm{d}^3 r_1 \ldots \mathrm{d}^3 r_N \, e^{-\beta U(r_1, \ldots, r_N)} \, . \tag{6.4}$$

To compute approximately this integral Q_N, one discretizes the volume V with a periodic lattice of $\mathcal{N}$ points. The total volume V is the product of the volume v of the elementary cell by $\mathcal{N}$ (for a cubic lattice). The molecules are located at the sites of this lattice. The interaction energy U contains a repulsion at short distance between the molecules, which we take into account by forbidding the occupation of a site of the lattice by more than one molecule. The sum over N in (6.2) is thus limited to $N \leq \mathcal{N}$.

The interaction energy also contains a short-range attraction between pairs of molecules. This may be modelled schematically by assigning an energy $-\epsilon$ to a pair of molecules that are nearest neighbours on the lattice; since the interaction is short ranged, we assume no potential energy for molecules which are not nearest neighbours. The total configuration energy is obtained by multiplying $-\epsilon$ by the number of pairs of nearest neighbours. A configuration is a choice of N sites among the $\mathcal{N}$. Once this choice is made, there are $N!$ ways of distributing the molecules over the N sites. It is convenient to define the occupation number of a site i, i.e., a number n_i which is equal to 1 if the site i is occupied and to 0 if it is empty. These n_i are not completely independent, since

$$\sum_{i=1}^{\mathcal{N}} n_i = N \, . \tag{6.5}$$

In terms of these occupation numbers, the interaction energy of a configuration $\mathcal{C}$ of the N molecules is given by

$$U(\mathcal{C}) = -\epsilon \sum_{\langle i,j \rangle} n_i n_j \, . \tag{6.6}$$

Since this energy is $N!$ times degenerate, one has then in this discretized formulation

$$Q_N = N! \sum_{\mathcal{C}} e^{-\beta U(\mathcal{C})} \, . \tag{6.7}$$

The constraint (6.5) makes the computation of the canonical partition function more difficult, since the configurations cannot be described by an independent assignment of the n_i to the whole lattice. However, the grand-canonical partition function is given simply by

$$Z = \sum_{\mathcal{C}} e^{(n_1 + \ldots n_{\mathcal{N}}) \ln z + \beta\epsilon \sum_{\langle i,j \rangle} n_i n_j} \, , \tag{6.8}$$

in which $\mathcal{C}$ is an assignment of 0 and 1 independently for each site over the whole lattice of $\mathcal{N}$ sites. To complete the identification with the Ising model one introduces a 'spin' for each occupation number:

$$n_i = \frac{1}{2}(1 + \sigma_i) \qquad \text{with} \quad \sigma_i = \pm 1\,. \tag{6.9}$$

This leads to

$$Z = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}\,, \tag{6.10}$$

with

$$E(\mathcal{C}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i + \text{constant}\,, \tag{6.11}$$

in which

$$J = \frac{1}{4} \qquad \text{and} \qquad h = \frac{\ln z}{2\beta} + \frac{c\epsilon}{4} \tag{6.12}$$

(c is the number of neighbours of a given site). If the Ising model presents a ferromagnetic order at temperatures $T < T_c$, it means that a change of h from $+0$ to -0 implies a transition between two phases of opposite magnetizations, which coexist along the line $h = 0$. How can one translate this into molecular language? The magnetization $\langle \sigma \rangle$ is related to the mean occupation of the sites. A jump of the magnetization from $+m_s$ to $-m_s$ is thus a discontinuity that is equivalent to the mean number of molecules per unit volume from a dense liquid phase to a gaseous phase. These two phases are spatially homogeneous ($\langle n_i \rangle$ is the same for all sites). The coexistence curve of the two phases in the plane (chemical potential versus temperature) is given by $h = 0$, $T \leq T_c$, i.e.,

$$\mu + \frac{c\epsilon}{2} + kT \ln \left(\left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right) = 0\,. \tag{6.13}$$

Chapter 7

The Transfer Matrix

The exact solution of the one-dimensional Ising model can be obtained by using the alternative method of the transfer matrix. This method presents a series of advantages: unlike the recursive method, it also can be applied when there is an external magnetic field. Moreover, it has many points in common with a discrete formulation of quantum mechanics, in particular the Feynman formulation in terms of a path integral. The transfer matrix method relies on a set of ideas that go beyond the application to the one-dimensional case and permits us to show the remarkable relationship that links classical systems of statistical mechanics in D dimensions with quantum systems in $(D-1)$. In the two-dimensional case, for instance, it permits us to obtain the exact solution of the Ising model in the absence of an external magnetic field. To study the one-dimensional case, let us consider once again a chain of N spins. For simplicity, we consider here the *homogeneous* case, in which there is only one coupling constant J, with hamiltonian

$$H = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} - h \sum_{i=1}^{N} \sigma_i \,. \tag{7.1}$$

The model, as defined, is easy to solve for all kinds of boundary conditions, but the calculations are even simpler if the chain is closed as if the sites were on a circle, i.e., if we use *periodic boundary conditions* (p.b.c.):

$$H = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - h \sum_{i=1}^{N} \sigma_i \,, \tag{7.2}$$

with the identification $\sigma_{N+1} = \sigma_1$. Clearly, the addition of one extra term in the energy has no effect on the ther- modynamic limit (it is easy here to calculate the free energy for open or periodic boundary conditions and verify this statement).

7.1 Definition and the free energy

Let us define

$$T(\sigma, \sigma') = e^{\beta\left(J\sigma\sigma' + \frac{1}{2}h(\sigma+\sigma')\right)} \,. \tag{7.3}$$

Then for periodic boundary conditions, one verifies immediately that

$$e^{-\beta H} = T(\sigma_1, \sigma_2)\, T(\sigma_2, \sigma_3) \ldots T(\sigma_{N-1}, \sigma_N)\, T(\sigma_N, \sigma_1) = \prod_{i=1}^{N} T(\sigma_i, \sigma_{i+1}) \,. \tag{7.4}$$

If one displays the four possible values of $T(\sigma, \sigma')$ when σ, $\sigma' = \pm 1$, one reads:

$$\mathcal{T} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J+h)} \end{pmatrix}, \tag{7.5}$$

in which the rows are labelled by the successive values $+1$ and -1 of σ , and the columns by the values of σ'. Summing over all configurations means summing over $\sigma_1 = \pm 1$, $\sigma_2 = \pm 1$, … , $\sigma_N = \pm 1$. Let us take the Boltzmann factor as written in (7.4) and sum over σ_2: it occurs only in the product of the first two factors $T(\sigma_1, \sigma_2)\, T(\sigma_2, \sigma_3)$. If we return to the matrix expression (7.5), the first σ_2 is an index of a column, the second of a row: therefore, the sum over σ_2 gives simply a product of the two matrices:

$$\sum_{\sigma_2} T(\sigma_1, \sigma_2)\, T(\sigma_2, \sigma_3) = \mathcal{T}^2(\sigma_1, \sigma_3), \tag{7.6}$$

in which $\mathcal{T}^2$ is the square of the matrix (7.5). Repeating this argument for the summation over $\sigma_3, \ldots, \sigma_N$ one arrives at

$$\sum_{\sigma_2, \ldots, \sigma_N} e^{-\beta H} = \mathcal{T}^N(\sigma_1, \sigma_1). \tag{7.7}$$

One last sum remains:

$$Z = \sum_{\{\sigma\}} e^{-\beta H} = \sum_{\sigma_1} \mathcal{T}^N(\sigma_1, \sigma_1), \tag{7.8}$$

the sum over the diagonal elements of the matrix $\mathcal{T}^N$. Therefore, we end up with

$$Z = \operatorname{Tr} \mathcal{T}^N. \tag{7.9}$$

The trace of a matrix is the sum of its diagonal elements, it is also the sum of its eigenvalues. We have thus to find the two eigenvalues of the matrix $\mathcal{T}^N$ but these eigenvalues are simply the Nth powers of the eigenvalues of $\mathcal{T}$. Therefore, we end up with

$$Z = \lambda_1^N + \lambda_2^N, \tag{7.10}$$

in which λ_1 and λ_2 are the two roots of the quadratic equation

$$\begin{aligned} &\lambda^2 - \left(e^{\beta(J+h)} + e^{\beta(J-h)}\right)\lambda + e^{2\beta J} - e^{-2\beta J} = 0 \\ &\implies \lambda^2 - 2e^{\beta J}\cosh(\beta h)\lambda + 2\sinh(2\beta J) = 0, \end{aligned} \tag{7.11}$$

namely

$$\lambda_\pm = \lambda_{1,2} = e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\cosh^2(\beta h) - 2\sinh(2\beta J)} \tag{7.12}$$

By convention, we will denote by λ_1 the larger of the two eigenvalues (plus sign before the square root). Then the free energy is given by

$$F = -\frac{1}{\beta}\ln Z = -\frac{1}{\beta}\ln\left(\lambda_1^N + \lambda_2^N\right) = -\frac{N}{\beta}\ln\lambda_1 - \frac{1}{\beta}\ln\left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N\right). \tag{7.13}$$

In the thermodynamic limit in which N goes to infinity, the last term is exponentially small as long as $\lambda_1/\lambda_2 < 1$, which is always true (except at $h = 0$, $T = 0$). Therefore we find, as expected, an extensive free energy, i.e., a free energy per site

$$f = \lim_{N\to\infty} \frac{F}{N} = -\frac{1}{\beta} \ln \lambda_1 \tag{7.14}$$

Now that we have computed exactly the free energy in the presence of an external field, we can easily obtain the magnetization (remember that we are in the case of an homogeneus system)

$$m = \langle \sigma_i \rangle = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sigma_i \right\rangle = -\frac{1}{N} \frac{\partial F}{\partial h}\,, \tag{7.15}$$

and let N go to infinity

$$m = -\frac{\partial f}{\partial h} = \frac{1}{\beta} \frac{\partial \ln \lambda_1}{\partial h} = \frac{1}{\beta\, \lambda_1} \frac{\partial \lambda_1}{\partial h}\,. \tag{7.16}$$

Doing some algebra we find

$$\frac{\partial \lambda_1}{\partial h} = \beta\, e^{\beta J} \sinh(\beta h) + \frac{\beta\, e^{2\beta J} \cosh(\beta h) \sinh(\beta h)}{\sqrt{e^{2\beta J} \cosh^2(\beta h) - 2 \sinh(2\beta J)}}\,. \tag{7.17}$$

We can now let h go to zero and determine whether a spontaneous magnetization may exist for this one-dimensional model. One concludes immediately that, for any temperature

$$m_s = \lim_{h\to 0} m = 0\,. \tag{7.18}$$

The one-dimensional Ising model can be solved exactly, but it does not exhibit any broken symmetry. The system is paramagnetic at all temperatures, i.e., the magnetization vanishes with the applied field.

7.2 Correlation functions

To understand better the absence of broken symmetry in this one-dimensional model, it is instructive to analyze the correlations between two spins located at sites i and $i + n$. Indeed, remember that the existence of a spontaneous magnetization could also manifest itself if, in zero field, one had

$$\lim_{n\to\infty} \langle \sigma_i \sigma_{i+n} \rangle \neq 0\,. \tag{7.19}$$

Let us compute

$$\langle \sigma_i \sigma_{i+n} \rangle = \frac{1}{Z} \sum_{\{\sigma\}} \sigma_i \sigma_{i+n}\, e^{-\beta H}\,. \tag{7.20}$$

The same identity (7.4) and the same transfer matrix technique leads to

$$\langle \sigma_i \sigma_{i+n} \rangle = \frac{1}{Z} \sum_{\sigma,\sigma'} \sigma\, \mathcal{T}^{n}(\sigma,\sigma')\, \sigma'\, \mathcal{T}^{N-n}(\sigma',\sigma)\,. \tag{7.21}$$

Introducing the third Pauli matrix, the diagonal matrix

$$\tau = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{7.22}$$

which is such that

$$[\tau \cdot A] (\sigma, \sigma') = \sigma A(\sigma, \sigma') , \tag{7.23}$$

one finds

$$\langle \sigma_i \sigma_{i+n} \rangle = \frac{1}{Z} \mathrm{Tr} \left[\tau \mathcal{T}^n \tau \mathcal{T}^{N-n} \right] . \tag{7.24}$$

$\mathcal{T}^n$ can be written in this way

$$\mathcal{T}^n = \lambda_1^n \left|1\right\rangle \left\langle 1\right| + \lambda_2^n \left|2\right\rangle \left\langle 2\right| . \tag{7.25}$$

Taking the limit of large N, fixed n, one obtains

$$\begin{aligned} \langle \sigma_i \sigma_{i+n} \rangle &= \frac{\left\langle 1\right| \tau \mathcal{T}^n \tau \mathcal{T}^{N-n} \left|1\right\rangle + \left\langle 2\right| \tau \mathcal{T}^n \tau \mathcal{T}^{N-n} \left|2\right\rangle}{\lambda_1^N + \lambda_2^N} = \frac{1}{\lambda_1^n} \left\langle 1\right| \tau \mathcal{T}^n \tau \left|1\right\rangle \\ &= \left|\left\langle 1\right| \tau \left|1\right\rangle\right|^2 + \left(\frac{\lambda_2}{\lambda_1}\right)^n \left|\left\langle 1\right| \tau \left|2\right\rangle\right|^2 . \end{aligned} \tag{7.26}$$

For $h = 0$ the two eigenvectors are

$$\begin{aligned} \left|1\right\rangle &= \frac{1}{\sqrt{2}} \left(\left|+\right\rangle + \left|-\right\rangle\right) \\ \left|2\right\rangle &= \frac{1}{\sqrt{2}} \left(\left|+\right\rangle - \left|-\right\rangle\right) , \end{aligned} \tag{7.27}$$

where $\left|+\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\left|-\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We obtain that $\left\langle 1\right| \tau \left|1\right\rangle = 0$ and $\left\langle 1\right| \tau \left|2\right\rangle = 1$. Therefore, the correlation function decreases exponentially with the distance $r = na$ between the two sites. In general ($h \neq 0$) we must calculate

$$\langle \sigma_i \rangle = \frac{1}{Z} \mathrm{Tr} \left[\tau \mathcal{T}^N \right] = \left|\left\langle 1\right| \tau \left|1\right\rangle\right| . \tag{7.28}$$

Hence, the connected two-point correlation function is given by

$$\langle \sigma_i \sigma_{i+n} \rangle_c = \langle \sigma_i \sigma_{i+n} \rangle - \langle \sigma_i \rangle \langle \sigma_{i+n} \rangle = \left(\frac{\lambda_2}{\lambda_1}\right)^n \left|\left\langle 1\right| \tau \left|2\right\rangle\right|^2 \propto e^{-r/\xi} . \tag{7.29}$$

As we have just said, the correlation function decreases exponentially with the distance $r = na$ between the two sites with a characteristic correlation length

$$\xi = \frac{a}{\ln \frac{\lambda_1}{\lambda_2}} \tag{7.30}$$

Therefore, there is a possibility of having a spontaneous magnetization, i.e., a non-vanishing limit of the correlation function when the two sites are far apart, only if the correlation length diverges, i.e., only if the two eigenvalues are *degenerate*. We have seen that this does not occur in this model, since λ_1 is strictly superior to λ_2 at any temperature. We shall see that this is a general characteristic of one-dimensional models with short-range interactions.

7.3 Absence of phase transition in one dimension

Our conclusion of the previous section relied on an explicit exact solution. Is it conceivable that, within the same class of one-dimensional Z_2-invariant Hamiltonians, there might be other cases in which the symmetry could be spontaneously broken? Do we have to repeat the calculation for a system in which the spin takes three values $(+1, 0, -1)$ instead of two? Or if the first and second neighbours interact? Or if we have a ladder of two chains of spins? In all these cases one can define a transfer matrix (3×3 or 4×4 in the above examples) with real positive entries. (The matrix is not always real symmetric, and it may have complex eigenvalues.) Repeating the above analysis for a $k \times k$ transfer matrix (with k finite, independent of N), one finds again an exponential fall-off of the correlation function with distance with a correlation length given by (7.30), or rather $\xi = a \left[\ln \frac{|\lambda_1|}{|\lambda_2|}\right]^{-1}$, in which $|\lambda_1| \geq |\lambda_2|$ are the two largest eigenvalue. Is it possible that a degeneracy of the largest eigenvalue could lead to a long-range order? This possibility is ruled out by the *Perron–Frobenius* theorem, which states that, "A matrix with real positive entries posseses one real, non-degenerate, eigenvalue, which exceeds the modulus of all the other eigenvalues. The corresponding eigenvector has real components, all of the same sign". (In fact, it is a version of this theorem for operators rather than matrices that one uses when one asserts that the spatial ground-state wave function of a Hamiltonian is non-degenerate and that the corresponding wave function is nodeless). Note that it is essential to assume that the rank of this transfer matrix is fixed, independent of the number of sites N.

Chapter 8

On Spontaneous Simmetry Breaking

Few words about the notation of this chapter. We specify the system of interest as some sample region Ω, in which is defined the Hamiltonian H_Ω. The volume of the region is V and the surface area is $S(\Omega)$. Often it will be convenient to think of our system as having some characteristic linear dimension L, so that $V(\Omega) \propto L^D$, $S(\Omega) \propto L^{D-1}$, where D is the dimensionality of the system. Usually, there will be boundary conditions specified on the boundary of Ω. Often these will be periodic or hard wall (i.e. zero flux of particles through a wall). The system may exist as a continuum (e.g. a fluid) or on a lattice (e.g. a magnet); for now, $V(\Omega)$ is finite. We write the Hamiltonian for the system as

$$-\frac{H_\Omega}{k_B T} = \sum_n K_n \Theta_n \tag{8.1}$$

where K_n are the *coupling constants* and the Θ_n are combinations of the dynamical degrees of freedom, which are summed over in the partition function. We shall sometimes refer to the Θ_n as *local operators*. The so-called coupling constants K_n, as, we know, are the external parameters, such as fields, exchange interaction parameters, temperature.... So, for example, if we are dealing with a magnet, the degrees of freedom are the (vector) spins on the lattice sites $\mathbf{S}_i$, where $1 \leq i \leq N(\Omega)$ and $\mathbf{S}_i$ is a n-dimensional vector. Thus the Θ_n are built out of combinations of the $\mathbf{S}_i$: i.e. $\Theta_1 = \sum_{a=1}^{n} \sum_i S_i^a$, $\Theta_2 = \sum_{i,j} \mathbf{S}_i \cdot \mathbf{S}_j$, etc.

A simple example of a term in H_Ω is that responsible for the *Zeeman effect*. The coupling constant in this context is the external field $\mathbf{h}$, and the corresponding local operator is the magnetic moment at a lattice site i, $\mathbf{S}_i$. Then the contribution of these terms to H_Ω is $-\sum_i \mathbf{h} \cdot \mathbf{S}_i$.

Experience tells us that the free energy is extensive for a large system: $F_\Omega \propto V(\Omega)$. Thus, we expect that for a finite system, we can write

$$F_\Omega = V(\Omega) f_b + S(\Omega) f_s + O(L^{D-2}) \,, \tag{8.2}$$

where f_b is the bulk free energy per unit volume or bulk free energy density and f_s, is the surface free energy per unit area. We can give a precise definition of these important quantities as follows:

$$f_b[K] \equiv \lim_{V(\Omega)\to\infty} \frac{F_\Omega[K]}{V(\Omega)} \tag{8.3}$$

when the limit exists and is independent of Ω. For a system defined on a lattice, with $N(\Omega)$ lattice sites, the bulk free energy per site is

$$f_b[K] \equiv \lim_{N(\Omega)\to\infty} \frac{F_\Omega[K]}{N(\Omega)} \tag{8.4}$$

when the limit exists and is independent of Ω. The bulk free energy $f_b[K]$ describes extensive thermodynamic behaviour (proportional to $V(\Omega)$) but does not describe surface or finite size behaviour. This information may be computed from the surface free energy

$$f_s[K] = \lim_{S(\Omega)\to\infty} \frac{F_\Omega[K] - V(\Omega) f_b[K]}{S(\Omega)} \tag{8.5}$$

when the limit exists and is independent of Ω. The limit in these equations is known as the thermodynamic limit. Sometimes, an auxiliary constraint is imposed simultaneously with the limit: for example, in fluid systems, taking the limit $V(\Omega) \to \infty$ is senseless unless one simultaneously takes the limit that the number of particles $N(\Omega)$ in the system also tends to infinity, in such a way that the density $N(\Omega)/V(\Omega)$ stays constant. We will see that the concepts of phase and phase transition are only sharply defined in this limit. The existence of the thermodynamic limit is not trivial. In order for a uniform bulk behaviour to exist, the forces in the system must satisfy certain properties, and the thermodynamic limit must be taken carefully.

8.1 Phase Boundaries and Phase Transitions

When $f_b[K]$ exists, then a precise definition of a phase boundary follows. Let us suppose that there are d coupling constants. The axes of the phase diagram are $K_1, K_2, \ldots, K_d$, and hence the dimension of the phase diagram is d. As a function of $[K]$, $f_b[K]$ is analytic almost everywhere. The possible non-analyticities of $f_b[K]$ are points, lines, planes, hyperplanes, etc. in the phase diagram. These singular loci have a dimensionality associated with them ($D_s = 0, 1, 2, \ldots$ respectively), and the important invariant for each type of singular locus is the codimension, c:

$$c \equiv D - D_s \tag{8.6}$$

This is an invariant in the sense that if we decide to include an extra variable in $\{K_n\}$ (and thus to the phase diagram), both D and D_s increase by 1 so that c remains fixed. Regions of analyticity of $f_b[K]$ are called *phases*. Loci of codimension $c = 1$ (i.e. loci which separate phases), are called *phase boundaries*. Loci of codimension greater than one cannot possibly represent phase boundaries. To give a prosaic example, in order to partition a room into two, it is necessary to build a wall across it - a pole through the centre of the room will not suffice.

8.1.1 Ambiguity in the Definition of Phase Boundary

The definition of phase given above is ambiguous. There may exist a path along which $f_b[K]$ is analytic going from one side of a phase boundary to the other. Using our "room analogy", this is like having a wall which does not quite reach the ceiling. At floor level, the room is partitioned, but a flying insect may be able to pass from one side of the room to the other without encountering any impediment to its progress. An example is the liquid-gas-solid phase diagram, shown in Fig.(8.1).

Although it is not possible to pass from fluid to solid without encountering a phase transition, it is possible to choose a path in $P - T$ space which goes from liquid to gas without encountering any singular behaviour in thermodynamic quantities. We shall see later that this is a reflection of the fact that the liquid and gas states have the same degree of symmetry, whereas a fluid has a

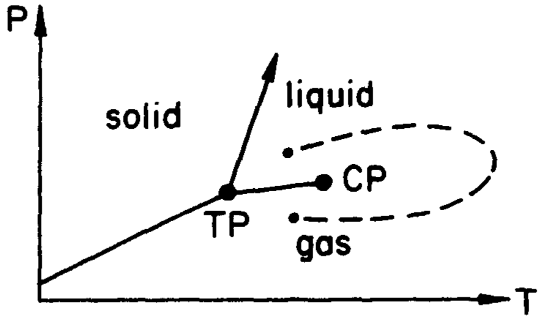


Figure 8.1. Phase diagram of a typical substance: P is pressure, T is temperature. The point TP is a triple point ($c = 2$) and the point CP is a critical point ($c = 2$). The solid-liquid phase boundary extends to arbitrarily high pressure (in principle), whilst the gas-liquid phase boundary ends at CP. The dashed curve represents a trajectory in the phase diagram along which no phase transition is encountered, even though a change of phase has apparently taken place.

higher degree of symmetry than a solid (at least if the solid is crystalline, which is the usual case in equilibrium)[1]

8.1.2 Types of Phase Transition

We shall show shortly that $f_b[K]$ is everywhere continuous: our demonstration will be specific to a particular model system, but the conclusion is true in general. This implies that where phase boundaries exist, they must come in two classes:

- $\partial f_b/\partial K_i$ is *discontinuous* across a phase boundary. It can be either one or more of the $\partial f_b/\partial K_i$ which is discontinuous. If this case occurs, then the transition is said to be a *first-order phase transition.*

- The only other remaining possibility for non-analytic behaviour is that all $\partial f_b/\partial K_i$ are continuous across the phase boundary. If this occurs, the transition is said to be a *continous phase transition.* Sometimes, this is referred to as a *second order transition* too, but for reasons already explained, this terminology is not encouraged.

8.1.3 Finite-Size Effects and the Correlation Length

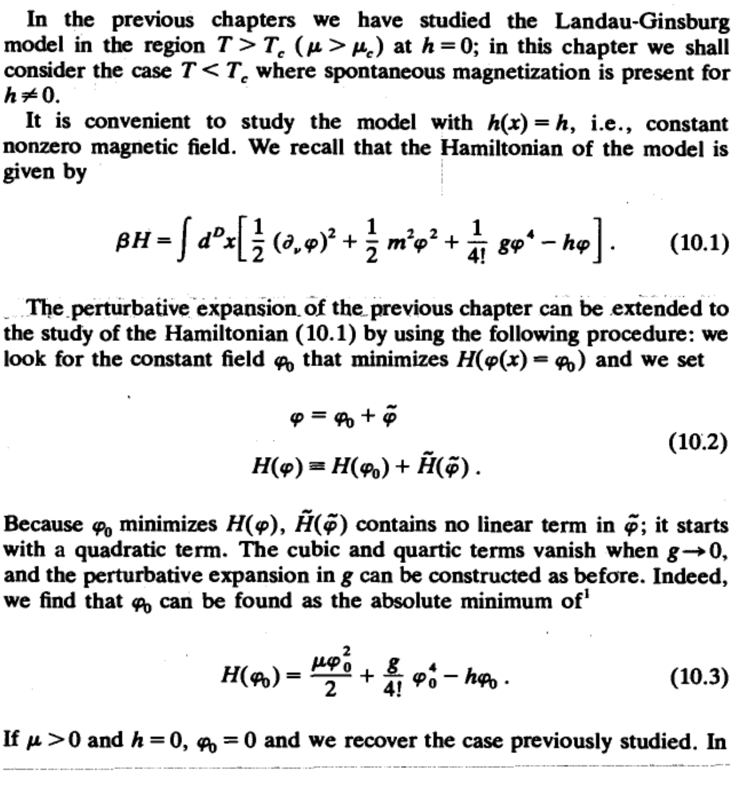
In practice, the thermodynamic limit is never attained: $10^{23} \neq \infty$. Thus it is reasonable to ask whether or not the thermodynamic limit is physically relevant. If there were perfect instrumental resolution, a change in the physical properties in a finite system would not occur over an infinitesimal interval of the relevant coupling constant, but would occur over some range. This phenomenon is an example of a *finite size effect.* In practice, however, instrumental resolution is usually the limiting factor. We can give a simple criterion for when predictions using $f_b[K]$ are not reliable, by introducing the concept of the correlation length ξ. Loosely speaking, the correlation length describes the spatial extent of fluctuations in a physical quantity about the average of that quantity.

---

[1] Even if the solid were in equilibrium and amorphous, it would still represent a tate of broken translations invariance, whereas in a fluid, this symmetry is unbroken.(See Goldenfeld Section 2.10)

For example, in a gas, there will be density fluctuations in thermal equilibrium. In a particular region of the sample, the density may be higher than the average density. We can choose to think of such a region as a droplet of near-liquid density floating in the gas. In thermal equilibrium, there is a distribution of droplet sizes, of course, but it turns out that there is a well-defined average size, at least away from the critical point itself. This characteristic size is, roughly speaking, what we mean by the term correlation length. It is not unreasonable that the droplet distribution should depend upon the position in the phase diagram: how close the system is to a phase boundary or critical point for example. Thus, we might expect that ξ depends upon the coupling constants, in particular temperature. The correlation length depends strongly on temperature near a continuous phase transition, diverging to infinity at the transition itself. This is what gives rise to scaling behaviour. Now we are ready to address the question of when the finite size of a real system is important. In a finite system, the correlation length is not able to diverge to infinity, since it cannot exceed L, the characteristic linear dimension of the system. Thus, for temperatures sufficiently close to T_c that the correlation length of an infinitely large system would exceed L, the behaviour of the finite system departs from the ideal behaviour described by f_b,. To make a rough estimate of when this occurs, let us assume that $\xi = \xi_0 t^{-2/3}$ where $t \equiv \frac{T-T_c}{T_c}$ with $\xi_0 = 10^-9m$ being the correlation length well away from the critical point. This form for ? is not unrealistic for magnetic or fluid systems: the value for probably overestimate calculation will be quite conservative. For a system with $L = 1cm$, we find that $\xi = L$ when the reduced temperature $t \approx 10^{-11}$. In this situation, finite size effects would be hard to observe.

8.2 Perturbation theory



In the previous chapters we have studied the Landau-Ginsburg model in the region $T > T_c$ ($\mu > \mu_c$) at $h = 0$; in this chapter we shall consider the case $T < T_c$ where spontaneous magnetization is present for $h \neq 0$.

It is convenient to study the model with $h(x) = h$, i.e., constant nonzero magnetic field. We recall that the Hamiltonian of the model is given by

$$\beta H = \int d^D x \left[\frac{1}{2} (\partial_\nu \varphi)^2 + \frac{1}{2} m^2 \varphi^2 + \frac{1}{4!} g \varphi^4 - h\varphi \right] . \qquad (10.1)$$

The perturbative expansion of the previous chapter can be extended to the study of the Hamiltonian (10.1) by using the following procedure: we look for the constant field φ_0 that minimizes $H(\varphi(x) = \varphi_0)$ and we set

$$\begin{aligned} \varphi &= \varphi_0 + \tilde{\varphi} \\ H(\varphi) &\equiv H(\varphi_0) + \tilde{H}(\tilde{\varphi}) . \end{aligned} \qquad (10.2)$$

Because φ_0 minimizes $H(\varphi)$, $\tilde{H}(\tilde{\varphi})$ contains no linear term in $\tilde{\varphi}$; it starts with a quadratic term. The cubic and quartic terms vanish when $g \to 0$, and the perturbative expansion in g can be constructed as before. Indeed, we find that φ_0 can be found as the absolute minimum of[1]

$$H(\varphi_0) = \frac{\mu \varphi_0^2}{2} + \frac{g}{4!} \varphi_0^4 - h\varphi_0 . \qquad (10.3)$$

If $\mu > 0$ and $h = 0$, $\varphi_0 = 0$ and we recover the case previously studied. In the general case we find

$$H(\varphi) = \int d^D x \left[H(\varphi_0) + \frac{1}{2} (\partial_\nu \tilde{\varphi})^2 + \frac{\tilde{\mu}}{2} \tilde{\varphi}^2 + \frac{1}{3!} g_3 \tilde{\varphi}^3 + \frac{1}{4!} g_4 \tilde{\varphi}^4 \right]$$

$$\mu\varphi_0 + \frac{g}{6} \varphi_0^3 = h \qquad (\varphi_0 h > 0) \tag{10.4}$$

$$\tilde{\mu} = \mu + \frac{1}{2} g \varphi_0^2 \qquad g_3 = g\varphi_0 \qquad g_4 = g \,.$$

When $g \to 0$ we have, for small h,

$$\begin{aligned} &\varphi_0 \sim \frac{h}{\mu}\,, \quad \mu > 0 \\ &\varphi_0 \sim \operatorname{sign}(h) \sqrt{\frac{-6\mu}{g}} + O(h)\,, \qquad \mu < 0\,. \end{aligned} \tag{10.5}$$

In this case the distribution probability of $\tilde{\varphi}$ becomes Gaussian: To leading order in g we get

$$\begin{aligned} &\langle \varphi \rangle = \varphi_0 \qquad f(h) \equiv -\frac{1}{V} \ln Z = H[\varphi_0] \\ &(\tilde{\mu} - \Delta_x) \langle \tilde{\varphi}(x) \tilde{\varphi}(y) \rangle_C = \delta(x-y)\,. \end{aligned} \tag{10.6}$$

The two-field connected correlation function in momentum space is $\tilde{G}_0(p) = 1/(\tilde{\mu} + p^2) = \int d^D x \exp(ipx) G_0(x)$; $G_0(x) = \langle \tilde{\varphi}(x) \tilde{\varphi}(0) \rangle$. In the limit $h \to 0^+$ for $\mu < 0$ we find

$$\tilde{\mu} = -2\mu \qquad \varphi_0 = \left(-\frac{6\mu}{g} \right)^{1/2} \qquad g_3 = \sqrt{-6\mu g}\,. \tag{10.7}$$

When $\mu < 0$, spontaneous symmetry breaking is present:

$$\begin{aligned} &\lim_{h \to 0^+} \langle \varphi \rangle = -\lim_{h \to 0^-} \langle \varphi \rangle \neq 0 \\ &f(h) = f(0) - |h| \left(-\frac{6\mu}{g} \right)^{1/2} + O(h^2)\,. \end{aligned} \tag{10.8}$$

Let us now see how to set up the perturbative expansion. For simplicity let us suppose that we are working in a theory with a cutoff where no ultraviolet divergences are present.[2] The diagrams contributing to $\langle \varphi \rangle$ and $\langle \varphi\varphi \rangle_C$ are shown in Fig. 10.1; the final result is, at the one-loop

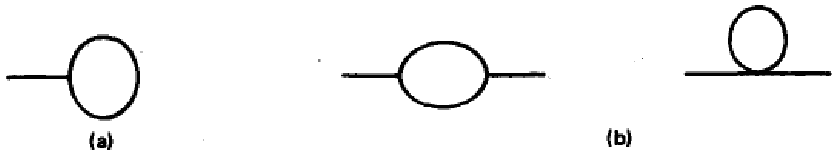


Figure 10.1. Diagrams contributing (a) to $\langle \varphi \rangle$ and (b) to $\langle \varphi\varphi \rangle_c$.

order,

$$\begin{aligned}\langle \varphi(0) \rangle &= \varphi_0 - g_3 \frac{1}{3!} \int \langle \tilde{\varphi}(0)\tilde{\varphi}^3(x) \rangle_c \, d^D x \\ &= \varphi_0 - g_3 \frac{1}{2} \int d^D x \langle \tilde{\varphi}(0)\tilde{\varphi}(x) \rangle_c \langle \tilde{\varphi}^2(x) \rangle_c \\ &= \varphi_0 - \frac{1}{2} g_3 \int d^D x \, G_0(x) G_0(0) = \varphi_0 - \frac{1}{2} g_3 / \tilde{\mu} \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \tilde{\mu}} \end{aligned}$$

$$\tilde{G}(p) = \frac{1}{p^2 + \tilde{\mu} - \Sigma(p)} \tag{10.9}$$

$$\Sigma(p) = \frac{1}{2} g_3^2 \int \frac{d^D k}{(2\pi)^D} \frac{1}{(p-k)^2 + \tilde{\mu}} \frac{1}{k^2 + \tilde{\mu}} - \frac{1}{2} g_4 \int \frac{d^D k}{(2\pi)^D} \frac{1}{k^2 + \tilde{\mu}} \,.$$

Similarly, the first correction to the free-energy density is given by the Gaussian integral over the field $\tilde{\varphi}$:

$$f(h) = H(\varphi_0) + \frac{1}{2} \int \frac{d^D k}{(2\pi)^D} \ln(k^2 + \tilde{\mu}) \,. \tag{10.10}$$

Equation (10.10) can also be written as

$$\begin{aligned} f(h) &= [\gamma(\varphi) - h\varphi]_{\varphi=\varphi_0} \\ \gamma(\varphi) &= \frac{\mu}{2} \varphi^2 + \frac{g}{4!} \varphi^4 + \frac{1}{2} \int \frac{d^D k}{(2\pi)^D} \ln(k^2 + \tilde{\mu}(\varphi)) \qquad (10.11) \\ \tilde{\mu}(\varphi) &= \mu + \frac{1}{2} g \varphi^2 \,, \end{aligned}$$

where $\gamma(\varphi)$ no longer depends on h.

Now it is important to note that although Eq. (10.10) has been derived here for $\langle \varphi \rangle = \varphi_0$, Eq. (10.9) for $\langle \varphi \rangle$ could have been obtained by setting $\langle \varphi \rangle = \bar{\varphi}$, where $\bar{\varphi}$ is the solution of the equation

$$\left. \frac{\partial \gamma}{\partial \varphi} \right|_{\varphi = \bar{\varphi}} = h \,, \tag{10.12}$$

as can be easily verified. Now it is possible to generalize Eq. (10.12) to all orders and to find simple rules for computing $\gamma(\varphi)$.[3] Toward this end it is convenient to define the magnetic field as a function of the expectation value [i.e., $h(\varphi)$] in such a way that

$$\langle \varphi \rangle |_{h=h(\varphi)} = \varphi \,. \tag{10.13}$$

In the most general case one can consider Eq. (10.12) for an x-dependent $\langle \varphi \rangle \equiv \varphi(x)$: h would also depend on x, and it would be a functional of φ; for simplicity here we study only the case in which neither h nor φ depends on x.

For later convenience we introduce a φ-dependent free-energy density $\tilde{f}(\varphi)$ and an auxiliary function $\gamma(\varphi)$ defined by[3]

$$\tilde{f}(\varphi) \equiv f(h(\varphi)) = \gamma(\varphi) - h\varphi \,. \tag{10.14}$$

In other words, $\tilde{f}(\varphi)$ is the value of the free-energy density when $\langle \varphi \rangle = \varphi$. In the zero-loop approximation we obviously have

$$\gamma(\varphi) = \frac{\mu^2}{2}\varphi^2 + \frac{g\varphi^4}{4!} \,, \tag{10.15}$$

while at one-loop order $\gamma(\varphi)$ is given in Eq. (10.11).

Intuitively the variational properties of the free energy suggest that, at a given h, $\langle \varphi \rangle$ is fixed by the condition

$$\langle \varphi \rangle = \varphi_m \,, \tag{10.16}$$

where φ_m is the minimum of $\tilde{f}(\varphi)$. Obviously φ_m must satisfy the equation

$$\left. \frac{d\gamma}{d\varphi} \right|_{\varphi=\varphi_m} = h \,. \tag{10.17}$$

The intuitive argument is indeed correct: using the known relation

$$\frac{df}{dh} = \langle \varphi \rangle \equiv \varphi \,, \tag{10.18}$$

one easily derives Eq. (10.17):

$$\frac{d\gamma}{d\varphi} = \frac{d}{d\varphi}[f(h) + \varphi h] = \frac{df}{dh}\frac{dh}{d\varphi} + \varphi\frac{dh}{d\varphi} + h = h(\varphi) \,. \tag{10.19}$$

It is also clear that, if Eq. (10.17) has many solutions, the one having the smallest value of $\tilde{f}$ will have minimal free energy and it will be the physical one.

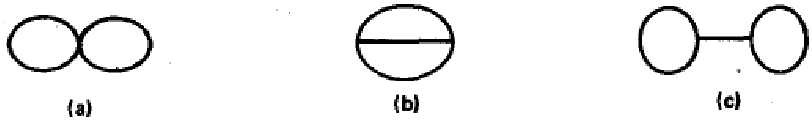


Figure 10.2. Two-loop contribution to the free energy; diagram (c) is one-line reducible and does not contribute to $\gamma(\varphi)$.

The transformation from an h-dependent free energy to a φ-dependent free energy (where $\varphi = -df/dh$) is a very common procedure in thermodynamics, which goes under the name of the Legendre transform;[4] it is also used for going from the Lagrangian to the Hamiltonian formalism.

There are graphical rules for computing $\gamma(\varphi)$ directly; it turns out that only one-line irreducible diagrams need be evaluated, greatly simplifying the computation. We shall not give the proof of this very useful result, neither shall we study further the properties of the function $\gamma(\varphi)$, as this would take too much space; in the rest of this chapter we shall consider the case $N \neq 1$, in which interesting new phenomena are present.

Before going on to look at $N \neq 1$, however, a few remarks are in order. The function $\gamma(\varphi)$ can be written as

$$\gamma(\varphi) = \sum_{0}^{\infty}{}_{n} \frac{1}{n!} \Gamma_n \varphi^n \,, \tag{10.20}$$

where $\Gamma_1 = 0$, $\Gamma_2 = [G_2(p)|_{p=0}]^{-1}$, and Γ_n is the sum of the amputated (i.e., without the contribution of external legs) one-particle-line irreducible diagrams. It is easy to convince ourselves of the correctness of Eq. (10.20) for small values of n (say, $n = 2, 3, 4$) by differentiating with respect to φ [Eq. (10.17)] $n-1$ times and reinterpreting the result in a graphical way (see Fig. 10.2). A full proof requires some tricky combinatorial analysis; in the simplest approach the proof is constructed by induction in n.

In order to compare the experimental data near the phase transition with the theory, it is particularly interesting to have a reliable result for the function $\gamma(\varphi)$ [or equivalently $h(\varphi)$] at the critical point. A few orders of the $4-D$ expansion for $\gamma(\varphi)$ have been computed.[5] Unfortunately we lack a high-loop computation done directly at three dimensions.

8.3 Continuous and Discrete Symmetries

The result that long range order is possible in the Ising model only for $D > 1$ relies on the fact that the symmetry group of H_Ω is Z_2, i.e. $H_\Omega[\{S_i\}] = H_\Omega[\{-S_i\}]$. However, things change if the spins S_i obey a *continuous symmetry* rather than a *discrete symmetry*. One way for this to occur is if the S_i are vectors, allowed to point in all directions (4π steradians) rather than just up or down. This occurs in the Heisenberg ferromagnet with Hamiltonian Heisenberg model. In general, this is a more realistic model of a ferromagnet than the Ising model, since there is no preferred direction for the spins to point in. In a real crystal, spin-orbit coupling is present in addition to the exchange interaction already considered. The spins will then couple to the electronic charge density, which reflects the presence and symmetry of the underlying crystal lattice. In the cases where this is significant, the rotational symmetry of H_Ω is broken, and there is a tendency for spins to align along crystallographically preferred directions. In extreme cases, this can be sufficiently strong that a better model Hamiltonian is not the Heisenberg Hamiltonian but the Ising Hamiltonian. The Heisenberg model has a continuous symmetry. Suppose that

$$\mathbf{S}_i = (S_i^x, S_i^y, S_i^z) \tag{8.7}$$

and $R(\theta)$ is a rotation matrix which rotates a vector in (x, y, z) space by an angle θ about direction y, for example. Then the energy is invariant under the simultaneous rotation of all the spins by an arbitrary angle θ;

$$H_\Omega[\{R(\theta)\,\mathbf{S}_i\}] = H_\Omega[\{\mathbf{S}_i\}]\,. \tag{8.8}$$

Of course, in general, we do not have to single out the y-axis. R can be an arbitrary rotation about an arbitrary direction. Thus, the Heisenberg model is rotationally invariant, and this symmetry is sometimes called $O(3)$ (in three dimensional space). It is very important that you realize that the rotations are of the spins, keeping the lattice fixed in space. The rotation operator $R(\theta)$ acts on the degrees of freedom, not the spatial co-ordinate system. The continuous rotational symmetry is spontaneously broken in the state of long range order. It is "easier" for thermal fluctuations to destroy long range order when there is a continuous symmetry as opposed to a discrete symmetry, because “there are more directions to point in (in spin space)”‘. To compensate for the increased entropy due to the larger dimensionality of the order parameter, an increased energy is required if there is to be long range order in the Heisenberg model at $T > 0$. This can be achieved by increasing the dimensionality of the lattice: spins have more neighbours with which to interact. We will see that to have long range order, we need $D > 2$.

8.3.1 The general case: $O(N)$ model

Another interesting generalization of the Ising model is provided by the $O(N)$ model, in which each spin $\mathbf{S}_i$ is a n-component vector associated to a point of the n-dimensional sphere. Thus the order parameter $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_N)$ and the statistical mechanics is governed by the generalized Hamiltonian

$$H[\varphi] = \int \mathrm{d}^D x \left[\frac{1}{2} \sum_{a=1}^{N} (\partial_\nu \varphi_a)^2 + \frac{1}{2}\,\mu \left(\sum_{a=1}^{N} \varphi_a^2 \right) + \frac{g}{4!} \left(\sum_{a=1}^{N} \varphi_a^2 \right)^2 - \sum_{a=1}^{N} h_a \varphi_a \right] \tag{8.9}$$

where **h** is the external field. Alternative notation are

$$\begin{aligned}
&\sum_{a=1}^{N} (\partial_\nu \varphi_a)^2 = (\nabla \boldsymbol{\varphi})^2 \equiv \sum_{\nu=1}^{D} \sum_{a=1}^{N} \left(\frac{\partial \varphi_a(\mathbf{x})}{\partial x_\nu} \right)^2 \\
&\sum_{a=1}^{N} \varphi_a^2 = \boldsymbol{\varphi}^2 \equiv \sum_{a=1}^{N} (\varphi_a(\mathbf{x}))^2 \\
&\left(\sum_{a=1}^{N} \varphi_a^2 \right)^2 = \boldsymbol{\varphi}^4 \equiv \left(\boldsymbol{\varphi}^2 \right)^2
\end{aligned} \tag{8.10}$$

The name of this model comes from its rotational symmetry. Let us consider the low temperature behavior of this system and assume that there exists a spontaneously broken, that is a phase where all spins point in the same direction, e.g. along the x-axis. Then the $O(N)$ rotational symmetry of the system is spontaneously broken, or rather reduced to the $O(N-1)$ symmetry under rotations around this direction. The homogeneous part of the Landau free energy density is[2]

$$V(\boldsymbol{\varphi}) = \frac{1}{2} \mu \left(\sum_{a=1}^{N} \varphi_a^2 \right) + \frac{g}{4!} \left(\sum_{a=1}^{N} \varphi_a^2 \right)^2 \tag{8.11}$$

which has the behaviour shown in Fig.(8.2), for $\mathbf{h} = 0$. For $\mu > 0$ there one minimum at $\boldsymbol{\varphi} = 0$, corresponding to the high temperature phase of the system which is not ordered, whereas the minimum at $\boldsymbol{\varphi} = -6\mu/g$ is infinitely degenerate for $\mu < 0$. The degeneracy corresponds to the directions in order parameter space (not real space) in which the system may order below T_c.

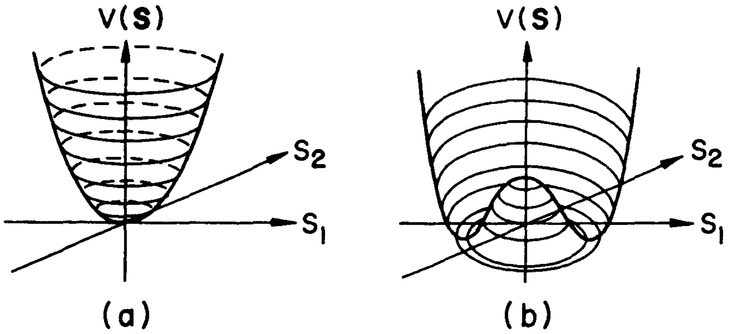


Figure 8.2. Homogeneous part of the Landau free energy for the $O(N)$ model. Sketched for simplicity is the case $N = 2$: (a) $T > T_c$. (b) $T < T_c$

8.3.2 Mermin-Wigner theorem

When T is greater than T_c, the Ising model ($N = 1$) and the Heisenberg model ($N = 3$) look very similar, the only difference being in the weights of the loop diagrams which are smooth functions of N. However, the theories with $N \neq 1$ have qualitatively new features as soon as

---

[2] Sometimes the homogeneous part of the Landau free energy is called the potential energy and the gradient part is sometime called the kinetic energy. These terms aremisleading in statistical mechanics, and allude to the similarity between the Hamiltonian and the Lagrangian of a self-interacting boson field theory.

$T < T_c$ at $h = 0$: *spin waves* are present, i.e., configurations whose energy differs from the ground state by an arbitrarily small amount. The $O(N)$ invariance of the theory implies that the spontaneous magnetization may point in an arbitrary direction on the N-dimensional sphere of radius φ_0 ($\sum_{a=1}^{N} \langle\varphi_a\rangle^2 = \varphi_0^2$) we have an infinite number of equilibrium, translational invariants, configurations which can be labeled by a continuous parameter (not only two as happens in the Ising model). The differences between discrete and continous simmetrt are sharpened by the *Mermin-Wigner theorem*: there is a critical dimension D_c such that for any short-range[3] Hamiltonian at $T \neq 0$, no spontaneous symmetry breaking is possible, D_c being 1 for discrete symmetries ($N = 1$) while $D_c = 2$ for continuous symmetries ($N > 1$).

8.4 Correlations in the ordered phase

8.4.1 The Susceptibility Tensor

We have just seen that the spontaneous breaking of a continuous symmetry is quite different from the discrete case. We now investigate how this difference shows up in the perturbative expansion. We consider an $O(N)$-invariant theory. If the external magnetic field is a constant [$h_a(x) = h_a$], the free energy may depend on h only through $|h| = \left|\sum_{a=1}^{N} h_a^2\right|^{1/2}$.
We define the *susceptibility tensor* by

$$\chi_{ab} \equiv -\frac{\partial^2 f}{\partial h_a \partial h_b} \equiv -\frac{\partial m_a}{\partial h_b}\,, \tag{8.12}$$

where f is the free energy per unit volume. The susceptibility tensor describes how the a^{th} component of the magnetisation is changed due to the b^{th} component of the external field. Since $H[\varphi]$ is $0(N)$ symmetric, f cannot depend on the direction of $\mathbf{h}$. Hence,

$$\begin{aligned} &f = f(h) \\ &h = |h| = \left(\sum_{a=1}^{N} h_a^2\right)^{1/2} . \end{aligned} \tag{8.13}$$

Thus differentiating (8.12) and using

$$\frac{\partial}{\partial h_a} = \frac{\partial h}{\partial h_a}\frac{\partial}{\partial h} = \frac{h_a}{|h|}\frac{\partial}{\partial |h|}\,, \tag{8.14}$$

we get

$$\begin{aligned} \chi_{ab} &= -\frac{\partial}{\partial h_a}\frac{h_b}{|h|}\frac{\partial f}{\partial |h|} = -\delta_{ab}\frac{1}{|h|}\frac{\partial f}{\partial |h|} + \frac{h_a h_b}{|h|}\frac{1}{|h|^2}\frac{\partial f}{\partial |h|} - \frac{h_a h_b}{|h|^2}\frac{\partial^2 f}{\partial |h|^2} \\ &= -\frac{h_a h_b}{|h|^2}\frac{\partial^2 f}{\partial |h|^2} - \frac{1}{|h|}\frac{\partial f}{\partial |h|}\left(\delta_{ab} - \frac{h_a h_b}{|h|^2}\right) . \end{aligned} \tag{8.15}$$

Now, $\mathbf{h} = h\mathbf{n}$ so that

$$\frac{h_a h_b}{|h|^2} = n_a n_b\,. \tag{8.16}$$

---

[3] If long-range forces are present, the theorem no longer holds. Spontaneous magnetization is present in the one-dimensional Ising model if $J_{ik} \sim |i-k|^{-2}$ at large $|i-k|$.

Thus

$$\chi_{ab} = n_a n_b \, \chi_L(h) + \chi_T(h) \, (\delta_{ab} - n_a n_b) \ , \tag{8.17}$$

with the longitudinal susceptibility

$$\chi_L(h) = -\frac{\partial^2 f}{\partial \, |h|^2} \, , \tag{8.18}$$

and the transverse susceptibility

$$\chi_T(h) = -\frac{1}{|h|} \frac{\partial f}{\partial \, |h|} = -\frac{m}{|h|} \, . \tag{8.19}$$

Now we can see the physical significance of these quantities: recall that

$$\chi_{ab} = -\frac{\partial m_a}{\partial h_b} \, . \tag{8.20}$$

Equation (8.17) implies that if the probe field **h** is in the same direction as m, then the susceptibility is given by χ_L. If the probe field is perpendicular to the direction of **m**, then the susceptibility is given by χ_T.

8.4.2 Excitations for $T < T_c$: Goldstone's Theorem

Suppose, for simplicity, that the point $h = 0$ is reached in the following way[4]

$$\begin{aligned} & h_N = h \\ & h_\alpha = 0 \qquad \alpha = 1, \ldots, N-1 \\ & h \to 0^+ \, , \end{aligned} \tag{8.21}$$

Equation (8.19) suggests that the transverse susceptibility $\chi_T = \chi_{aa}$ $(a < N)$ becomes infinite when $|h| \to 0$ if $m \neq 0$: for small $|h|$ the addition of a new external field may change completely the direction in which the magnetization points. The divergence of χ_T when $|h| \to 0$ implies the presence of long-range correlations.

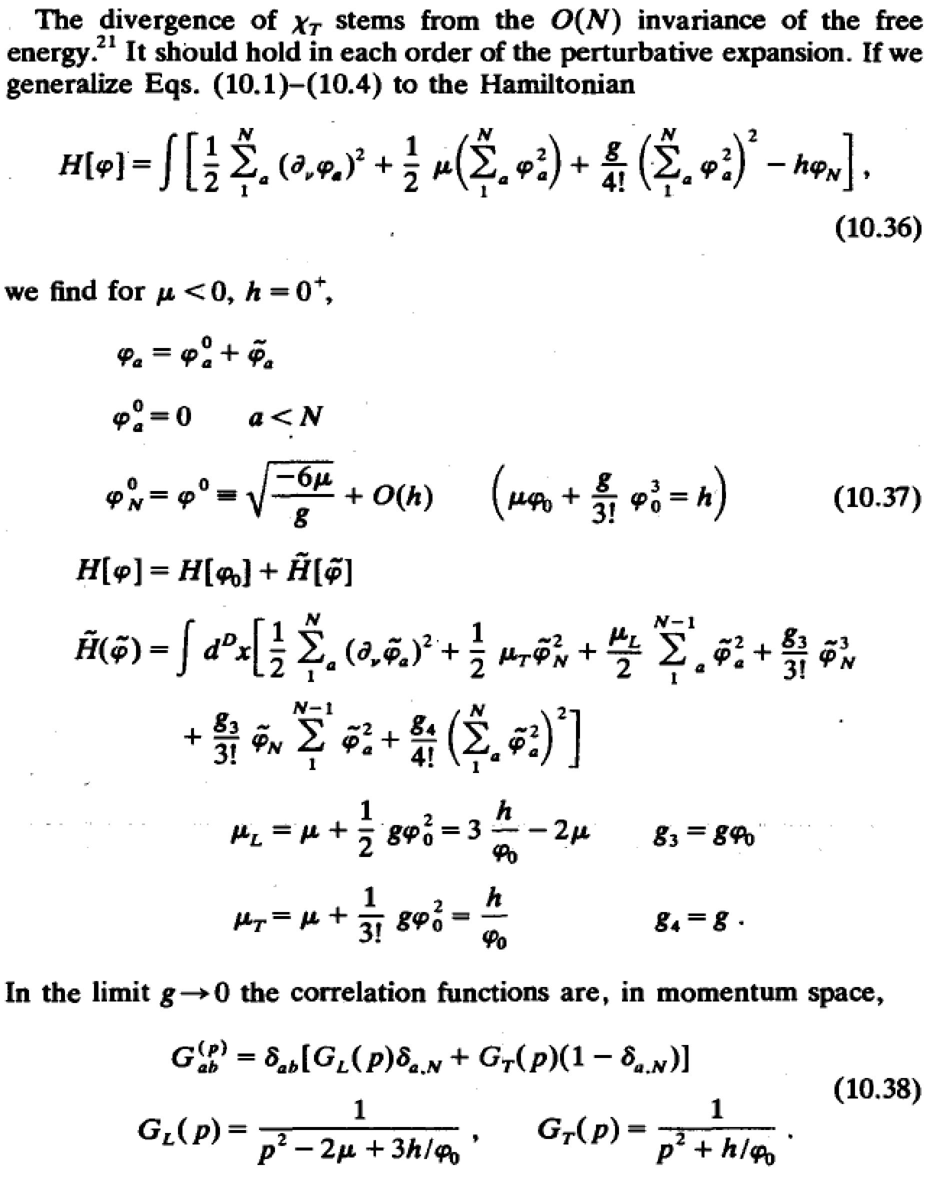
In other words, in the transverse direction, the system is infinitely susceptible - it requires an infinitesimal amount of field to rotate the direction of magnetisation by a finite, non-zero amount! This is quite reasonable, when we consider the potential $V(\boldsymbol{\varphi})$ for $\mu < 0$. There are an infinite number of minima of the effective Hamiltonian, each corresponding to the system spontaneously acquiring a magnetisation along a different direction. It requires no energy to go between these minima, since they are degenerate. We see that when the $O(N)$ symmetry of H is spontaneously broken there are two sorts of fluctuation: those *parallel* to the direction in which the system has ordered, and those *perpendicular* to it. These fluctuations cost different amounts of energy: a fluctuation parallel to the direction of ordering (i.e. a *longitudinal fluctuation*) is one which causes the magnitude m of **m**to change. By inspection of the potential $V(\boldsymbol{\varphi})$, we see that there is an energy penalty for increasing or decreasing m. On the other hand, fluctuations in the direction perpendicular to **m** only change the *direction* of **m**, and cost no energy, as we have already mentioned. So, if the symmetry is spontaneously broken

$$f(h) = -\,|h| \, m_s - \frac{1}{2} \, \chi_L \, |h|^2 \ . \tag{8.22}$$

---

[4] There is an infinite number of equivalent directions; to simplify the notation we have chosen one.

So far, we have only discussed the case of spatially uniform systems. What happens when we allow for spatial variations of the deviation of the order parameter from the equilibrium value in the ordered state? Now the gradient term in $H[\varphi]$ must also be considered. Reasoning by continuity, we might expect that longitudinal fluctuations still have high energy, whereas transverse fluctuations will now require a small but non-zero energy, the amount being proportional to $\sum_{a=1}^{N} (\partial_\nu \varphi_a)^2$. Hence, the longer the wavelength of the fluctuation, the smaller the associated free energy. We have made a rather useful deduction: purely because of the ordering, a system with an $O(N)$ symmetry should have an excitation spectrum with modes of arbitrarily low free energy. In fact, there will be one such mode associated with each transverse direction, which makes $N-1$ low energy modes. These modes are called *Goldstone modes.* Another striking consequence of spontaneous symmetry breaking is the form of the correlation functions below T_c.



The divergence of χ_T stems from the $O(N)$ invariance of the free energy.[21] It should hold in each order of the perturbative expansion. If we generalize Eqs. (10.1)–(10.4) to the Hamiltonian

$$H[\varphi] = \int \left[\frac{1}{2} \sum_{1}^{N}{}_a (\partial_\nu \varphi_a)^2 + \frac{1}{2} \mu \left(\sum_{1}^{N}{}_a \varphi_a^2 \right) + \frac{g}{4!} \left(\sum_{1}^{N}{}_a \varphi_a^2 \right)^2 - h\varphi_N \right], \tag{10.36}$$

we find for $\mu < 0$, $h = 0^+$,

$$\begin{aligned}
&\varphi_a = \varphi_a^0 + \tilde{\varphi}_a \\
&\varphi_a^0 = 0 \qquad a < N \\
&\varphi_N^0 = \varphi^0 \equiv \sqrt{\frac{-6\mu}{g}} + O(h) \qquad \left(\mu\varphi_0 + \frac{g}{3!} \varphi_0^3 = h \right)
\end{aligned} \tag{10.37}$$

$$H[\varphi] = H[\varphi_0] + \tilde{H}[\tilde{\varphi}]$$

$$\begin{aligned}
\tilde{H}(\tilde{\varphi}) = \int d^D x \Big[&\frac{1}{2} \sum_{1}^{N}{}_a (\partial_\nu \tilde{\varphi}_a)^2 + \frac{1}{2} \mu_T \tilde{\varphi}_N^2 + \frac{\mu_L}{2} \sum_{1}^{N-1}{}_a \tilde{\varphi}_a^2 + \frac{g_3}{3!} \tilde{\varphi}_N^3 \\
&+ \frac{g_3}{3!} \tilde{\varphi}_N \sum_{1}^{N-1} \tilde{\varphi}_a^2 + \frac{g_4}{4!} \left(\sum_{1}^{N}{}_a \tilde{\varphi}_a^2 \right)^2 \Big]
\end{aligned}$$

$$\mu_L = \mu + \frac{1}{2} g\varphi_0^2 = 3\,\frac{h}{\varphi_0} - 2\mu \qquad g_3 = g\varphi_0$$

$$\mu_T = \mu + \frac{1}{3!} g\varphi_0^2 = \frac{h}{\varphi_0} \qquad g_4 = g\,.$$

In the limit $g \rightarrow 0$ the correlation functions are, in momentum space,

$$\begin{aligned}
&G_{ab}^{(p)} = \delta_{ab}[G_L(p)\delta_{a,N} + G_T(p)(1 - \delta_{a,N})] \\
&G_L(p) = \frac{1}{p^2 - 2\mu + 3h/\varphi_0}\,, \qquad G_T(p) = \frac{1}{p^2 + h/\varphi_0}\,.
\end{aligned} \tag{10.38}$$

$G_L(p)$ is the *longitudinal correlation function*, measuring the correlations in the order parameter components parallel to the direction of ordering, whilst $G_T(p)$ is the *transverse correlation function*, measuring the correlations between components of the order parameter which are orthogonal to the direction of ordering. The transverse correlation function has no μ_T term in the denominator (when $h \to 0$), and thus the transverse correlations have *power-law* decay for $T < T_c$. The longitudinal correlation function decays exponentially below T_c, with a correlation length $\xi^2 = (2\mu)^{-1}$. Very often the pole at $p^2 = 0$ for $h = 0$ in the transverse correlation function is called the *Goldstone boson.*[5] We see that for spontaneous symmetry breaking in the $O(N)$ model there is one mode - the longitudinal mode - with a finite correlation length, and $N - 1$ modes - the transverse modes φ_a, $a < N$ with an infinite correlation length.[6] Other systems with different continuous symmetries may have different numbers of Goldstone modes.

Although our calculation was performed only at the first order, the form of is preserved to all orders in perturbation theory; this result - *Goldstone's theorem* - is a direct consequence of the original $O(N)$ symmetry. Goldstone's theorem is generally true when there is a spontaneously broken symmetry, and evasions of the theorem when there are gauge fields present are important in superconductivity and in the electroweak theory, giving rise to a finite electromagnetic penetration depth and intermediate vector bosons respectively. Everyday examples of Goldstone modes include *spin waves* (in magnets) and *phonons* (spontaneous breaking of translational invariance).

In terms of Goldstone bosons the Mermin-Wigner theorem states that continuous symmetries cannot be spontaneously broken at finite temperature in systems with sufficiently short-range interactions in dimensions $D \leq 2$. Intuitively, this means that long-range fluctuations can be created with little energy cost and since they increase the entropy they are favored. This is because if such a spontaneous symmetry breaking occurred, then the corresponding Goldstone bosons, being massless, would have an infrared divergent correlation function (at finite temperature, thermally activated fluctuations of Nambu–Goldstone modes destroy the long-range order, and prevent spontaneous symmetry breaking in one- and two-dimensional systems.).

---

[5] A telegraphic proof of the no-go theorem (Mermin-Wigner theorem) of the previous section runs as follows: the spontaneous breaking of a continuous symmetry implies the presence of Goldstone bosons. Goldstone bosons cannot exist in two dimensions ($1/p^2$ is not a Fourier transform of a function), ergo continuous symmetries are not spontaneously broken in two dimensions.

[6] In quantum field theory, the bare propagator of the boson field φ has the form $\widetilde{G}_0(k) = 1/(k^2 + m^2)$, where $m$is the mass of the particle represented by φ. Thus m is analogous to ξ^-1 in statistical mechanics. Goldstone modes are therefore referred to sometimes as *massless* because their propagator is the same as that for massless bosons in quantum field theory.