to determine the nature of the corresponding quantum theory. This new action still has the same symmetry of the action of the free theory, namely, it is invariant by the sign inversion of the fields. In addition to this, it has a lower bound so long as the parameter λ, which we call the coupling constant, is positive and not zero. On the other hand, unlike what happened in the case of the free theory, the parameter α may be negative in this model, so long as λ satisfies these conditions.

Note that the addition to the action of the free theory of a single cubic term is out of the question for two reasons: it would break the symmetry and, more importantly, would cause the action not to have a lower bound. This second problem is much more serious than the failure of the action to be invariant by the symmetry transformations, since it would imply the non-existence of the corresponding quantum theory. We could, on the other hand, include a cubic term together with the quartic term, thus obtaining a non-symmetrical but stable theory. If we want to have a stable theory and keep the symmetry, we should restrict the discussion to terms with even powers of the field. We will do this here, for simplicity and ease of presentation, and motivated by the fact that, in general, symmetries have an important role to play in physics. We will discuss explicitly the case φ^4, but almost everything that we will do can also be done for the cases φ^{2p}, $p = 2, 3, 4, \ldots$, with analogous results.

We say, in the classical theory defined by the action given above, that the field φ is *self-interacting*. As we may see in future volumes of this series, it is also possible to define models with fields having several components that interact with each other, and that involve invariance by groups of symmetry transformations which are larger and more complex than the simple sign reflections that we have in the model with a single component. It is also possible to define manageable models with different types of field that interact with one another, which are, of course, the most important models for real physics. However, for our objectives here we may limit ourselves to the model with a single field component, postponing to a future opportunity the discussion of the more complex models.

Unlike what happened in the case of the free field, in the non-linear models there is no known way to calculate the predictions of the quantum theory in exact analytical form. In this section we will limit ourselves to the qualitative description of the behavior of the model by means of heuristic arguments based on extensive experience with its numerical treatment. Later on we will develop a technique of approximate perturbative calculations that will allow us to determine in a quantitative and fairly reliable way some of the main characteristics of the model. In general, in the case of the non-linear models it will always be necessary to make use of some approximation technique or of computer simulations in order to determine the behavior of the models.

In our approach to the subject, the computer simulations will often be the main tool for the exploration of the models. Once one becomes well acquainted with the technique of stochastic simulation, it can become a language for the understanding of the models, sometimes leading one to the solution of problems, sometimes suggesting new ideas, new observables and even new models. The ideas and techniques involved

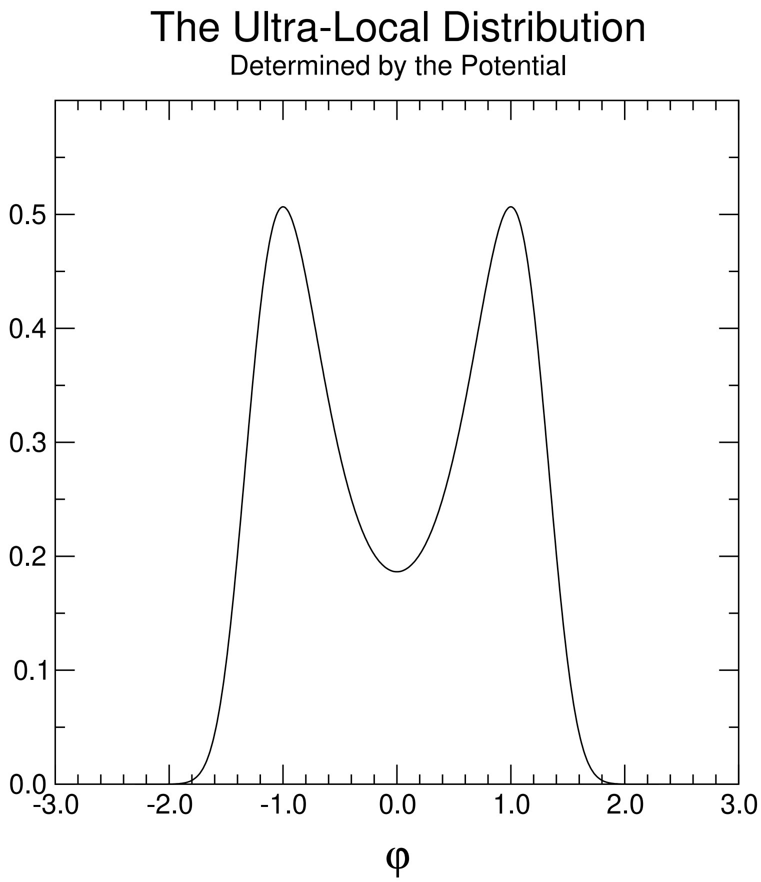

Figure 1.1.1: A typical ultra-local distribution of the fields, given by the potential, that is, by the polynomial terms of the action, in the case $(\alpha < 0, \lambda > 0)$.

In this way the theory is reduced to the study of the effect of the derivative term on the local distributions of either χ or φ at each site. The inverse of relation (1.1.2), which usually can only be obtained numerically, gives us $\varphi(\chi)$ and enables us to obtain φ with the correct non-uniform distribution, starting from the variable χ with an uniform probability distribution within a closed interval, which is not difficult to generate numerically. In this way a part of the structure of the model, the part of the distribution given by the ultra-local terms of the action, is implemented in an exact way. This is, in fact, one of the ways in which one can simulate this model in practice, by producing values of χ at each site with the correct distribution, getting from them the corresponding values of φ, and simulating the dynamics of the derivative term by the use of stochastic techniques. The use in stochastic simulations is the main application of this decomposition, which usually is not very useful as an analytical approach.

In a very general way, the complete local distribution that rules the fluctuating values assumed by the fields at an arbitrarily given site is given by the combination of the effects of the potential and of the derivative term. In order to discuss the behavior not be equal to α, and in addition to this neither will λ_R be equal do λ. In fact, let us recall the fact that the parameter α can be negative in this model, while the parameter $\alpha_R = (m_R a)^2$ is necessarily non-negative, and should tend to zero in the continuum limit. The rule of the game now is that neither α nor λ have any direct physical meaning and that we are free to do with them whatever is necessary, within the constraints of the stability of the theory, in order to have α_R and λ_R assume physically acceptable and significant values in the limit $N \to \infty$.

Since we have two free parameters to adjust in the model, that is, two functions $\alpha(N)$ and $\lambda(N)$ of the increasing size N of the lattice that we may define, it may seem at first sight that we may always choose these functions so as to obtain any physically acceptable values of $\alpha_R(\alpha, \lambda, N)$ and $\lambda_R(\alpha, \lambda, N)$ in the limit $N \to \infty$. However, this is not necessarily so because, besides the stability constraints that we must impose on the basic parameters of the theory, it may be that the dynamics of the theory itself imposes over the renormalized parameters α_R and λ_R other constraints, with the consequence that not all the possibilities are actually realized in practice. In an extreme case, it is possible that there are *no* choices of the functions $\alpha(N)$ and $\lambda(N)$ for which the values of α_R and λ_R are physically acceptable in the limit, in which case we say that the quantum theory of the model *does not exist*. In a more general way, it may be that not all pairs of physically acceptable values for α_R and λ_R are reachable by means of some path $[\alpha(N), \lambda(N)]$ with increasing N, in the space of parameters of the theory. For example, it may be that a constraint between α_R and λ_R is established in the limit, preventing us from choosing both of them freely, in which case we may say that the parameters α and λ become *degenerate* in the limit.

In what follows we will describe a qualitative way of understanding the behavior of the model which, despite the fact that it is purely intuitive and heuristic, based on the phenomenology of computer simulations, will give us qualitatively correct results, as we will verify later on by means of approximate calculations[1]. In order to do this we will, as was mentioned above, represent the effect of the derivative term of the action by the fluctuations that it implies for the values of the field φ at a given site, resulting on a Gaussian local distribution of values with a width of the order of 1. Let us recall that in the case of the free field the width σ_0 did not depend on m_0 in the continuum limit. On finite lattices the width did depend on α_0, but not very strongly, so long as α_0 was not zero on finite lattices. In an analogous way, in our case here we expect the width σ not to depend on m_R in the limit, while on finite lattices it should not depend too strongly on either α or λ.

In this way, in first approximation we may imagine that the width of the local distribution behaves like a semi-rigid body with finite dimensions which are almost constant along the continuum limit. If the width is "squeezed" to any value below its normal size, this gives rise to a non-zero value for the renormalized mass m_R. A zero squeezing force corresponds to zero m_R, and the larger the squeeze, the larger the renormalized mass. As the lattice size N increases the width becomes more "rigid", in the sense that the same squeezing force corresponds to a larger value of m_R, until it becomes infinitely rigid in the continuum limit, in which any non-zero

[1]This argument was originally developed in collaboration with Dr. Timothy Edward Gallivan.

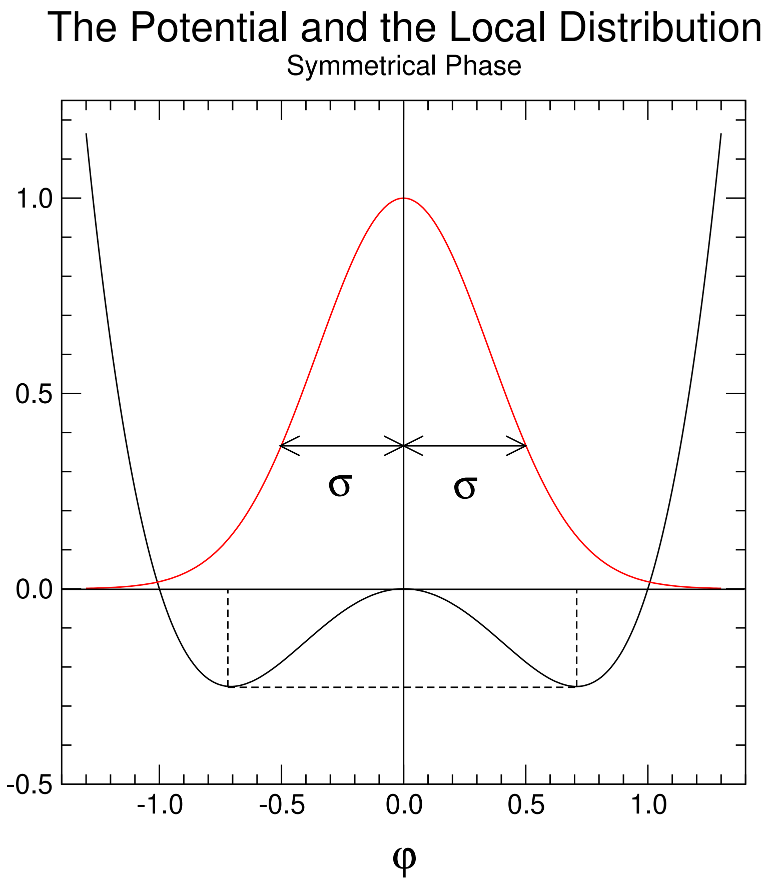

Figure 1.1.3: The potential and the local distribution due to the derivative term, in the case of the non-linear theory, in the symmetrical phase.

potential is larger. This squeeze of the width of the distribution gives rise to a finite and non-zero value for α_0 and hence for the renormalized mass. The decrease in the width of the distribution is never very large, and it is still a quantity of the order of 1. The difference in width due to the squeeze goes to zero in the continuum limit because, as we saw in the first volume of this series, in this limit it is necessary that α_0 go to zero in the free theory, making the potential well become infinitely wide and flat at the position of the minimum. The fact is that such a vanishing effect over the width is sufficient to give to the renormalized mass, in the limit, any positive value we wish.

In our model with a quartic term, in the case where we have both α and λ positive or zero, we should expect a qualitatively similar behavior, since we have the same derivative term in the action and a potential well with a similar form, although the detailed format of the curve is not exactly the same. In this case, in order for the potential well to become infinitely wide in the continuum limit, thus allowing α_R to go to zero and m_R to approach a finite value, it is necessary that both α and λ tend to zero in the limit. In this way we have just made, without too *parameter α fill be necessarily negative*, in any continuum limit of this model that has any chance of not being trivial. In fact, in this case something very interesting happens, because the potential of the model acquires a double well, as shown in figure 1.1.3, which alters completely the behavior of the model, since now a new relevant parameter related to the potential arises, given by the distance between the two minima, which can be easily calculated from the potential. We see that we now have two different widths at play in the problem, the width of the local distribution and the distance between the two minima. We also have two widths related to the potential, the width of each one of the two wells and the total width of the two wells, which are related by a factor of approximately two. The positions of the two minima of the potential are given by $\varphi = \pm\sqrt{-\alpha/\lambda}$, while the value of the potential at the minima is given by $-\alpha^2/(4\lambda)$.

We see at once that now the statistical disadvantage of the rise of the potential at each side of the double well, which tend to squeeze the distribution, can be compensated by the statistical advantage due to the two local minima of the potential, which may tend to widen the distribution. Another way to put it is to say that the central bump of the potential tends to "un-squeeze" the distribution, working against the squeezing tendency of the potential rises at the two sides of the double well. If we tune our parameters in an appropriate way, it may be possible to end up with a vanishing squeeze in the limit, without the need for an infinitely wide well. In this way the possibility arises that we may have in this case $\alpha_R = 0$ without it being necessary that α or λ approach zero in the continuum limit. In other words, the possibility arises that there are certain non-zero pairs of values (α_c, λ_c) that we can approach in the continuum limit so that $\alpha_R \to 0$ in the limit, a behavior which, as we discussed before in the section in [5], is typical of second-order phase transitions. Refining a little our analysis we verify that indeed such a phase transition happens in this model, related to a process of *spontaneous symmetry breaking*.

As we saw, both the distance between the local minima and the width of each one of the two wells around them are proportional to $\sqrt{-\alpha/\lambda}$, which may be made as large as we wish by choosing α negative and with absolute value much larger than λ. In this way, by adjusting the parameters we can make the two potential wells much wider than the width of the local distribution, which is always of the order of 1, thus making it no longer statistically favorable for the distribution to stay centered around $\varphi = 0$. The depth of the two wells is given by $\alpha^2/(4\lambda)$ and also increases when we make the absolute value of α larger than λ, contributing to make it statistically favorable for the distribution to shift to one of the two sides, thus falling into one of the two wells. Since the two wells are identical, this happens in a random way, spontaneously to one of the two sides, which therefore spontaneously *breaks the symmetry* which so far implied that the expectation value of the field had to be zero, $\langle\varphi\rangle = 0$. Note that the local distribution must fall to the *same* side at all sites, otherwise the derivative term would make a huge unfavorable contribution to the statistical weights. The situation of broken symmetry is illustrated qualitatively in figure 1.1.4.

We discover in this way that a process of *spontaneous symmetry breaking* occurs only be obtained by a more complete calculation, and which means that $\beta_c = C_0^2\sigma_0^2$. We have therefore an equation determining the pairs of values (α_c, λ_c),

$$C_0^2\sigma_0^2\lambda_c + \alpha_c = 0,$$

indicating that the locus of the phase transition is a critical line with a negative slope, which extends from the Gaussian point $(\alpha = 0, \lambda = 0)$ all the way to infinity, within the quadrant $(\alpha < 0, \lambda > 0)$ in the parameter plane of the model. Of course it is unlikely that the critical curve is exactly a straight line, because we did not take into consideration, in this qualitative argument, the changes in the depth of the wells due to the variation of the parameters, but we will see later that a straight line is in fact quite a reasonable approximation. Observe that the symmetrical phase occupies all the quadrant $(\alpha \geq 0, \lambda \geq 0)$ and part of the quadrant $(\alpha \leq 0, \lambda \geq 0)$, differing therefore from the classical expectation that making $\alpha < 0$ would always break the symmetry. This is, of course, a direct consequence of the exchange of the classical point body by an extended quantum object within the potential well.

We may now draw a *critical diagram* for the model, illustrating in this way the two phases and the critical curve, like the one that can be seen in figure 1.1.5. The half-axis $(\lambda = 0, \alpha < 0)$ and the lower half-plane $\lambda < 0$ are not included in the diagram, of course, since the model is unstable in these regions. The choice of two functions $\alpha(N)$ and $\lambda(N)$ that determines a particular continuum limit of the model corresponds to a path drawn in this diagram, which may start at any point within the stable region but which must necessarily end at some point of the critical curve, which is the locus where we have $\alpha_R(\alpha, \lambda) = 0$ in the limit. These paths are called *flows*, or renormalization flows of the model. The Gaussian point is the critical point of the theory of the free field and the continuum limits of that model are represented by flows that go along the semi-axis $(\lambda = 0, \alpha > 0)$ in the direction of $\alpha = 0$. We can see here, once again, that any limits staying within the quadrant $(\alpha \geq 0, \lambda \geq 0)$ must approach the Gaussian point. Flows can approach the same point of the critical line from either the symmetrical phase or the broken-symmetrical phase, possibly producing different results.

The slope of the critical curve at the Gaussian point is finite and non-zero and can be calculated by a perturbative approximation, as we shall see later in this chapter. The slope of the curve at the asymptotic region is also finite and non-zero, and can be related to critical points of other models of scalar fields, the so-called non-linear sigma models, as we shall also see later on. In addition to this, the qualitative properties of this curve can also be confirmed by means of another process of approximation that we will examine in detail later on, namely the so-called mean-field techniques. By and large the nature of the critical curve is rather well established and understood in any dimension $d \geq 3$, and the analysis can be extended without any important qualitative changes to the models $\lambda\varphi^{2p}$ for $p \geq 3$, as well as to multi-component models which are invariant under larger symmetry groups. In this last case the presence of more field components does introduce some new elements into the structure, of course. Usually more precise calculations of the

$$\lambda v_R^2 = -(\alpha + C_0^2 \lambda \sigma_0^2),$$

whenever the quantity in parenthesis is negative, and zero otherwise. Note that the quantity in parenthesis coincides with the equation of the critical line, thus showing that indeed v_R^2 is zero over that line. Note also that the result for v_R^2 contains $1/\lambda$, possibly indicating that a calculation which is purely perturbative in λ may not be sufficient to obtain this result.

We may also obtain estimates for the dimensionless squared mass α_R, using arguments similar to the ones above. Starting with the symmetrical phase, which contains the possibility that $\lambda = 0$, with $\alpha > 0$, we know two things about α_R: first, it must be zero over the critical line, and second it must be equal to α when $\lambda = 0$. Since the equation of the critical line contains a term linear in α, it is clear that in order to satisfy both these criteria we must make α_R equal to that equation, thus obtaining

$$\alpha_R = \alpha + C_0^2 \lambda \sigma_0^2 \quad \text{for} \quad \alpha + C_0^2 \lambda \sigma_0^2 \geq 0,$$

where the condition is the same as before, characterizing the symmetrical phase. In the broken-symmetrical phase we must work a little more to get the result. First of all, let us discuss why there should be a non-zero physical mass in this case. This is so because, after the symmetry breaks and the local distribution falls within one of the two wells, it will become squeezed by it, leading to an increase in the renormalized mass, is a way similar to what happens in the free theory. So it follows that α_R should have its minimum of zero at the critical line and increase when one goes away from it on either side.

In order to estimate the value that α_R should have in the broken-symmetrical phase, let us go deeply into it, making $-\alpha \gg \lambda$, so that the potential is very wide and the local distribution is sitting around one of the two local minima. Under these conditions we may approximate the potential in the relatively small region where the local distribution is significantly different from zero by a parabola. If we calculate the second derivative of the potential at the minimum, which gives the curvature of this parabola, we get for it the value -2α, which is positive because α is negative. By comparison with the situation in free theory we see now that deep in the broken-symmetrical phase we should have $\alpha_R = -2\alpha$. Adding to this that in this phase α_R still must be zero over the critical line, we see that we must make α_R proportional to the equation of the line, with a constant of proportionality that will bring about the correct value in the deeply broken regime. With all this it is not difficult to see that we must have

$$\alpha_R = -2(\alpha + C_0^2 \lambda \sigma_0^2) \quad \text{for} \quad \alpha + C_0^2 \lambda \sigma_0^2 \leq 0,$$

where the condition is once again the same as before, characterizing the broken-symmetrical phase. Note that the value -2α for the curvature of the parabola near the minimum only goes to zero is we make $\alpha = 0$ and hence go back to the Gaussian quantities. If we recall our discussion about the physical significance of the block variables in the section in [6], we will see that it is the *dimensionfull* variables and parameters that have a more direct physical relevance in the quantum theory. While the dimensionless local variables and parameters, usually numbers of the order of 1 that do not scale significantly in the continuum limit, are convenient both for establishing mathematical facts about the internal structure of the models and for dealing with them in a practical way in computer simulations, the dimensionfull variables include scale factors that cause them to scale in the continuum limit in the correct way in order to represent the superpositions of the dimensionless observables over the large and increasing numbers of sites contained within the blocks. As we discussed in that section, such superpositions constitute the only type of quantity within the theory that can in fact be directly observed.

The definition of the dimensionfull field in terms of the dimensionless field in the $\lambda\varphi^4$ model is the same as in the theory of the free field, $\phi = a^{(2-d)/2}\varphi$, since it is determined only by the derivative term. We do not have in this case a mass term properly speaking, since α can be both positive and negative, but in a way similar to that of the theory of the free field we may introduce a parameter m with dimensions of mass by means of the relation $m^2 = |\alpha|/a^2$. A simple analysis of the quartic term gives us, finally, the definition of the dimensionfull coupling constant $\Lambda = a^{d-4}\lambda$. The treatment of the sums over links and sites and of the finite differences in the continuum limit of the classical theory, in terms of integrals and derivatives, is identical to the one discussed in the case of the free theory, so that we obtain for the action $S[\phi]$ in the continuum limit, in terms of the dimensionfull quantities,

$$S[\phi] = \int_V \mathrm{d}^d x \left\{ \frac{1}{2}\sum_\mu [\partial_\mu \phi(\vec{x})]^2 \pm \frac{m^2}{2}\phi^2(\vec{x}) + \frac{\Lambda}{4}\phi^4(\vec{x}) \right\},$$

where the sign of the quadratic term depends on the sign of α. Observe that the relation existing between the parameters λ and Λ of the classical theory implies that, since Λ must remain finite in the limit, λ must behave in different and definite ways in each space-time dimension. For $d \geq 5$ it is necessary that λ diverge to infinity in the limit in order that Λ be different from zero, which shows that these classical theories have a rather singular behavior in this case. For $d \leq 3$, on the other hand, it is necessary that $\lambda \to 0$ in the limit in order for Λ to remain finite, showing that in this case the behavior is the reverse of that of the previous case. For $d = 4$ we have that $\Lambda = \lambda$ and therefore in this case it is not possible to make any definite statement of this type. Given these scaling relations between λ and Λ it is reasonable to think that the dimensionfull renormalized coupling constant Λ_R should be defined in terms of λ_R in an analogous way.

Our expectation is that, just as it is the constant Λ that has physical relevance in the classical theory, the constant Λ_R should play the same role in the quantum theory. As was already pointed out, the analysis of the block propagator of the free theory in the section in [6] indicated that it is the dimensionfull quantities, based on

1.1.8. Derive the scaling relations between λ and Λ for the classical theory of the $\lambda\varphi^{2p}$, $p = 3, 4, \ldots$ models, in dimensions d from 3 to 5. Assuming only that λ remains finite in the limit, verify in which cases it is possible to have non-trivial $N \to \infty$ limits, in which Λ is finite and non-zero.

1.2 Perturbation Theory

In this section we will develop in detail a perturbative approximation technique for the $\lambda\varphi^4$ model which we introduced in section 1.1. As we shall see later on, it will allow us to confirm the qualitative behavior of the model, which was described in a heuristic way in that section. Let us recall that the model is defined by the action

$$S[\varphi] = \frac{1}{2}\sum_{\ell}(\Delta_\ell\varphi)^2 + \frac{\alpha}{2}\sum_{s}\varphi^2(s) + \frac{\lambda}{4}\sum_{s}\varphi^4(s),$$

containing a quartic interaction term. Due to the presence of the quartic term we do not know how to solve the model analytically. However, without this term the model becomes Gaussian and then we are able to solve it completely. It becomes clear then that the results of the complete model should converge to the corresponding results of the free theory when we make $\lambda \to 0$ since, in the continuum limit, this implies that we must approach the Gaussian point in the parameter plane of the model.

The main idea of perturbation theory is to develop an expansion for the complete model around the soluble Gaussian model. Presumably, for small values of the coupling constant the results of the complete model are not very different from the results of the free theory and hence we may understand the interaction term as a small perturbation applied to the Gaussian model. In this way, maybe we will be able to use the expansion in order to obtain useful approximations for the complete model near the Gaussian point in the critical diagram. This is just the usual expectation that one has for an approximation scheme, but a word of warning is in order here. Although we will see that it is in fact possible to calculate some useful approximations, things are not as simple as one may think at first, and the approximation scheme does not work quite in the way that one would expect.

The first step in the development of the perturbative technique is the separation of the action in two parts, which we shall denominate S_0 and S_I,

$$S = S_0 + S_I,$$

where S_0 is a purely Gaussian action. For the time being we will not be very specific about the detailed form of each one of the two parts. We have, for an arbitrary observable $\mathcal{O}$ of the complete model,

$$\langle\mathcal{O}\rangle = \frac{\displaystyle\int[\mathbf{d}\varphi]\mathcal{O}[\varphi]e^{-S_0}e^{-S_I}}{\displaystyle\int[\mathbf{d}\varphi]e^{-S_0}e^{-S_I}}. \tag{1.2.1}$$

$$S_M = \frac{\alpha_0}{2} \sum_s \varphi^2(s),$$

diverge as powers of N in the continuum limit, even if we keep the models within boxes with finite volumes. In the case of S_K we have $\langle S_K \rangle \sim N^d/2$, while in the case of S_M we have $\langle S_M \rangle = (m_0 L)^2 N^{d-2} \sigma_0^2/2$. In addition to this, it is possible to show that in the free theory the following relation holds,

$$\langle \varphi^4 \rangle_0 = 3 \langle \varphi^2 \rangle_0^2,$$

which is the result indicated in the problem in [9]. From these consideration it follows that, assuming that the general form of S_I is given by S_V,

$$S_V = \sum_s \left(\frac{\alpha}{2} \varphi^2 + \frac{\lambda}{4} \varphi^4 \right),$$

where $S = S_K + S_V$, we have for its expectation value

$$\langle S_V \rangle_0 = \frac{\sigma_0^2}{2} \left(\alpha + \frac{3}{2} \sigma_0^2 \lambda \right) N^d.$$

This means that, so long as the factor within parenthesis is not zero in the limit, $\langle S_V \rangle_0$ diverges as N^d in the continuum limit.

At first sight it may seem that the expression in parenthesis may indeed vanish in the limit, since we must remember that, as was discussed in section 1.1, α is necessarily negative in the limit, while the factors contained in the second term of the expression are all positive. In fact, this expression is similar to our heuristic estimate for the equation of the critical curve, which was $\alpha + C_0^2 \sigma_0^2 \lambda = 0$. However, one can verify a-posteriori that the expression is not identical to the equation of the critical curve, either by numerical means or by the the approximations in which we will calculate the equation of the curve later on. For example, in the case of the perturbative approximation we will verify that the two expressions differ by the extra factor of $1/2$ that appears in the second term in the parenthesis in the expression of $\langle S_V \rangle_0$.

In any case, even if the expression in parenthesis did coincide with the equation of the critical line, it would not be equal to zero on finite lattices, but would only approach zero in the $N \to \infty$ limit, with some inverse power of N. Since the expression is multiplied by a factor of N^d, it would have to go to zero very fast in order to avoid the divergence. As we saw in section 1.1 and will confirm quantitatively later on, the equation of the critical curve is directly related to the value of α_R, so that it must go to zero exactly as N^{-2}, which is not enough to eliminate the divergence in the dimensions of interest, $d \geq 3$. Furthermore, even if everything worked out and $\langle S_V \rangle_0$ did go to zero in the limit, if we consider that we also have that $\langle S_K \rangle_0$ diverges as N^d in the limit, we see that the resulting theory could not possibly fail to become make both α and λ approach zero very quickly, thus making the model return to the Gaussian point.

This behavior of S_I is the basic cause that is behind all the divergences that appear in the perturbative expansion of the model. It is directly related to the strong fluctuations undergone by the fields in the continuum limit, as well as with the fact that the dominating field configuration are discontinuous in the limit, as we studied in the section in [8]. Despite all this, it is still very reasonable to think that the observables $\langle\mathcal{O}\rangle$ of the complete model are continuous functions of the parameters of the model, because the observables are defined by means of statistical averages that eliminate the fluctuations and discontinuities which are characteristic of the fundamental field. In other words, it is reasonable to think that $f(\varepsilon)$ is at least a continuous and differentiable function of ε, so that there should be at least a reasonable first-order approximation for f near $\varepsilon = 0$, and it could even be that f is an analytical function of ε (problem 1.2.1).

We are faced here by a rather strange situation: on the one hand, it is reasonable to think that there is an approximation up to some order for the observables of the complete model in the vicinity of the Gaussian point but, on the other hand, we see that this approximation may not be accessible by means of the perturbative expansion starting from the definition of the quantum theory, due to the divergences that appear. Observe that this apparent conflict is related to a *exchange of order* of two limits, involving the continuum limit and the limit of the summation of the perturbative series. We may argue that on finite lattices the perturbative series can be summed, since all the quantities involved are finite and well-behaved in this case. Hence, in principle we may sum the perturbative series on finite lattices and after that take the continuum limit. However, when we write the series only up to a certain term of finite order and then take the continuum limit, we are inverting the order of the two limits. Although it is reasonable to think that, once the continuum limit is taken, the resulting observables should have convergent expansions in terms of the parameters of the model, there is no guarantee that these expansions are those obtained by the exchange of the order of the limits. In fact, the divergences that appear show us that the two procedures must have very different results.

At this point it is important to observe that the equation (1.2.1) which defines the observables of the quantum theory is a ratio of two quantities involving S_I and that, due to this, it is possible that some or even all the divergences due to this quantity end up by cancelling each other, between those coming from the numerator and those coming from the denominator, if we make a careful expansion of the ratio, that is, a careful expansion of $f(\varepsilon)$. We will verify later on that it is indeed possible to obtain in this way a useful approximation for some of the observables of the complete model, despite the divergences that are involved in the limit, but we should keep in mind that we are dealing with a singular expansion, so that it should come as no surprise it not everything works out perfectly as expected. It is in this context that the idea of *renormalization* appears for the first time with a recognizable meaning. Unfortunately, this term is used for several different things in the structure of the theory, but here it really has to do with renormalizing something in the usual sense.

parameter α will become strictly negative. Therefore we cannot include the α term in S_0, because this quadratic action would become unbounded from below and the corresponding measure well be ill-defined even on finite lattices. The alternative of including only the derivative term in S_0 and of simply including the α term in S_I is also not adequate, since the free massless theory that results from this has a zero mode that could be absent from the complete model, leading to the possibility of the appearance of spurious infrared divergences.

In order to avoid possible infrared problems we will introduce into the model a new parameter $\alpha_0 \geq 0$ associated to a quadratic term containing φ^2, in such a way that the model is not actually changed. Dealing first with the case in which we are in the symmetrical phase, we will choose for S_0 the action of the free theory as we have studied it since the section in [10],

$$S_0[\varphi] = \frac{1}{2}\sum_{\ell}(\Delta_\ell\varphi)^2 + \frac{\alpha_0}{2}\sum_{s}\varphi^2(s).$$

The interaction part S_I of the action will contain the remaining terms of the original action and a term containing α_0 with the opposite sign, so that the sum of S_0 and S_I continues equal to the original action. It follows that in this symmetrical phase we will have for S_I

$$S_I = \sum_{s}\left[\frac{\alpha - \alpha_0}{2}\varphi^2(s) + \frac{\lambda}{4}\varphi^4(s)\right].$$

The parameter α_0 is clearly irrelevant in the exact model and the final results should be independent of it. We will see later on that this is indeed the case but, since α_0 appears both in S_0 and in S_I, which will be treated in very different ways during the development of the approximation technique, we will also see that there are some subtleties relating to the role played by α_0. Up to this point it seems that we are free to keep the parameter α_0 finite and non-zero in the $N \to \infty$ limit, but it is not very reasonable to do this because this procedure would correspond to a diverging mass m_0 for the distribution defined by S_0 in the limit. Instead of that, we will choose $\alpha_0 = m_0^2/N^2$ and work with an m_0 which is kept finite in the limit, rather than diverging. What we are hinting at here is that perhaps it is possible to improve the quality of the approximation by a suitable choice of the free parameter α_0. If we knew beforehand the value m_R of the renormalized (physical) mass of the complete model in the limit, we could even consider making $m_0 = m_R$. Although it is not apparent at this moment that we should do this, or that we could do it, since we do not yet know m_R, we will see later on that this is, in fact, a natural and very convenient choice.

In the broken-symmetrical phase we expect that the expectation value of the field $\langle\varphi\rangle$ will be different from zero and, in order to enable us to develop the perturbative approximation is a simpler way, it is convenient to first rewrite the model in terms of a shifted field φ' given by

$$\varphi' = \varphi - v_R, \quad \varphi = \varphi' + v_R, \ \langle\varphi'\rangle = 0, \quad \langle\varphi\rangle = v_R,$$

1.2.1. ($\star$) Determine whether the function $f(\varepsilon)$ defined in equation (1.2.2) is or is not analytical as a function of ε. In order to do this, first extend ε to the complex plane, $\varepsilon = x + \imath y$ with real x and y, writing the function f, now complex, as $f = u(x, y) + \imath v(x, y)$. Verify then whether $u(x, y)$ and $v(x, y)$ satisfy the two Cauchy-Riemann conditions: $\partial_x u(x, y) = \partial_y v(x, y)$ and $\partial_y u(x, y) = -\partial_x v(x, y)$. Perform the verification both on finite lattices and in the continuum limit.

1.2.2. ($\star\star$) It is argued in the text that the problems with the perturbative expansion originate from the fact that $\langle S_I \rangle_0$ diverges as N^d in the continuum limit. This causes, for example, the denominator of equation (1.2.1), which defines the observables, to behave in the limit as

$$\left\langle e^{-S_I} \right\rangle_0 \to 0.$$

One could imagine that one way to try to get around this problem is to add to the action a field-independent term $\zeta(\alpha, \lambda, N)$, which corresponds to multiplying both the numerator and the denominator of equation (1.2.1) by a number $Z(\alpha, \lambda, N) = \exp[\zeta(\alpha, \lambda, N)]$. This corresponds to a *renormalization* of the statistical averages that define the expectation values of the complete model in terms of the expectation values of the free theory, leading to

$$\langle \mathcal{O} \rangle = \frac{\displaystyle\int [\mathbf{d}\varphi] \mathcal{O}[\varphi] e^{-S_0} e^{\zeta - S_I}}{\displaystyle\int [\mathbf{d}\varphi] e^{-S_0} e^{\zeta - S_I}} = \frac{\left\langle \mathcal{O}[\varphi] e^{\zeta - S_I} \right\rangle_0}{\left\langle e^{\zeta - S_I} \right\rangle_0}.$$

Naturally, this does not change the observables. However, we are now free to choose ζ in any way we choose, and we may consider choosing it so that the quantity $\zeta - S_I$ acquires a small or even a vanishing average value, rather than diverging as N^d in the limit. It is clear that in this case ζ will have to be chosen so as to diverge in the limit and hence cancel the divergence of the average value of S_I. Observe however that in this way we can control only the *average value* of the difference $\zeta - S_I$, we cannot control the *fluctuations* of this quantity, because ζ cannot depend on the fields.

If we recall that, as was seen in the text, the large-N limit of equation (1.2.1) is of the type $0/0$, it is reasonable to think that a general criterion or *renormalization condition* for the choice of ζ would be

$$\left\langle e^{\zeta - S_I} \right\rangle_0 = 1,$$

which causes the limit to cease to be of the type $0/0$, but which is a very complicated condition to implement. To first order, we may think that the condition $\langle \zeta - S_I \rangle_0 = 0$ should be sufficient, and it is a condition which is much simpler to deal with. About this type of renormalization procedure we have the following tasks to propose:

In the symmetrical phase we necessarily have that $v_R = 0$, while in the broken-symmetrical phase we may have $v_R \neq 0$. If the phase transition is of second order with respect to this parameter, as it is to be expected, then the critical curve is the geometrical locus in the parameter plane (α, λ) of the model where the solution $v_R = 0$ becomes the only possibility, when we move from the broken-symmetrical phase to the symmetrical phase in the parameter plane. What we will do is to determine the values of (α, λ) for which $v_R \neq 0$ is a possibility and then impose that $v_R = 0$ be the only solution, so as to determine the critical curve. Along the process, a trivial $v_R = 0$ solution that exists in all the parameter plane will be factored out and eliminated. In the broken-symmetrical phase this solution corresponds, to make an analogy with the classical case, to the unstable solution in which the system is at the local maximum of the potential at $\varphi = 0$.

In order to perform this calculation we must use the separation of the action in the free and interaction parts given in equations (1.2.4) and (1.2.5), which are those that should be used in the broken-symmetrical phase. First of all we write the definition of v_R, that is, that it is the expectation value of the original field $\varphi = \varphi' + v_R$. Next we use the perturbative expansion given in equation (1.2.3) in order to write the expectation values involved, limiting ourselves to the terms of order zero and one. We choose arbitrarily the site of the lattice with integer coordinates $\vec{n} = \vec{0}$ in order to do the calculation, a choice which is possible due to the discrete translation invariance of the lattice. Doing all this we obtain

$$v_R = \langle \varphi' + v_R \rangle_0 - \left[\langle (\varphi' + v_R) S_I \rangle_0 - \langle (\varphi' + v_R) \rangle_0 \langle S_I \rangle_0 \right].$$

Since $\langle \varphi' \rangle_0 = 0$ by construction, several terms vanish and we obtain, up to this order, a very simple equation,

$$\langle \varphi' S_I \rangle_0 = 0.$$

Is we write this in detail, substituting the expression for S_I and then using all the available symmetries in order to simplify the expression (problem 1.3.1), in particular the fact that the expectation values of odd powers of the field are zero due to the fact that S_0 is symmetrical by reflection of the fields, we obtain

$$\sum_s \left[v_R \left(\alpha + \lambda v_R^2 \right) \langle \varphi'(s) \varphi'(0) \rangle_0 + \lambda v_R \left\langle \varphi'^3(s) \varphi'(0) \right\rangle_0 \right] = 0. \tag{1.3.1}$$

This equation is simply the lattice version of the equation known as the *"tadpole" equation* in one-loop order. Since all the terms contain at least one factor of v_R, we may now cancel out one factor of v_R, which is the trivial $v_R = 0$ solution which we mentioned before, obtaining

$$\left(\alpha + \lambda v_R^2 \right) \left\langle \varphi'(0) \sum_s \varphi'(s) \right\rangle_0 + \lambda \left\langle \varphi'(0) \sum_s \varphi'^3(s) \right\rangle_0 = 0.$$

The calculation of the remaining expectation values involves only Gaussian integrals and we obtain for the first term (problem 1.3.2)

$$-\alpha_c = 3\lambda_c\sigma_0^2, \tag{1.3.5}$$

where $\alpha > -3\lambda\sigma_0^2$ corresponds to the symmetrical phase and $\alpha < -3\lambda\sigma_0^2$ to the broken-symmetrical phase. We see that this equation has the same form of the heuristic estimate that we proposed in section 1.1, differing from it only by the numerical factor $\sqrt{3} \sim 1.73$ involved in the evaluation of the relation between the width σ_0 of the local distribution and the parameter $\sqrt{-\alpha/\lambda}$ of the potential well. In other words, the result coincides with our heuristic estimate if we choose for the numerical constant C_0 introduced in that section the value $C_0 = \sqrt{3}$. We may now write our perturbative result for v_R in terms of the expression in the equation of the critical line as

$$v_R = \sqrt{\frac{-(\alpha + 3\lambda\sigma_0^2)}{\lambda}}, \tag{1.3.6}$$

which is only real in the broken-symmetrical phase, as expected, and which shows explicitly how v_R goes to zero when one approaches the critical line from the broken-symmetrical phase.

At this point it is important to point out, quite emphatically, that we have just found one more worrisome property of the perturbative approximation technique. We have found here a definite result for the position of the critical curve for the model in a box with periodical boundary conditions, for any value of N, either finite or not. In addition to this, this position of the critical curve has the curious property of depending weakly on the irrelevant parameter α_0 if N is finite, and of becoming independent of the same parameter if $N \to \infty$. Taken in this superficial way, our result seems to indicate that, given a value of α_0, the system displays a completely well-defined phase transition on finite lattices with periodical boundary conditions.

However, it is a well-known fact that there is no possibility of existence of a phase transition on finite lattices with periodical boundary conditions in systems of the type that we are examining here. In this kind of system, with couplings only between next-neighbors and without external borders, the phase transition can be realized only in the $N \to \infty$ limit. We can only presume that the curious dependence on α_0 for finite N is somehow related to this fact, effectively indicating, at best, that there can be a kind of "approximate critical behavior" for finite N. This is one more circumstance in which we verify that this method of approximation has rather singular properties and that it should only be used with the greatest care.

A particularly interesting aspect of the structure of the model that we can obtain from equation (1.3.5) is the slope $\partial\lambda_c/\partial\alpha_c$ of the critical curve near the Gaussian point, which is given by

$$\frac{\partial\lambda_c}{\partial\alpha_c} = -\frac{1}{3\sigma_0^2}. \tag{1.3.7}$$

We may ask here how close to the truth this result can be. Note that it depends neither on α nor on λ, and let us recall that the dependence on α_0 vanishes in the and where

$$g_1(\vec{n}_1,\vec{n}_2) = g_0(\vec{n}_1,\vec{n}_2) - \left[\langle\varphi(\vec{n}_1)\varphi(\vec{n}_2)S_I\rangle_0 - g_0(\vec{n}_1,\vec{n}_2)\langle S_I\rangle_0\right].$$

The expectation values that appear here are the zero-order propagator, as we calculated it before in the theory of the free field,

$$g_0(\vec{n}_1,\vec{n}_2) = \frac{1}{N^d}\sum_{\vec{k}} e^{\imath\frac{2\pi}{N}\vec{k}\cdot(\vec{n}_1-\vec{n}_2)}\frac{1}{\rho^2(\vec{k})+\alpha_0},$$

and the expectation values containing S_I. The first one of these can be easily calculated (problem 1.3.5) in terms of expectation values that we have discussed and calculated before in the sections in [12] and [11], yielding

$$\langle S_I\rangle_0 = \frac{1}{2}\left(\alpha-\alpha_0+\frac{3}{2}\lambda\sigma_0^2\right)\sigma_0^2 N^d. \tag{1.3.8}$$

Observe that all the terms diverge strongly in the continuum limit, containing factors of N^d. The calculation of the last expectation value (problem 1.3.6) is longer and, after some work, we may write it in the form

$$\begin{aligned}\langle\varphi(\vec{n}_1)\varphi(\vec{n}_2)S_I\rangle_0 &= \frac{1}{2}\left(\alpha-\alpha_0+\frac{3}{2}\lambda\sigma_0^2\right)\sigma_0^2 N^d\, g_0(\vec{n}_1,\vec{n}_2)\\ &\quad+(\alpha-\alpha_0+3\lambda\sigma_0^2)\sum_{\vec{n}} g_0(\vec{n}_1,\vec{n})\, g_0(\vec{n},\vec{n}_2).\end{aligned} \tag{1.3.9}$$

At this point we have everything written in terms of the propagator of the free theory. Observe that here also we have terms with strong divergences, involving factors of N^d. The sums over position space may be rewritten in momentum space and manipulated in such a way that, when all the terms are brought together, one verifies that all the terms with strong divergences cancel out, resulting in the final expression for the first-order propagator in position space,

$$g_1(\vec{n}_1,\vec{n}_2) = \frac{1}{N^d}\sum_{\vec{k}} e^{\imath\frac{2\pi}{N}\vec{k}\cdot(\vec{n}_1-\vec{n}_2)}\left\{\frac{1}{\rho^2(\vec{k})+\alpha_0} - \frac{\alpha-\alpha_0+3\lambda\sigma_0^2}{[\rho^2(\vec{k})+\alpha_0]^2}\right\}.$$

The expression within braces is the form of the propagator in momentum space. Observe that this time the result depends significantly on α_0. On the other hand, we may use our freedom in principle, of choosing α_0 in any way we wish within the stability bounds, in order to simplify this expression, by eliminating the second term, which contains a double pole. In order to do this is suffices to choose

$$\alpha_0 = \alpha + 3\lambda\sigma_0^2.$$

We may do this only so long as the resulting α_0 remains positive and so long as it goes to zero in the continuum limit. Examining the expression in the right-hand side approximate the expectation values of the complete model by expectation values of a Gaussian model, which is characterized by only two independent quantities, the expectation value of the field v_R, which is related to the first-order moment (observables with a single power of the field) of the statistical distribution of the model, and the renormalized mass m_0, which is related to the second-order moment (observables with two powers of the field). In the case in which there is a non-vanishing v_R in the complete model, the shift from the field φ to the field φ' can be understood as a way to make identical the first-order moments of the two distributions, that of the complete model and that of the Gaussian model used for the approximation. In a similar way, the choice $\alpha_0 = \alpha_R$ can be understood as a way to make identical the second-order moments of the two distributions. Both are implicit conditions which are resolved in a self-consistent way at the end of the calculations.

We see therefore that what we are dealing with here is, much more than part of a perturbative expansion, a *Gaussian approximation* technique, which is not at all an expansion, but rather a single-step self-consistent type of approximation. Since the Gaussian does not have any moments with order greater than two, we cannot expect that this technique can be successfully used to approximate observables that are related to the higher moments of the distribution of the complete model. In particular, we should not expect that it will be useful to examine the issue of the renormalized coupling constant and the phenomenon of the interaction between particles within the structure of the quantum theory, which are related to the moments of order four and larger. In addition to this, we should not expect that it will be possible to improve on the results obtained here by the inclusion in the calculations of the terms of higher order of the expansion given in equation (1.2.3) since, when we adjust the only two independent moments existing in the Gaussian distribution so as to make them identical to the corresponding moments of the distribution of the complete model, we are already doing the best that can be done in terms of approximate a non-Gaussian distribution by a Gaussian distribution.

As our last objective in this section, we calculate the propagator of the model in the broken-symmetrical phase. The calculations are all very similar to the corresponding calculations in the symmetrical phase, except for the need of the use in this case of the shifted field φ'. In particular, in this case the same type of cancellation of all the terms with strong divergences takes place. After some work (problem 1.3.7) we obtain in this phase for the first-order propagator, which we denote by $g_1'(\vec{n}_1, \vec{n}_2)$, with

$$\langle \varphi'(\vec{n}_1)\varphi'(\vec{n}_2)\rangle \approx g_1'(\vec{n}_1, \vec{n}_2),$$

the result

$$g_1'(\vec{n}_1, \vec{n}_2) = \frac{1}{N^d}\sum_{\vec{k}} e^{\imath\frac{2\pi}{N}\vec{k}\cdot(\vec{n}_1-\vec{n}_2)}\frac{1}{\rho^2(\vec{k}) + \alpha_R},$$

where the renormalized mass is now defined in terms of the dimensionless parameter

$$\alpha_R = -2\left(\alpha + 3\lambda\sigma_0^2\right), \tag{1.3.11}$$

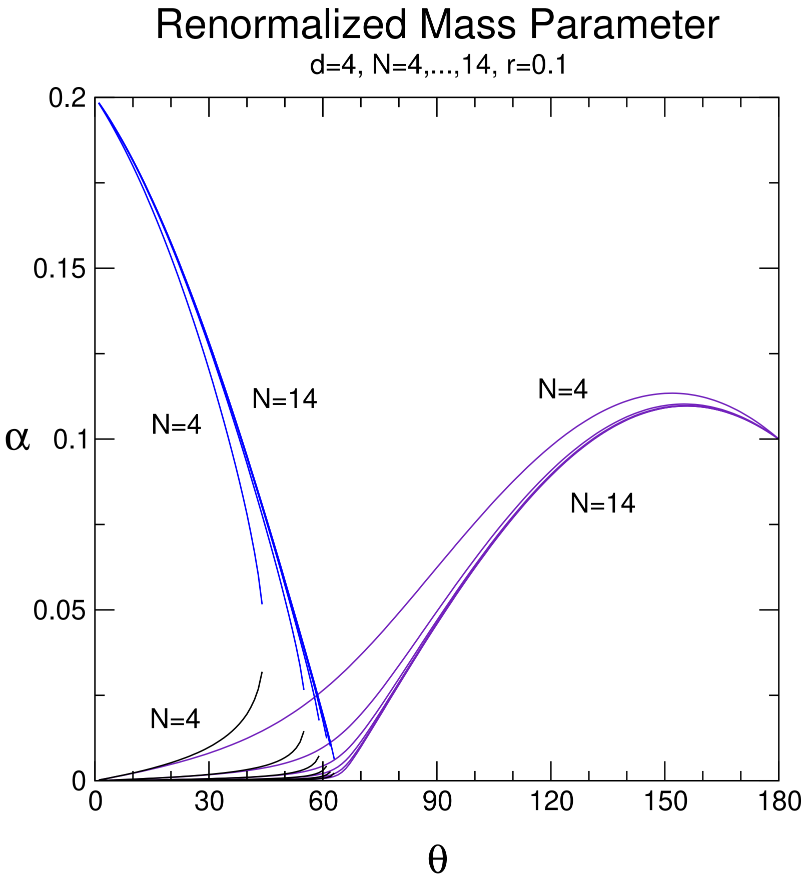

Figure 1.3.1: Perturbative results for the renormalized mas parameter α_R. The curves on the left part are in the broken-symmetrical phase. In this graph α and λ are represented by the equivalent parameters r and θ, which are defined in the text. The values of the angle θ are given in degrees.

can be determined numerically, and we might call this curve the perturbative critical curve. The small solution becomes identically zero in the $N \to \infty$ limit, showing that we should not attribute to it any physical meaning. Once more this seems to be just a perturbative ghost associated to an unstable solution sitting at the maximum that the potential has at the origin. The curves corresponding to the large solution are the ones with their maximums at the left in figure 1.3.1.

One can see that it is the large positive solutions in either phase that carry the expected physical meaning by noting that for θ equal to 180° we are over the positive α semi-axis and therefore have the result for the free theory, $\alpha_R = \alpha$, since in this case $\alpha = r = 0.1$. For θ close to 0° we approach the negative α semi-axis where, as we discussed before, the two potential wells acquire a large separation from each other and the local distribution of the fields sits at the minimum of one of them. As we saw before, the minimum of the potential can then be approximated by a parabola with a positive quadratic coefficient -2α, and the model once more breaking in the quantum theory of the $\lambda\varphi^4$ model. As we shall see later on, these results are surprisingly precise in some cases and, by and large, give us a qualitatively correct picture of the critical behavior of the model. On the other hand, it is doubtful that the technique can be extended in an effective way to other observables and aspects of the model. As far as one can verify up to this point, the model seems to contain particles of mass m_R, which we may adjust freely, in addition to being able to generate a non-vanishing expectation value $v_R = \langle\varphi\rangle$ for the dimensionless field. In the continuum limit v_R vanishes, since we must approach the critical curve where the phase transition is of second order, with $v_R = 0$ over the curve, but it is possible to adjust things so that the dimensionfull field has a non-vanishing expectation value $V_R = \langle\varphi\rangle$ in the limit (problem 1.3.12). Hence, up to this point the model seems to contain only the phenomena of propagation and of spontaneous symmetry breaking. Whether or not it contains anything beyond this is an issue for further exploration and discussion (problems 1.3.13 and 1.3.14).

Problems

1.3.1. Write the expectation value $\langle\varphi' S_I\rangle_0$ in detail and derive equation (1.3.1).

1.3.2. Calculate in detail the expectation value shown in equation (1.3.2).

1.3.3. Calculate in detail the expectation value shown in equation (1.3.3).

1.3.4. Show that the result expressed by equation (1.3.5) implies that the result in equation (1.3.7) for the slope of the critical curve at the origin is in fact exact. In order to show this, take the limits involved with due care: take first the limit $N \to \infty$ under the condition $\alpha_0 = m_0^2/N^2$ for finite m_0, and then take the limit in which $\alpha_c \to 0$ and $\lambda_c \to 0$ along the critical curve, and in which the ratio λ_c/α_c is kept finite and non-zero. Obtain the final result in the form

$$\tan(\theta) = \lim_{N\to\infty} \frac{1}{3\sigma_0^2(\alpha_0)},$$

where θ is the angle between the negative α semi-axis and the tangent to the critical curve at the Gaussian point.

1.3.5. Calculate in detail the expectation value shown in equation (1.3.8).

1.3.6. ($\star$) Calculate in detail the expectation value shown in equation (1.3.9).

1.3.7. ($\star$) Calculate in detail the propagator in the broken-symmetrical phase, arriving at the result shown in equation (1.3.11).

1.3.8. Show that the perturbative equation which determines the renormalized mass parameter α_R in the symmetrical phase, as a function of α and λ, 1.3.11. ($\star$) For given α, λ and N, write a program to solve numerically the equation

$$\alpha_R = -2[\alpha + 3\lambda\sigma_0^2(\alpha_R)]$$

for $\alpha_R(\alpha, \lambda)$.

1.3.12. Verify, in dimensions from $d = 3$ to $d = 5$, whether or not there are any continuum limits in which $V_R = \langle\varphi\rangle$ is finite. If there are, identify them and verify what values the renormalized mass m_R can have in such limits. In particular, consider limits in which it is required that both V_R and m_R remain finite. Show that in $d = 3$ this requirement forces us to go to the Gaussian point in the limit, that in $d = 4$ we can satisfy it at any point along the critical line, and that in $d = 5$ it forces us to make λ tend to infinity, a limit which is also known as the sigma-model limit.

1.3.13. ($\star$) Calculate, using the first-order perturbative approximation scheme presented in the text, and making the choice $\alpha_0 = \alpha_R$, in each one of the two phases of the model, the quantity σ_4 given by $\sigma_4^4 = \langle\varphi^4\rangle$ in the symmetrical phase and by $\sigma_4^4 = \langle\varphi'^4\rangle$ in the broken-symmetrical phase. Show that, in either case, one obtains

$$\sigma_4^4 \simeq 3\sigma_0^4 - 6\lambda S_4,$$

where the sum S_4 is given in terms of the free propagator by

$$S_4 = \sum_{\vec{n}} g_0^4(\vec{0}, \vec{n}).$$

1.3.14. ($\star$) Evaluate, in each one of the dimensions $d = 3$ to $d = 5$, the behavior of the sum S_4 that appears in problem 1.3.13, using approximations by integrals or numerical methods. Determine the conditions under which S_4 goes to zero in the continuum limit, which causes the factorization rule $\sigma_4^4 = 3(\sigma_0^2)^2$ to hold, just as is the case of the free theory. Observe that this implies that, to first order, $\langle S_I\rangle = \langle S_I\rangle_0$ in the continuum limit, thus showing that the exchange of the complete distribution by the Gaussian distribution does not affect appreciably the singular character of the action. Observe also that this factorization shows that the local distribution of values of the field at a site tends to become Gaussian in the continuum limit, that is, the model becomes progressively more similar to the free theory.

1.3.15. Analyze the behavior in the continuum limit of the equations that determine α_R in the two phases of the model, discussed in problems 1.3.8 and 1.3.10, verifying that both lead to the same critical curve.

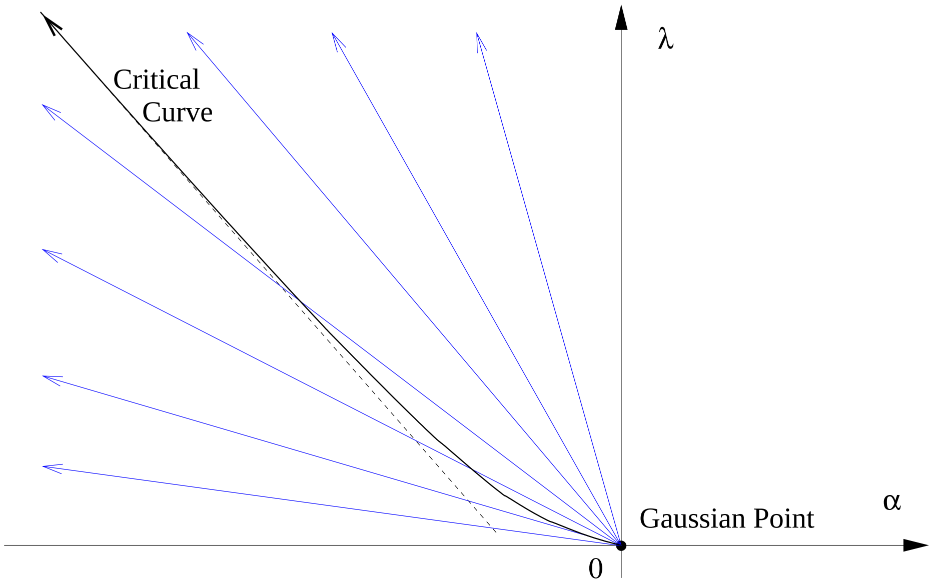

Figure 2.1.1: Limits leading from the polynomial models to the sigma models by means of negative-slope straight lines departing from the Gaussian point. Note that lines parallel to the α and λ coordinate axes are excluded.

be strictly negative and finite, in these limits we will have $\alpha \to -\infty$ and $\lambda \to \infty$ in such a way that $-\alpha/\lambda$ is a positive constant. Figure 2.1.1 illustrates the situation for lines starting from the Gaussian point. Observe that in these limits only the slope of the lines really matters. It makes no difference whether we use $\lambda = -C_1\alpha$ or $\lambda = -C_1\alpha + C_2$ for some finite constant C_2, because the finite additional term becomes irrelevant in the limit. In fact, we may take the limit over any curve that approaches asymptotically a negative-slope straight line. An important example of this is the critical curve of the $\lambda\varphi^4$ model, which in the $\lambda \to \infty$ limit approaches the critical point of the corresponding Ising model.

In order to establish this connection between the two models we start by recalling that the action of the $\lambda\varphi^4$ model without external sources, as it was defined in section 1.1, is given by

$$S[\varphi] = \frac{1}{2}\sum_{\ell}(\Delta_\ell\varphi)^2 + \frac{\alpha}{2}\sum_{s}\varphi^2(s) + \frac{\lambda}{4}\sum_{s}\varphi^4(s).$$

As we already discussed in section 1.1, we may now separate the action of the model in two parts, a kinetic part S_K containing only the derivative terms,

$$S_K[\varphi] = \frac{1}{2}\sum_{\ell}(\Delta_\ell\varphi)^2,$$

and a potential part S_V containing the polynomial terms,

$$S_V[\varphi] = \sum_{s} V(\varphi), \quad \text{where} \quad V(\varphi) = \frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4.$$

where $\sum[\psi = \pm 1]$ represents the sum over all the possibilities for combinations of the sign of the field at each site, over all sites. Writing this functional integral explicitly this once, for clarity, we have

$$\int [\mathbf{d}\varphi \; \varrho(\alpha, \lambda, \varphi)] \; e^{-S_K[\varphi]} = \prod_s \left\{ \sum_{\psi(s)=\pm 1} \int_0^\infty \mathrm{d}\varphi(s) \; \varrho[\alpha, \lambda, \varphi(s)] \right\} \; e^{-S_K[\psi\varphi]},$$

where $\psi(s)$ is a new variable holding the sign of the field at the site s, while φ assumes only positive values, due to the limits of integration adopted.

Let us now examine the behavior of $\varrho(\alpha, \lambda, \varphi)$ when $\lambda \to \infty$ and $\alpha = -\beta\lambda$, for some positive β. Note that, since lines making angles 0 and $\pi/2$ with the α coordinate axis are excluded, so are the corresponding values $\beta = \infty$ and $\beta = 0$. Executing the calculation of the integral in the denominator we obtain (problem 2.1.1), in terms of the parabolic cylinder functions $\mathbf{D}_\nu$,

$$\int_0^\infty \mathrm{d}\varphi \; e^{-V(\varphi)} = \frac{\sqrt{\pi}}{2} \left(\frac{\lambda}{2} \right)^{-\frac{1}{4}} e^{\frac{\alpha^2}{8\lambda}} \, \mathbf{D}_{-\frac{1}{2}} \left(\frac{\alpha}{\sqrt{2\lambda}} \right). \tag{2.1.2}$$

Using the asymptotic form of $\mathbf{D}_{-\frac{1}{2}}$ (problem 2.1.2) and substituting α in terms of λ, we may write the distribution $\varrho(\lambda, \varphi) = \varrho(\alpha = -\beta\lambda, \lambda, \varphi)$, for large values of λ, as

$$\varrho(\lambda, \varphi) \simeq \sqrt{\frac{\beta\lambda}{\pi}} \; e^{-\frac{\lambda}{4}(\varphi^2 - \beta)^2}. \tag{2.1.3}$$

We see here that indeed we cannot have either $\beta = 0$ or $\beta = \infty$, because in either case ϱ would vanish identically and hence would cease to be a normalizable statistical distribution. Given a finite and non-zero value of β we also see that, when $\lambda \to \infty$, ϱ tends to zero for all φ except for $\varphi = \sqrt{\beta}$, where it diverges as $\sqrt{\lambda}$. Hence, given a continuous and limited function $f(\varphi)$ and considering the normalization of ϱ, one can verify that, in the $\lambda \to \infty$ limit (problem 2.1.3),

$$\int_0^\infty \mathrm{d}\varphi \; f(\varphi) \; \varrho(\lambda, \varphi) = f\left(\sqrt{\beta}\right) \int_0^\infty \mathrm{d}\varphi \; \varrho(\lambda, \varphi) = f\left(\sqrt{\beta}\right). \tag{2.1.4}$$

In other words, the distribution $\varrho(\lambda, \varphi)$ tends to a Dirac delta function,

$$\lim_{\lambda \to \infty} \varrho(\lambda, \varphi) = \delta\left(\varphi - \sqrt{\beta}\right) = 2\sqrt{\beta} \; \delta(\varphi^2 - \beta).$$

The conclusion is that in this limit the expectation values of the polynomial model may be written as

$$\langle \mathcal{O} \rangle_N = \frac{\sum[\psi = \pm 1] \int [\mathbf{d}|\varphi| \; \delta(\varphi^2 - \beta)] \; \mathcal{O}[\psi\varphi] \; e^{-S_K[\psi\varphi]}}{\sum[\psi = \pm 1] \int [\mathbf{d}|\varphi| \; \delta(\varphi^2 - \beta)] \; e^{-S_K[\psi\varphi]}},$$

action without changing the observables. We are left, therefore, with the bilinear term, and we write the action as (problem 2.1.4)

$$S_K[\psi] = -\beta \sum_{\ell} \psi(\ell_-)\psi(\ell_+). \tag{2.1.5}$$

We have here an interaction between next neighbors involving the product of unit spins, exactly like in the Ising model. Hence we see that the infinite coupling limit of the $O(1)$ polynomial model is indeed the Ising model. Therefore, the expectation values of the polynomial model can be written as expectation values in this model, by means of a simple rescaling of the variable appearing within the observable,

$$\langle \mathcal{O}[\varphi] \rangle_N = \frac{\sum [\psi = \pm 1] \; \mathcal{O}[\sqrt{\beta}\psi] \; e^{\beta \sum_\ell \psi(\ell_-)\psi(\ell_+)}}{\sum [\psi = \pm 1] \; e^{\beta \sum_\ell \psi(\ell_-)\psi(\ell_+)}}.$$

Observe that, once the $\lambda \to \infty$ limit is taken in the way explained here, this relationship between the two classes of models is exact and involves no approximations of any kind.

It is important to discuss here the situation regarding the introduction of external sources into the model in this limit. The Ising model inherits from the polynomial model the introduction of external sources by means of a linear term in the action,

$$-\sum_s j(s)\varphi(s) = -\sqrt{\beta} \sum_s j(s)\psi(s).$$

It follows therefore that, apart from a rescaling of the sources by $\sqrt{\beta}$, the introduction of external sources is to be done in the usual way. In order to write this term in the form which is customary in statistical mechanics, we define the external sources $\eta(s)$ for the Ising model as $j(s) = \sqrt{\beta}\eta(s)$, so that the external-source term of the action acquires the form

$$-\sum_s j(s)\varphi(s) = -\beta \sum_s \eta(s)\psi(s).$$

At first sight it might seem natural to include the source term of the polynomial model in the potential part of the action, together with the α and λ terms, and then to rework the derivation of the large-coupling limit. However, this should not be done, for two reasons: first, the external source term does not change in the limit and has in fact no role to play in it; second, it is not simply a polynomial term in φ, because its coefficient j is not a constant like α or λ, but rather an arbitrary function of the sites. The external-source term should therefore be left in the action, together with the kinetic term S_K, and should not be included in the measure with the potential term S_V.

When the external-source term is treated in this way, the derivation of the large-coupling limit proceeds exactly as before, nothing changes in the derivation because no steps in it depend on other terms that the complete action may contain, besides both confirmed qualitatively by the computer simulations. In each dimension $d \geq 3$ we have therefore the same situation, a critical line in the parameter plane of the polynomial model, connecting at one end with the Gaussian point at the origin, and connecting at the other end with the critical point of the Ising model, over the arc at infinity.

The situation for $d = 2$ is very peculiar in the case of the $O(1)$ models and deserves to be mentioned here. In this case the polynomial model does not exist in the vicinity of the Gaussian point, except for the Gaussian point itself. There is, therefore, no critical line connecting the Gaussian point to the critical point of the two-dimensional Ising model, which does exist, however, as is well known. In the corresponding $\lambda\varphi^4$ model there is a convergent expansion near the Gaussian point [28], which shows that the observables are analytical functions of the parameters and that there is therefore no critical behavior, as our perturbative results indicate. However, it has also been shown that there is a phase transition in the polynomial model for sufficiently large λ [28], indicating that there should be a critical line starting somewhere within the critical diagram, away from the Gaussian point, and extending from there to the critical point of the Ising model over the arc at infinity. The details regarding this peculiar situation are currently unknown.

Going back to the cases $d \geq 3$, besides the indications that we saw here, the computer simulations also indicate that the Ising models have the same triviality behavior of the polynomial models, in the sense that λ_R vanishes in the continuum limit. In all these models it seems that the kinetic term of the action completely dominates the dynamics, and that the potential terms are not sufficient to change qualitatively the behavior dictated by the kinetic term. If even making the coupling parameter λ tend to infinity we cannot obtain truly interacting models, it becomes clear that a deeper structural change of the models is necessary. This is exactly what one does when one discusses gauge theories involving vector fields, in which the interactions are introduced precisely through the kinetic term. There still is, however, a rather long path to follow before we get to that.

Problems

2.1.1. Calculate the integral in the denominator of equation (2.1.1), obtaining the result given in equation (2.1.2); see for example [15].

2.1.2. Use the asymptotic form of the parabolic cylinder functions $\mathbf{D}_\nu$ in order to write the local distribution of the Ising models in the form given in equation (2.1.3); see for example [25].

2.1.3. Show that the equation (2.1.4) for the local distribution of the Ising models is valid in the $\lambda \to \infty$ limit. In order to do that, show that the exchange of $f(\varphi)$ for $f(\sqrt{\beta})$ in the left-hand side does not change the limit.

2.1.4. Expand the squares of the finite differences in the action S_K, use the condition of constraint and neglect field-independent constants, in order to write the action in the form given in equation (2.1.5).

originates from the usual kinetic term and which can be written in the form of a coupling term containing products of fields at neighboring sites. This is the type of separation of the action that will be of interest in this section.

In its original formulation, applied to a statistical-mechanic system that does not necessarily have to involve only interaction between next neighbors, the mean-filed method consists of the replacement of the interactions of a given site with all the others by an interaction between that given site and a background field that does not undergo statistical fluctuations. In a situation like this we would be typically dealing with the electromagnetic interactions between the charges located at a given site and all the other charges distributed across the crystalline lattice of a solid, whose effects on the site at issue are felt through the electromagnetic fields that each charge gives rise to. What one does in this type of approximation is to replace the fluctuating electromagnetic field generated by the set of all the other charges by a *mean field* that does not fluctuate. This mean field is defined at each site, representing the average collective effect of all the other sites over the charges located at that point. Naturally, in order for this scheme to be useful it is necessary that we be able to calculate the mean field in terms of the charges distributed across the crystalline lattice. The calculation of this mean field clearly involves two aspects: first, there is a sum over the volume of the lattice, in order to take into account the effect of all the other sites, which are at various distances from the site at issue; second, there is a temporal average in order to eliminate the statistical fluctuations of the field, which can be exchanged for an ensemble average, according to the usual procedure of statistical mechanics.

A realization of this idea in a model defined on the lattice, like the ones we want to deal with here, must take into account only the interactions of a given site with its next neighbors. On a cubical lattice, like the ones we have been using, we can imagine that we define at each site an average field that represents the effect of the $2d$ next neighbors of the site. Of course in this case we are not dealing with electromagnetic interactions but with the self-interactions of the scalar field. Since this mean field does not fluctuate, from the point of view of quantum field theory it is not dynamical, and hence it should be treated like an external field $j_{\rm MF}$, that couples to the field φ of the site at issue by means of an action term of the type $j_{\rm MF}\varphi$. Naturally, in order for this scheme to be useful it is necessary to adopt some criterion to allow the calculation of the value of $j_{\rm MF}$ in terms of the collection of fields, now uncoupled, that exist at the neighboring sites. The usual mean-field method on a lattice of *arbitrary size* consists of the replacement, in the action, of the interaction terms of each site with its next neighbors by an interaction term of the site with a non-dynamical external field, whose value is equal to the sum of the ensemble averages of the dynamical field at the neighboring sites,

$$\sum_{\vec{n}_\ell}^{2d} \varphi(\vec{n})\varphi(\vec{n}_\ell) \rightarrow \ \varphi(\vec{n}) \sum_{\vec{n}_\ell}^{2d} \langle \varphi(\vec{n}_\ell) \rangle = 2d\ \varphi(\vec{n}) \langle \varphi \rangle,$$

where the sum runs over the $2d$ links ℓ that connect the site at the position $\vec{n}$ to its next neighbors at the positions $\vec{n}_\ell$, and where $\langle\varphi\rangle$ is the average value of the field

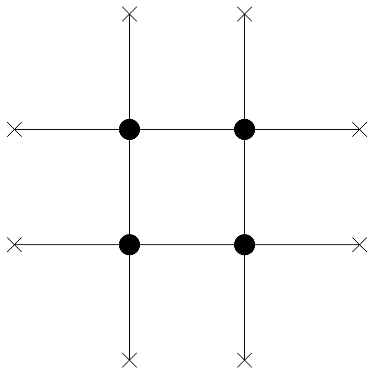

Figure 2.2.2: Lattice with $N = 2$ and fixed boundary conditions, which constitutes the first extension of the mean-field method.

each other, any result obtained for one of them is valid for all the others. Therefore, in the traditional method it is sufficient to keep a single active site, without any changes in the results, which establishes the equivalence of the two interpretations. In any of the two interpretations the dynamical fields interact only with a constant background field, independently of how we interpret this non-dynamical field, either as an external field or as the field of the neighboring sites.

The resulting structure, in this second interpretation, is a lattice with $N = 1$ and a border where the field is kept fixed at its average value, just like the lattices with fixed boundary conditions that we have seen before in [32], [33] and [34], which are represented as in the diagram in figure 2.2.1, including the active site and the border sites. In any of the two interpretations the mathematical consequence of the approximation is that the infinite-dimensional functional integral on the lattice is replaced by a one-dimensional integral over the dynamical field at the only remaining active site,

$$\frac{\displaystyle\int [\mathbf{d}\varphi]\ \mathcal{O}[\varphi]\ e^{-S[\varphi]}}{\displaystyle\int [\mathbf{d}\varphi]\ e^{-S[\varphi]}} \quad\longrightarrow\quad \frac{\displaystyle\int \mathrm{d}\varphi\ \mathcal{O}(\varphi)\ e^{-S_{\mathrm{MF}}(\varphi)}}{\displaystyle\int \mathrm{d}\varphi\ e^{-S_{\mathrm{MF}}(\varphi)}}, \tag{2.2.1}$$

where S_{MF} is the mean-field approximation for the lattice action and $\mathcal{O}(\varphi)$ is some observable that depends only on the field at the single remaining active site. It is usually possible to calculate analytically the resulting integral, which establishes the usefulness of the method in its conventional form.

This second interpretation of the method suggests at once the definition of a series of approximations of the continuous system, of which the usual mean-field method is the first. Just consider lattices in which more than one site is left active, within a central cluster, while the fields at the borders are kept fixed. For example, we may consider a sequence $N = 1, 2, 3, \ldots$ of cubical lattices, such that the second approximation, with $N = 2$, is given by the lattice illustrated in figure 2.2.2, with 2^d sites, all in direct contact with the border.

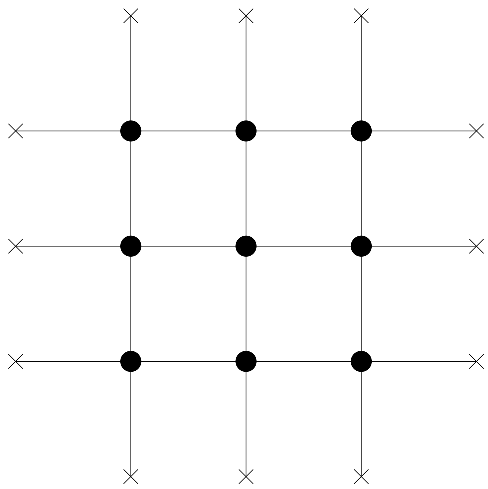

Figure 2.2.4: Lattice with $N = 3$ and fixed boundary conditions, which constitutes the second extension of the mean-field method.

a single site. These calculations may always be performed by freezing the fields at the sites of the border at some arbitrary value, thus leaving a single active site. One calculates then the average value of the field at the active site, by means of the integral that appear in equation (2.2.1), using $\mathcal{O}(\varphi) = \varphi$. Having done this, one compares the result obtained, which will depend on the value that was chosen for the fields at the border, with that value. Of course in general they will be different and the self-consistency problem is to find the value to be used at the border that reproduces exactly the same value for the average value of the dynamical field at the active site. The determination of the value of the mean field by a self-consistent procedure like this was first introduced in the case of the constant coupling method of statistical mechanics.

In an analytical calculation one can simply impose this condition a posteriori, so that it results in an algebraic equation for the average value of the field. Once this equation is solved and the average value of the field is found in terms of the parameters of the model, mean-field approximations for other quantities may also be obtained. The same self-consistency condition can also be imposed in the context of a Monte Carlo simulation in which the filed at the border is kept at a constant value, without fluctuations, either for $N = 1$ or for larger values of N. In this numerical approach a negative feedback mechanism can be used to slowly adjust the value of the field at the border so that it and the average value of the field measured in the interior of the lattice converge to a common limiting value. We denominate this special type of fixed boundary conditions *self-consistent boundary conditions.*

The stochastic simulation with $N = 1$ is equivalent to a Monte Carlo calculation of the integral that appears in equation (2.2.1). In this case the feedback mechanism can be implemented is a very simple way. One puts at the border fields a tentative value and lets the dynamical field fluctuate. One then measures the average value of the fluctuating field at the active site. If this average value differs from the tentative to be paid for the sharpness of the transitions obtained in this way. The feedback mechanism can consume a large amount of computer resources if we want really very sharp results, specially in simulations that already suffer from the notorious critical slowing down problems near the critical region. Fortunately, there exist currently algorithms that avoid completely this kind of problem for the scalar models that we are examining here.

Observe that we have here a set of systems with a *finite* number of degrees of freedom that still display *complete* critical behavior. How to reconcile this fact with the previously mentioned fact that there is no true critical behavior in systems with a finite number of degrees of freedom? What happens is that the relevant results contained in [30] are not relevant for the self-consistent systems, because they assume that one is discussing systems with a finite number of degrees of freedom and no external couplings, which implies that one must use periodical boundary conditions in order to avoid the border. In our fixed-boundary systems there is an additional element, which is precisely the interaction with the border sites, for which there is a self-consistent condition. In a heuristic and intuitive way, we may think of these systems as finite systems that have, however, a "window to infinity". The self-consistent boundary conditions act in fact as a kind of semi-transparent window opening onto an infinite outer lattice that surrounds our finite lattice, letting in some information about the infinite lattice of which our finite lattice is a cutout. It would be possible, in fact, to consider other types of boundary fixed conditions, more complex, sophisticated and transparent than the one we are considering here. For example, instead of keeping the boundary fields completely fixed at their mean values, we might consider letting them fluctuate around the mean value in some controllable way. In the future we may discuss in more detail a proposal along these general lines. But before that we must illustrate the method by means of some specific calculations with self-consistent boundary conditions.

Our extension of the method to lattices of arbitrary size provides us with an explanation of why the usual mean-field method fails completely for most models in dimensions $d = 1$ and $d = 2$. For this purpose it is necessary to remember that these are models that do not display phase transition in the $N \to \infty$ limit. In fact, there are theorems [31] that show that models with only next-neighbor couplings cannot have ordered phases, with oriented fields, in the $N \to \infty$ limit, in dimension $d = 1$, for any symmetry groups they may be invariant by. For dimension $d = 2$ the same is true for models invariant by continuous symmetry groups, such as $SO(\mathfrak{N})$ for $\mathfrak{N} > 1$. For discrete symmetry groups the existence of oriented phases is possible in $d = 2$, as shown by the Ising model, with the discrete symmetry group $O(1) = \mathbb{Z}_2$. There are no theorems like these for $d \geq 3$, cases in which the models usually display well-defined phase transitions in the $N \to \infty$ limit.

Let us see how these theorems are realized for the models defined on finite lattices of increasing size. The situation is similar in the cases of periodical boundary conditions and of fixed boundary conditions, but it is easier to give the explanation in the case of fixed boundary conditions. In this case one verifies that the models always display a phase transition on finite lattices, for any model and any space-

Getting back to our explanation of the realization of the theorems for the models defined on finite periodical lattices, in this case the system is always in a single broken-symmetrical phase, and taking due care with the drift of the zero-mode one can measure v_R as a function of β. The function $v_R(\beta)$ turns out to be a continuous and differentiable function, which is never zero, being typically small in the range of the parameters of the model where the symmetrical phase will appear in the $N \rightarrow \infty$ limit, and typically large in the complementary range. What happens in the $N \rightarrow \infty$ limit, when there is a phase transition, is that the curve $v_R(\beta)$ gradually changes and thus approaches a continuous but non-differentiable curve in the limit. The point where the curve becomes non-differentiable is the critical point and, in this case, it appears at finite values of the parameters. In the cases in which the system does not display a phase transition in the limit the curve not only changes shape, but also moves to arbitrarily large values of beta, so that once more all that remains in the limit is a symmetrical, non-oriented phase.

Finally, observe that we are *not* stating here that the continuum limits of the periodical systems and of the self-consistent systems are completely identical, because some of the observables may depend on the boundary conditions adopted, which are different in either case. These are two different classes of continuum limits, whose properties can be somewhat different. Usually the more basic observables of the system, such as the values of the parameters at the critical points, the renormalized masses and the expectation values of the fields, will not depend significantly on the boundary conditions, but observables with a subtler type of behavior, such as the critical exponents, may very well depend strongly on the boundary conditions. In fact, one can show that the critical exponents differ significantly in the two cases that we are discussing here. Besides the fact that fixing the field at the border does not make much physical sense in the context of quantum field theory, this is another reason why it is important to think about generalizations of the self-consistent boundary conditions described here, for example in order to allow the border fields to fluctuate, as was mentioned before in this section. This subject may be discussed in more detail in a future volume of this series.

2.3 Some Mean Field Results

We will perform in this section a few analytical calculations, in the usual mean-field approximation with $N = 1$, of some observables of the scalar field models, using the ideas presented in section 2.2. The approximations with $N \geq 2$, involving the extension of the method which was also discussed in that section, are usually too complex for an analytical treatment and are better characterized, therefore, as material for performing stochastic simulations, which will be discussed in a future volume. The calculations we will present here can be understood as exact analytical solutions for the lattice systems with a single active site. We will start with the Ising models, in which the calculations are somewhat simpler.

A quantity of particular interest in the Ising models is the magnetization, which

$$\nu_R = 2d\,\beta\nu_R.$$

A factor of ν_R cancels off and we are left with

$$1 = 2d\,\beta.$$

After the cancellation of the trivial solution $\nu_R = 0$ we may impose the *condition* $\nu_R = 0$ in order to find the region where *only* this zero solution is possible. Observe that the equation above does not depend on ν_R. Had we kept the higher-order terms of the expansion, the condition $\nu_R = 0$ would eliminate them at this point, leaving only the equation above. Hence we obtain the mean-field approximation for the critical points

$$\beta_c(d) = \frac{1}{2d}. \tag{2.3.2}$$

This technique of expansion around the critical point is useful in cases in which the mean-field results cannot be obtained explicitly, or in which the exact solutions of the resulting equations are difficult to determine analytically.

In four dimensions this $N = 1$ result is in fairly good agreement with the numerical results obtained for larger values of N. The agreement is poorer in three dimensions, very poor is two dimensions, and the result fails completely in one dimension, in which there is no critical point at all. Observe that the success or lack thereof on the $N = 1$ approximation in reproducing well the results for large values of N is not a diagnostic about the validity of the method itself, bur rather an indication of whether or not the results for increasing values of N accumulate around some finite value of β. For $d = 1$ the values $\beta_c(N)$ for each lattice size diverge to $+\infty$ when $N \to \infty$, so that no finite lattice can represent well the limit. In larger dimensions the critical points $\beta_c(N)$ of the finite lattices converge to finite values in the $N \to \infty$ limit, so that in these cases the approximation of the limiting result by the results for finite lattices makes some sense, being better or worse depending on the speed of convergence.

Table 2.3.1 contains results for the critical points of the Ising models for dimensions from $d = 1$ up to $d = 4$. For $d = 4$ the result was obtained by means of extrapolations to the $N \to \infty$ limit of the results of stochastic simulations with periodical boundary conditions on lattices with N from 4 to 10. For $d = 3$ we quote the most precise result that we know about [36]. For $d = 2$ we quote the well-known result for the two-dimensional Ising model [37]. Since the estimate from periodical simulations is obtained from the continuous and differentiable curves of the magnetization as a function of β, the error bars indicated are only approximate estimates. The entry ∞ indicates the case in which the theorems about long-distance order [31] imply that there is no phase transition, which is the same in which $\beta_c(N) \to \infty$ when $N \to \infty$ in the simulations with self-consistent fixed boundary conditions. The results of equation (2.3.2) are also included, for comparison. As one can see, in $d = 4$

$$v_R = \frac{\displaystyle\int_{-\infty}^{\infty} \mathrm{d}\varphi\ \varphi\ e^{2d\,v_R\varphi}\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle\int_{-\infty}^{\infty} \mathrm{d}\varphi\ e^{2d\,v_R\varphi}\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}.$$

If we use here our "radial-angular" decomposition of φ into its absolute value and its sign, $\varphi = \psi|\varphi|$, with $\psi = \pm 1$, as was done in section 2.1, we may write this as

$$v_R = \frac{\displaystyle\sum[\psi=\pm 1]\int_{0}^{\infty} \mathrm{d}\varphi\ \psi\varphi\ e^{2d\,v_R\psi\varphi}\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle\sum[\psi=\pm 1]\int_{0}^{\infty} \mathrm{d}\varphi\ e^{2d\,v_R\psi\varphi}\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}.$$

We may now execute the sum over the single active site and hence get

$$v_R = \frac{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ \varphi\ \sinh(2d\,v_R\varphi)\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ \cosh(2d\,v_R\varphi)\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}. \tag{2.3.5}$$

In this case it is also possible to perform the integrations analytically in terms of special functions (problem 2.3.5) but, close to the critical curve, it suffices to calculate the right-hand side of this equation for small values of v_R. Therefore, we expand equation (2.3.5) to first order in v_R and obtain

$$v_R = 2d\,v_R\,\frac{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ \varphi^2\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}.$$

Just like before, a factor of v_R cancels out and we obtain as our mean-field result an equation giving implicitly the critical curve $\lambda = f(\alpha)$,

$$\frac{1}{2d} = \frac{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ \varphi^2\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle\int_{0}^{\infty} \mathrm{d}\varphi\ e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}. \tag{2.3.6}$$

Once more the integrations can be done, this time in terms of the parabolic cylinder functions $\mathbf{D}_\nu$ (problem 2.3.6), but this does not help us to solve this equation in order to write the equation of the critical curve in explicit form. The fact that the left-hand side of this equation is equal to the mean-field value of the critical point β_c of the corresponding Ising model is not an accident, it is clearly related to the fact that the $\lambda \to \infty$ limits of the $\lambda\varphi^4$ models converge to the Ising models (problem 2.3.7).

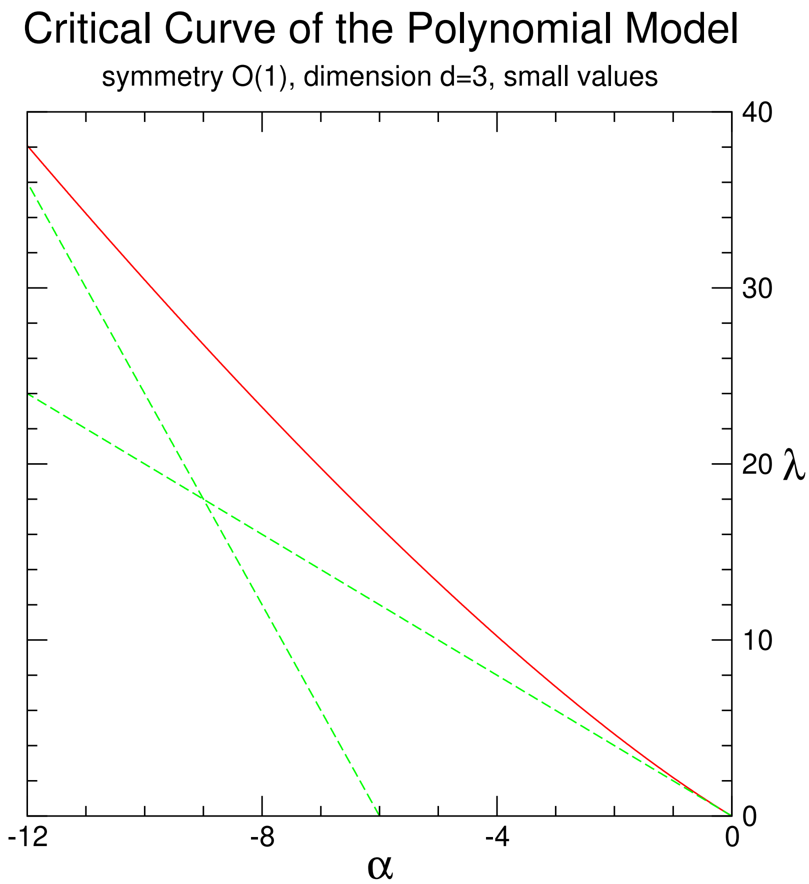

Figure 2.3.1: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 3$, for the smaller values of the parameters.

We are led to conclude that the mean-field approximation confirms the basic properties of the critical curves that we discussed before in sections 1.1 and 1.3. One can improve on the comparison with the perturbative results by calculating the mean-field approximation for the quantity $\sigma^2 = \langle\varphi^2\rangle$ over the critical curve, where $v_R = \langle\varphi\rangle$ is zero. The mean-field approximation for σ^2 is given by

$$\sigma^2_{\rm MF} = \frac{\sum[\psi = \pm 1] \int_0^\infty {\rm d}\varphi\; \varphi^2\; e^{2d\, v_R \psi\varphi}\; e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}}{\sum[\psi = \pm 1] \int_0^\infty {\rm d}\varphi\; e^{2d\, v_R \psi\varphi}\; e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}},$$

and if we execute the sum over the signs at the single active site we get

$$\sigma^2_{\rm MF} = \frac{\int_0^\infty {\rm d}\varphi\; \varphi^2\;\; \cosh(2d\, v_R\varphi)\; e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}}{\int_0^\infty {\rm d}\varphi\;\; \cosh(2d\, v_R\varphi)\; e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}}.$$

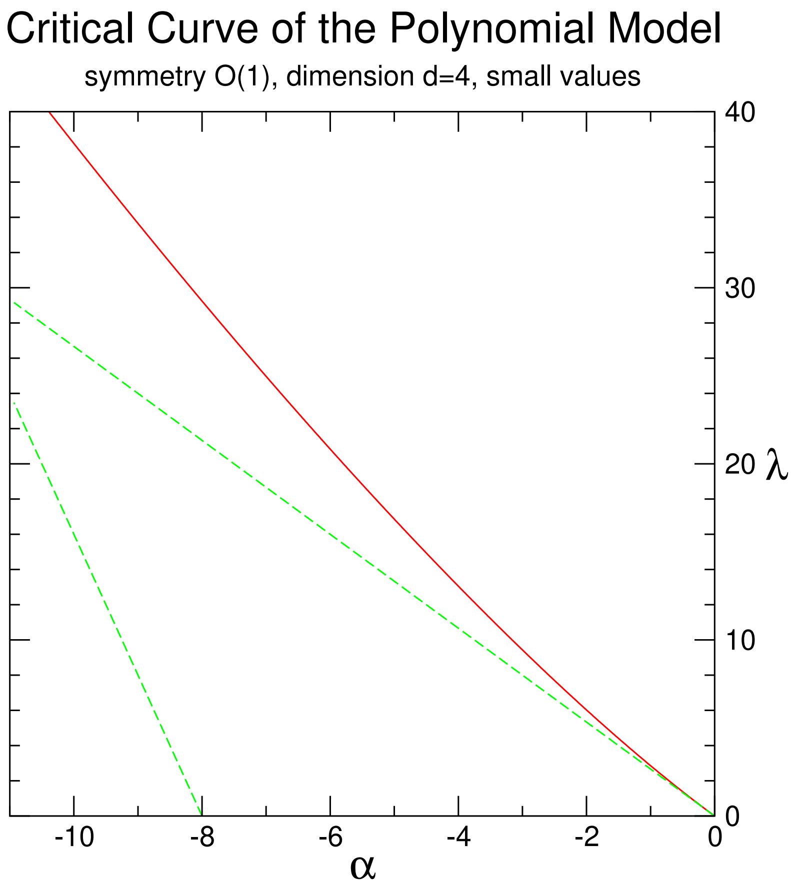

Figure 2.3.3: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 4$, for the smaller values of the parameters.

If we now recall the perturbative result for the equation of the tangent line to the critical curve at the Gaussian point, which can be obtained from equation (1.3.7), and can be written as

$$\lambda(\alpha) = -\frac{1}{\sigma_0^2}\,\frac{\alpha}{3},$$

we see that the two results are identical except for the exchange of σ_0^2 and $\sigma_{\rm MF}^2$. Therefore, the two results for the slope of the critical curve at the Gaussian point coincide within the expected level of precision of the mean-field approximation, since the mean-field result is just an approximation, while the perturbative result of the Gaussian approximation is, for this particular quantity, presumably exact.

We can also calculate the critical exponent of v_R in the case of the polynomial models. In order to do this it is necessary to expand equation (2.3.5) up to a higher order in v_R, so as to allow us to write the differential of $v_R^2(\alpha, \lambda)$ as a function of $\mathrm{d}\alpha$ e $\mathrm{d}\lambda$. This is the work proposed in problem 2.3.10, and it can be shown that it is possible to write the differential of v_R^2 as points of the Ising models is due to the fact that we are using the $N = 1$ mean-field approximation in the right-hand column. Another part of the discrepancies may be due to the fact that we are comparing results for systems with very different boundary conditions, since the left-hand column refers to systems with periodical boundary conditions. A situation like this is probably more likely to be realized for the finer, more delicate observables, such as the critical exponents and correlation functions, than for the more basic objects such as the critical points.

The choice of boundary conditions is an important subject within the structure of the quantum theory, just as it is in the classical theory. It should be noted that there are other ways, besides the one that we examined here, to implement fixed boundary conditions, some of which may be physically more natural and compelling from the point of view of quantum field theory. One example of this kind of thing was proposed in the last chapter of a previous volume of this series of books [35]. Other ideas related to that one may be discussed in future volumes.

Problems

2.3.1. Verify that $\nu_R = 0$ always satisfies equation (2.3.1). In order to discover whether there are values of β for which it is possible to have a solution $\nu_R \neq 0$, use the series expansions of the hyperbolic functions to obtain the relation

$$\sum_{k=0}^{\infty} \frac{(2d\,\beta\nu_R)^{2k+1}}{(2k+1)!}\,[(2k+1) - 2d\,\beta] = 0.$$

Observe that all the factors in each term of this series are always positive except for the last one on the right. Use this fact to determine the interval of values of β for which it is possible to have $2d\,\beta\nu_R$ and therefore ν_R different from zero as a solution of this equation, and thus determine the mean-field critical points β_c of the Ising models.

2.3.2. Write a program to solve numerically equation (2.3.1) and plot graphs of ν_R as a function of β for some values of $d \geq 3$. Consider using the exponential bisection method and consider the results of problem 2.3.1. Verify how close to the exact result in the broken-symmetrical phase is the ansatz

$$\nu_R = \sqrt{\frac{3(\beta - \beta_c)}{3\beta - 2\beta_c}},$$

which has the correct behavior for β close to β_c and that tends to 1 for $\beta \to \infty$.

2.3.3. Expand the right-hand side of equation (2.3.1) up to the third order on the variable $A = 2d\,\beta\nu_R$ and show that the terms with even powers of the variable vanish, thus obtaining the result

$$\nu_R = A - \frac{A^3}{3}.$$

The integrals may now be written in terms of the parabolic cylinder functions $\mathbf{D}_\nu$ [15], so use them to write the final form of the equation

$$\frac{\sqrt{2\lambda}}{2d} = \frac{\sum_{k=0}^{\infty} \frac{(d\,v)^{2k}}{k!} \left(\frac{\lambda}{2}\right)^{-k/2} \mathbf{D}_{-\left(k+\frac{3}{2}\right)} \left(\frac{2d+\alpha}{\sqrt{2\lambda}}\right)}{\sum_{k=0}^{\infty} \frac{(d\,v)^{2k}}{k!} \left(\frac{\lambda}{2}\right)^{-k/2} \mathbf{D}_{-\left(k+\frac{1}{2}\right)} \left(\frac{2d+\alpha}{\sqrt{2\lambda}}\right)}.$$

2.3.6. Use the same techniques of problem 2.3.5 in order to calculate the integrals that appear in equation (2.3.6), and thus obtain

$$\frac{\sqrt{2\lambda}}{2d} = \frac{\mathbf{D}_{-\left(\frac{3}{2}\right)} \left(\frac{2d+\alpha}{\sqrt{2\lambda}}\right)}{\mathbf{D}_{-\left(\frac{1}{2}\right)} \left(\frac{2d+\alpha}{\sqrt{2\lambda}}\right)}. \tag{2.3.8}$$

Observe that this result is the same of problem 2.3.5 if we truncate the two series that appear there, leaving only their first terms, those with $k = 0$.

2.3.7. Take explicitly the Ising-model limit of equation (2.3.6), making $\lambda \to \infty$ and $\alpha \to -\infty$ with $\alpha = -\beta\lambda$, and show that it reduces to the known result for the value of the critical point of the Ising model.

2.3.8. Obtain the asymptotic form of the critical curve, for large values of $-\alpha$ and λ, using in equation (2.3.8) the asymptotic expansion of the parabolic cylinder functions $\mathbf{D}_\nu$ [25], to the lowest non-vanishing order, thus obtaining the result

$$\lambda\left(\alpha\right) = -\frac{1}{\beta_c}\left(\alpha + 2d\right).$$

2.3.9. In order to obtain the behavior of the critical curve for small values of $-\alpha$ and λ, write equation (2.3.6) in the form

$$\int_0^\infty \mathrm{d}\varphi\, \varphi^2\, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]} = \frac{1}{2d}\int_0^\infty \mathrm{d}\varphi\, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}$$

and differentiate implicitly in terms of $\mathrm{d}\alpha$ e $\mathrm{d}\lambda$, applying the resulting coefficients at the Gaussian point. The integrals that appear in these coefficients are expressible in terms of the Γ function [14], so use them to obtain

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\alpha} = -\frac{2d}{3}.$$

Integrate this first-order differential equation for $\lambda(\alpha)$ with the boundary condition $\lambda(0) = 0$, thus obtaining the final result

$$\lambda(\alpha) = -\frac{1}{\beta_c}\,\frac{\alpha}{3}.$$

$$
\begin{aligned}
C_\alpha &= \frac{\displaystyle \frac{1}{3d}\int_0^\infty \mathrm{d}\varphi\,\varphi^2\left(\frac{1}{2d}-\varphi^2\right)e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle \int_0^\infty \mathrm{d}\varphi\,\varphi^2\left(1-\frac{2d\,\varphi^2}{3}\right)e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}},\\
C_\lambda &= \frac{\displaystyle \frac{1}{6d}\int_0^\infty \mathrm{d}\varphi\,\varphi^4\left(\frac{1}{2d}-\varphi^2\right)e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}{\displaystyle \int_0^\infty \mathrm{d}\varphi\,\varphi^2\left(1-\frac{2d\,\varphi^2}{3}\right)e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}}.
\end{aligned}
$$

Stating from these expressions verify that C_α and C_λ are finite and non-vanishing numbers in the vicinity of the critical curve. This suffices to show that the critical exponent of v_R is once again $\gamma = 1/2$ when we approach the critical curve in the parameter plane of the model, from any direction within the broken-symmetrical phase. Verify also that these two coefficients are negative, showing that the gradient of v_R^2 is oriented in the expected direction.

$j = 0$. In addition to this, the connected Green functions $g_{(c,j)1,\ldots,n}$ in the presence of external sources can be obtained by means of multiple functional differentiations with respect to j of the functional $W[j] = \ln(Z[j])$. It is in these connected functions that the true correlations of the theory are encoded, including in the case in which we have non-vanishing external sources. The connected two-point function was calculated explicitly in the equation in reference [42] and the calculation of the corresponding three-point function was proposed in the problem in reference [43]. In that problem, starting from the definition of the three-point function in terms of $W[j]$,

$$\frac{\partial^3 W[j]}{\partial j_1 \partial j_2 \partial j_3} = g_{(c,j)1,2,3},$$

one shows that this functions can be written in terms of the complete three-point function as

$$g_{(c,j)1,2,3} = g_{(j)1,2,3} - g_{(c,j)1,2}\,\varphi_{(c)3} - g_{(c,j)2,3}\,\varphi_{(c)1} - g_{(c,j)3,1}\,\varphi_{(c)2} - \varphi_{(c)1}\,\varphi_{(c)2}\,\varphi_{(c)3},$$

which corresponds to the subtraction from the complete function of all possible factorizations in terms of connected functions with a smaller number of points. Note that we returned here to the notation of the section in reference [39], denoting the expected value of the field by $\varphi_{(c)}$ instead of v_R and the dependencies on the positions $\vec{n}_i$ of the sites by indices, $\varphi_{(c)}(\vec{n}_1) = \varphi_{(c)1}$. Observe that in any circumstances in which $\varphi_{(c)} = 0$, which naturally implies that $j = 0$, we have for this connected function $g_{(c)1,2,3} = g_{1,2,3}$, which makes the definition of the renormalized coupling constant considerably simpler in models based on three-point interactions, as is the case, for example, for electrodynamics.

However, the polynomial models of scalar fields, such as the $\lambda\varphi^4$ model that we are studying here, are based on interactions involving four or more points, so that we must go at least up to the four-point function in order to be able to define the renormalized coupling constant. It is necessary, therefore, to take a fourth and last derivative of $W[j]$ (problem 3.1.1), which results, after a long but straightforward algebraic calculation, in a relation between the connected four-point function and the corresponding complete function,

$$\begin{aligned}
g_{(c,j)1,2,3,4} &= g_{(j)1,2,3,4} \\
&- \left[g_{(c,j)2,3,4}\,\varphi_{(c)1} + g_{(c,j)1,3,4}\,\varphi_{(c)2} + g_{(c,j)1,2,4}\,\varphi_{(c)3} + g_{(c,j)1,2,3}\,\varphi_{(c)4}\right] \\
&- \left[g_{(c,j)1,2}\,g_{(c,j)3,4} + g_{(c,j)1,3}\,g_{(c,j)2,4} + g_{(c,j)1,4}\,g_{(c,j)2,3}\right] \\
&- \left[g_{(c,j)1,2}\,\varphi_{(c)3}\,\varphi_{(c)4} + g_{(c,j)1,3}\,\varphi_{(c)2}\,\varphi_{(c)4} + g_{(c,j)1,4}\,\varphi_{(c)2}\,\varphi_{(c)3}\right. \\
&\quad \left. + g_{(c,j)2,3}\,\varphi_{(c)1}\,\varphi_{(c)4} + g_{(c,j)2,4}\,\varphi_{(c)1}\,\varphi_{(c)3} + g_{(c,j)3,4}\,\varphi_{(c)1}\,\varphi_{(c)2}\right] \\
&- \varphi_{(c)1}\,\varphi_{(c)2}\,\varphi_{(c)3}\,\varphi_{(c)4}.
\end{aligned} \qquad (3.1.1)$$

As was the case for the three-point function, this time we also obtain, as one can see, the subtraction from the complete function of all the possible factorizations in terms of connected functions with a smaller number of points. The expression

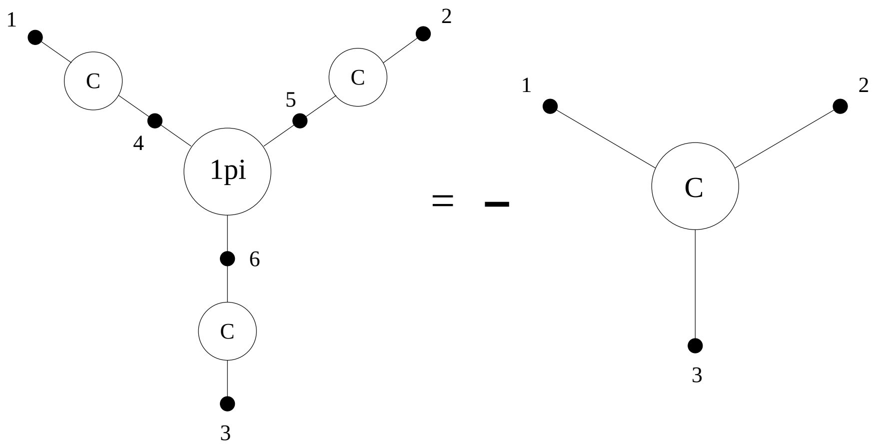

Figure 3.1.1: Diagrammatic representation of the equation for the three-point function.

This shows that the triple functional derivative of $\Gamma[\varphi_{(c)}]$ is related to the three-point connected function. We can see that the connected three-point function is obtained from the triple functional derivative of $\Gamma[\varphi_{(c)}]$ by means of a type of triple transformation in which the transformation function is the propagator. In a diagrammatic language, we can say that the triple functional derivative is a *vertex* to which are connected three *external legs* representing the three propagators that act as transformation functions, this whole set of elements being equivalent to the connected three-point function. This is therefore a new type of decomposition, a way of decomposing the connected three-point function into simpler, more fundamental parts. These simpler parts are denominated *one-particle irreducible* or "1pi" functions. The corresponding diagram is illustrated in figure 3.1.1. If we differentiate this expression a fourth and last time, using the same techniques and ideas, we obtain, after long algebraic passages (problem 3.1.4), the relation

$$\begin{aligned}
&\sum_{5,6,7,8} g_{(c,j)1,5}\ g_{(c,j)2,6}\ g_{(c,j)3,7}\ g_{(c,j)4,8} \frac{\delta^4\Gamma[\varphi_{(c)}]}{\delta\varphi_{(c)5}\delta\varphi_{(c)6}\delta\varphi_{(c)7}\delta\varphi_{(c)8}} \\
&= -g_{(c,j)1,2,3,4} \quad + \sum_{5,6} g_{(c,j)1,2,5} \frac{\delta^2\Gamma[\varphi_{(c)}]}{\delta\varphi_{(c)5}\delta\varphi_{(c)6}} g_{(c,j)3,4,6} \\
&\qquad\qquad\qquad\quad + \sum_{5,6} g_{(c,j)1,3,5} \frac{\delta^2\Gamma[\varphi_{(c)}]}{\delta\varphi_{(c)5}\delta\varphi_{(c)6}} g_{(c,j)2,4,6} \\
&\qquad\qquad\qquad\quad + \sum_{5,6} g_{(c,j)1,4,5} \frac{\delta^2\Gamma[\varphi_{(c)}]}{\delta\varphi_{(c)5}\delta\varphi_{(c)6}} g_{(c,j)2,3,6}.
\end{aligned} \tag{3.1.3}$$

This relation has an interesting diagrammatic representation, which we illustrate in figure 3.1.2, where the symbol $\Box_{(c)}$ represents the inverse of the propagator, as defined in the equation in reference [48]. The last three parts of this diagram correspond to all the possible ways to build a four-point process with the connected three-point functions and at most a single internal $\Box_{(c)}$. We see therefore that the left side of the equation corresponds to the difference between the connected four-point function and these constructions. It is due to this that the functions generated directly by $\Gamma[\varphi_{(c)}]$ are called "one-particle irreducible" or "1pi", that is, they are irreducible functions in the sense that they cannot be separated into other functions by the elimination of an internal line of the corresponding the diagram. In somewhat more physical terms, the 1pi four-point function is the part of the four-point interaction that cannot be built out of two three-point interactions. The dimensionless versions of these 1pi functions will be denoted by γ, with position variables as arguments,

$$\frac{\mathfrak{d}^n\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)1}\dots\mathfrak{d}\varphi_{(c)n}} = \gamma(\vec{n}_1,\dots,\vec{n}_n) = \gamma_{1,\dots,n}.$$

These functions are also referred to as *truncated*, meaning that the propagators corresponding to the external legs are absent from the functions $\gamma_{1,\dots,n}$. It is for this reason that they are also called "vertices", meaning that they represent only the central vertices that connect together the external legs of the diagrams, without the inclusion of the external legs themselves. All this diagrammatic nomenclature is mentioned here only to make contact with what one sees in the more traditional ways to approach the theory, since it will not have much importance for our approach in these notes.

For models with $j = 0$ and in which there is symmetry by the reflection of the fields, not only we have $\varphi_{(c)} = 0$ but the symmetry also implies that all the functions with an odd number of points vanish. In particular, $g_{(c)1,2,3} = 0$ and there are, therefore, no three-point interactions. In this case equation (3.1.3) can be simplified to

$$\sum_{5,6,7,8} g_{(c)1,5}\, g_{(c)2,6}\, g_{(c)3,7}\, g_{(c)4,8} \frac{\mathfrak{d}^4\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)5}\mathfrak{d}\varphi_{(c)6}\mathfrak{d}\varphi_{(c)7}\mathfrak{d}\varphi_{(c)8}} = -g_{(c)1,2,3,4}. \tag{3.1.4}$$

It is clear that the four-fold functional derivative has the effect of extracting from the quartic term of $\Gamma[\varphi_{(c)}]$ its coefficient, which will be proportional to the renormalized coupling constant. Therefore, the last step we must take in this sequence of calculations it to isolate $\gamma_{5,6,7,8}$ in the relation above between this 1pi functions and the connected correlation function $g_{(c)1,2,3,4}$, which we already know how to write directly in terms of expectation values of products of the fields. In order to do this we will rewrite this equation in momentum space, performing Fourier transforms on the four variables $\vec{n}_1,\dots,\vec{n}_4$ which are not added over. Taking the four-fold Fourier transform of equation (3.1.4) and using the discrete translation invariance of the lattice (problem 3.1.5), we obtain

$$\widetilde{g}_{(c)1}\,\widetilde{g}_{(c)2}\,\widetilde{g}_{(c)3}\,\widetilde{g}_{(c)4}\,N^{4d}\,\widetilde{\gamma}_{1,2,3,4} = -\widetilde{g}_{(c)1,2,3,4},$$

where we are denoting the momentum coordinates by indices and where $\widetilde{g}_{(c)1} = \widetilde{g}_{(c)}(\vec{k}_1)$ is the momentum-space propagator, which depends on only a single momentum coordinate, due to the discrete translation invariance. Since the propagators in momentum space are never zero, we may now isolate the 1pi function, writing it as

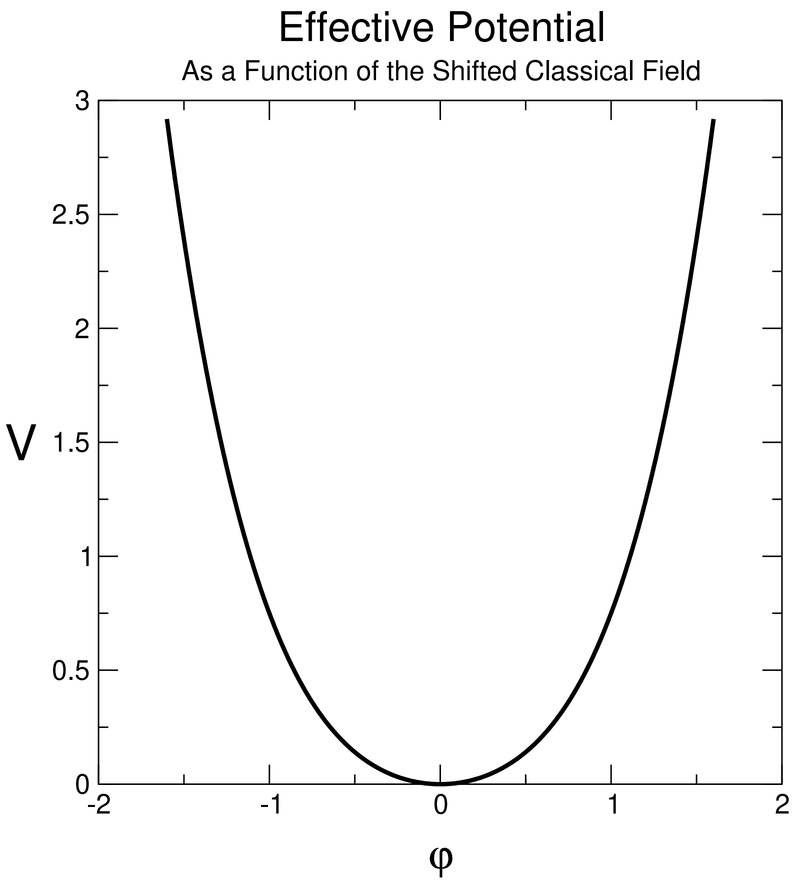

Figure 3.1.4: The effective potential as a function of the shifted classical field $\varphi'_{(c)}$.

symmetries that the fundamental action which defines the model has. Naturally, at this point it is necessary to consider for a while the issue of the possibility of spontaneous symmetry breaking that we know to exist in this model.

Let us recall that on finite lattices the symmetry is always broken, with $v_R \neq 0$. If we introduce into the model an infinitesimal constant external source δj, the field will spontaneously orient itself in the direction of the external source, be it positive or negative, without the system presenting any resistance to this change. For a sufficiently small δj, this happens without any significant change in the "energy" (in fact, the action) of the system, so that the *effective potential* of the theory, the part of the effective action that does not involve derivatives of $\varphi_{(c)}$, must be completely flat in a region around the point $\varphi_{(c)} = 0$, as figure 3.1.3 illustrates. We use in this figure the quantity v_R without arguments as the value of $v_R(\alpha, \lambda, j)$ for $j \to 0$ by positive values. Both α_R and λ_R are renormalized parameters that are related to the form of the graph of the effective potential in the regions $\varphi_{(c)} \geq v_R$ and $\varphi_{(c)} \leq -v_R$.

If we rewrite the effective potential in terms of the shifted classical field $\varphi'_{(c)} = \varphi_{(c)} - v_R$, recalling that v_R changes sign when δj changes sign, then we can represent then adding over the variable $\vec{n}_1$ and differentiating a second time we get

$$\sum_1 f_1(1) \frac{\partial^2 V_{(4)}}{\partial \varphi'_{(c)1} \partial \varphi'_{(c)2}} = \lambda_R \sum_1 f_1(1) 3\varphi'^2_{(c)1} \delta_{1,2} = 3\lambda_R f_2(1) \varphi'^2_{(c)2}.$$

Multiplying now by $f_2(2)$, adding over the variable $\vec{n}_2$ and differentiating a third time we get

$$\begin{aligned} \sum_{1,2} f_1(1) f_2(2) \frac{\partial^3 V_{(4)}}{\partial \varphi'_{(c)1} \partial \varphi'_{(c)2} \partial \varphi'_{(c)3}} &= 3\lambda_R \sum_2 f_2(2) f_2(1) 2\varphi'_{(c)2} \delta_{2,3} \\ &= 6\lambda_R f_3(2) f_3(1) \varphi'_{(c)3}. \end{aligned}$$

Repeating the procedure a fourth and last time we get

$$\begin{aligned} \sum_{1,2,3} f_1(1) f_2(2) f_3(3) \frac{\partial^4 V_{(4)}}{\partial \varphi'_{(c)1} \partial \varphi'_{(c)2} \partial \varphi'_{(c)3} \partial \varphi'_{(c)4}} &= 6\lambda_R \sum_3 f_3(3) f_3(2) f_3(1) \delta_{3,4} \\ &= 6\lambda_R f_4(3) f_4(2) f_4(1). \end{aligned}$$

Multiplying by $f_4(4)$ and adding over $\vec{n}_4$ we obtain for the fourth functional derivative of V_4

$$\sum_{1,2,3,4} f_1(1) f_2(2) f_3(3) f_4(4) \frac{\partial^4 V_{(4)}}{\partial \varphi'_{(c)1} \partial \varphi'_{(c)2} \partial \varphi'_{(c)3} \partial \varphi'_{(c)4}} = 6\lambda_R \sum_4 f_4(4) f_4(3) f_4(2) f_4(1).$$

Since this fourth functional derivative is equal to the 1pi four-point function, we obtain

$$\sum_{1,2,3,4} f_1(1) f_2(2) f_3(3) f_4(4) \gamma_{1,2,3,4} = 6\lambda_R \sum_4 f_4(4) f_4(3) f_4(2) f_4(1).$$

In the left-hand side of this equation we have N^{4d} times the four-fold Fourier transform of $\gamma_{1,2,3,4}$, while in the right-hand side, recalling that the mode functions $f_i(j)$ are exponentials that satisfy orthogonality and completeness relations, we have N^d times a Kronecker delta function that expresses the conservation of momentum, that is,

$$N^{4d}\, \widetilde{\gamma}_{1,2,3,4} = 6\lambda_R N^d \delta^d_{1+2+3+4}.$$

Combining now this equation with equation (3.1.5) we obtain the final relation between the renormalized coupling constant and the connected correlation functions,

$$\lambda_R \delta^d_{1+2+3+4} = -\frac{1}{6N^d} \, \frac{\widetilde{g}_{(c)1,2,3,4}}{\widetilde{g}_{(c)1}\, \widetilde{g}_{(c)2}\, \widetilde{g}_{(c)3}\, \widetilde{g}_{(c)4}}. \tag{3.1.6}$$

For combinations of momenta that satisfy the conservation condition $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 + \vec{k}_4 = \vec{0}$ the delta function is simply 1 and, assuming implicitly the conservation of momentum, we may write

Substituting the complete functions of three and two points in terms of the corresponding connected functions, show that the connected four-point function is related to the complete four-point function by means of

$$\begin{aligned}
g_{(c,j)1,2,3,4} &= g_{(j)1,2,3,4} \\
&- \left[g_{(c,j)2,3,4}\,\varphi_{(c)1} + g_{(c,j)1,3,4}\,\varphi_{(c)2} + g_{(c,j)1,2,4}\,\varphi_{(c)3} + g_{(c,j)1,2,3}\,\varphi_{(c)4}\right] \\
&- \left[g_{(c,j)1,2}\,g_{(c,j)3,4} + g_{(c,j)1,3}\,g_{(c,j)2,4} + g_{(c,j)1,4}\,g_{(c,j)2,3}\right] \\
&- \left[g_{(c,j)1,2}\,\varphi_{(c)3}\,\varphi_{(c)4} + g_{(c,j)1,3}\,\varphi_{(c)2}\,\varphi_{(c)4} + g_{(c,j)1,4}\,\varphi_{(c)2}\,\varphi_{(c)3}\right. \\
&\quad \left. + g_{(c,j)2,3}\,\varphi_{(c)1}\,\varphi_{(c)4} + g_{(c,j)2,4}\,\varphi_{(c)1}\,\varphi_{(c)3} + g_{(c,j)3,4}\,\varphi_{(c)1}\,\varphi_{(c)2}\right] \\
&- \varphi_{(c)1}\,\varphi_{(c)2}\,\varphi_{(c)3}\,\varphi_{(c)4},
\end{aligned}$$

which corresponds to the subtraction from the complete function of all the possible factorizations in terms of connected functions with a smaller number of points.

Observe that a significant part of the long algebraic passages involved in this problem has already been executed before in the problem in reference [40], relative to the three-point function. A simpler alternative way to obtain the results shown above is to start from the final result of that problem, doing an additional differentiation and using once more the same result to substitute the complete three-point function where necessary.

3.1.2. Show that in the theory of the free scalar field, that is, in the $\lambda\varphi^4$ model for the case $\lambda = 0$ and $\alpha \geq 0$, the connected four-point function given in equation (3.1.1) vanishes identically. Recall the results related to the factorization of the correlation functions of the free theory in momentum space, discussed in the section in reference [44], they will be very useful here.

3.1.3. Starting from the equation in reference [46] show that we can write that equation in the form

$$\sum_{3,4} g_{(c,j)1,3}\; g_{(c,j)2,4} \frac{\partial^2 \Gamma[\varphi_{(c)}]}{\partial\varphi_{(c)3}\partial\varphi_{(c)4}} = g_{(c,j)1,2}.$$

Then differentiate this equation once more with respect to j, using the chain rule to rewrite the derivatives as derivatives with respect to $\varphi_{(c)}$, thus obtaining

$$\begin{aligned}
\sum_{4,5,6} g_{(c,j)1,4}\; g_{(c,j)2,5}\; g_{(c,j)3,6} \frac{\partial^3 \Gamma[\varphi_{(c)}]}{\partial\varphi_{(c)4}\partial\varphi_{(c)5}\partial\varphi_{(c)6}} &+ \\
+ \sum_{4,5} g_{(c,j)1,3,4}\; g_{(c,j)2,5} \frac{\partial^2 \Gamma[\varphi_{(c)}]}{\partial\varphi_{(c)4}\partial\varphi_{(c)5}} &+ \\
+ \sum_{4,5} g_{(c,j)1,4}\; g_{(c,j)2,3,5} \frac{\partial^2 \Gamma[\varphi_{(c)}]}{\partial\varphi_{(c)4}\partial\varphi_{(c)5}} &= g_{(c,j)1,2,3}.
\end{aligned}$$

$$\frac{1}{N^{nd}}\sum_{1,\ldots,n} f_1(1)\ldots\ f_n(n)\ F_{1,\ldots,n}=\widetilde{F}_{1,\ldots,n},$$

where, to simplify the notation, we are denoting the mode functions of the Fourier basis as

$$f_i(j)=e^{\imath\frac{2\pi}{N}\vec{n}_i\cdot\vec{k}_j},$$

besides indicating the momentum coordinates by indices on the remaining functions in momentum space, as we have done before for the position coordinates. Use also the fact that, due to the discrete translation invariance of the lattice, we have

$$\frac{1}{N^d}\sum_{1} f_1(1)\ g_{(c)1,2}=f_2(1)\ \widetilde{g}_{(c)1},$$

where $\widetilde{g}_{(c)1}$ is the propagator in momentum space, which depends only on a single momentum coordinate, due the the discrete translation invariance, in order to write the final result

$$\widetilde{g}_{(c)1}\ \widetilde{g}_{(c)2}\ \widetilde{g}_{(c)3}\ \widetilde{g}_{(c)4}\ N^{4d}\ \widetilde{\Gamma}_{1,2,3,4}=-\widetilde{g}_{(c)1,2,3,4}.$$

3.1.6. Derive the expression for λ_R in terms of the zero-momentum correlation functions of the model, for the general case in which $j\neq 0$ and $v_R\neq 0$.

3.1.7. Show that the expression of the coupling constant in terms of correlation functions with a given constant momentum $\vec{k}$ that enters in the direction of the vertex in two of the four external legs and goes out in the opposite direction in the other two legs is

$$\lambda_R(\vec{k})=\frac{1}{6N^d}\,\frac{2\langle|\widetilde{\varphi}(\vec{k})|^2\rangle^2-\langle|\widetilde{\varphi}(\vec{k})|^4\rangle}{\langle|\widetilde{\varphi}(\vec{k})|^2\rangle^4},$$

both in the case in which $j=0$ and $v_R=0$ and in the general case.

3.2 Critique of Perturbative Renormalization

Using the techniques and ideas developed in section 1.2 we may try to calculate perturbatively the renormalized coupling constant λ_R, which we wrote in terms of observables of the model in section 3.1. Using equation (3.1.7) and calculating the observables involved to first order, with the same choice $\alpha_0=\alpha_R$ that we used

$$s_2(\alpha_0) = \sum_{\vec{n}} g_0^2(\vec{0}, \vec{n}) = \frac{1}{N^d} \sum_{\vec{k}} \frac{1}{(\rho_k^2 + \alpha_0)^2}. \tag{3.2.1}$$

In principle it is also possible to do the calculation in the broken-symmetrical phase (problem 3.2.3), but currently we do not yet know the answer for that case. Making the choice $\alpha_0 = \alpha_R$ as before, we obtain

$$\lambda_R = \lambda[1 - 9\lambda s_2(\alpha_R)]. \tag{3.2.2}$$

Observe that, since the sum is a positive quantity, the correction is always negative, tending to decrease the magnitude of the positive classical result. We may try to evaluate the behavior of the sum s_2 for large values of N by means of approximations by integrals, as we did before in other cases. Doing this (problem 3.2.4) we obtain the following results, for the usual values of the dimension d:

$$\begin{aligned}
d = 1: &\quad s_2 \approx \frac{1}{3\pi (m_R L)^3} N^3, \\
d = 2: &\quad s_2 \approx \frac{1}{4\pi (m_R L)^2} N^2, \\
d = 3: &\quad s_2 \approx \frac{1}{2\pi^2 m_R L} N, \\
d = 4: &\quad s_2 \approx \frac{1}{8\pi^2} \ln(N), \\
d \geq 5: &\quad s_2 \approx \frac{\Omega_d}{(2\pi)^4 (d-4)},
\end{aligned} \tag{3.2.3}$$

where Ω_d is the total solid angle of d-dimensional space. We see therefore that the second-order result is divergent for $d = 3$ and $d = 4$, and finite for $d \geq 5$. Observe that, in order for the results in $d = 3$ and $d = 4$ to make any sense, it is necessary that we make $\lambda \to 0$ in the continuum limit, thus forcing us to return to the Gaussian point. In $d = 4$ this takes us back to the free theory but, curiously, despite this limitation the result is still of some interest in the case $d = 3$, because in this case the dimensionfull coupling constant is given by $\Lambda_R = \lambda_R / a = N\lambda_R / L$ and we therefore have, in terms of the dimensionfull quantities, a finite expression,

$$\Lambda_R = \Lambda[1 - 9\Lambda S_2], \tag{3.2.4}$$

where the dimensionfull quantity $S_2 = s_2 L/N$ has a finite limit, so long as we have a finite and non-vanishing renormalized mass m_R,

$$S_2 = \frac{s_2 L}{N} \approx \frac{1}{2\pi^2 m_R}.$$

This seems to indicate that in $d = 3$ there are non-trivial limits, with $\Lambda_R \neq 0$, that approach the Gaussian point, where both λ and λ_R are zero. We can use this $d = 3$ result to exhibit explicitly renormalization flows $[\alpha(N), \lambda(N)]$, in the parameter space of the corresponding model, that approach the Gaussian point in such a way that both m_R and Λ_R are different from zero in the limit (problem 3.2.5).

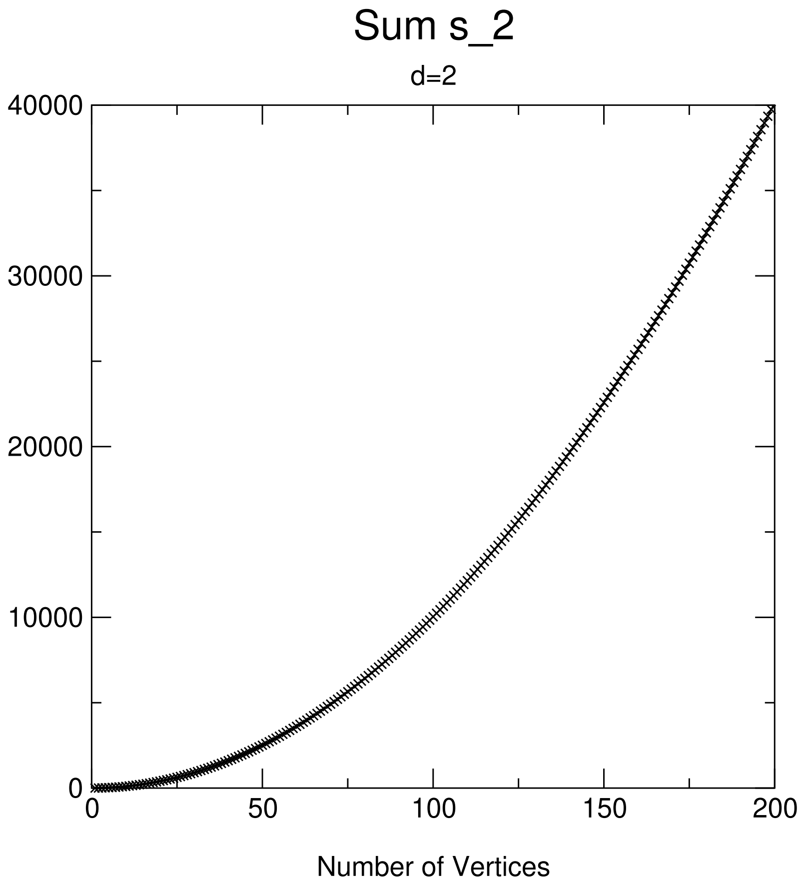

Figure 3.2.2: Behavior of the sum s_2 with N in the case $d = 2$.

non-vanishing. It is very reasonable to think that the dimensionless quantities like λ_R will always be finite in the limit, when we make $N \to \infty$, since they do not scale with N. Since we have that $\Lambda_R = a^{d-4}\lambda_R$, for dimensions $d \geq 5$ we have that a finite λ_R implies a vanishing Λ_R in the limit. We establish in this way the expectation that the model is completely trivial in $d \geq 5$, with λ_R finite and Λ_R vanishing over the critical curve. Of course this conclusion depends on the higher-order terms of the perturbative series of λ_R being all finite, besides the series being convergent. Obviously, none of these two things is guaranteed. One might even imagine that the series could end up converging to zero, a hypothesis which would not change the physical meaning of the theory, since that meaning is defined in terms of Λ_R, which would still vanish in the limit. However, that hypothesis would make sense in terms of the critical behavior of the dimensionless quantity λ_R, as we discussed above in the case $d = 3$.

In the case $d = 4$ we cannot conclude anything about the behavior of the theory from our results, since in this case s_2 diverges logarithmically while $\Lambda_R = \lambda_R$. We can confirm our analytical estimates of the dependencies of s_2 on N and calculate approximately the relevant coefficients, performing numerically the sum on finite

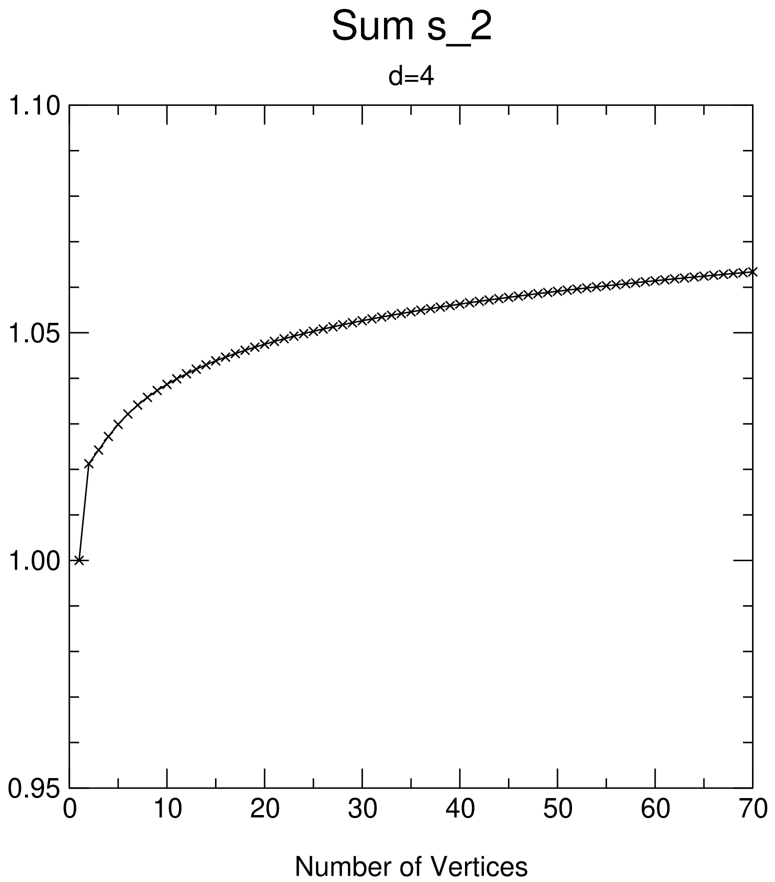

Figure 3.2.4: Behavior of the sum s_2 with N in the case $d = 4$.

divergent sums and it is not possible to use them in a systematic way to establish a system of successive approximations, which hopefully would be increasingly precise, for all the observables of the theory.

Therefore, we see that we do not really have a complete expansion for the $\lambda\varphi^4$ model, but only a set of approximations that work reasonably well in some cases. What we have here is not a consistent and systematic series development, but only a set of isolated approximations whose validity can only be verified, ultimately, by direct comparison with numerical results or other non-perturbative approximations. Observe that this behavior of the perturbative expansion is due to the exchange of the order of the two limits involved, the continuum limit and the series summation limit. For finite N all the terms of the expansion are finite and the series may be convergent, or at least asymptotic, but in the $N \to \infty$ limit the individual terms become infinite and nothing can be done to salvage the situation in general, except making the parameters of the model converge sufficiently fast to the Gaussian point. In order to make use of the series it is imperative to first sum it on finite lattices and only then take the continuum limit, one cannot invert the order of the limits.

In conclusion, we verify that in the case of the coupling constant we do not have whose coefficients may be adjusted so as to reproduce faithfully the characteristics of the ensemble of the complete model. When we choose $\alpha_0 = \alpha_R$ we are making the moment of order two of the Gaussian ensemble reproduce in a perfect way the corresponding moment of the ensemble of the complete model, so that the difference between the two distributions can in fact be considered a small perturbation of the Gaussian distribution, in so far as that observable is concerned. However, independently of any choices of parameters, the fourth-order moment of the Gaussian ensemble is always zero and cannot be adjusted to reproduce the corresponding moment of the complete ensemble. Therefore, the fourth-order moment can never be understood as a small perturbation when we are dealing with observables that only exist when this moment is not zero.

It is important that we discuss here the role of the traditional scheme of perturbative renormalization, in the context of the calculations on the lattice[1]. First of all, there can be no doubt that the definition of the model on the lattice implies the existence of well-defined relations $\alpha_R(\alpha, \lambda)$ and $\lambda_R(\alpha, \lambda)$ between these renormalized quantities and the parameters of the model, for any values of these parameters within the stable region of the parameter plane of the models and in any dimension d, for finite lattices, and over the critical curve in any dimension $d \geq 3$, in the continuum limit. Analogously, for any observable $\mathcal{O}$ that is physically relevant we have a well-defined relation $\mathcal{O}(\alpha, \lambda)$, under the same conditions. What we discover when we work out the development of perturbation theory is that the perturbative approximations for the relations between the observables $\mathcal{O}$, α_R and λ_R and the parameters α and λ are singular, in the sense that they contain quantities that diverge in the continuum limit.

The traditional perturbative renormalization scheme consists of giving up, at this point, any effort of extracting from the theory the relations $\alpha_R(\alpha, \lambda)$, $\lambda_R(\alpha, \lambda)$ and $\mathcal{O}(\alpha, \lambda)$ that it contains. In addition to this, on each finite lattice, where all quantities are finite, we may consider rewriting $\mathcal{O}$ directly in terms of α_R and λ_R, eliminating the parameters α and λ from the picture in favor of their renormalized counterparts, and thus obtaining a relation $\mathcal{O}(\alpha_R, \lambda_R)$ that, possibly, will not contain any quantities that diverge in the continuum limit. If it is possible to do this, then the relation obtained is a well-behaved perturbative approximation of $\mathcal{O}$ in terms of α_R and λ_R, and we may then take the continuum limit without stumbling on any singularities. If it is possible to do this for the perturbative calculations to all orders, then we say that the theory is perturbatively renormalizable, and the scheme produces a complete perturbative series in the continuum limit, with finite individual terms, which may or may not be convergent. Since α and λ are not directly observable, while α_R and λ_R presumably are, the resulting function $\mathcal{O}(\alpha_R, \lambda_R)$ is a direct relation between observables of the theory, so that not much seems to be lost when one does this.

Of course, rewriting $\mathcal{O}$ in terms of α_R and λ_R on finite lattices may not be easy, in fact, it may not be possible at all in closed form, so one may be compelled to

[1]These arguments were developed in discussions with Dr. Timothy E. Gallivan and with Prof. Henrique Fleming.

we see that λ should diverge in order for this relation to be valid. This is the opposite of what we saw in the original perturbative expansion, where we verified that λ must go to zero in order for the perturbative approximation to be valid. However, since by now we have given up obtaining from the theory the relations between the basic quantities and the renormalized quantities, we might as well just not worry about this any more, and simply disregard the equation above. Of course, one cannot avoid the strong impression that this whole procedure is mired with guesswork and arbitrariness. It certainly looks like it would be very difficult to show that all the facts assumed do indeed hold to all orders and hence to establish the results of this procedure on a firm logical basis.

Anyway, up to this point the change of variables has not really been of any use, since it simply introduced another parameter λ_0 that ended up being another name for λ_R. No additional information about the relation between λ_R and the basic parameters of the model was obtained. It is important to emphasize that this fact is no more than a limitation of the perturbative method and that this relation undoubtedly exists in the model defined by means of the lattice. In order to show the possible usefulness of the perturbative renormalization scheme, we may now consider the calculation of a third observable $\mathcal{O}$, at first in terms of α and λ, resulting in a relation of the type

$$\mathcal{O} = f_{\mathcal{O}}(\alpha, \lambda),$$

which presumably contains some terms with divergent factors. We may now substitute λ for λ_0, re-expanding the resulting expression and neglecting once more the higher-order terms that appear. With some more manipulation we may also substitute α for α_R, thus obtaining a new relation

$$\mathcal{O} = \bar{f}_{\mathcal{O}}(\alpha_R, \lambda_R),$$

which, so long as the model is perturbatively renormalizable, should not contain any divergences. In this way we extract from the model a well-behaved relation between α_R, λ_R and $\mathcal{O}$, although the fundamental perturbative expansion in terms of α and λ is not well behaved. As an example of such an observable, we may consider the coupling constant for a non-vanishing momentum $\vec{k}$, a quantity which is related in a direct way to the scattering cross-sections. Calculating the coupling constant for the same momentum $\vec{k}$ on all the four external legs, up to order λ^2, we obtain (problem 3.2.9)

$$\lambda_R(\vec{k}) = \lambda \left\{ 1 - 3\lambda \left[2s_2(\vec{0}, \alpha_0) + s_2(\vec{k}, \alpha_0) \right] \right\},$$

where the sum $s_2(\vec{k}, \alpha_0)$ is given by

$$s_2(\vec{k}, \alpha_0) = \frac{1}{N^d} \sum_{\vec{k}_1} \frac{1}{[\rho^2(\vec{k}_1 + \vec{k}) + \alpha_0][\rho^2(\vec{k}_1 - \vec{k}) + \alpha_0]}.$$

actually possible for the renormalized parameters, in terms of which one chooses to write the renormalized perturbative expansion. In regard to this aspect of the structure of the theory it is important to emphasize the profound difference between a truly physical theory, such as quantum electrodynamics, and models that have only the role of mathematical laboratories, such as the polynomial models. In quantum electrodynamics we can go to the laboratory and determine experimentally the values of the renormalized mass and of the renormalized coupling constant, that is, of the mass and charge of the electron, thus establishing that certain values are possible for these quantities. On the other hand, in the laboratory models we are limited to what we can calculate analytically or numerically and we must extract this type of information from the relations that the models establish between the renormalized quantities and the parameters involved in their definitions. Since perturbation theory is not able to give us these relations in a complete form, it only remains for us to try non-perturbative methods, such as computational stochastic simulations, as tools to establish the possible values for the renormalized parameters. Another way to characterize this profound difference is to say that, in the case of a truly physical theory, we have access to the use of the ultimate computer: the fundamental laws of physics at play in nature.

Problems

Note: Some of the calculations contained in some of these problems are really *very* long, and a considerable amount of organization and care is needed to get to the end without errors. The problems containing such long calculations are marked with three stars.

3.2.1. Calculate λ_R using equation (3.1.7) and calculating the observables involved to first order, with the choice $\alpha_0 = \alpha_R$, thus obtaining the classical result $\lambda_R = \lambda$. Do the calculations both in the symmetrical phase and in the broken-symmetrical phase. Note that, since λ_R itself is already a first-order quantity in λ, in order to keep consistent orders of the expansion parameter the numerator of equation (3.1.7) should be calculated to first order, while it is enough to calculate the denominator to order zero.

3.2.2. ($\star$) Calculate λ_R to second order, using equation (3.1.7), with the choice $\alpha_0 = \alpha_R$, in the symmetrical phase, obtaining the result quoted in the text,

$$\lambda_R = \frac{\lambda\left[1 - 4\dfrac{\alpha_R - \alpha_0}{\alpha_0} - 9\lambda s_2(\alpha_0)\right]}{\left[1 - \dfrac{\alpha_R - \alpha_0}{\alpha_0}\right]^4}.$$

Note that, since λ_R itself is already a first-order quantity in λ, in order to keep consistent orders of the expansion parameter the numerator of equation (3.1.7)

Finally, recalling that α_0 must go to zero as $1/N^2$ in the limit, make $v_R = 0$ and show that, in the continuum limit, one recovers the one-loop result for the equation of the critical curve,

$$\left(\alpha + 3\lambda\sigma_0^2\right)^2 = 0.$$

Therefore, we conclude that the equation of the critical curve does not contain corrections of order λ^2 and that any correction to the order-λ result must be at least of order λ^3.

3.2.8. $(\star\star\star)$ Try to calculate the renormalized mass m_R in the symmetrical phase up to order ε^2. Start by calculating the propagator to this order, obtaining the result

$$\begin{aligned} g_2(\vec{n}_1,\vec{n}_2) \;=\;& g_0(\vec{n}_1,\vec{n}_2) - \left(\alpha - \alpha_0 + 3\lambda\sigma_0^2\right)\sum_{\vec{n}_3} g_0(\vec{n}_1,\vec{n}_3)g_0(\vec{n}_3,\vec{n}_2) \\ &+ \left(\alpha - \alpha_0 + 3\lambda\sigma_0^2\right)^2 \sum_{\vec{n}_3}\sum_{\vec{n}_4} g_0(\vec{n}_1,\vec{n}_3)g_0(\vec{n}_3,\vec{n}_4)g_0(\vec{n}_4,\vec{n}_2) \\ &+3\lambda\left(\alpha - \alpha_0 + 3\lambda\sigma_0^2\right)\sum_{\vec{n}_4} g_0^2(0,\vec{n}_4)\sum_{\vec{n}_3} g_0(\vec{n}_1,\vec{n}_3)g_0(\vec{n}_3,\vec{n}_2) \\ &+6\lambda^2\sum_{\vec{n}_3}\sum_{\vec{n}_4} g_0(\vec{n}_1,\vec{n}_3)g_0^3(\vec{n}_3,\vec{n}_4)g_0(\vec{n}_4,\vec{n}_2), \end{aligned}$$

which, in momentum space, can be written as

$$\begin{aligned} N^d\widetilde{g}_2(\vec{k}) \;=\;& \frac{1}{\rho^2(\vec{k}) + \alpha_0} - \frac{\alpha - \alpha_0 + 3\lambda\sigma_0^2}{[\rho^2(\vec{k}) + \alpha_0]^2} + \frac{(\alpha - \alpha_0 + 3\lambda\sigma_0^2)^2}{[\rho^2(\vec{k}) + \alpha_0]^3} \\ &+\frac{3\lambda(\alpha - \alpha_0 + 3\lambda\sigma_0^2)}{[\rho^2(\vec{k}) + \alpha_0]^2}s_2(\alpha_0) + \frac{6\lambda^2}{[\rho^2(\vec{k}) + \alpha_0]^2}s_3(\vec{k},\alpha_0), \end{aligned}$$

where the sum s_3 is given by

$$s_3(\vec{k},\alpha_0) = \frac{1}{N^{2d}}\sum_{\vec{k}_1,\vec{k}_2}\frac{1}{[\rho^2(\vec{k}_1) + \alpha_0][\rho^2(\vec{k}_2) + \alpha_0][\rho^2(\vec{k} - \vec{k}_1 - \vec{k}_2) + \alpha_0]}.$$

Observe that, since α_R must still vanish over the critical curve, and the equation of that curve did not change up to the order ε^2, as we saw in problem 3.2.7, we should expect that the renormalized mass also does not change up to this order. Discover whether or not it is possible to choose α_0 in an appropriate way and thus show that the renormalized mass also does not change up to this order, establishing therefore the consistency of the two calculations[3].

[3]Note: the answer to this last question is currently unknown.

[20] *Ibid.*[3], formula 8.445, page 961.

[21] *Ibid.*[3], formula 8.451, equation 5, page 962.

[22] *Ibid.*[3], formula 8.467, page 967.

[23] *Ibid.*[3], formula 8.471, equations 1 and 2, page 967.

[24] *Ibid.*[3], formula 8.486, equations 1 to 4, page 970.

[25] *Ibid.*[3], formula 9.246, equation 2, pages 1065-1066.

[26] *Ibid.*[3], formula 9.247, equation 1, page 1066.

[27] *Ibid.*[3], formula 9.254, equations 1 and 2 (Attention: in old editions, the sign of this second formula is wrong!), page 1067.

[28] It has been shown that there is a phase transition for sufficiently large λ, see J. Glimm, A. Jaffe and T. Spencer, "Phase Transitions for φ_2^4 Quantum Fields", Commun. Math. Phys. **45,** (1975), 203–216. This, plus the fact that the exact solution for critical point of the two-dimensional Ising model is known, guarantees that there is a critical curve starting at some sufficiently large value of λ. There are also expansions that are both convergent and analytical near the Gaussian point, see J. Glimm, A. Jaffe and T. Spencer, "A Convergent Expansion about Mean Field Theory I. The Expansion", Ann. Physics, **101,** (1976), 610–630, and J. Glimm, A. Jaffe and T. Spencer, "A Convergent Expansion about Mean Field Theory II. Convergence of the Expansion", Ann. Physics, **101,** (1976), 631–669. This indicates that there is no critical behavior of the model in a neighborhood of the Gaussian point, and that therefore the critical curve does not connect with that point.

[29] For the standard treatment in statistical mechanics of the one-site and several-site approximations for ferromagnetic systems, see: J. S. Smart, "Effective Field Theories of Magnetism", Saunders, (1966), and references therein. For a discussion of the method in the context of quantum field theory, see: E. Brézin, J. C. Le Guillou and J. Zinn-Justin, "Field Theoretical Approach to Critical Phenomena", in "Phase Transitions and Critical Phenomena", volume **6,** C. Domb and M. S. Green, editors, Academic Press, (1976).

[30] The basic idea here is that the partition function of finite systems is an analytic function and that, therefore, the systems cannot display the behaviors of non-differentiability and even of discontinuity that are characteristic of phase transitions, see: C. N. Yang and T. D. Lee, "Statistical Theory of Equations of State and Phase Transitions I. Theory of Condensation", Phys. Rev. **87,** (1952), 404–409; T. D. Lee and C. N. Yang, "Statistical Theory of Equations of State and Phase Transitions II. Lattice Gas and Ising Model", Phys. Rev. **87,** (1952), 410–419.