Non-Linear Scalar Models

A Continued Exploration of Quantum Field Theory

Note: this book is still incomplete; the first two chapters and two thirds of the third chapter can be found in this version; the first two chapters are complete and reasonably self-contained; the remaining material depends on stochastic simulations that are currently being worked on.

Jorge L. deLyra
Department of Mathematical Physics
Physics Institute
University of São Paulo

Version 0.5, May 2006

Preface

This is the second in a series of a few short books about the foundations of quantum field theory.

We would like to acknowledge here the contribution to parts of the material contained within this book by former students and collaborators who were active in this subject in the past: Dr. Timothy Edward Gallivan and Dr. André Cavalcanti Rocha Martins. When appropriate, due credit will be given in footnotes along the text. We would also like to thank those that helped by reading and correcting the manuscript at various stages of its production.

Contents

Chapter 1

The Polynomial Model

In this chapter we will examine the polynomial models of scalar fields. These are generalizations of the Gaussian model in which one tries to introduce interactions into the model by the addition of a new term to the action, containing a power of the field which is greater than the power two that appears in the Gaussian model.

1.1 Definition of the Model

In the previous volume of this series [2] we studied in detail the theory of the free scalar field. That model was sufficiently simple to allow us to calculate analytically all the predictions of the theory. As we saw, both in the case of the classical theory and in the case of the quantum theory this simplicity follows from the linearity of the model. We also saw that this same linearity is responsible for the fact that the model does not contain the concept of interactions between particles, and hence that the only physics that it does contain is the propagation of free particles. This was shown by the factorization of all the correlation functions in terms of the propagator, and also by the fact that the energies of the particles are simply additive, that is, the energy of a state containing two particles is the sum of the energies of the two corresponding single-particle states, implying the absence of any kind of interaction energy.

We will make here a first trial at including interactions in the theory, for which it will be necessary to break the linearity of the model, including in the action terms with more that two powers of the field. We will therefore examine the model defined by

$$S[\varphi] = \frac{1}{2}\sum_{\ell}(\Delta_\ell\varphi)^2 + \frac{\alpha}{2}\sum_{s}\varphi^2(s) + \frac{\lambda}{4}\sum_{s}\varphi^4(s), \tag{1.1.1}$$

which we denominate the $\lambda\varphi^4$ *polynomial model*. We therefore choose to break the linearity of the model by the introduction of a new ultra-local term into the action, leaving untouched the term containing the derivatives. This is the simplest example of a model that, in the classical theory, contains interacting fields. Our task here is

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

in the methods of stochastic simulation constitute a rather extended topic with a very technical character, and will be developed in detail in a separate volume of this small series of books.

The character of the classical theory determined by our action is clear, and it is not necessary to examine it in detail. The definition of the classical theory is the same as before, the classical solution is the configuration φ that minimizes the action. The fact that it exists is guaranteed by the conditions we impose on the parameters of the model: $\lambda > 0$ with any α or $\lambda = 0$ with $\alpha \geq 0$. We may derive, in a way analogous to the one used before for the free theory, the corresponding classical equation of motion, which will be, of course, a non-linear differential equation for φ (problem 1.1.1). In order to begin the examination of the behavior of the quantum theory, we recall that it is defined by the probability distribution over all possible configurations of φ, given by

$$\frac{[\mathbf{d}\varphi]\; e^{-\sum_\ell \frac{1}{2}(\Delta_\ell \varphi)^2 - \sum_s \left(\frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)}}{\int [\mathbf{d}\varphi]\; e^{-\sum_\ell \frac{1}{2}(\Delta_\ell \varphi)^2 - \sum_s \left(\frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)}},$$

where we grouped separately the term containing the derivatives and the ultra-local part, containing the polynomial terms, both the quadratic one and the quartic one, which we call the interaction term. If we recall that the measure $[\mathbf{d}\varphi]$ is a product of differentials over all the sites s, representing the fact that in this measure the stochastic variables φ have uniform probability distributions, we see that we can include the ultra-local terms in the measure, writing the distribution as

$$\frac{\left[\mathbf{d}\varphi\; e^{-\left(\frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)}\right]\; e^{-\sum_\ell \frac{1}{2}(\Delta_\ell \varphi)^2}}{\int \left[\mathbf{d}\varphi\; e^{-\left(\frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)}\right]\; e^{-\sum_\ell \frac{1}{2}(\Delta_\ell \varphi)^2}}.$$

In this new measure the variables φ no longer have uniform probability distributions, but have instead the probability distribution given by the exponential of the potential. A typical example of such a distribution can be seen in figure 1.1.1. We see in this way that a possible way to understand our model is to think of it as constituted of the dynamics implemented by the derivative term, but applied indirectly to new random variables χ located at the sites, with uniform probability distributions, which are given in terms of the variables φ by the differential relation

$$\mathrm{d}\chi = \mathrm{d}\varphi\; e^{-\left(\frac{\alpha}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)}.$$

This means that we may write the stochastic variable χ, which has an uniform probability distribution within a closed interval, in terms of the stochastic variable φ, which has a non-uniform probability distribution over the whole real line, as

$$\chi(\varphi) = \int_0^\varphi \mathrm{d}\varphi'\; e^{-\left(\frac{\alpha}{2}\varphi'^2 + \frac{\lambda}{4}\varphi'^4\right)}. \tag{1.1.2}$$

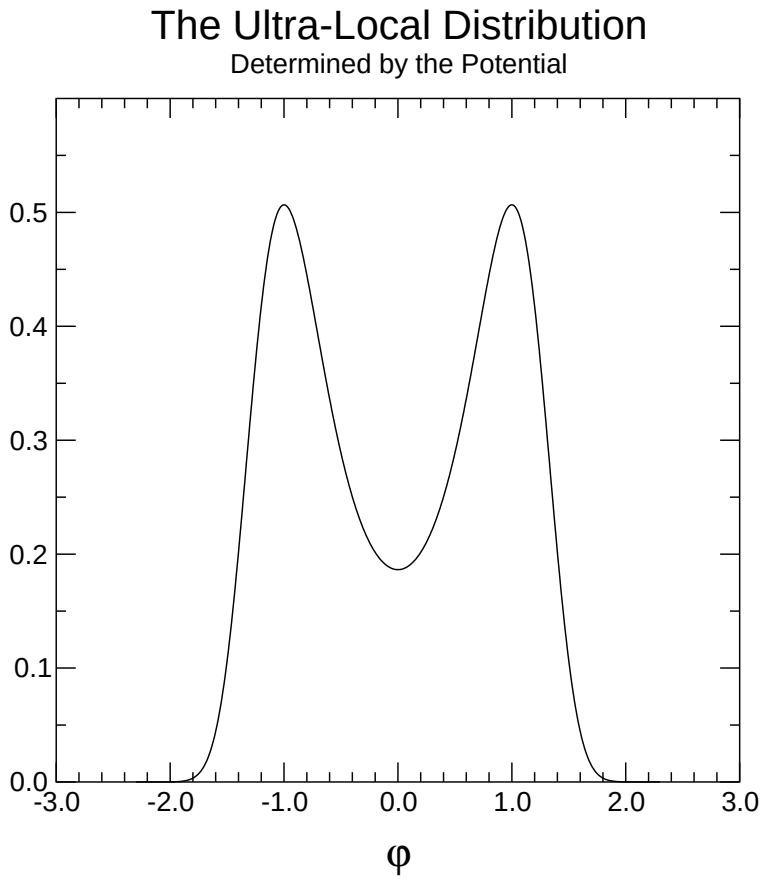

Figure 1.1.1: A typical ultra-local distribution of the fields, given by the potential, that is, by the polynomial terms of the action, in the case ($\alpha < 0, \lambda > 0$).

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

of the model at an intuitive level, it is very useful to invert the decomposition described above, representing the effect of the derivative term by the Gaussian local distribution that it implies, which was studied in detail in the first volume of this series of books, and considering directly the effects that the potential term may have on it. This allows us to use in the analysis of the model our classical intuition concerning the behavior of an object within the potential, by extending the object from a simple point body to a fluctuating statistical distribution. In this way we are able to build an almost-classical intuition in order to understand the behavior of the model.

We can imagine that we draw a copy of the potential at each site and that we put inside it the value of the field at that site, while the terms $(\Delta_\ell \varphi)^2$ interconnect each pair of neighboring sites. The stochastic dynamics implemented by these derivative terms will cause the values of the fields to fluctuate at each site, so that we will actually have a distribution of values within each potential. In the theory of the free field these local distributions are simple Gaussians, whose width σ_0 is a number of the order of 1 for any lattice size, a number that determines the value of the physical mass (also called the renormalized mass), which in the case of the free theory is simply $m_R = m_0$. In our current model the complete local distributions also have a width $\sigma \approx \sigma_0$ of the order of 1, but their format may no longer be exactly Gaussian. The renormalized mass is defined in this model, just as in the free theory, as the inverse of the correlation length of the model, measured through the two-point correlation function, and presumably it is also related to σ, as is the case for the free theory. What we propose to do in this section is to understand the complete dynamics of the model by considering directly the influence of the potential over the Gaussian local distribution implemented by the derivative term.

Let us recall that the relation between the two-point function and the renormalized mass was studied in detail, in the case of the free field, in a section of the first volume [4]. The exponential decay of the two-point function for large distances r, given by $\exp(-m_0 r)$ as was studied in that section, is a general property of all massive theories, either free or otherwise. The behavior of the two-point function of the interacting theories for large distances is in general of this same type, except for the exchange of the parameter m_0 for another parameter m_R which usually differs from m_0. In the numerical approach we usually measure this new parameter by means of a curve-fitting process applied to the numerical propagator of the theory in momentum space rather than position space, that is, to the Fourier transform of the two-point function, which is technically easier to do, and also more efficient. We do this in the expectation, to be confirmed a posteriori, that the form of this function in the interacting models is not very different from its format in the free theory, and it usually works very well.

In the model that we are introducing here, given values of N, α and λ, we will have not only a resulting value for the parameter α_R related to the renormalized mass, but also some resulting value for the *renormalized coupling constant* λ_R, which is the physical coupling constant whose nature and precise definition we will examine in more detail later on. Unlike what happened in the free theory, in general α_R will

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.

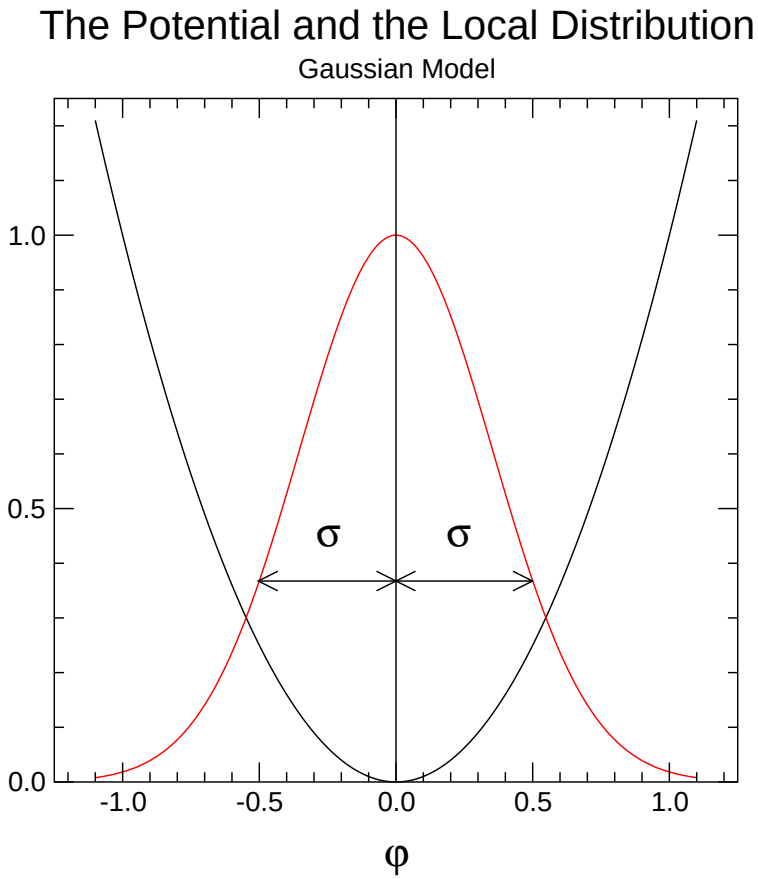

Figure 1.1.2: The potential and the local distribution due to the derivative term, in the case of the free theory.

squeezing force gives rise to an infinite m_R. Let us imagine now that we insert this local distribution inside the potential defined by the ultra-local terms of the action at each site, as shown in figure 1.1.2 for the case of the free theory, with a Gaussian distribution and a quadratic potential. If we were examining the classical theory, we would put inside the potential a point body representing the value of the classical solution for the field, and it would come to rest at the minimum of the potential. The examination of the behavior of the quantum theory corresponds to the introduction into the potential well of an extended object which can be represented heuristically by our semi-rigid body, which becomes rigid in the continuum limit in the sense explained above.

The width of the local distribution in the absence of the potential is determined only by the derivative term and corresponds to the value $\alpha_R = 0$, that is, to a zero renormalized mass. When we put the distribution under the action of the potential on a finite lattice what happens is that it tends to concentrate the values of the field around the minimum, and hence squeezes the distribution, decreasing its width, because it is statistically unfavorable for the field to exist in the positions where the

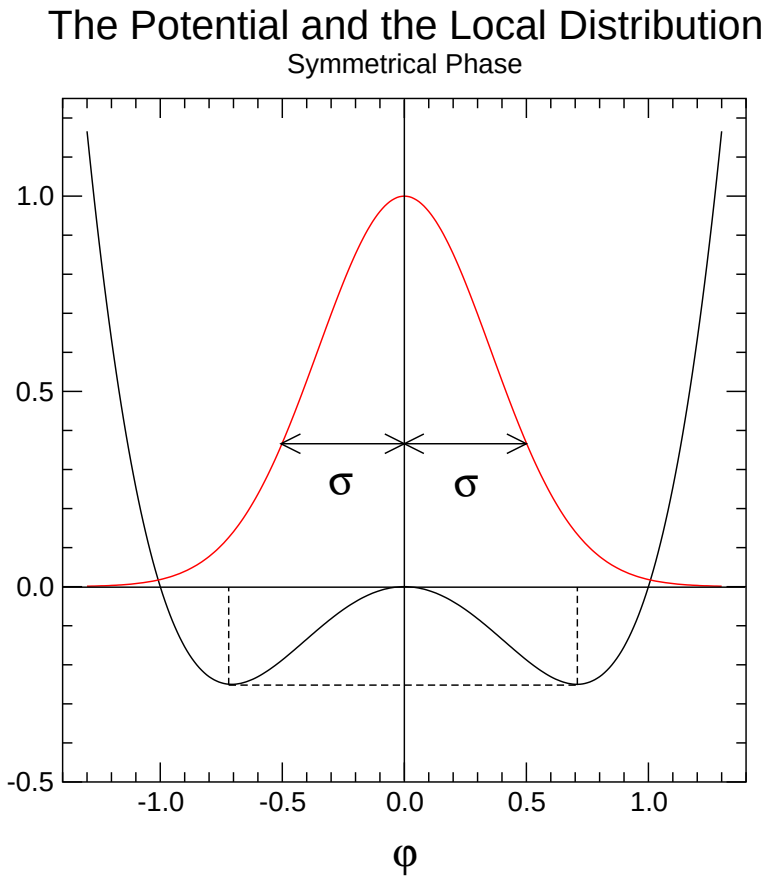

Figure 1.1.3: The potential and the local distribution due to the derivative term, in the case of the non-linear theory, in the symmetrical phase.

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

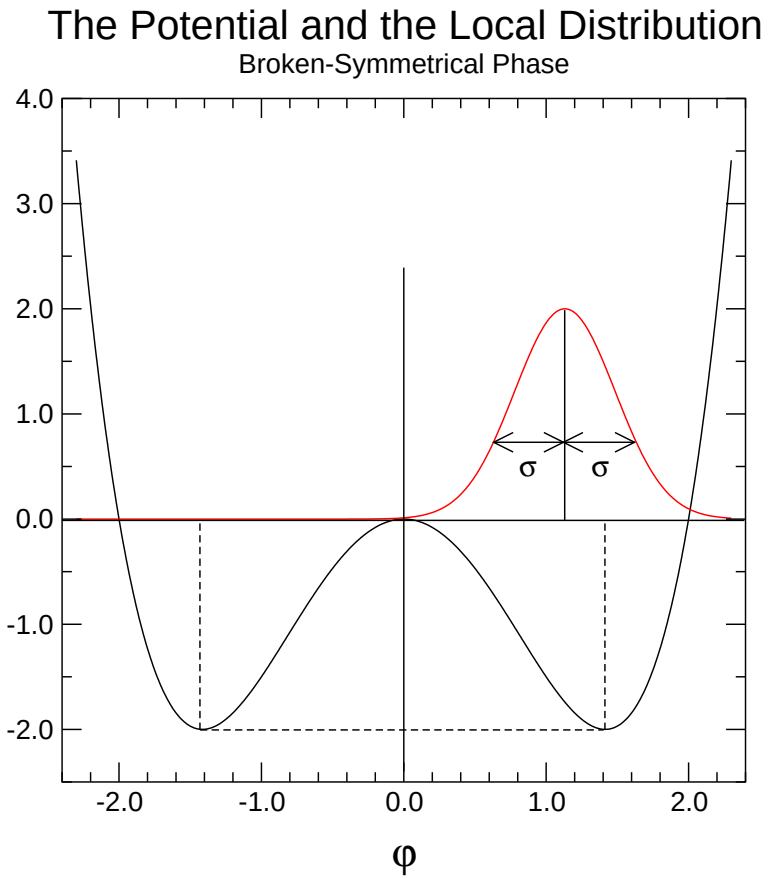

Figure 1.1.4: The potential and the local distribution due to the derivative term, in the case of the non-linear theory, in the broken-symmetrical phase.

much effort, a prediction with very serious consequences regarding the behavior of the model: if we limit ourselves to the case in which $\alpha \geq 0$ and $\lambda \geq 0$ for all N, it will be necessary to make both $\alpha \to 0$ and $\lambda \to 0$ in the limit, *which takes the model back to the critical point of the Gaussian model and therefore eliminates any possibility that* λ_R *be different from zero in limits of this type.*

Except for the case $d = 3$, this implies that there are in fact no interactions between particles in the quantum theory of this model, in any limits that stay within the quadrant given by $\alpha \geq 0$ and $\lambda \geq 0$. We say that in this case the model has only the *trivial* limit, leading to the theory of the free field, or that the theory is *trivial* in this sector of the space of parameters of the model. The case $d = 3$ is a little different because, since in this case the physical coupling constant has dimensions of mass, it is possible that there are interactions even if the model approaches the Gaussian point, e phenomenon that we will discuss later on.

We conclude that, if we are to have any chance of finding an interesting limit in this model, it will be necessary for at least one of the two parameters to be negative. Since we cannot make λ negative due to the stability constraints, it follows that *the*

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

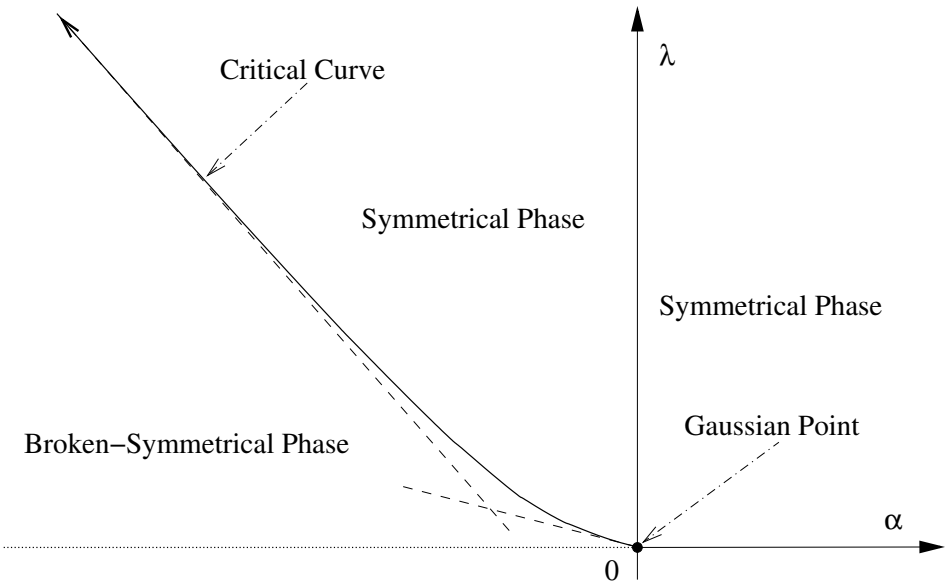

Figure 1.1.5: Critical diagram of the non-linear $\lambda\varphi^4$ model.

in this model, giving origin to two *phases* in the space of parameters of the model, in each one of which the behavior of the model is of a certain type, different from the other one. The expectation value $v_R = \langle\varphi\rangle$ is the *order parameter* of the transition, being equal to zero in one of the phases, the symmetrical one, and different from zero in the other phase, the broken-symmetrical one. Since the parameter space is part of a two-dimensional plane, we expect that the two phases be separated by a one-dimensional curve. In fact, we can easily estimate the locus of this phase-transition curve. The argument that leads us to verify that the two phases exist depends crucially on the width of the potential, which is proportional to $\sqrt{-\alpha/\lambda}$. This quantity does not change so long as α and λ are proportional to each other, $\alpha = -\beta\lambda$, so that the location of the points where the transition occurs, separating the two phases, should depend only on the proportionality constant β.

We can estimate this quantity assuming that at the transition the distance between the two wells is of the order of the width σ_0 of the local distribution, whose value is determined predominantly by the derivative term of the action. It is clear that, if the distance between the wells is significantly smaller than σ_0, the local distribution will tend to remain centered around $\varphi = 0$, with its width somewhat reduced, while if the distance is significantly larger than σ_0, the local distribution will tend to shift sideways and fall into one of the wells. In order to better understand this argument it is useful to think about the extreme cases, the one in which the total width of the potential is much smaller than σ_0 and the local distribution is highly squeezed within it, and the one in which the total width of the potential is much larger than σ_0 and the local distribution is completely free to move within the potential and therefore to fall into one of the two wells. This estimate gives us the relation $\sigma_0 \approx \sqrt{-\alpha_c/\lambda_c}$ for the critical values of the parameters or, more precisely, $C_0\sigma_0 = \sqrt{-\alpha_c/\lambda_c}$, where C_0 is some positive constant of the order of 1 that can

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

position of the critical lines will involve some rather intensive and possibly difficult computer work.

Observe that there are many other flows that approach the Gaussian point, besides those defined directly by the free-field model. For example, we have a class of flows that go along the vertical half-axis ($\lambda > 0, \alpha = 0$) in the direction of $\lambda = 0$. This class of flows can produce trivial limits in which the resulting renormalized mass is determined by the dimensionless coupling parameter λ, instead of by α. This phenomenon, which we can see here in a very simple way, is also known by the name of "dimensional transmutation". It is also possible to approach the Gaussian point from the broken-symmetrical phase, from under the critical curve. In this way we may define trivial limits in which the field, although free, has non-zero expectation values $V_R = \langle\phi\rangle$. Except in the case $d = 3$ any limits that are candidates to not being trivial must approach some other point of the critical curve, and not the Gaussian point. If all possible flows of the model turn out to be trivial then we say that the model is trivial. Then, except for the introduction of the concept of spontaneous symmetry breaking, such models are just another way to produce the theory of the free scalar field in the continuum limit. They may still be useful test models on finite lattices, though.

Observe that what we have obtained here is a type of critical behavior just like the one described in the section in [5], where the quantity $v_R = \langle\varphi\rangle$ plays the role played by the magnetization in the case of statistical mechanics, as shown in the first figure of that section. However, we can do a little better here, and continue our heuristic argument in order to get an estimate for v_R as a function of α and λ. First of all let us point out that we should be able to get only the absolute value of v_R and not its sign, since the symmetry can break to either side. Let us therefore estimate v_R^2 and not v_R. It must be zero at the critical line, that is, when the squared width of the potential, $-\alpha/\lambda$, is equal to the squared width of the local distribution, $C_0^2\sigma_0^2$. It must also be equal to zero whenever the potential is less wide than that, that is,

$$v_R^2 = 0 \quad \text{for} \quad -\frac{\alpha}{\lambda} \leq C_0^2\sigma_0^2,$$

characterizing the symmetrical phase. To this we may add that, if the potential is wider than the width of the local distribution, then the distribution should be able to shift to one side by something like the difference between the two, giving as the corresponding estimate for v_R^2 the value

$$v_R^2 = -\frac{\alpha}{\lambda} - C_0^2\sigma_0^2 \quad \text{for} \quad -\frac{\alpha}{\lambda} \geq C_0^2\sigma_0^2,$$

characterizing the broken-symmetrical phase. Note that this formula gives the correct value $v_R^2 = 0$ at the critical line. It also gives the correct value in the case of an extremely wide potential, in which case $C_0^2\sigma_0^2$, which is always of the order of 1, can be neglected by comparison with $-\alpha/\lambda$, and we should have the average v_R of the local distribution sitting at the minimum of the potential, that is, $v_R^2 = -\alpha/\lambda$. In short, we have for v_R^2 the result

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

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

point. Therefore the potential never becomes infinitely flat around its local minima in the broken-symmetrical phase, except in limits that approach the Gaussian point. There are, therefore, no other possible continuum limits with finite masses in this phase, except those that tend to the Gaussian point or to some other point of the critical line.

As a last refining touch of our argument, we may point out its relation to the question of the triviality of the model. Let us look back at figure 1.1.3 and imagine that it represents the critical situation, in which the potential is just wide enough not to squeeze the local distribution. This means that the local distribution is almost free to move, but there is no space for it to actually do so. In this situation the squeezing action of the outer walls of the potential and the spreading action of the central bump are exactly balanced, so that the net action over the width of the local distribution vanishes. However, even without changing the width, the potential can tend to change the *shape* of the distribution. In fact, if we consider in which parts of the local distribution the central bump and the outer walls act, we realize that the bump tends to flatten and spread the top of the Gaussian, while the two outer walls tend to increase the slopes on the two sides of the distribution. This change of shape is exactly what one would expect if the local distribution tended to be more like the function $\exp(-\lambda_R\varphi^4/4)$ than like a Gaussian.

In fact, one would expect λ_R to manifest itself by affecting the form of the complete local distribution. Since α_R is related to the second moment of the distribution, its width, it is reasonable to expect that λ_R is correspondingly related to the fourth moment of the distribution, and hence to the shape of the curve that describes it. However, our experience with α_R shows that, due to the derivative term of the action, the local distribution has a very rigid character in the continuum limit, as a consequence of the requirements of propagation, which requires α_R to go to zero in the limit. There is therefore legitimate doubt that we can have this distribution significantly changed in shape in the continuum limit, to allow for a non-zero value of λ_R, without disturbing the dynamics of propagation and thus ending up with infinitely massive particles and no propagation. Note that if it is true that we must have $\lambda_R \to 0$ in the limit, then it is imperative that λ_R be zero exactly over the critical line, where we already know that α_R is zero, otherwise there would be no possible continuum limits except those going to the Gaussian point.

Of course we cannot resolve this difficult matter with only heuristic arguments. In fact, we will see that perturbation theory is also not enough to handle this issue, and we will have to use computer simulations in order to explore it. However, we can say that, if indeed it turns out that we must have $\lambda_R \to 0$ in the limit, then only theories that can be non-trivial with λ_R going to zero still have a chance of being truly interacting quantum theories. This involves the scaling relations between λ_R and its dimensionfull version Λ_R. As is discussed in problem 1.1.8, in the classical case, this leaves, of all polynomial models φ^{2p}, $p = 2, 3, 4, \ldots$, in all dimensions $d = 3, 4, 5, \ldots$, a single possibility: the $\lambda\varphi^4$ model in $d = 3$.

We end this section with a short discussion of the continuum limit of the classical theory, which requires rewriting the action of the model in terms of dimensionfull

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,

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

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

the dimensionfull field ϕ, that have direct physical relevance, being directly related to the quantities which are observable in the quantum theory. Although we have only shown this fact in the case of the free theory, we will assume that it is true in general, a working hypothesis that we will only be able to confirm a posteriori by the accumulation of calculational experience, numerical or otherwise. We will see that this relation of scale between λ_R and Λ_R will be very useful to enable us to understand heuristically the behavior of the quantum models. Before anything else is done, however, it will be necessary to define in a more precise way the constant λ_R of the quantum theory, which we will do in the third chapter of this book.

Problems

1.1.1. Derive the classical equation of motion for the $\lambda\varphi^4$ model, showing that it is a non-linear equation. Write the equation both on finite lattices, in terms of the dimensionless field, and in the continuum limit, using the dimensionfull field.

1.1.2. Calculate the position, the depth and the bottom curvature of the potential wells in the models $\lambda\varphi^{2p}$, $p = 3, 4, \ldots$, which are defined by the action

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

verifying that the spontaneous symmetry breaking situation is qualitatively similar to that of the $\lambda\varphi^4$ model. Sketch qualitatively the phase diagrams of these models and estimate whether the critical curves in these cases are more or less steep than the critical curve of the quartic model.

1.1.3. Show that the stochastic variable χ introduced in the text for the $\lambda\varphi^4$ model is bound within the interval $[0, \chi_{\max}]$, where $\chi_{\max}$ is a finite quantity which depends on α and λ.

1.1.4. Define a generalization of the stochastic variable χ, which was introduced in the text for the case of the $\lambda\varphi^4$ model, for the case of the $\lambda\varphi^{2p}$, $p = 3, 4, \ldots$ models.

1.1.5. Estimate and sketch the positions, on the phase diagram of the $\lambda\varphi^4$ model, of the curves defined by $\alpha_R(\alpha, \lambda) = C$, where C is a constant. Do it on both sides of the critical curve. Remember that $\alpha_R = 0$ over the critical curve, that $\alpha_R = \alpha$ for the free theory, and that $\alpha_R \approx -2\alpha$ for negative α and $\lambda \approx 0$.

1.1.6. Write the action of the $\lambda\varphi^4$ model in terms of the Fourier transforms $\widetilde{\varphi}$ of the fields and show that, due to the presence of the quartic term, the Fourier modes in momentum space do not decouple from each other when $\lambda \neq 0$.

1.1.7. Show that, in the case $d = 1$, the model $\lambda\varphi^4$ is identical to the quantum mechanics of an anharmonic oscillator.

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,

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

We may now write this in terms of the measure of the free theory defined by S_0, dividing both numerator and denominator by $\int [\mathbf{d}\varphi] e^{-S_0}$ and thus obtaining

$$\langle \mathcal{O} \rangle = \frac{\left\langle \mathcal{O}\, e^{-S_I} \right\rangle_0}{\left\langle e^{-S_I} \right\rangle_0},$$

where the subscript 0 denotes expectation values of the theory defined by S_0,

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

The term S_I of the action is the one that contains the parameter λ, that we presume to be small. However, in general S_I may contain also other parameters, so that in order to enable us to do the development of the perturbation theory in a more organized and explicit fashion it is convenient to use, instead of λ, a new expansion parameter ε that we introduce as follows,

$$f(\varepsilon) = \frac{\left\langle \mathcal{O}\, e^{-\varepsilon S_I} \right\rangle_0}{\left\langle e^{-\varepsilon S_I} \right\rangle_0}. \tag{1.2.2}$$

We have therefore that $f(0) = \langle \mathcal{O} \rangle_0$ and $f(1) = \langle \mathcal{O} \rangle$. Perturbation theory consists of making a series expansion, which we denominate the *perturbative expansion*, of $f(\varepsilon)$ around $\varepsilon = 0$, up to a certain order, followed by the use of the resulting expressions at the point $\varepsilon = 1$. Of course this can only be a good approximation to the complete theory if S_I is a small quantity. Classically we can make S_I small by adjusting the values of λ and any other parameters that it may contain but, as we shall see in what follows, *this is not possible in the quantum theory*. This is the basic fact that is at the root of all difficulties with the perturbative approach to quantum field theory.

In order to understand the origin of the difficulties it is necessary to recall some important properties of the theory of the free scalar field, since we are writing our quantities here in terms of the expectation values of that theory. As we saw in the section in [7], in the case of the dimensions $d \geq 3$ which are the ones of interest for quantum field theory, the quantity $\sigma^2 = \langle \varphi^2 \rangle$, which we denote here by σ_0^2 to record the fact that it is a quantity relating to the free theory, is a finite and non-zero quantity both on finite lattices and in the continuum limit. In addition to this, we showed in the section in [8] that the quantity $\langle [\Delta_\mu \varphi]^2 \rangle$ is also finite and non-zero both on finite lattices and in the limit, in which case it has the value $1/d$. Still in the section in [8] these facts were used to show that both the expectation value of the kinetic part S_K of the action,

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

and the expectation value of the part S_M of the action containing the mass term,

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

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

trivial in that case, since the interaction term would then become vanishingly small in the continuum limit, when compared to the remaining part of the action.

The conclusion to which we are forced is that $\langle S_V \rangle_0$ in fact diverges in the continuum limit as N^d. It is important to observe once more that this divergence is not due to an integration over an infinite volume, because we can do the complete development of the theory within a finite box without any change in this result. This divergence is a property of the continuum limit, an *ultraviolet* characteristic of the theory, in which the influence of the high-frequency and short-wavelength modes of momentum space predominate. It is *not* a property of the infinite-volume limit, that is, of the *infrared* regime of the theory, in which the low-frequency and long-wavelength modes predominate. It becomes clear, therefore, that it is not possible to keep S_V small by mere changes in the parameters α and λ, except if we make them converge rapidly to zero in the continuum limit, which takes us back to the Gaussian point, where all the results are already known, constituting the theory of the free scalar field.

At this point it does not seem that this perturbative technique can end up having any practical use but, in any case, let us proceed with our analysis of the situation. If we consider for a moment the denominator of equation (1.2.1) it is clear that we will have, in the continuum limit,

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

while the perturbative expansion of this quantity, obtained by the series expansion of the exponential function, will contain divergent terms if we keep ε finite and non-zero when we take the limit,

$$\left\langle e^{-\varepsilon S_I} \right\rangle_0 \approx 1 - \varepsilon \langle S_I \rangle_0 + \ldots, \quad \text{where} \quad \langle S_I \rangle_0 \to \infty.$$

We see here that a simple and naive expansion within such a singular structure can make a vanishing quantity appear as a collection of infinities in the terms of the expansion. We can now see that the limit of equation (1.2.1) for $N \to \infty$ is a limit of the form $0/0$. However, it certainly exists, so long as the theory is well defined, which we expect to be true so long as we keep the parameters of the theory within the stable region of the critical diagram. The denominator can be understood as the ratio of the measures of the interacting model and of the free theory,

$$\left\langle e^{-S_I} \right\rangle_0 = \frac{\displaystyle\int [\mathbf{d}\varphi] e^{-(S_0 + S_I)}}{\displaystyle\int [\mathbf{d}\varphi] e^{-S_0}},$$

so that the conclusion we arrive at is that these two measures are related in a singular way in the continuum limit. On any finite lattice the expectation value $\langle S_I \rangle_0$ is finite and we can improve the approximation by decreasing somewhat the parameters α and λ. However, in the continuum limit the only form to avoid the divergence is to

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.

In fact, one can treat the problem at hand by making a change in the normalization of both the numerator and the denominator of equation (1.2.1), eventually obtaining the same results that we will obtain here in a more direct way (problem 1.2.2).

We will examine here the first-order and second-order terms in ε for the expansion of $f(\varepsilon)$, for which we obtain

$$f(\varepsilon) = f(0) + \varepsilon f'(0) + \frac{1}{2}\varepsilon^2 f''(0) + \ldots,$$

where the first three terms contain (problem 1.2.3)

$$\begin{aligned} f(0) &= \langle \mathcal{O} \rangle_0, \\ f'(0) &= -\left[\langle \mathcal{O} S_I \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I \rangle_0\right], \\ f''(0) &= \left[\langle \mathcal{O} S_I^2 \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I^2 \rangle_0\right] - 2\langle S_I \rangle_0 \left[\langle \mathcal{O} S_I \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I \rangle_0\right]. \end{aligned}$$

Making $\varepsilon = 1$ we obtain

$$\begin{aligned} \langle \mathcal{O} \rangle \approx\ & \langle \mathcal{O} \rangle_0 - \left[\langle \mathcal{O} S_I \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I \rangle_0\right] \\ & + \frac{1}{2}\left\{\left[\langle \mathcal{O} S_I^2 \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I^2 \rangle_0\right] - 2\langle S_I \rangle_0 \left[\langle \mathcal{O} S_I \rangle_0 - \langle \mathcal{O} \rangle_0 \langle S_I \rangle_0\right]\right\}. \quad (1.2.3) \end{aligned}$$

This is the approximation for $\langle \mathcal{O} \rangle$ up to the order ε^2, that is, effectively up to the order λ^2. We will use it later on to calculate perturbative approximations for some of the observables of the model.

Observe that it is not to be expected that this expansion may produce a convergent series for the observables of the model. An alternative way to see this is to observe that there cannot be a non-vanishing convergence radius for the series of $f(\varepsilon)$ around $\varepsilon = 0$ in the complex ε plane, because a non-vanishing convergence disk around zero would include negative values of ε, which correspond to points in the unstable region of the parameter plane of the model, where we know that it does not exist. At most what we can hope to obtain are reasonable approximations up to a certain order, which hopefully will be good enough to allow us to form a correct qualitative idea about the behavior of the model. Note that the model would be clearly more useful if it did not cease to exist when we exchange the sign of the coupling constant. One is led to recall that this is the expected situation in electrodynamics, in which we can have charges of either sign.

In order to complete the development of our perturbative ideas, we must now return to the issue of the separation of the action S in parts S_0 and S_I. This separation will depend on whether we want to perform calculations in one or the other of the two phases of the model, the symmetrical phase or the broken-symmetrical phase, whose existence and nature we discussed in section 1.1. In any case S_0 must satisfy the two essential conditions: it must be no more than quadratic on the fields and it must be stable, which means that it must correspond to a well-behaved theory of free fields, having therefore a lower bound.

The issue of stability must be examined carefully at this point. As we saw in section 1.1, in any continuum limit that does not approach the Gaussian point the

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

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

where v_R is the expectation value of the field which, in the absence of any external sources breaking the discrete translation invariance of the lattice, as is our case here, should be constant, having the same values at all the sites. Since v_R is a constant it follows that the derivative term of the action remains unchanged. The polynomial terms which are quadratic and quartic on the fields, however, are transformed according to the relations

$$\begin{aligned}\varphi^2 &= \varphi'^2 + 2v_R\varphi' + v_R^2, \\ \varphi^4 &= \varphi'^4 + 4v_R\varphi'^3 + 6v_R^2\varphi'^2 + 4v_R^3\varphi' + v_R^4.\end{aligned}$$

We may neglect the constant terms, that do not depend on the field, since the exponentials of these terms are constant factors that appear both in the numerator and in the denominator of the ratio that defines the observables, thus cancelling off and not affecting in any way the statistical distribution of the model. Doing this we obtain for the complete action of the model

$$S = \sum_s \left[\frac{1}{2} \sum_\mu (\Delta_\mu \varphi')^2 + v_R \left(\alpha + \lambda v_R^2\right) \varphi' + \frac{\alpha + 3\lambda v_R^2}{2} \varphi'^2 + \lambda v_R \varphi'^3 + \frac{\lambda}{4} \varphi'^4 \right].$$

Since we know that α will always be strictly negative, we introduce now the parameter $\alpha_0 \geq 0$ and separate the action into a free part

$$S_0 = \frac{1}{2} \sum_\ell (\Delta_\ell \varphi')^2 + \frac{\alpha_0}{2} \sum_s \varphi'^2(s) \tag{1.2.4}$$

and an interaction part

$$\begin{aligned} S_I &= \sum_s \left[v_R \left(\alpha + \lambda v_R^2\right) \varphi'(s) + \frac{\alpha - \alpha_0 + 3\lambda v_R^2}{2} \varphi'^2(s) \right. \\ &\qquad \left. + \lambda v_R \varphi'^3(s) + \frac{\lambda}{4} \varphi'^4(s) \right]. \end{aligned} \tag{1.2.5}$$

This is the form of the interaction term to be used in the broken-symmetrical phase of the model. We have therefore a completely well-defined scheme for trying to obtain approximations for the observables of the complete model in the vicinity of the Gaussian point, both in the symmetrical phase and in the broken-symmetrical phase. We must now perform in detail the calculation for some particular observables of the model, always keeping in mind that this is a very singular approximation scheme and that it may turn out that not everything will work as we might hope, in order to verify what we may learn about the structure of the model by means of the use of this technique.

Problems

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:

(a) Show that this scheme would be sufficient to make the perturbative series well-behaved, with finite terms in the continuum limit, so long as, besides keeping at zero the average value of the quantity $\zeta - S_I$, we could also keep the *fluctuations* of this quantity at some finite average sizes around zero. Please note that we are *not* talking about the series being convergent, but only about its individual terms not diverging in the limit.

(b) Show that it is not possible to satisfy this condition in this model. In order to do this consider the calculation of $\langle(\zeta - S_I)^2\rangle_0$ under the condition that $\zeta = \langle S_I\rangle_0$, that is, calculate the quantity

$$\langle S_I^2\rangle_0 - \langle S_I\rangle_0^2$$

and show that it does not have a finite limit when $N \to \infty$.

(c) Repeat the first-order calculations presented in the text, using these ideas and the condition $\langle\zeta - S_I\rangle_0 = 0$ in order to determine ζ, thus showing that exactly the same results presented in the text are obtained in this context.

1.2.3. Perform explicitly the expansion of $f(\varepsilon)$ up to the order ε^2 and derive the form of the three terms that appear in equation (1.2.3).

1.3 Spontaneous Symmetry Breaking

Having developed in section 1.2 the ideas about the perturbative approximation for the observables of the $\lambda\varphi^4$ model, we will now discuss the calculation of some of the observables of the model to first order in ε which, in the cases to be examined here, is also known as the "one-loop" order[2]. The first thing that we will try to calculate will be the position of the critical curve near the Gaussian point. In order to to this we will examine the expectation value of the field,

$$v_R = \langle\varphi\rangle,$$

which functions as an order parameter for the phase transition that exists in the model. Of course, if we have in the model a non-vanishing external source j, then we should expect that v_R is also non-vanishing. The situation of spontaneous symmetry breaking is that in which we have $v_R \neq 0$ even when $j = 0$. Therefore, we will consider here the case $j = 0$ and try to verify whether or not it is possible to obtain solutions of the model with $v_R \neq 0$ in the limit in which $N \to \infty$. Observe that only in this limit of large lattices one can expect to obtain a situation of phase transition, with the existence of two distinct phases in the parameter plane of the model, separated by a phase-transition curve.

[2]These calculations were developed in collaboration with Dr. André Cavalcanti Rocha Martins.

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)

$$\left\langle \varphi'(0) \sum_s \varphi'(s) \right\rangle_0 = \frac{1}{\alpha_0}. \tag{1.3.2}$$

For the other expectation value we obtain (problem 1.3.3)

$$\left\langle \varphi'(0) \sum_s \varphi'^3(s) \right\rangle_0 = \frac{3}{N^d \alpha_0} \sum_{\vec{k}} \frac{1}{\rho^2(\vec{k}) + \alpha_0}. \tag{1.3.3}$$

Observe how we avoided infrared problems in both cases, by the introduction of the non-zero parameter α_0. In these calculations all the strong divergences due to the behavior of S_I, which consist of terms proportional to N^d, cancel out. This fact corresponds, in the usual language of the traditional approach to the theory directly in the continuum, to the cancellation of the so-called "vacuum bubbles", and is a direct consequence of the fact that we are expanding a ratio of two functional integrals. We have therefore for our tadpole equation

$$\frac{(\alpha + \lambda v_R^2)}{\alpha_0} + \frac{3\lambda}{N^d \alpha_0} \sum_{\vec{k}} \frac{1}{\rho^2(\vec{k}) + \alpha_0} = 0.$$

We recognize now that the sum over the momenta is our already well-known quantity σ_0^2, the square of the width of the local distribution of the fields in the measure of S_0. We obtain therefore, substituting in terms of σ_0^2 and cancelling the factor of $1/\alpha_0$,

$$\lambda v_R^2 + \alpha + 3\lambda \sigma_0^2 = 0. \tag{1.3.4}$$

This equation gives us v_R for small values of $-\alpha$ and λ in the broken-symmetrical phase.

Let us consider here the issue of the dependence of this result on α_0. Observe that the result does not depend explicitly on α_0, but it may depend on this parameter through the squared width σ_0^2. For finite N the width does indeed depend on α_0 but, as was shows in the section in [11], in the continuum limit it does not depend on this parameter, so long as we make it go to zero sufficiently fast. More precisely, it suffices that we make $\alpha_0 = m_0^2 L^2/N^2$, for some finite m_0, for the limit to be completely independent of the value of m_0. The mass parameter m_0 could even be chosen to have the same value as the renormalized mass m_R of the model, but there is no need for this coming from this calculation, all we know up to now is that m_0 must be finite. Note that the need to choose α_0 dependent on N in a certain way in order to make the results independent of m_0 is already a first indication that the perturbative expansion is not completely well-behaved, since there should be no dependence at all on α_0.

Going back to the analysis of the critical behavior of the model, if we impose now that the *only* possible value for v_R be zero, we obtain from equation (1.3.4), by setting $v_R = 0$ in it, the equation of the critical curve, to wit

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

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

d	$\tan(\theta)$	θ (degrees)
3	$\simeq 1.3189$	$\simeq 52.83$
4	$\simeq 2.1515$	$\simeq 65.07$
5	$\simeq 2.8828$	$\simeq 70.87$

Table 1.3.1: Table of the slopes of the critical curves at the Gaussian point.

continuum limit. Hence, if the perturbative technique establishes at least a first-order approximation for the result of the complete model, then this result should be exact at the Gaussian point (problem 1.3.4). We will see later on that it is consistent with the results obtained by means of mean-field techniques and of stochastic simulations. In the case of the stochastic simulations realized so far, it has been verified that it is particularly difficult to execute them close to the Gaussian point, due to the fact that in that region the potential wells of the model become very shallow, which makes it more difficult to control the statistical errors. Therefore, up to now it has not been possible to do more than to confirm qualitatively this result with the stochastic simulations.

Using the asymptotic values of $\sigma_0^2(N)$ in the dimensions $d = 3$ to $d = 5$, presented in the section in [11], we obtain the results shown in table 1.3.1 for the slope, where θ is the smaller angle that the tangent line to the critical curve at the Gaussian point makes with the negative α semi-axis. It is interesting to observe that in the cases $d = 1$ and $d = 2$, since $\sigma_0(N)$ diverges, the slopes go to zero in the limit and the critical curve collapse onto the negative α semi-axis, where the model does not exist because this semi-axis is part of the unstable region. One might consider the interpretation that this is the perturbative way of verifying that the $\lambda\varphi^4$ model does not really exist as a quantum field theory for $d < 3$.

Our next objective is to calculate the propagator of the model, which we will do first in the symmetrical phase. We will denote the dimensionless two-point function of the complete model by

$$g(\vec{n}_1, \vec{n}_2) = \langle \varphi(\vec{n}_1)\varphi(\vec{n}_2) \rangle.$$

To order zero we simply have that $g(\vec{n}_1, \vec{n}_2) = g_0(\vec{n}_1, \vec{n}_2)$, where

$$g_0(\vec{n}_1, \vec{n}_2) = \langle \varphi(\vec{n}_1)\varphi(\vec{n}_2) \rangle_0,$$

so that in this order we have the results of the free theory, $\alpha_R = \alpha$ and a simple pole with its residue equal to one (apart from the normalization factor of $1/N^d$) in the region of imaginary momenta $\rho^2(\vec{k}) = -\alpha_R$. Note that this first-order result is not sufficient to allow us to take the continuum limit, because we know that α must become negative in the limit, while α_R cannot be negative. In the next-order approximation, using equation (1.2.3) up to first order, we will have a result that we shall denote by $g_1(\vec{n}_1, \vec{n}_2)$, with

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

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

of this equation we recognize it as the expression in the equation of the critical curve that we calculated before, which shows that it does in fact go to zero in the $N \to \infty$ limit, so long as we take the limit in such a way that the parameters of the theory approach the critical curve. We therefore have here some perturbative evidence that the phase transition of the model is indeed of second order and we see once more why it is necessary to take the system to the critical situation in the continuum limit. In addition to this, as we already discussed before in this section, the expression in the right-hand side of the equation is positive in the symmetrical phase, showing that α_0 will be approaching zero by positive values and thus establishing the consistency of this choice for α_0. Observe that, with this choice for α_0, we may write the result for the propagator as

$$g_1(\vec{n}_1, \vec{n}_2) = \frac{1}{N^d} \sum_{\vec{k}} e^{\imath \frac{2\pi}{N} \vec{k} \cdot (\vec{n}_1 - \vec{n}_2)} \frac{1}{\rho^2(\vec{k}) + \alpha_R}, \tag{1.3.10}$$

that is, we get a propagator with form identical to that of the propagator of the free theory, with a renormalized mass m_R, where we see that $\alpha_R = \alpha_0$ and the renormalized mass is given by

$$m_R^2 = \lim_{N \to \infty} N^2 \alpha_R / L^2.$$

Formally, we may try to understand the expression $\alpha_R = \alpha + 3\lambda\sigma_0^2$ for the renormalized mass parameter α_R as the sum of a zero-order term α and a first-order term proportional to λ. However, in truth this is misleading, because we must recall that the parameter α is in fact negative in any continuum limit and hence that the first-order term cannot be considered as a small correction to the situation in the theory of the free field, in which α must be positive. We see that, in spite of the fact that we have developed this approximation technique in the lines of an expansion in a perturbative series, the resulting object has a character rather different from the expected.

We will see later on that the results of this process of approximation for the renormalized mass agree surprisingly well with the results of the stochastic simulations. In particular, note that the result indicates a unit residue for the pole of the propagator, exactly as in the free theory, α_R being the only non-trivial parameter that appears. This unit residue is also found in all the stochastic simulations, within the statistical errors. Judging by the form of the propagator, one would say that the spectrum of the theory seems to be that of free particles with mass m_R. One might consider interpreting this as an indirect perturbative indication related to the underlying triviality of the model. At least, the result for the residue is compatible with it.

It is interesting to try to understand in clearer physical terms the nature of the approximation technique that we have developed. The crucial point for the success of the technique is the choice of α_0, which ends up being equivalent to a preliminary implicit choice $\alpha_0 = \alpha_R$, to be resolved after the end of the calculation, a possibility that was suggested in section 1.2. From the very beginning we are trying to

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

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

which is a positive quantity in this phase. Once again the expression in the equation of the critical curve appears, showing once more that α_R will go to zero when we approach this curve in the continuum limit, this time by the other side, from the broken-symmetrical phase. The factor of 2 that appears in this result confirms once more our heuristic expectations and, as we will see later on, it also matches with surprising precision the numerical results in this phase.

Observe that, since σ_0^2 is a function of α_R, both this result and the result for the symmetrical phase are not explicit solutions for α_R but rather equations that determine α_R in an implicit way,

$$\alpha_R - \alpha - 3\lambda\sigma_0^2(\alpha_R) = 0$$

in the symmetrical phase and

$$\alpha_R + 2\alpha + 6\lambda\sigma_0^2(\alpha_R) = 0$$

in the broken-symmetrical phase, where

$$\sigma_0^2(\alpha_R) = \frac{1}{N^d}\sum_{\vec{k}} \frac{1}{\rho^2(\vec{k}) + \alpha_R}.$$

It is not difficult to determine the existence, the number and the character of the solutions of these equations, if one separates from the sum in σ_0^2 the term containing the zero mode, and to find these solutions on finite lattices by numerical means (problems 1.3.8, 1.3.9, 1.3.10 and 1.3.11). We show in figure 1.3.1 a graph with some of the numerical solutions, illustrating their behavior for lattices of increasing size. In this graph, instead of the usual Cartesian coordinates α and λ in the parameter plane of the model, we use polar coordinates centered at the Gaussian point, with the radius r given by $\sqrt{\alpha^2 + \lambda^2}$, and the angle θ defined as the angle between the radius vector (α, λ) and the negative α semi-axis.

The equation for the symmetrical phase has two solutions, but only one of them is positive and hence physically meaningful. While the positive solution remains finite and non-vanishing in the $N \to \infty$ limit, the negative solution becomes identically zero in the limit. Note that this equation has solutions over the whole parameter plane and not only in the symmetrical phase. The curves corresponding to this solution are the ones with their maximums at the right in figure 1.3.1. Of course this solution only has meaning in the symmetrical phase, but since this is not a well-defined concept on finite lattices, in order to determine the range of validity of the solution on finite lattices, from the point of view of perturbation theory, we must first discuss the solutions in the broken-symmetrical phase. The equation for the broken-symmetrical phase has two positive solutions, a small one and a large one, but only for certain values of parameters, thus defining a perturbative broken-symmetrical phase even on finite lattices. In the complementary region of the parameter plane, which we might call the perturbative symmetrical phase, the equation has no real solutions. On each finite lattice the position of the curve separating these two regions

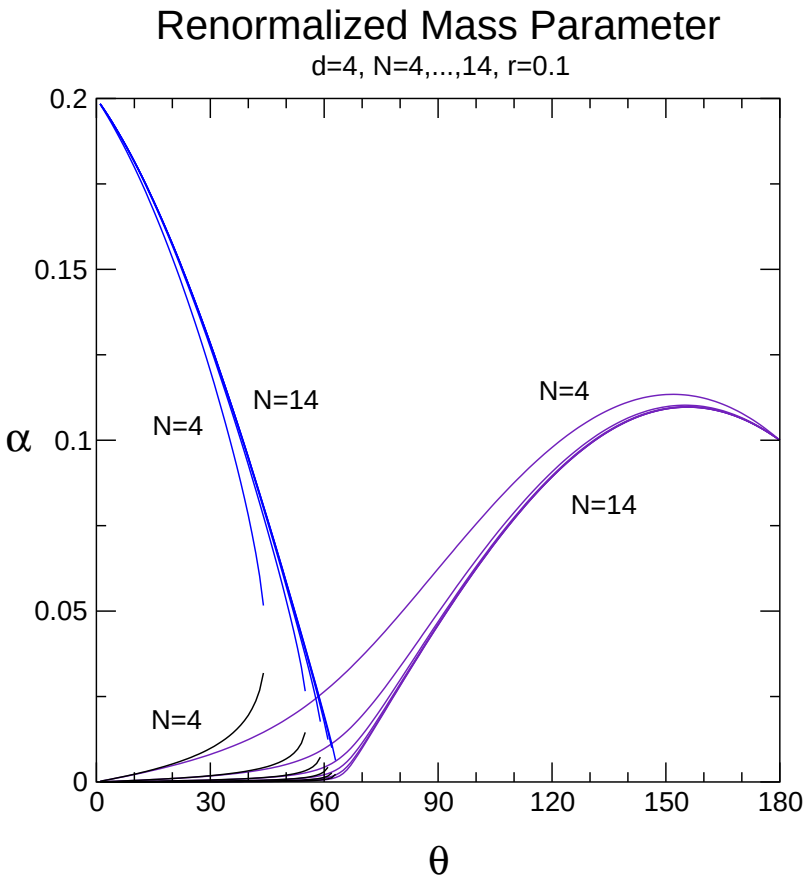

Figure 1.3.1: Perturbative results for the renormalized mas parameter α_R. The curves on the left part are in the broken-symmetrical phase. In this graph α and λ are represented by the equivalent parameters r and θ, which are defined in the text. The values of the angle θ are given in degrees.

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^o 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^o 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

approaches the free theory on finite lattices, with this new value for the dimensionless mass parameter. As one can see in figure 1.3.1 in this case the large solution does indeed approach $\alpha_R = -2\alpha$ as expected, where $-2\alpha = 2r = 0.2$. Just as in the case $\theta = 180^o$ the solution seems to be exact in this case, because for θ approaching 0^o we have a vanishing λ and the distribution tends to become purely Gaussian, so that the Gaussian approximation tends to become a perfect one. It is therefore to the two large solutions that we should attribute physical meaning, using the broken-symmetrical solution where it exists and using the symmetrical solution where the broken-symmetrical solution does not exist, and therefore truncating in this way the symmetrical solution. Note that there is a discontinuity between the two solutions at the transition point, but this discontinuity vanishes in the $N \to \infty$ limit. In this limit the edges of the two curves approach the value $\alpha_R = 0$ at the critical line.

In the $N \to \infty$ limit σ_0^2 becomes independent of α_R, the critical transition becomes completely well-defined and the relations between α_R and the parameters (α, λ) become linear, in either phase. If we define $\sigma_\infty^2 = \sigma_0^2(N \to \infty)$, then we have in the symmetrical phase

$$\alpha_R = \alpha + 3\sigma_\infty^2 \lambda,$$

and in the broken-symmetrical phase

$$\alpha_R = -2\alpha - 6\sigma_\infty^2 \lambda,$$

which are equations of planes over the (α, λ) parameter plane. The three-dimensional graph of $\alpha_R(\alpha, \lambda)$ over the parameter plane is composed of pieces of two planes that intersect within the (α, λ) plane at the critical line. The first one is a piece of the plane defined by the critical line and the line $\alpha_R = \alpha$ within the vertical (α, α_R) plane, the second one is a piece of the plane defined by the critical line and the line $\alpha_R = -2\alpha$ in that same vertical (α, α_R) plane. The relevant part of the first plane is the part that stands over the symmetrical phase, the relevant part of the second plane is the part that stands over the broken-symmetrical phase.

Note that both on finite lattices and in the $N \to \infty$ limit the perturbative solution for the expectation value v_R of the field is proportional to the value of the renormalized mass in the broken-symmetrical phase, that is, we have

$$v_R^2 = \frac{\alpha_R}{2\lambda}.$$

Due to the extra dependence on λ, we can have independent values of the two dimensionless renormalized quantities, v_R and α_R. Whether or not the same is true for the corresponding dimensionfull quantities V_R and m_R depends on the dimension d of space-time, because the dimensions of the field and therefore of V_R depend on it (problem 1.3.12).

The final conclusion of this effort is that the perturbative technique of Gaussian approximation allows us to calculate in a useful way the observables related to the aspects of propagation of particles and to the aspects of spontaneous symmetry

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,

$$\alpha_R - \alpha - 3\lambda\sigma_0^2(\alpha_R) = 0,$$

were we recall that

$$\sigma_0^2(\alpha_R) = \frac{1}{N^d}\sum_{\vec{k}} \frac{1}{\rho^2(\vec{k}) + \alpha_R},$$

has a single positive solution for each pair of values (α, λ) in the stable region of the parameter plane of the model. In order to do this, remember that the parameter α_R has to be positive or zero and consider the behavior of the left side of the equation when $\alpha_R \to 0$ and when $\alpha_R \to \infty$. Remember that the sum that defines the quantity σ_0^2 includes the zero mode and write it as

$$\sigma_0^2 = \sigma_0'^2 + \frac{1}{N^d \alpha_R},$$

where $\sigma_0'^2$ has a finite limit for $\alpha_R \to 0$. Show also that there is a second solution, which is negative (and hence destitute of any physical meaning) and which becomes identically zero in the $N \to \infty$ limit.

1.3.9. ($\star$) For given α, λ and N, write a program to solve numerically the equation

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

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

1.3.10. Show that the perturbative equation which determines the renormalized mass parameter α_R in the broken-symmetrical phase, as a function of α and λ,

$$\alpha_R + 2\alpha + 6\lambda\sigma_0^2(\alpha_R) = 0,$$

has two different real and positive solutions for some pairs of values (α, λ) in the stable region of the parameter plane of the model, and no real solutions for other pairs of values. Use the same ideas and techniques that were suggested in problem 1.3.8. Show that the condition on α and λ for the existence of solutions can be written in an implicit way, which depends on α_R on finite lattices, as

$$(\alpha + 3\lambda\sigma_0'^2)^2 \geq \frac{6\lambda}{N^d},$$

and interpret the meaning of this condition on the continuum limit. Show that when the equality holds in the condition above the renormalized mass parameter is given by $\alpha_R = \sqrt{6\lambda/N^d}$, and hence that it goes to zero at the critical curve in the continuum limit. Show also that the smaller of the two solutions becomes identically zero in the $N \to \infty$ limit.

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.

1.3.16. Show, using the first-order perturbative results obtained in the text, that in the complete model the observable σ_1 given by $\sigma_1^2 = \langle\varphi^2\rangle$ in the symmetrical phase and by $\sigma_1^2 = \langle\varphi'^2\rangle$ in the broken-symmetrical phase, is equal, in either case, to the observable σ_0^2 of the free theory. Observe that this shows that the field fluctuates in a similar way in either model, undergoing fluctuations with the same typical size.

1.3.17. Show that it is possible to take continuum limits to the Gaussian point over the positive λ semi-axis, that is, keeping $\alpha = 0$ constant during the limit. Determine how to take this type of limit so that the renormalized mass is finite and non-zero, that is, discover how $\lambda(N)$ must be so that m_R has a finite and non-zero limit. This type of limit, which produces a non-zero renormalized mass without involving any parameters with dimensions of mass from the corresponding classical theory, is related to what has been conventionally called the phenomenon of "dimensional transmutation".

Chapter 2

The Sigma-Model Limit

In this chapter we will discuss the infinite-coupling limit of the polynomial models, which is, in a way, the antithesis of the perturbative approach. We will see that in the $\lambda \to \infty$ limit there is an *exact* representation of the polynomial models, which is another class of non-linear models of scalar fields, known as the *sigma models*. As we will see in the simple case we will deal with here, in this limit the quantum $\lambda\varphi^4$ models can be identified with a simple example of these sigma models. In the one-component models that we are examining here these sigma models reduce to the well-known Ising models. Note that we are dealing here with an equivalence between the *quantum* versions of the models, not between the corresponding classical field theories, which are quite different from one another.

2.1 Derivation of the Ising Model

In this section we will take the $\lambda \to \infty$ limit of the $\lambda\varphi^4$ model. This can be done for more general models than the one-component $\lambda\varphi^4$ model we are examining here, namely the multi-component $\lambda\varphi^4$ models which are invariant by the $SO(\mathfrak{N})$ groups of transformations. In general the $\lambda \to \infty$ limit of these models will take us to the corresponding $SO(\mathfrak{N})$-invariant sigma models. In our case here, however, we will deal only with the one-component $\lambda\varphi^4$, which is invariant by the sign reflections. The simple discrete set of transformations given by the identity and the reflection also forms a group of transformations, a discrete group which is known by either $O(1)$ or $\mathbb{Z}_2$. In this case the corresponding sigma model is simply the Ising model which was mentioned in [13]. In this way we will establish that we can use the Ising model as a direct representation of the infinite-coupling limit of the $O(1)$-symmetrical $\lambda\varphi^4$ model.

The Ising model can be obtained from the $\lambda\varphi^4$ polynomial model in the limit in which the coupling parameter λ tends to positive infinity over negative-slope straight lines in the parameter plane of the $\lambda\varphi^4$ model. These lines must exist only for $\lambda > 0$, because otherwise they would cross the region where the model is unstable. Also, their slopes must be strictly negative (not zero) and finite, which rules out horizontal and vertical lines in the parameter plane. Since the slopes must

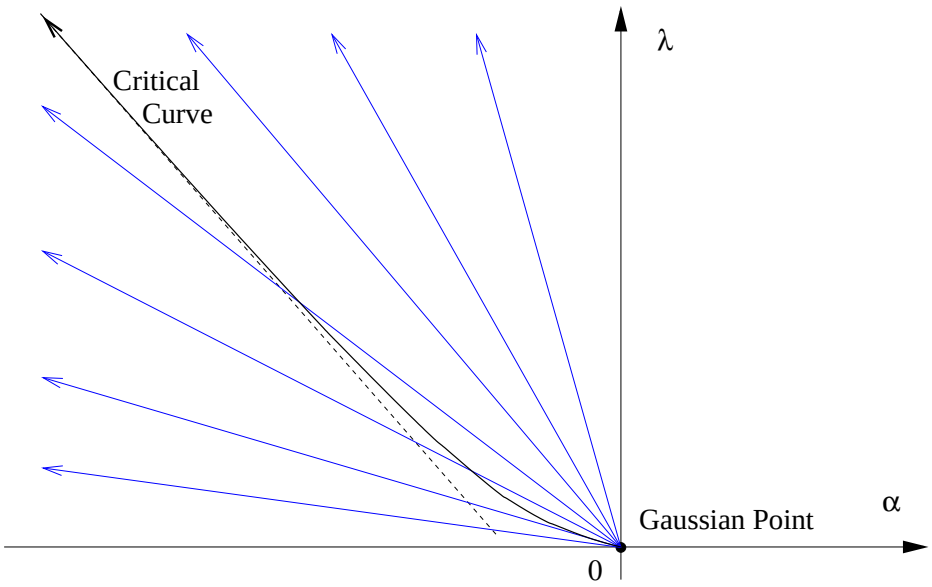

Figure 2.1.1: Limits leading from the polynomial models to the sigma models by means of negative-slope straight lines departing from the Gaussian point. Note that lines parallel to the α and λ coordinate axes are excluded.

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,

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

The functional integrals that appear in the definition of the observables of the quantum theory may now be written as

$$\int [\mathbf{d}\varphi]\; e^{-S[\varphi]} = \int [\mathbf{d}\varphi]\; e^{-S_V[\varphi]}\; e^{-S_K[\varphi]},$$

so that, including the exponential of S_V in the measure of the distribution of the model, we may write this distribution as

$$[\mathbf{d}\varphi]\; e^{-S[\varphi]} = \left[\mathbf{d}\varphi\; e^{-V(\varphi)}\right] e^{-S_K[\varphi]}.$$

Since we may multiply this quantity by any constant independent of the fields without changing the observables, we may write the measure as

$$[\mathbf{d}\varphi\; \varrho(\alpha, \lambda, \varphi)],$$

where $\varrho(\alpha, \lambda, \varphi)$ is the local part of the distribution, which we thus include in the measure, normalized as

$$\varrho(\alpha, \lambda, \varphi) = \frac{e^{-V(\varphi)}}{\displaystyle\int_0^\infty \mathrm{d}\varphi\; e^{-V(\varphi)}}. \tag{2.1.1}$$

Observe that the integration runs only over the positive values of φ and hence can be interpreted as an integration over the absolute value of the field. With this the following normalization condition for ϱ holds,

$$\int_0^\infty \mathrm{d}\varphi\; \varrho(\alpha, \lambda, \varphi) = 1.$$

The integration over the absolute value of the field, when generalized to the more complex models having $SO(\mathfrak{N})$ symmetry, corresponds to the separation of the integration variables into a radial part and an angular part. In this way the arguments being presented here can be easily generalized to that case. In general the radial variable will be the modulus of the $\mathfrak{N}$-dimensional field vector in the internal symmetry space, which in our case here reduces to the absolute value of φ. Also, in our discrete $O(1)$-symmetrical case the integral over the angular part reduces to a sum over only two points, corresponding to rotations by each one of the two angles 0 and π, and hence to the two possible signs, $\cos(0) = 1$ and $\cos(\pi) = -1$. Therefore, in our present case the integral over all the values of φ at each site is being decomposed in this way into an integral over the positive values of φ and a sum over the two possible signs. Since both $\varrho(\alpha, \lambda, \varphi)$ and $V(\varphi)$ are even functions of φ, they are in fact functions only of the absolute value of φ, and independent of its sign, so it is not necessary to make their dependence on the absolute value explicit. Given all this, we may write the functional integrals as

$$\int [\mathbf{d}\varphi\; \varrho(\alpha, \lambda, \varphi)]\; e^{-S_K[\varphi]} = \sum [\psi = \pm 1] \int [\mathbf{d}|\varphi|\; \varrho(\alpha, \lambda, \varphi)]\; e^{-S_K[\psi\varphi]},$$

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

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

where the remaining part of the measure can be written explicitly as

$$[\mathbf{d}|\varphi|\ \delta(\varphi^2-\beta)] = \prod_s \mathrm{d}|\varphi(s)|\ \delta[\varphi^2(s)-\beta],$$

where the Dirac delta functions imply a condition of constraint on the fields, $\varphi^2 = \beta$, or $\varphi = \sqrt{\beta}$, since the sign of φ is being considered explicitly and hence φ is positive. We may now use the Dirac delta functions to perform all the integrations over $|\varphi|$, thus obtaining for the expectation values

$$\langle \mathcal{O} \rangle_N = \frac{\sum [\psi = \pm 1]\ \mathcal{O}[\sqrt{\beta}\psi]\ e^{-S_K[\sqrt{\beta}\psi]}}{\sum [\psi = \pm 1]\ e^{-S_K[\sqrt{\beta}\psi]}}.$$

We may now examine the form of the action S_K under these conditions, in order to simplify it and exhibit it in a more familiar form. In terms of the new variables ψ and the parameter β this action can be written as

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

where the new field variables $\psi = \varphi/\sqrt{\beta}$ satisfy the constraint $\psi^2 = 1$, since they are just signs, and the parameter β appears now multiplying the action, just like the parameter $\beta = 1/(kT)$ of statistical mechanics. Although we may want to think of β as the inverse of a fictitious temperature, in order to guide our intuition about the behavior of the models, we should remember that our β is, in truth, related to the mass parameter α and the coupling constant λ, and *not* to any truly physical temperature related to the dynamical system we are studying. The phenomenon that something relating to a coupling constant appears multiplying the action due to a scaling redefinition of the fields is typical of the gauge theories, as we may see in future volumes.

Note that, although $S_K[\psi]$ is a purely quadratic functional of ψ, the model is *not* the free theory, due to the fact that the field ψ satisfies an equation of constraint, and is not, therefore, a free real variable. This is a situation in which the non-linearities, instead of being introduced by a term in the action, are introduced instead by the measure of the functional integral, which is where the constraint is implemented in the quantum theory. We may now perform one more transformation of the action of the model, with the intention of showing in a clearer way its relation with the Ising model of statistical mechanics. If we expand the squares of the derivatives contained in the action, we get

$$(\Delta_\ell \psi)^2 = \psi^2(\ell_-) - 2\psi(\ell_-)\psi(\ell_+) + \psi^2(\ell_+),$$

where $\psi(\ell_-)$ and $\psi(\ell_+)$ are the fields at the two ends of the link ℓ. Using now the equation of constraint $\psi^2 = 1$ we see that the two terms containing the squares are constant, independent of the fields, which means that they can be neglected in the

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

the potential term S_V. We may therefore introduce external sources into the resulting Ising model exactly as we would do in the original polynomial model, and the whole functional generator formalism is made available for the analysis of the Ising model, without any change. Hence the Ising model is in fact an exact and complete direct representation of the infinite-coupling limit of the corresponding $\lambda\varphi^4$ polynomial model. This relationship can be generalized to the multicomponent $SO(\mathfrak{N})$-invariant polynomial models and the corresponding sigma models. It can also be generalized to models with larger powers of the fields. It is possible to show (problem 2.1.5) that the Ising model can also be obtained as the $\lambda \to \infty$ limit of the $\lambda\varphi^{2p}$ polynomial models, with $p = 3, 4, 5, \ldots$, in a way which is completely analogous to the $p = 2$ case that we examined here.

Another aspect which we must examine here is the one relating to the superposition process involved in the definition of the block variables. Once again this is inherited by the Ising model from the polynomial model, so we still have a simple linear superposition of the fields ψ at the various sites within the block. Note that although these fundamental fields satisfy the constraint $\psi^2 = 1$, the same will not be true for the block variables. If we consider the process of linear superposition of the fields within a block, in order to define a block variable, it is evident (problem 2.1.6) that the sum and the average of a collection of signs $\psi = \pm 1$ will not itself have unit absolute value. If the fields are distributed is a very random way, without any appreciable alignment, their average will tend to have an absolute value much smaller than 1. Only in the opposite case, when the fields are highly aligned, the absolute value of the average will tend to 1. The absolute value of the sum may have any value, either larger or smaller than 1.

We see therefore that in general the introduction of external sources will cause the block variables, which are the variables in terms of which we should interpret the theory, to deform to any value, without respecting an equation of constraint. In fact, they will behave much like the corresponding variables of the polynomial models. In short, we see that the constraint that appears for the fundamental field in the large-coupling limit does not survive the block-variable superposition process and that the Ising model we get in the limit is not fundamentally different from the $\lambda\varphi^4$ polynomial model it derives from. Hence we confirm that the Ising models are not to be seen as a completely different class of models, but as a way to examine directly the behavior of the polynomial models in the $\lambda \to \infty$ limit. Since, as was mentioned before, the Ising models can also be obtained as the limits of the $\lambda\varphi^{2p}$ polynomial models for any $p \geq 2$, they become a tool for the examination of the large-coupling limit of a large class of models.

We will finish this section with some comments about the critical behavior of the models. As we shall see later, approximate calculations based on the mean-field technique show that the Ising models have well-defined critical points β_c for dimensions $d \geq 3$. Since they are the $\lambda \to \infty$ limits of the polynomial models, we see that the perturbative Gaussian approximation also predicts well-defined critical points for the Ising models, although they are infinitely distant from the Gaussian point. In fact, the two predictions do not differ very much from each other, and are

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).

2.1.5. Repeat in a qualitative way the derivation presented in the text for the case of the $\lambda\varphi^4$ models, in order to show that the Ising models are obtained, in limits in which $\alpha \to -\infty$ and $\lambda \to \infty$, with constant $-\alpha/\lambda$, from the corresponding $\lambda\varphi^{2p}$ models, for any $p \geq 3$.

2.1.6. Show that the average of the field ψ over any number of sites has an absolute value smaller than or equal to 1. Show also that the absolute value of the sum of ψ over a block of sites may have any positive value in the continuum limit.

2.2 The Mean Field Method

In this section we will introduce the mean-field approximation method. This is an approximation process which is not perturbative, not being based on the Gaussian model. It is a traditional approximation method of statistical mechanics, which is also useful for obtaining non-perturbative approximations in quantum field theory. It is easy to use it to obtain approximations for local quantities defined on lattice sites, such as, for example the expectation value of the field. It is a well-known fact that the results of the method tend to improve with the increasing dimension of the space where the models are defined. Usually the results are reasonably good in three dimensions and even better in four dimensions, while the method often fails completely in one and two dimensions. There are even some speculations that, for some quantities, the method becomes exact in sufficiently large dimensions. In its usual formulation this method does not establish a series of successive approximations, but rather a single approximation, making it difficult to evaluate the errors involved in this approximation. We will introduce here an extension of the method, that improves the situation and allows us to understand its bad performance for small dimensions.

The formulation we will present is specifically for systems defined on the Euclidean lattice, with interactions only between next neighbors. As we will see, it is intimately related to stochastic simulations of the systems on finite lattices in which one uses a certain type of fixed boundary conditions, which we denominate *self-consistent* boundary conditions. These structures on finite lattices constitute an extension of the usual mean-field method and, unlike the usual method, they give us a whole series of successive approximations. The first of them will be the usual approximation, while the subsequent approximations converge to the exact solution of the models within a finite box with fixed boundary conditions, as the lattice spacing is decreased. This extension of the method is similar, but not identical, to other well-known extensions of the traditional method in statistical mechanics, such as the Oguchi method and the Bethe, Peierls, Weiss and Kikuchi [29] method. This approximation method can be used both in the polynomial models and in the sigma models, we will assume only that we have an $O(1)$ model of scalar fields, on a lattice of dimension d, with the usual forms of the action. As we saw in sections 1.1 and 2.1, the action of any of these models can always be separated in two parts, a strictly local one and one involving only interactions between next neighbors, which

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

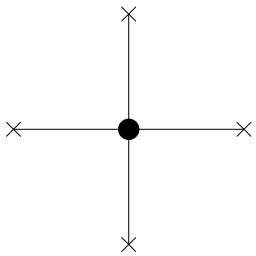

Figure 2.2.1: Lattice with $N = 1$ and fixed boundary conditions, related to the usual mean-field approximation.

at any of the neighboring sites, assuming that they are all equivalent by discrete translation invariance. This means that we are using for the external field the value $j_{\mathrm{MF}} = 2d\langle\varphi\rangle$. The calculation of this average value is made, in the context of this method, a-posteriori and in a self-consistent way: one calculates the average value at the active site located at $\vec{n}$ and imposes that this value be equal to the average value at the neighboring sites.

In this way we replace the detailed interactions between the field at each site and the fields at the neighboring sites by an interaction at each site with a background field, which does not fluctuate, thus rendering the problem mathematically more tractable. It is interesting to observe that the spirit of this approximation is somewhat different from the spirit of the usual approximation in statistical mechanics, in which we consider the interaction of the dynamical variables with an external mean field, ignoring completely that the interactions are established through the links of the lattice. Observe that, in our case here, the dimension of the lattice appears explicitly in the approximation. However, this distinction will only be really relevant when we consider the extension of the method to clusters of sites. For the case discussed so far, in which only the field at a single site is kept active, that is, undergoing statistical fluctuations, the two methods are identical. They are known in statistical mechanics as the *constant coupling method*, which was developed by Yvon, Nakamura, Kasteleijn, Van Kranendonk, Kikuchi and Callen [29].

This replacement of the interactions between next neighbors by an interaction with a non-dynamical field that undergoes no fluctuations is clearly a very radical change and it is rather surprising that it can produce good results, even if only for some observables. In particular, since the dynamical fields at each site interact only with the constant background field and no longer with each other, it is clear that the fields at the various sites will become completely uncorrelated from each other in this approximation, so that the calculation of correlation functions is out of the question. There is, however, an alternative interpretation of the method, which will allow us to extend it to clusters of active sites and hence to recover the correlations among sites. This alternative interpretation, which changes nothing in the mathematics involved, is that the fields at all the sites are frozen at their average values, except for a single arbitrarily chosen site, which remains active. Since in the traditional mean-field method all the sites are equivalent and uncorrelated from

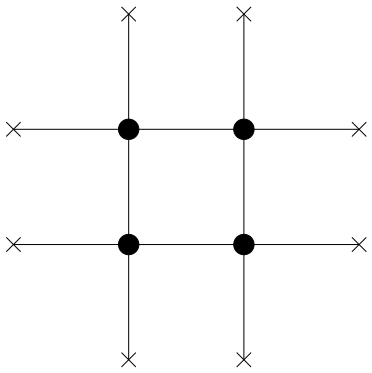

Figure 2.2.2: Lattice with $N = 2$ and fixed boundary conditions, which constitutes the first extension of the mean-field method.

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,

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

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.

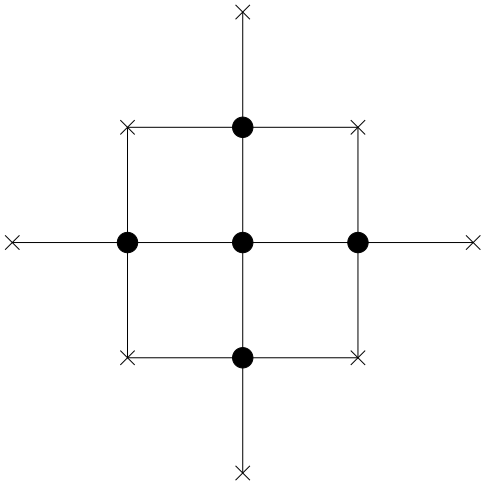

Figure 2.2.3: Modified lattice used in the Bethe-Peierls cluster method.

This extension of the mean-field method is similar to the method of clusters of Bethe, Peierls, Weiss and Kikuchi in statistical mechanics, in which groups of connected sites are considered. The first cluster considered in this method is the Bethe-Peierls cluster, which is shaped like a diamond as shown, within the context of our lattice representation, by the diagram in figure 2.2.3, including the border sites, which are not part of the original cluster method. Larger clusters with formats similar to this one may also be considered and used for analytical calculations. However, the amount of analytical work involved is usually quite large, to achieve only modest gains in the quality of the results obtained.

In the general case our extension is not identical to the traditional cluster method, because in our case the interaction of the cluster with the mean field is established only through the border, not by means of an external mean field that acts also on the internal sites of the cluster, which have no direct contact with the border. This kind of internal site appears in the Bethe-Peierls cluster and also in the cubical clusters starting from $N = 3$, as illustrated in the diagram in figure 2.2.4. The two methods also differ regarding the type of self-consistency condition which is imposed. In the case of the Bethe-Peierls cluster, rather than adjusting the external mean field so that it becomes consistent with the average value of the fields at the active sites, what is done is to adjust it so that the average of the field at the central site is identical to the average of the field at the other $2d$ sites of the cluster, which are all equivalent to each other due to the symmetry of the cluster. Hence, what one actually imposes in this case is that the normal derivative of the average value of the field vanish at the border.

It is clear that the $N \to \infty$ limit of our sequence of cubical clusters produces exactly the continuum limit within a finite cubical box with a certain type of self-consistent boundary conditions. Let us now discuss, in greater detail, the self-consistency condition to be imposed on the border sites. The usual $N = 1$ mean-field approximation is sufficient for the calculation of strictly local quantities, defined at

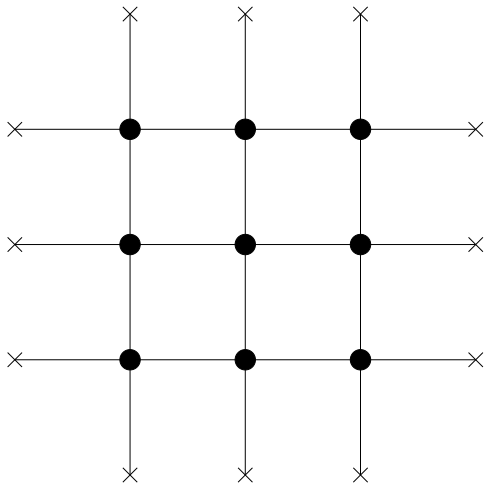

Figure 2.2.4: Lattice with $N = 3$ and fixed boundary conditions, which constitutes the second extension of the mean-field method.

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

value at the border, the value at the border is modified so as to coincide with the value measured in the interior. This is done many times at regular intervals, so that eventually the adjustment of the value at the border becomes negligible and the border fields stay at the desired value. From this moment on one can start measuring whatever observables one may be interested in. For lattices larger than that of the case $N = 1$ a similar mechanism may be used, but this time there are several possible variations of the procedure. For example, we may measure and feed back to the border the average value calculated for the spacial average of the fields over all the internal sites or, alternatively, we may use a spacial average over only the internal sites which are in direct contact with the border, thus implementing a self-consistency condition in the spirit of the Bethe-Peierls condition, involving the normal derivative at the border.

One of the most interesting properties of these systems with fixed but self-consistent boundary conditions relates to their characteristics of critical behavior. Usually we build the models and their corresponding stochastic simulations on finite lattices with periodical boundary conditions[1]. These systems suffer from the inconvenience that there is no true critical behavior on finite lattices of this kind, that is, in systems with a finite number of degrees of freedom and no external boundary [30]. For example, if we calculate by means of stochastic simulations on finite lattices the expectation value v_R of the field in the Ising models, a quantity which is analogous to the magnetization, as a function of the inverse temperature β, we typically obtain curves $v_R(\beta)$ that are continuous, differentiable and monotonically increasing. There are no sharp transitions except in the $N \to \infty$ limit, which makes it considerably more difficult to extract from these simulations the critical values of the parameters of the models by means of extrapolations of the finite-lattice results to the $N \to \infty$ limit.

In contrast to this, the self-consistent lattice systems display sharp transitions and complete critical behavior even on finite lattices. In the case $N = 1$ the fact that the curve of the magnetization displays a sharp transition at a certain critical value of β, a point where it is not differentiable, can be verified analytically. In numerical simulations the sharpness of the transitions is limited, of course, by the technical and numerical limitations of the computer simulations but, with increasing expenditure of computer resources, the transitions can, at least in principle, be made as sharp as one desires for any given N, quite unlike the case of periodical simulations. These self-consistent simulations are, therefore, potentially better for the calculation of critical quantities. With simulations for larger values of N we can not only improve the calculation of local quantities such as the expectation value of the field, we can also calculate significant approximations for non-local quantities, such as the correlation functions for the theory defined within a finite box. This is, therefore, a very useful extension of the mean-field method. Clearly, there is a numerical price

[1]For a more complete understanding of the rest of this section it is useful, although not completely essential, to have some knowledge of the simulation techniques and some experience with stochastic simulations. This subject will be examined in detail in another volume of this series of books.

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-

time dimension. In this way a certain $\beta_c(N)$ is defined on each finite lattice, which is the critical point for that lattice size, assuming that we use as an example a sigma or Ising model. In the cases in which there is a phase transition in the $N \to \infty$ limit $\beta_c(N)$ tends to a finite value β_c^* when $N \to \infty$. In the cases in which there is no phase transition in the limit what happens is that $\beta_c(N)$ increases without limit when N goes to infinity. Since in all cases we have the non-oriented phase for $\beta < \beta_c$, in these cases this random phase is the only one that remains in the $N \to \infty$ limit.

Thus we see that, if there is a phase transition for some well-defined finite β_c^* in the $N \to \infty$ limit, then the $N = 1$ approximation, which also displays a well-defined phase transition, will not be qualitatively different from the limit, although it may be quantitatively quite different, thus giving rise to an approximation that is interpreted as successful. In contrast to this, in the case in which there is no phase transition in the limit $N \to \infty$ the $N = 1$ approximation, because it always displays a well-defined phase transition, becomes qualitatively different from the limit and is thus interpreted as a complete failure. Observe that the $N = 1$ approximation fails precisely in the cases in which the system does *not* display critical behavior in the limit of large values of N, cases which are not, therefore, of much interest for us.

The situation regarding the realization of the theorems is not very different for periodical boundary conditions, except for the fact that in this case there are no well-defined critical points on finite lattices. However, in order to tackle this case, we must first dispel a common misconception regarding finite lattice systems with periodical boundary conditions. Although it is true that if one measures the expectation value of the field v_R in such systems one gets zero within errors, this does not really mean that the single phase existing on finite lattices is the symmetrical phase. The reason why one gets zero for v_R in these circumstances is not that the field configurations are typically non-oriented, but rather that the average value us washed out by the wandering of the direction of alignment. The best way to describe what happens is to say that the system is always in the broken-symmetrical phase on finite lattices, and that it only becomes symmetrical in a certain region of its parameter space in the $N \to \infty$ limit.

One can verify this fact in at least three ways, which we now describe shortly. First, one can include in the action a constant external action j and verify that the resulting value of v_R does not vanish in the limit in which $j \to 0$. Second, one can consider looking at the expectation value of the average of the field over the lattice, which is just the zero mode, the zero-momentum Fourier transform $\widetilde{\varphi}(\vec{0})$ which is like the magnetization in statistical mechanics; if one measures both its expectation value and the expectation value of its square, one gets zero for the first but not for the second, meaning that the field configurations are typically oriented, and that the direction of this orientation drifts. Third, one can eliminate the drift on finite lattices by hand by freezing the zero mode of the field in an arbitrarily chosen direction, and then verifying that one gets explicitly a non-vanishing average magnetization; this changes nothing in the $N \to \infty$ limit, since in this limit the drift is frozen in any case.

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 $N \to \infty$ limit, and typically large in the complementary range. What happens in the $N \to \infty$ 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

is the order parameter of these models. It can be defined as the expectation value v_R of the field, which will be pointing predominantly in the direction that we choose arbitrarily for the symmetry breaking. This is the direction in which we will keep oriented the fields at the border. We will calculate v_R in the mean-field approximation as a function of the parameter β and of the dimension d of space-time. As we saw before in section 2.2, the parameter β can be understood as the inverse of a fictitious temperature. For a given value of d, the critical point β_c is the value of the inverse temperature for which the magnetization $v_R(d,\beta)$ ceases to be zero, when we "cool" the system and therefore increase β. For the calculations it is convenient to use the dimensionless field ψ that was introduced in section 2.2, so that for the purposes of this section we will use, rather than $v_R = \langle\varphi\rangle$, the quantity $\nu_R = \langle\psi\rangle$, where $v_R = \sqrt{\beta}\nu_R$. It is ν_R rather than v_R that is more closely related to the standard definition of the magnetization of statistical mechanics. As was derived in that section, the action can be written in terms of ψ as

$$S[\psi] = -\beta \sum_{\ell} \psi(\ell_-)\psi(\ell_+),$$

while the constraint is written as

$$\psi^2 = 1,$$

where the field ψ is, therefore, either $+1$ or -1. We now freeze all the sites except one at the value $\psi = \nu_R$, arbitrarily choosing the positive direction as the direction of orientation of the fields. Therefore in our $N = 1$ mean-field approximation the relation $\nu_R = \langle\psi\rangle$ can be written as

$$\nu_R = \frac{\sum [\psi = \pm 1]\ \psi\ e^{2d\,\beta\nu_R\psi}}{\sum [\psi = \pm 1]\ e^{2d\,\beta\nu_R\psi}}.$$

Since the sum is over a single site, in this simple $O(1)$ case we can immediately write the result as

$$\nu_R = \frac{\sinh(2d\,\beta\nu_R)}{\cosh(2d\,\beta\nu_R)}, \tag{2.3.1}$$

thus obtaining an algebraic equation that in principle gives us the complete solution for ν_R in the $N = 1$ system. This equation can be solved numerically in order to provide us with graphs of ν_R as a function of β (problem 2.3.2). It can also be used to determine the value of β at which $\nu_R = 0$ becomes the only possible solution (problem 2.3.1), by the use of series expansions. However, the simplest way to obtain the critical points is to simply expand its right-hand side around $\nu_R = 0$. This is sufficient to determine the critical points because the phase transition is of second order in these models and, therefore, we have that $\nu_R \approx 0$ for $\beta \approx \beta_c$. Expanding and keeping only up to first-order terms (problem 2.3.3) we get

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

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 $d = 4$

d	$\mathfrak{N}$	β_c	
		periodical	mean-field
1	1	∞	0.500
2	1	0.4406868	0.250
3	1	0.22165(4)	$0.16\overline{6}$
4	1	0.15 ± 0.02	0.125

Table 2.3.1: Critical points of the Ising models in dimensions from $d = 1$ to $d = 4$.

the mean-field result deviates about 17% from the numerical results, while for $d = 3$ it deviates by about 25%.

Calculating the magnetization in an infinitesimal neighborhood of β_c one can also obtain the mean-field approximation for the corresponding critical exponent. The definition of this critical exponent is as follows. If we have, close to the critical point in the broken-symmetrical phase,

$$\nu_R \approx C\,(\beta - \beta_c)^{\gamma} \tag{2.3.3}$$

for some non-vanishing constant C, then γ is the critical exponent of ν_R. Once more, this can be done using the complete solution in terms of the hyperbolic functions (problem 2.3.4), but the simplest way to obtain the result is to expand equation (2.3.1) to third order in ν_R (problem 2.3.3), thus obtaining

$$\nu_R = 2d\,\beta\nu_R - \frac{8}{3}d^3\beta^3\nu_R^3.$$

Once more a factor of ν_R cancels out and we obtain

$$1 = 2d\,\beta\left[1 - \frac{4}{3}d^2\beta^2\nu_R^2\right].$$

Remembering that $\beta_c = 1/(2d)$ and considering that we are in the vicinity of the critical point, with $\beta \approx \beta_c$, we may write this as

$$\nu_R = \sqrt{3}\,\frac{\beta_c}{\beta}\sqrt{\frac{\beta - \beta_c}{\beta}} \approx \sqrt{3}\sqrt{\frac{\beta - \beta_c}{\beta_c}} = C(\beta - \beta_c)^{1/2}. \tag{2.3.4}$$

This shows that the mean-field approximation for the critical exponent is $\gamma = 1/2$, and determines the value of the constant $C = \sqrt{3/\beta_c}$. This value for the critical exponent of the magnetization is characteristic of the mean-field calculations with $N = 1$.

Another mean-field calculation of interest is that of the critical curves of the $\lambda\varphi^4$ polynomial models, which are given by equation of the type $\lambda = f(\alpha)$. In this case we will use the usual dimensionless field φ as our variable. If we define the magnetization for this case as $v_R = \langle\varphi\rangle$ we may write, in the $N = 1$ mean-field approximation,

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

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).

Although we are not able to solve equation (2.3.6) analytically in order to write the equation $\lambda = f(\alpha)$ of the critical curve in explicit form, it is possible to solve the asymptotic form of the equation, for large values of λ (problem 2.3.8). Doing this we discover that in this limit the critical curve is asymptotic to the straight line defined by the equation

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

where $\beta_c = 1/(2d)$ are the critical points of the corresponding Ising models, so that the asymptote cuts the λ axis at the point $-4d^2$ and forms with the negative α semi-axis an angle θ such that $\tan(\theta) = 1/\beta_c$. Expanding equation (2.3.6) for small values of $-\alpha$ and λ one can also obtain the slope of the tangent line to the critical curve at the Gaussian point (problem 2.3.9). Doing this we obtain for this tangent line the equation

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

We see that the asymptotic slope is $-1/\beta_c$, larger therefore than the slope at the Gaussian point, which is $-1/(3\beta_c)$, by a factor of 3, thus showing that the critical curve has its concavity turned mostly upwards.

In order to get the graphs of the critical curve we are compelled to solve the equation by numerical means. In fact, this can be done not only for our $O(1)$ case here, but for the $SO(\mathfrak{N})$ generalizations as well. A curious fact is that it is in fact easier to do this for $\mathfrak{N} = 2$ and the other even-$\mathfrak{N}$ cases than for $\mathfrak{N} = 1$ and the other odd-$\mathfrak{N}$ cases, because in the even-$\mathfrak{N}$ cases it turns out that the integrals can be written in terms of the error function. For the odd-$\mathfrak{N}$ cases we must use direct numerical integration in order to solve the equation, which is a technique that can be used in all cases [38]. The graphs in figures from 2.3.1 to 2.3.4 show the critical curves obtained by such numerical means in a few cases. Each graph shows also the tangent line at the Gaussian point and the asymptotic line for large values of λ.

Figure 2.3.1 shows the solution in the case $d = 3$, for fairly small values of the parameters $-\alpha$ and λ, while figure 2.3.2 shows the same solution for larger values of the parameters. Figures 2.3.3 and 2.3.4 show the corresponding data for the case $d = 4$. It should be noted that, as one can see in the two graphs with the larger values of the parameters, the asymptotic lines actually cross the critical curves. This implies that at some location for even larger values of the parameters the critical curves must have inflection points and reverse their concavities, presumably approaching their asymptotes from below rather than from above. One can see, looking directly at the data shown in the graphs, that the critical curves do in fact slowly approach the asymptotes, but it seems that the location of the inflection point is at very large values of the parameters, that so far have not been probed numerically. It is interesting to observe that this behavior seems to be characteristic of the $O(1)$ models. In the $SO(\mathfrak{N})$ models with $\mathfrak{N} > 1$ one does not see this type of crossing, and the critical curves seem to approach their asymptotes from above.

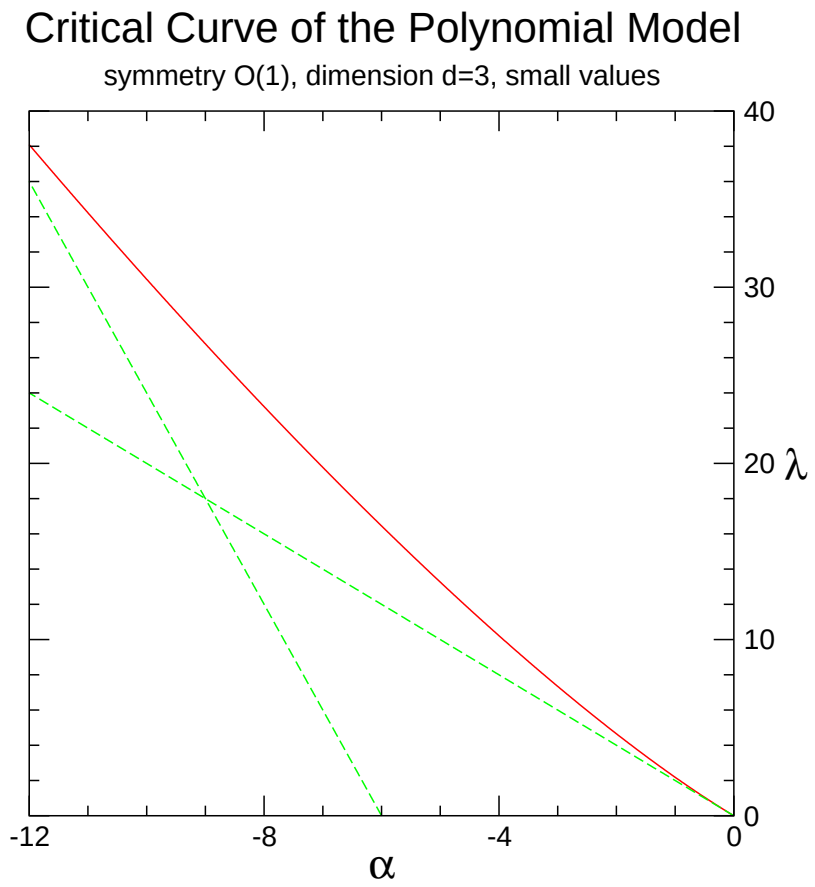

Figure 2.3.1: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 3$, for the smaller values of the parameters.

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

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

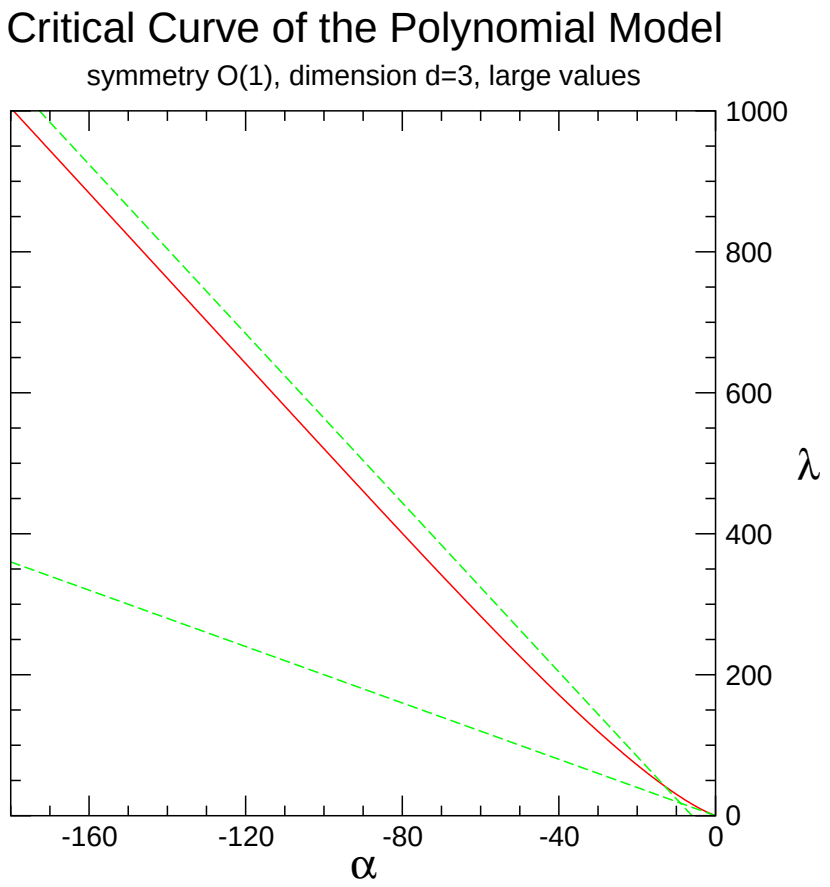

Figure 2.3.2: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 3$, for the larger values of the parameters.

At the critical curve, where $v_R = 0$, this reduces to

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

Note that the right-hand side of this equation is exactly equal to the right-hand side of equation (2.3.6), which determines the critical line, and that therefore we have

$$\sigma^2_{\rm MF} = \frac{1}{2d} = \beta_c,$$

so that the equation of the tangent line to the mean-field critical curve at the Gaussian point can be written as

$$\lambda(\alpha) = -\frac{1}{\sigma^2_{\rm MF}}\ \frac{\alpha}{3}.$$

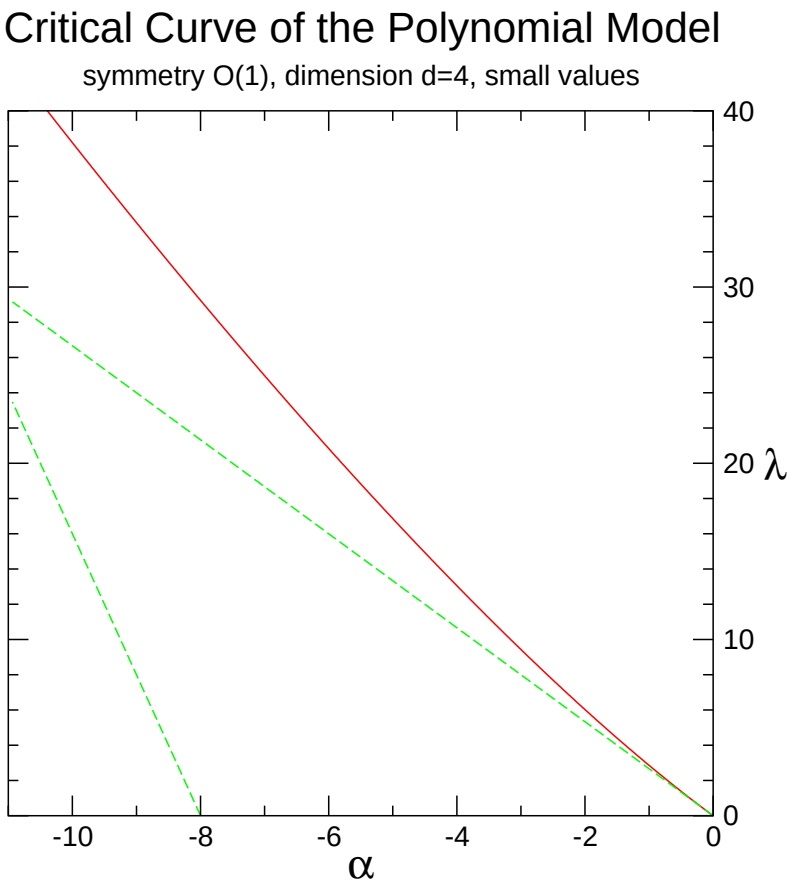

Figure 2.3.3: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 4$, for the smaller values of the parameters.

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 σ_{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 dα e dλ. 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

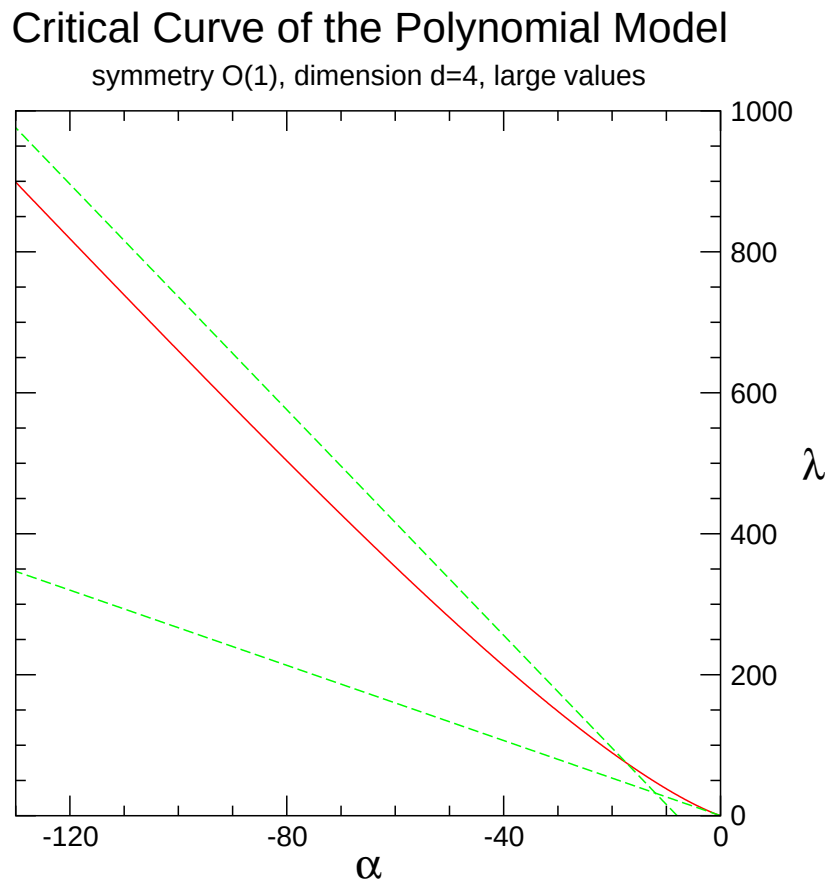

Figure 2.3.4: The mean-field critical curve of the $\lambda\varphi^4$ model with $O(1)$ symmetry, in $d = 4$, for the larger values of the parameters.

$$\mathrm{d}(v_R^2) = C_\alpha \, \mathrm{d}\alpha + C_\lambda \, \mathrm{d}\lambda,$$

where the coefficients C_α e C_λ are finite and non-vanishing expressions in the vicinity of the critical curve. The fact that we are able to write the differential of v_R^2 directly in terms of the differentials of α and λ with coefficients that are finite and not zero over the critical curve is sufficient to show that the critical exponent of v_R is, once more, $\gamma = 1/2$.

In conclusion, we have discovered that the mean-field method is related, through its generalizations, to systems on finite lattices with fixed and self-consistent boundary conditions. These systems define a second family of continuum limits for the models, which does not necessarily have to be identical to the one defined by the systems with periodical boundary conditions. Although the two families of limits have, by and large, similar properties, many of the details are not identical and some quantities of interest may, in fact, depend on the boundary conditions. For example, it may be that only part of the discrepancies shown in table 2.3.1 for the critical

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

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

2.3.4. Starting from the relation $\nu_R \approx C\,(\beta - \beta_c)^{\gamma}$ that defines the critical exponent γ of ν_R, write the corresponding relation for the parameter $A = 2d\,\beta\nu_R$ and differentiate the resulting equation with respect to β, thus obtaining an expression for the critical exponent,

$$\gamma = \frac{\beta - \beta_c}{A}\,\frac{\mathrm{d}A}{\mathrm{d}\beta} - \frac{\beta - \beta_c}{\beta}.$$

In the vicinity of the critical point we have $\beta \approx \beta_c$ and $A \approx 0$, so that we may write for γ

$$\gamma = \frac{\beta - \beta_c}{A}\,\frac{\mathrm{d}A}{\mathrm{d}\beta}. \tag{2.3.7}$$

In order to calculate $\mathrm{d}A/\mathrm{d}\beta$, differentiate equation (2.3.1) with respect to A, and thus obtain the result

$$\frac{\mathrm{d}A}{\mathrm{d}\beta} = \frac{2d\,\sinh^2(A)}{\sinh(A)\cosh(A) - A}.$$

Use this in the expression for γ and expand it to second order around the critical point, then use the result $\beta_c = 1/(2d)$, in order to obtain for γ the expression

$$\gamma = \frac{3}{2C^2\beta_c}(\beta - \beta_c)^{1-2\gamma}.$$

Note that, given the second-order nature of the phase transition in these models, we know that γ must be within the interval $(0, 1]$. Examine the behavior of this expression in the cases $1/2 < \gamma \leq 1$ and $0 < \gamma < 1/2$, and show by reduction to absurd that the only possible value for the critical exponent is $\gamma = 1/2$. Substitute this value in the equation above and obtain the value of the constant C, thus reproducing equation (2.3.4) which was derived in the text.

2.3.5. Use the series expansions of the hyperbolic functions $\sinh(B)$ and $\cosh(B)$, with $B = 2d\,v_R\varphi$, that appear in equation (2.3.5), in order to rewrite that equation in the form

$$\frac{1}{2d} = \frac{\displaystyle\sum_{k=0}^{\infty} \frac{(2d\;v_R)^{2k}}{(2k+1)!} \int_0^{\infty} \mathrm{d}\varphi\;\varphi^{2k+2}\;e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}}{\displaystyle\sum_{k=0}^{\infty} \frac{(2d\;v_R)^{2k}}{(2k)!} \int_0^{\infty} \mathrm{d}\varphi\;\varphi^{2k}\;e^{-\left[(d+\alpha/2)\varphi^2 + (\lambda/4)\varphi^4\right]}}.$$

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

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

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

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

Obtain the same result in another way, using in equation (2.3.8) the asymptotic expansion of $\mathbf{D}_\nu$ [25], since for $\lambda \to 0$ the arguments of the $\mathbf{D}_\nu$ functions that appear in that equation go to infinity.

2.3.10. In order to determine the mean-field value of the critical exponent of v_R in the $\lambda\varphi^4$ model, write equation (2.3.5) in the form

$$\begin{aligned} v_R \int_0^\infty \mathrm{d}\varphi \, \cosh(B) \, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]} \\ = \int_0^\infty \mathrm{d}\varphi \, \varphi \, \sinh(B) \, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]}, \end{aligned}$$

where $B = 2d\, v_R\varphi$. Then differentiate in terms of $\mathrm{d}\alpha$ and $\mathrm{d}\lambda$, keeping in mind that v_R is a function of α and λ, in order to write the differential of v_R in terms of its gradient as

$$\mathrm{d}v = \frac{C_1}{C_0}\,\mathrm{d}\alpha + \frac{C_2}{C_0}\,\mathrm{d}\lambda,$$

where the coefficients are given by

$$\begin{aligned} C_0 &= \int_0^\infty \mathrm{d}\varphi \, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]} \\ &\quad \times \left\{ \left[1 - 2d\,\varphi^2\right] \cosh(B) + B \, \sinh(B) \right\}, \\ C_1 &= \frac{1}{2} \int_0^\infty \mathrm{d}\varphi \, \varphi^2 \, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]} \\ &\quad \times \left[v \, \cosh(B) - \, \varphi \, \sinh(B) \right], \\ C_2 &= \frac{1}{4} \int_0^\infty \mathrm{d}\varphi \, \varphi^4 \, e^{-\left[(d+\alpha/2)\varphi^2+(\lambda/4)\varphi^4\right]} \\ &\quad \times \left[v \, \cosh(B) - \, \varphi \, \sinh(B) \right]. \end{aligned}$$

Next expand the hyperbolic functions in each one of these coefficients for small values of B, that is, in the vicinity of the critical curve. Observe that it is enough to expand C_1 e C_2 to first order but that C_0 must be expanded to the next non-vanishing order, because the equation of the critical curve implies that the terms of orders zero and one of its expansion cancel each other. In this way, obtain the differential of v_R^2 as

$$\mathrm{d}(v^2) = C_\alpha \, \mathrm{d}\alpha + C_\lambda \, \mathrm{d}\lambda,$$

where the coefficients, which define the gradient of v_R^2, are given by

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

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

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.

Chapter 3

Interactions Between Particles

3.1 The Coupling Constant

Having discussed in sections 1.2 and 1.3 the behavior of the $\lambda\varphi^4$ model with respect to the one-point and two-point correlation functions, which are related respectively to the phenomena of spontaneous symmetry breaking and of the propagation of waves and particles, we will now consider the three-point and four-point functions, which are related to the phenomena of interaction between waves or between particles within the model. Our first task will be to discuss the nature of the renormalized (or physical) coupling constant λ_R, and of its dimensionfull version Λ_R, relating them to expectation values of observables of the model. In this way we will define these quantities and determine, at least in principle, the way to calculate them.

As we will see, the renormalized coupling constant is a quantity that vanishes in the Gaussian model and whose value measures how non-Gaussian the renormalized ensemble of the model under study is, thus determining its true degree of non-linearity and the existence or not, within the structure of the model, of phenomena of interaction between waves or between particles. Note that we are not talking here about the ensemble of the fundamental field, but rather about the ensemble of the physical variables associated to blocks, as was discussed in the section in reference [6], since these are the variables that are directly associated to the actual physical observables of the theory. Hence, we should expect the quantity of greater interest in this discussion to be the *dimensionfull* renormalized coupling constant Λ_R, since it is the dimensionfull quantities that scale in the correct way and thus are related to the block-variable observables, as we saw explicitly in the case of the propagator in the section in reference [6].

In order to be able to write the renormalized coupling constant in terms of observables of the model, we return to the discussion of the formalism of the generating functionals and of the effective action, which were introduced in the sections in references [39] and [45]. We saw in the section in reference [39] that the complete Green functions $g_{1,\ldots,n} = \langle\varphi_1 \ldots \varphi_n\rangle$ of the theory in the absence of external sources can be obtained by means of multiple functional differentiations with respect to j of the functional $Z[j]$ defined in the equation in reference [41], after which one makes

$j = 0$. 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 $W[j]$,

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

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,

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

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

consists of a relatively small number of terms with different structures, each one accompanied of all the possible permutations of the position indices. Observe that in any circumstances in which $\varphi_{(c)} = 0$, corresponding necessarily to $j = 0$, we obtain a much simpler relation, that can be written as

$$g_{(c)1,2,3,4} = g_{1,2,3,4} - [g_{1,2}\, g_{3,4} + g_{1,3}\, g_{2,4} + g_{1,4}\, g_{2,3}], \tag{3.1.2}$$

since for $j = 0$ we have that $g_{(c)i,j} = g_{i,j}$. Just as $g_{(c,j)i,j}$ gives us the true two-point correlations, $g_{(c,j)i,j,k,l}$ gives us the true four-point correlations of the model, that is, those which are not just superpositions of two-point correlations. While the two-point function is related to the propagation of waves and particles, the four-point function is related to the interaction between these waves and between these particles. It gives us the part of the complete four-point function which is not just a product of two-point functions. The part of the complete four-point function that can be decomposed into such a product of two-point functions corresponds to two waves or particles propagating together in a region of space-time, but that superpose linearly, passing transparently through each other, without interacting with one another. This is in fact all that happens in the theory of the free scaler field. One can show (problem 3.1.2) that in that theory, where everything is linear, the connected four-point function vanishes identically, a fact which corresponds to the lack of interactions between waves or between particles in that theory.

In order to relate this function with the renormalized coupling constant we must go back to the discussion of the concept of the effective action, which was introduced in the section in reference [39] and discussed in detail in the section in reference [45]. The renormalized coupling constant is one of the parameters that appears in the expression of the effective action, it is the parameter that relates most directly to the connected four-point function and that encodes in the most concise way the structure of interactions of the theory. As we saw, the effective action $\Gamma[\varphi_{(c)}]$ is a functional of $\varphi_{(c)}$ defined from $W[j]$ by a Legendre transformation. As we saw in the equation in reference [47], the double functional derivative of $\Gamma[\varphi_{(c)}]$ with respect to $\varphi_{(c)}$ is related to the inverse of the propagator. Starting from that equation we may write, with some changes of indices and an additional sum over the lattice, the equation

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

If we differentiate this once more with respect to j, using then the chain rule in order to rewrite the derivatives as derivatives with respect to $\varphi_{(c)}$, as we did in the derivation of the equation in reference [47], and recalling also that the double functional derivative of $\Gamma[\varphi_{(c)}]$ is the inverse of the propagator, we obtain (problem 3.1.3)

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

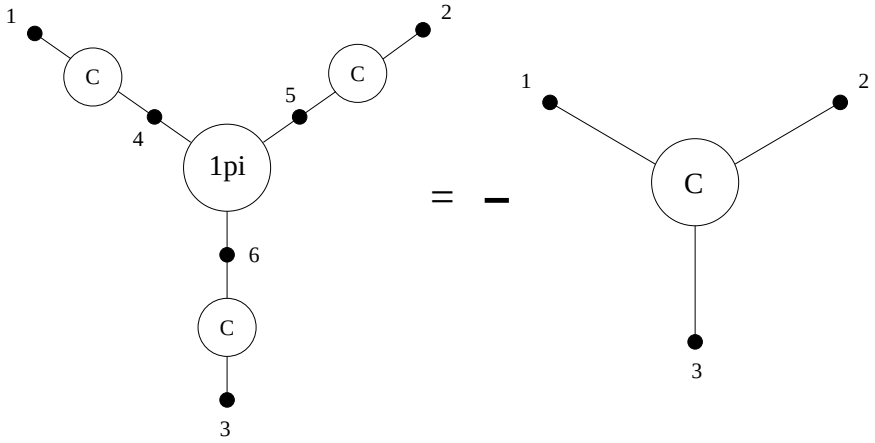

Figure 3.1.1: Diagrammatic representation of the equation for the three-point function.

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

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

This relation has an interesting diagrammatic representation, which we illustrate in figure 3.1.2, where the symbol $\square_{(c)}$ 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

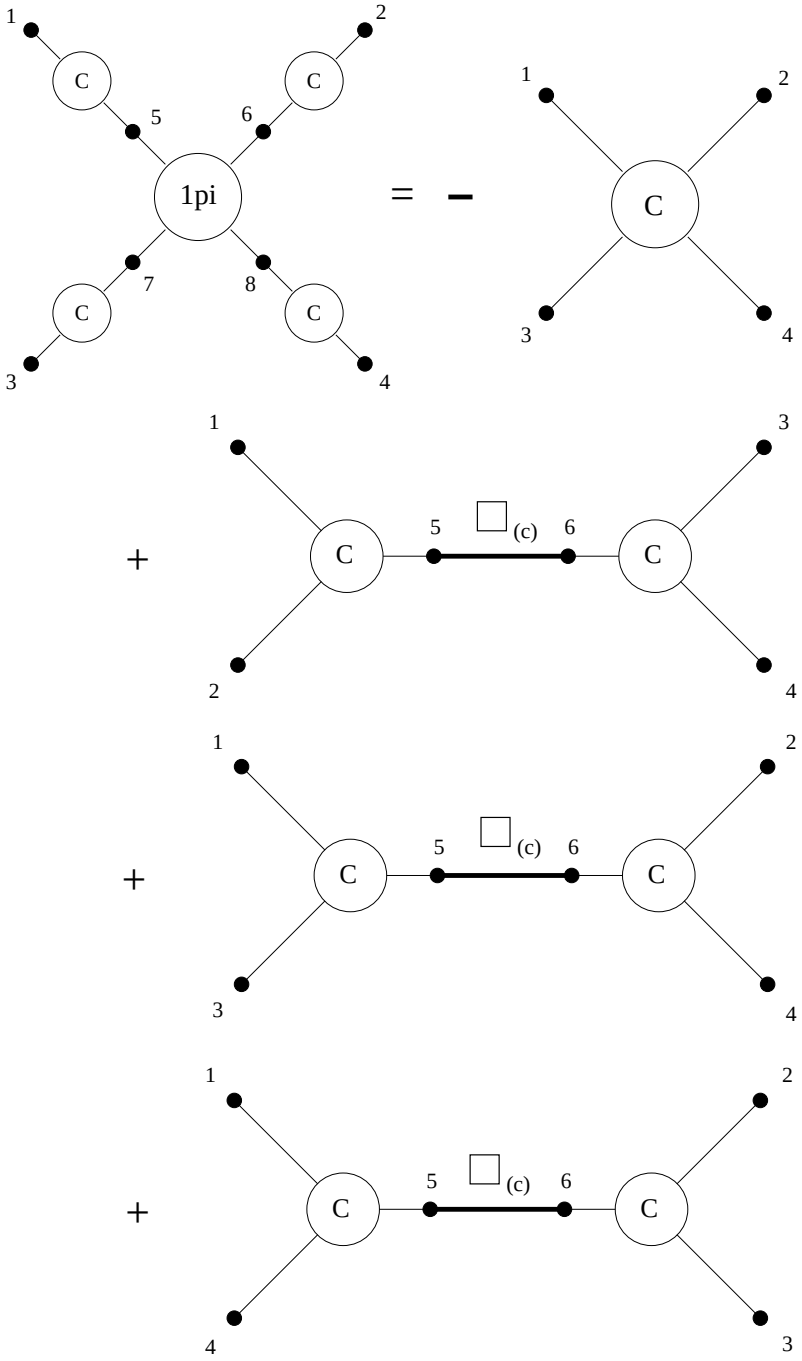

Figure 3.1.2: Diagrammatic representation of the equation for the four-point function.

functions and at most a single internal $\square_{(c)}$. 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

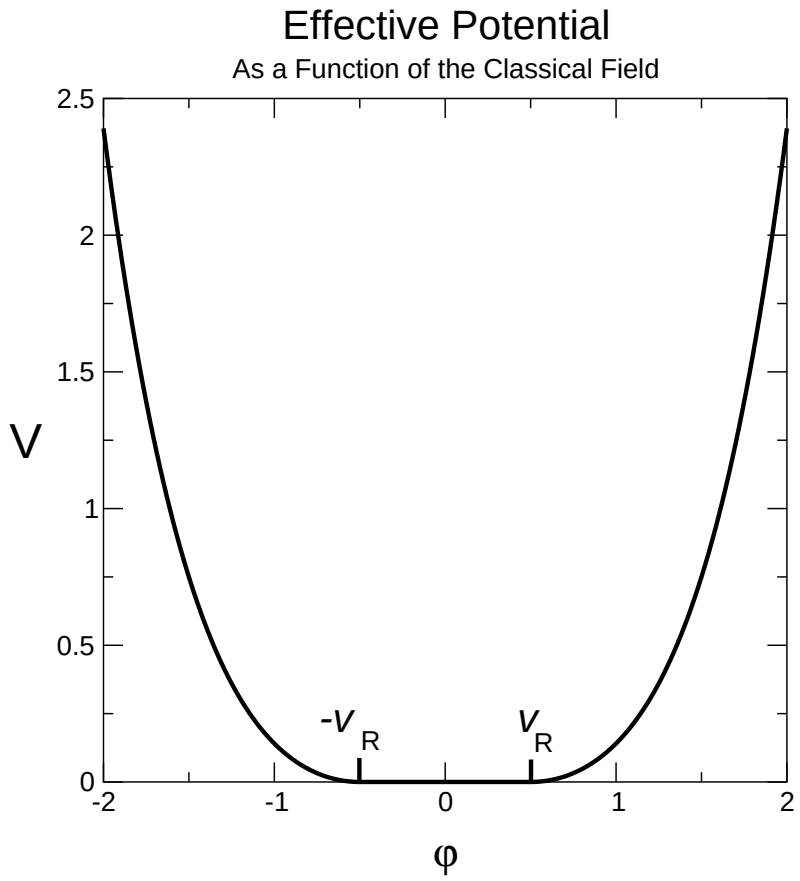

Figure 3.1.3: The effective potential as a function of the classical field $\varphi_{(c)}$.

$$\widetilde{\gamma}_{1,2,3,4} = \frac{-1}{N^{4d}} \frac{\widetilde{g}_{(c)1,2,3,4}}{\widetilde{g}_{(c)1}\,\widetilde{g}_{(c)2}\,\widetilde{g}_{(c)3}\,\widetilde{g}_{(c)4}}. \tag{3.1.5}$$

We have here the 1pi function written in terms of expectation values of the model in momentum space.

In order to relate this to the renormalized coupling constant, we will have to make some assumptions about the form that the effective action $\Gamma[\varphi_{(c)}]$ may have, which will be based on the symmetries of the model. Let us recall then that our polynomial model is defined by the action given in equation (1.1.1), which we reproduce here,

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

We will be interested primarily in analyzing the low-momentum regime of the model, because this is enough for obtaining the value of the renormalized coupling constant, since it appears as part of a local potential, which exists even for fields which are constant over the lattice, having therefore infinite wavelength and vanishing momenta. In addition to this, we will assume that the effective action has the same

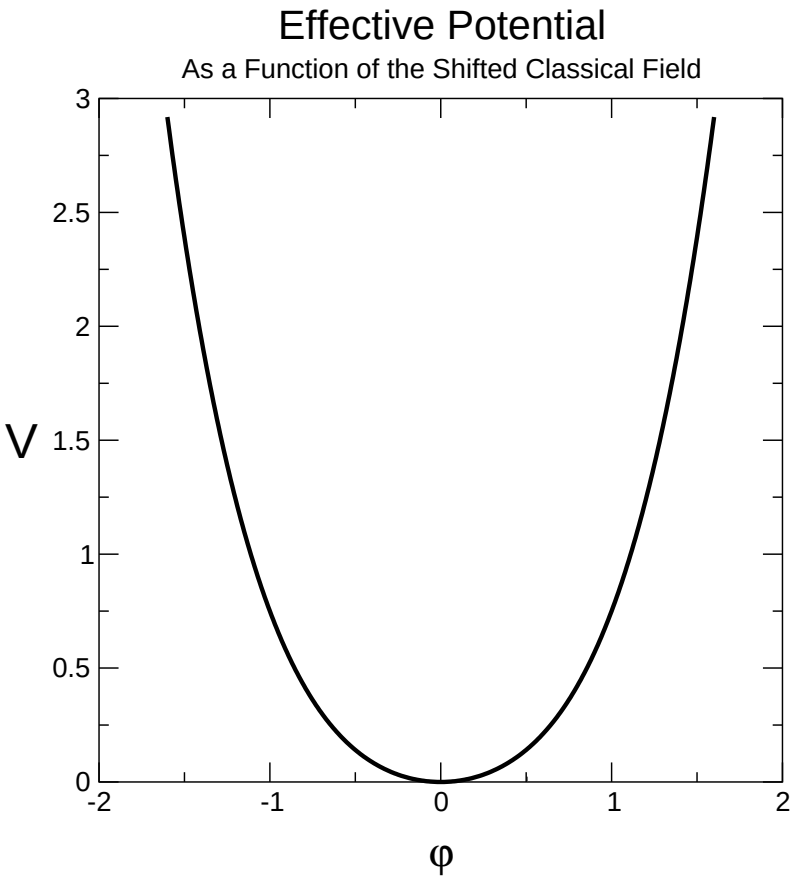

Figure 3.1.4: The effective potential as a function of the shifted classical field $\varphi'_{(c)}$.

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

the effective potential as shown in figure 3.1.4. Observe that in the continuum limit we have $v_R \to 0$, since it is necessary that we approach the critical curve in the limit, where $v_R = 0$ since the phase transition is second-order. Hence, if we will end up by taking the limit, we can do the analysis either in terms of $\varphi'_{(c)}$ or in terms of $\varphi_{(c)}$. For simplicity, we will limit ourselves here to the derivation of the relation between λ_R and the observables of the model in the case in which $j = 0$, but it is not difficult to generalize the result (problem 3.1.6).

Since we will assume that the effective action, when written in terms of $\varphi'_{(c)}$, has the same symmetries that the fundamental action which defines the model has, it follows that $\Gamma[\varphi'_{(c)}]$ must be composed of terms that have the same symmetries of the terms existing in $S[\varphi]$, that is, that it must have the general form

$$\Gamma[\varphi'_{(c)}] = \frac{1}{\zeta}\left\{\frac{1}{2}\sum_{\ell}[\Delta_\ell \varphi'_{(c)}]^2 + \frac{\alpha_R}{2}\sum_s \varphi'^2_{(c)}(s) + \frac{\lambda_R}{4}\sum_s \varphi'^4_{(c)}(s) + \text{ (others)}\right\},$$

where we wrote explicitly the terms which are relevant for the analysis of the low-momentum regime of the model, ζ is the residue of the pole of the propagator and "others" indicates terms with more than four powers of the field and terms with more than two derivatives. Terms with many derivatives do not contribute significantly to the low-momentum regime and terms with more than four powers of the field do not contribute to the four-point function. Based on the numerical experience with this model, we may assume that $\zeta = 1$, which seems to be true with significant precision in all cases examined so far.

When taking the functional derivatives of $\Gamma[\varphi'_{(c)}]$, and considering that we are interested in the case $j = 0$, we should realize that it is implicit that we should put $\varphi'_{(c)} = 0$ at the end of the calculations, because this is the value of $\varphi'_{(c)}$ that corresponds to the condition $j = 0$ in this model. It is clear that the quadratic terms will vanish anyway when we take the derivatives, while the terms with powers larger than four will vanish due to the condition $\varphi'_{(c)} = 0$. Therefore, we may consider only the terms of the effective action that contain exactly four powers of the field and no derivatives, and so we are reduced to considering only the term

$$V_{(4)}[\varphi'_{(c)}] = \frac{\lambda_R}{4}\sum_0 \varphi'^4_{(c)0},$$

which is the term of the effective potential which is relevant for zero momentum. Taking the first derivative we get

$$\frac{\mathfrak{d}V_{(4)}}{\mathfrak{d}\varphi'_{(c)1}} = \lambda_R \sum_0 \varphi'^3_{(c)0}\delta_{0,1} = \lambda_R \varphi'^3_{(c)1}.$$

Multiplying this equation by $f_1(1)$, where, in order to simplify the notation, we are denoting the mode functions of the Fourier basis as

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

then adding over the variable $\vec{n}_1$ and differentiating a second time we get

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

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

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

Repeating the procedure a fourth and last time we get

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

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

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

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

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

Naturally, since we neglected the terms in $\Gamma[\varphi'_{(c)}]$ with larger powers of the momenta, this relation only makes sense for small or vanishing momenta. We take, therefore, the zero-momentum case $\vec{k}_1 = \vec{k}_2 = \vec{k}_3 = \vec{k}_4 = \vec{0}$ in order to obtain, substituting the connected functions in terms of the complete functions,

$$\lambda_R = \frac{1}{6N^d}\,\frac{3\langle|\widetilde{\varphi}(\vec{0})|^2\rangle^2 - \langle|\widetilde{\varphi}(\vec{0})|^4\rangle}{\langle|\widetilde{\varphi}(\vec{0})|^2\rangle^4}. \tag{3.1.7}$$

If we recall the factorization relations of the free theory for the correlation functions in momentum space, that were introduced in the section of reference [44], we immediately see that this quantity vanishes identically in the free theory.

We may also write this result in terms of the dimensionfull quantities, using the appropriate scaling relations to transform φ in ϕ and λ_R in Λ_R, thus obtaining

$$\Lambda_R = \frac{1}{6L^d}\,\frac{3\langle|\widetilde{\phi}(\vec{0})|^2\rangle^2 - \langle|\widetilde{\phi}(\vec{0})|^4\rangle}{\langle|\widetilde{\phi}(\vec{0})|^2\rangle^4}.$$

This is the quantity whose value determines whether or not there exists in this model the phenomenon of non-linear interaction between waves, or between particles. Naturally, this quantity is of great physical interest and we will dedicate some time to the examination of its properties.

Problems

3.1.1. Using the definition of the connected four-point correlation function

$$g_{(c,j)1,2,3,4} = \frac{\mathfrak{d}^4 W[j]}{\mathfrak{d}j_1\mathfrak{d}j_2\mathfrak{d}j_3\mathfrak{d}j_4},$$

in a theory with a non-vanishing external source j, show that it is related to the complete functions of four, three and two points by the formula

$$\begin{aligned}
g_{(c,j)1,2,3,4} &= g_{(j)1,2,3,4} \\
&- [g_{(j)2,3,4}\,\varphi_{(c)1} + g_{(j)1,3,4}\,\varphi_{(c)2} + g_{(j)1,2,4}\,\varphi_{(c)3} + g_{(j)1,2,3}\,\varphi_{(c)4}] \\
&- [g_{(j)1,2}\,g_{(j)3,4} + g_{(j)1,3}\,g_{(j)2,4} + g_{(j)1,4}\,g_{(j)2,3}] \\
&+ 2[g_{(j)1,2}\,\varphi_{(c)3}\,\varphi_{(c)4} + g_{(j)1,3}\,\varphi_{(c)2}\,\varphi_{(c)4} + g_{(j)1,4}\,\varphi_{(c)2}\,\varphi_{(c)3} \\
&\quad + g_{(j)2,3}\,\varphi_{(c)1}\,\varphi_{(c)4} + g_{(j)2,4}\,\varphi_{(c)1}\,\varphi_{(c)3} + g_{(j)3,4}\,\varphi_{(c)1}\,\varphi_{(c)2}] \\
&- 6\,\varphi_{(c)1}\,\varphi_{(c)2}\,\varphi_{(c)3}\,\varphi_{(c)4}.
\end{aligned}$$

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

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

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

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

Next, use the fact that the second functional derivative of $\Gamma[\varphi_{(c)}]$ is the inverse of the propagator and rearrange the terms in order to obtain the final relation in the form

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

3.1.4. Starting from the final result of equation (3.1.3), differentiate it once again with respect to j, using the chain rule to rewrite the derivatives as derivatives with respect to $\varphi_{(c)}$, thus obtaining

$$\begin{aligned}
\sum_{5,6,7,8} g_{(c,j)1,5}\; g_{(c,j)2,6}\; g_{(c,j)3,7}\; g_{(c,j)4,8} \frac{\mathfrak{d}^4\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)5}\mathfrak{d}\varphi_{(c)6}\mathfrak{d}\varphi_{(c)7}\mathfrak{d}\varphi_{(c)8}} \quad + \\
+ \sum_{5,6,7} g_{(c,j)1,4,5}\; g_{(c,j)2,6}\; g_{(c,j)3,7} \frac{\mathfrak{d}^3\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)5}\mathfrak{d}\varphi_{(c)6}\mathfrak{d}\varphi_{(c)7}} \quad + \\
+ \sum_{5,6,7} g_{(c,j)1,5}\; g_{(c,j)2,4,6}\; g_{(c,j)3,7} \frac{\mathfrak{d}^3\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)5}\mathfrak{d}\varphi_{(c)6}\mathfrak{d}\varphi_{(c)7}} \quad + \\
+ \sum_{5,6,7} g_{(c,j)1,5}\; g_{(c,j)2,6}\; g_{(c,j)3,4,7} \frac{\mathfrak{d}^3\Gamma[\varphi_{(c)}]}{\mathfrak{d}\varphi_{(c)5}\mathfrak{d}\varphi_{(c)6}\mathfrak{d}\varphi_{(c)7}} \quad = \quad g_{(c,j)1,2,3,4}.
\end{aligned}$$

Using again the final result of problem 3.1.3 and, once more, the fact that the second functional derivative of $\Gamma[\varphi_{(c)}]$ is the inverse of the propagator, obtain the final relation

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

3.1.5. Starting from equation (3.1.4), write it in the form

$$\sum_{5,6,7,8} g_{(c)1,5}\; g_{(c)2,6}\; g_{(c)3,7}\; g_{(c)4,8}\; \Gamma_{5,6,7,8} = -g_{(c)1,2,3,4}$$

and execute four Fourier transforms on the external variables $\vec{n}_1, \ldots, \vec{n}_4$, using the corresponding variables $\vec{k}_1, \ldots, \vec{k}_4$ in momentum space, and recalling that, for a function F of n position variables $\vec{n}_i$,

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

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

for the calculation of the propagator, we obtain (problem 3.2.1) the classical result $\lambda_R = \lambda$. Naturally, this implies that $\Lambda_R = \Lambda$ to first order in ε, which means to first order in λ, which seems to indicate, in a superficial way, that we may have a non-vanishing renormalized coupling constant in the quantum theory. We will see, however, that this is an excessively superficial analysis and that the real situation is much more complex than what it indicates.

Let us recall that, in the calculation of the propagator, the calculation of order zero in ε resulted in the classical propagator, that is, in the propagator of the Gaussian ensemble that we are using to approximate the ensemble of the complete theory, while the calculation to first order in ε gave us quantum corrections to the classical result. Also, this first-order result was not a small correction of the zero-order result, but instead was qualitatively different from it. In the case of the coupling constant it is clear that the zero-order calculation results in $\lambda_R = 0$, which is the value of this quantity in the Gaussian ensemble, while the first-order calculation gives us the classical result. This situation is to be expected, since we are now calculating a quantity which is, by definition, at least proportional to the expansion parameter λ, and that vanishes when the expansion parameter vanishes, unlike what was the case for α_R. Unlike what happened in the case of α_R, we are discussing here a quantity that does not exist at all in the Gaussian ensemble of the free theory. Just like what happened in the case of the propagator, it is possible that passing to the next order, which takes us away from the classical results, will make a qualitative difference.

These results that we refer to as "classical" correspond, in the traditional language, to Feynman diagrams with zero loops, that is, to the "tree" approximation. These results do not include the effects of the quantum fluctuations of the fields on the observables. Hence, the zero-loop approximations do not include the quantum effects contained in the theory, but only the effects of the classical dynamics of the fields. In order to include the effect of the quantum fluctuations it is necessary to do the calculations up to the lowest order of ε which includes diagrams with one loop. In the case of m_R this meant doing calculations up to the first order in ε, but in order to explore the effects of the quantum fluctuations of the fields on the renormalized coupling constant it is necessary to calculate λ_R to second order in ε, thus including diagrams with up to one loop. These calculations to order ε^2 are considerably longer and more complex than those to the first order, and involve quantities with strong divergences that behave as either N^d or as N^{2d}, all of which cancel out completely from the final results. Doing the calculation in the symmetrical phase, with vanishing momenta on the four external legs, we obtain (problem 3.2.2)

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

where the dimensionless sum $s_2(\alpha_0)$ that appears here is given, in terms of the dimensionless free propagator $g_0(\vec{n}, \vec{n}')$, by

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

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).

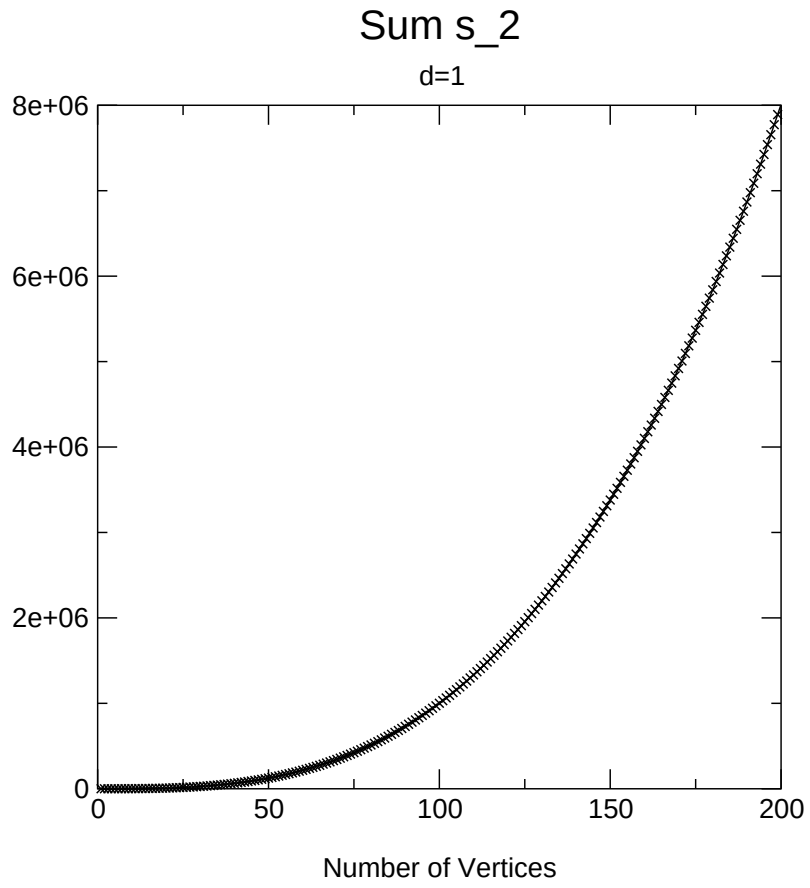

Figure 3.2.1: Behavior of the sum s_2 with N in the case $d = 1$.

Of course this type of limit constitutes a small subset of all possible alternatives, in which we approach, in the limit, some other arbitrary point of the critical curve, rather than the Gaussian point. About these other possibilities our perturbative approximation has nothing to say, but note that, is we assume that there is at least one such limit for each one of these points, in which Λ_R is finite, it follows immediately that it is necessary that $\lambda_R \to 0$ over the whole critical curve, when we make $N \to \infty$. This means that it is highly likely that the ensemble of the renormalized theory becomes Gaussian over the critical curve in the continuum limit. Since λ_R is a dimensionless quantity that measures, just like α_R, a moment of the distribution of the renormalized model, it is very reasonable to think that both should have the same particular type of behavior in the locus of the parameter plane of the model where the critical transition takes place. In other words, it is reasonable to think that λ_R should always go to zero over the critical curve in the continuum limit, as part of the critical behavior of the model.

We may also try to extract some information of the result in equation (3.2.2) in the case $d \geq 5$. In this case our result seems to indicate that we will have a finite and non-vanishing λ_R as a function of α and λ, since the sum s_2 is finite and

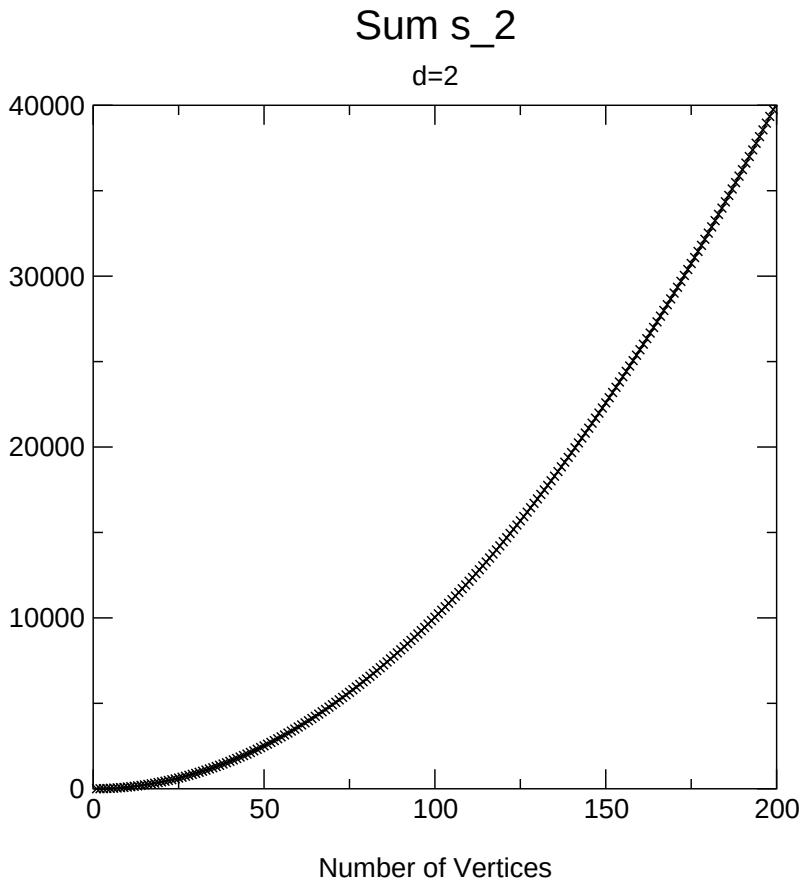

Figure 3.2.2: Behavior of the sum s_2 with N in the case $d = 2$.

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

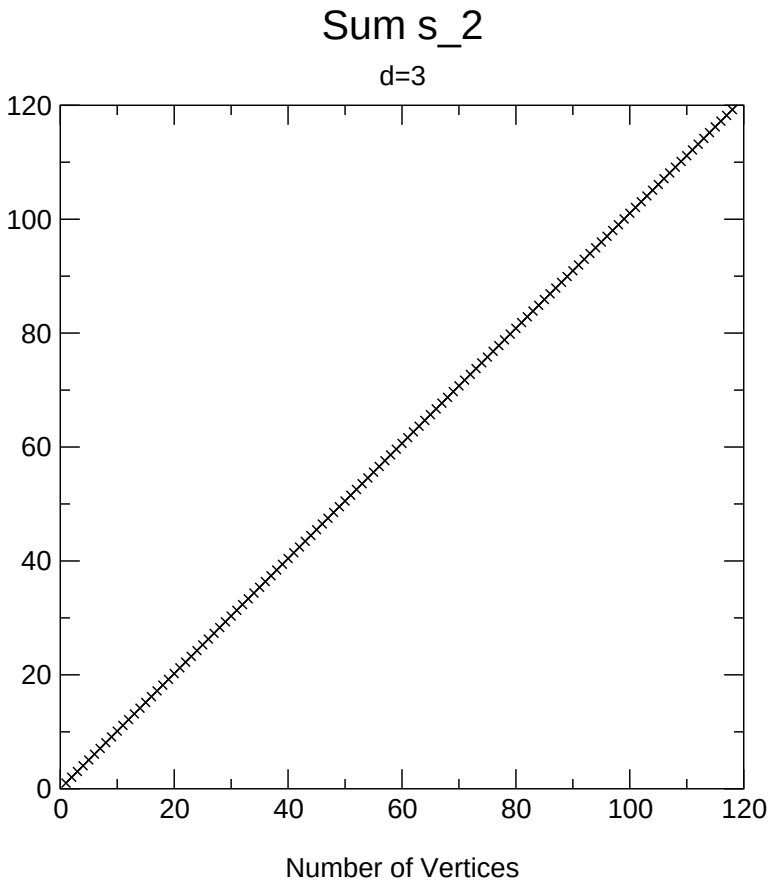

Figure 3.2.3: Behavior of the sum s_2 with N in the case $d = 3$.

lattices (problem 3.2.6). The results obtained in this way, for $m_R L = 1$, are shown in the graphs found in figures from 3.2.1 to 3.2.5. Simple but good-quality curve fittings, with only a single parameter in the cases $d = 1$, $d = 2$ and $d = 3$, and with three parameters in the cases $d = 4$ and $d = 5$, give the approximate results

$$\begin{aligned} s_2(d=1) &\approx 1.0N^3, \\ s_2(d=2) &\approx 1.0N^2, \\ s_2(d=3) &\approx 1.0N, \\ s_2(d=4) &\approx 1.0 + 0.013\ln(0.93N), \\ s_2(d=5) &\approx 0.19 + 1.0\frac{1}{N} - 0.014\frac{1}{N^2}. \end{aligned}$$

One may also consider calculating the expectation value of the field and of the propagator up to order ε^2, which corresponds to the inclusion of diagrams with up to two loops. In some cases one may still be able to extract from these calculations some useful information such as, for example, for the determination of the critical curve (problem 3.2.7) and, in some cases, for the determination of the renormalized mass (problem 3.2.8). However, in general the results of these calculations include

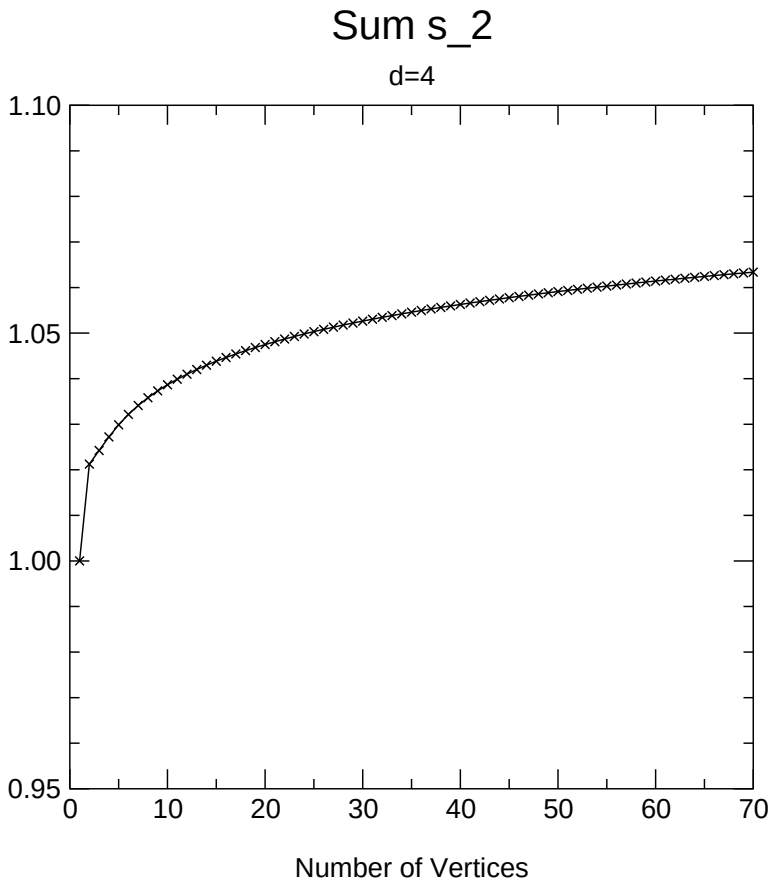

Figure 3.2.4: Behavior of the sum s_2 with N in the case $d = 4$.

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

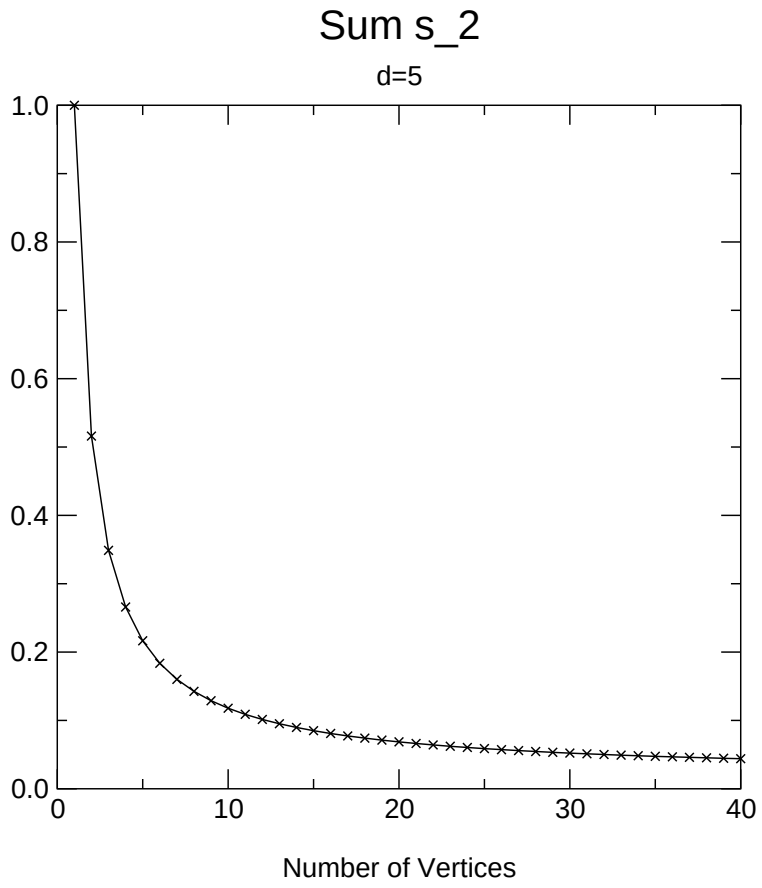

Figure 3.2.5: Behavior of the sum s_2 with N in the case $d = 5$.

a perturbative approximation as successful as in the cases of the renormalized mass and of the expectation value of the field. It allows us to go so far as to formulate the *conjecture* that, in any dimension $d \geq 3$, in any limit where $N \to \infty$, $\alpha(N) \to \alpha_{(c)}$ and $\lambda(N) \to \lambda_{(c)}$, where $(\alpha_{(c)}, \lambda_{(c)})$ is a point over the critical curve, we have $\lambda_R \to 0$. However, it does not allow us to make concrete predictions about the behavior of the model away from the Gaussian point, as we were able to do for the renormalized mass and the expectation value of the field. It seems, therefore, that the discussion of the perturbative approximation should be separated in two parts, the first one being relative to the calculation of the quantities involving, at most, the second moment of the distribution of the model, as is the case for v_R and α_R, while the second one relates to the quantities involving the higher-order moments, such as λ_R.

We see that the reason why the approximation does not work so well for the coupling constant is the fact that it is not possible to impose, in this case, the equivalent of the condition $\alpha_0 = \alpha_R$ used in the case of the renormalized mass, which transformed the first-order perturbative approximation into a self-consistent Gaussian approximation. This is due, of course, to the fact that the Gaussian model, the only one we know how to solve exactly, has no moments of order greater than two,

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.

re-expand the expressions that appear when one tries to do this, possibly neglecting higher-order terms in order to keep consistent powers of the expansion parameters. All this considerably complicates the whole argument and makes it more difficult to understand what is really going on when one does all this. Let us try to exemplify this in the context of the calculations that we made here, at first in a very simple and possibly incomplete way. We have calculated one-loop approximations for α_R and λ_R, obtaining expressions of the form

$$\alpha_R = f_\alpha(\alpha, \lambda) \quad \text{and} \quad \lambda_R = f_\lambda(\alpha, \lambda).$$

We saw that, while the one-loop propagator is entirely finite, the renormalized mas parameter being given by

$$\alpha_R = \alpha + 3\sigma_0^2\lambda,$$

the one-loop coupling constant contains the divergent sum s_2, being given by

$$\lambda_R = \lambda[1 - 8s_2\lambda].$$

We proceed then to a change of variables, introducing a new parameter λ_0 in place of λ, defined by the relation

$$\lambda = \lambda_0 + 9s_2\lambda_0^2.$$

Note that this mixes powers of λ and corresponds, therefore, to a reorganization of the perturbative expansion. Substituting this expression for λ in the result for λ_R we verify that the divergent terms of order λ_0^2 cancel out, so that we obtain

$$\lambda_R = \lambda_0 - 162s_2^2\lambda_0^3 - 729s_2^3\lambda_0^4.$$

We now argue that we can neglect in this equation the terms of orders λ_0^3 and λ_0^4, not because they are small, since they are clearly divergent in the limit, but under the allegation that they will cancel out with the remaining higher-order terms that have not yet been explicitly included in this analysis. This is the first condition involved in the criterion of perturbative renormalizability in a weak sense, term by term in the perturbative expansion, without preoccupation with its convergence. What we are requiring here is that the divergent terms cancel out, not in the original series, but after its reorganization by the change of variables from λ to λ_0. Under these conditions we have

$$\lambda_R = \lambda_0,$$

which shows that our change of variables is in fact a change from an expansion in terms of the basic parameter λ to another expansion in terms of the renormalized parameter λ_R. If we now write λ in terms of λ_R,

$$\lambda = \lambda_R + 9s_2\lambda_R^2,$$

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

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

Proceeding with the substitution of λ by $\lambda_0 = \lambda_R = \lambda_R(\vec{k} = \vec{0})$ we obtain

$$\lambda_R(\vec{k}) = \lambda_0 \left\{ 1 + 3\lambda_0 \left[s_2(\vec{0}, \alpha_0) - s_2(\vec{k}, \alpha_0) \right] \right\}. \tag{3.2.5}$$

The difference of the two sums can be evaluated for small values of the momentum $\vec{k}$ with the help of approximations by integrals and, doing this in $d = 4$ and for large values of N (problem 3.2.10), we obtain a finite result,

$$\lambda_R(\vec{k}) = \lambda_R \left[1 + 24\pi^2 \lambda_R \frac{k^2}{L^6 m_R^6} \right].$$

We say that the theory is perturbatively renormalizable if it is possible to do this in each order of perturbation theory, and hence to obtain predictions with arbitrarily high precision for $\mathcal{O}$, given values of α_R and λ_R. Note that, if we imagine that the theory is in fact trivial, then we see that this result is not wrong, but that is is simply rather irrelevant, because in this case the only possible value for λ_R is zero and the relation simply shows that $\lambda_R(\vec{k}) = 0$ for any $\vec{k}$. We can see now that there is in fact a rather subtle problem behind all this. When we do this kind of manipulation we are giving up obtaining from the theory the relations between (α_R, λ_R) and (α, λ) and, instead of that, *we implicitly assume that certain values of α_R and λ_R are possible* in the context of the model defined in a non-perturbative way by means of the lattice. This seems to be a very reasonable thing to do in a model which is defined with two free parameters, and we certainly know which values are or are not possible for α and λ. However, we do not have now any information about *which values are in fact possible* for the renormalized parameter λ_R, according to the non-perturbative definition of the model. Therefore, we do now know which values we may in fact use for λ_R in this perturbative renormalization scheme.

It is implicitly assumed, in the traditional perturbative renormalization scheme, that the possible values for (α_R, λ_R) are the same which are possible for (α, λ). However, in general it is possible that this is not true, and that there are restrictions for the images of the relations $\alpha_R(\alpha, \lambda)$ and $\lambda_R(\alpha, \lambda)$ determined by the non-perturbative definition of the models. One restriction that we already know to exist in this model is that $\alpha_R \geq 0$, while the parameter α can be either positive or negative on finite lattices, and must become negative in the continuum limit, as we saw in section 1.3. Another fact, which is even more important than this one, is that there certainly are important restrictions for λ_R in a model that ends up being trivial, in which the only possible value for λ_R in the $N \to \infty$ limit is zero. We can always determine beforehand which values are possible for (α, λ), but we cannot do the same for (α_R, λ_R). Triviality implies that the usual implicit hypothesis, that the possible values for (α_R, λ_R) are the same which are possible for (α, λ), is false. To continue with the usual perturbative renormalization scheme under these conditions can only produce fictitious results, without any physical or mathematical relevance.

The conclusion is that a model satisfying the criterion of perturbative renormalizability is not sufficient to guarantee the usefulness of its perturbative expansion, renormalized in the usual way. It is also necessary to determine the values which are

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)

should be calculated to second order, while it is enough to calculate the denominator to first order.

3.2.3. ($\star\star\star$) Repeat the calculation of λ_R to second order proposed in problem 3.2.2, this time in the broken-symmetrical phase[2].

3.2.4. Evaluate the asymptotic behavior of the sum s_2 given in equation (3.2.1), for large values of N, approximating it by integrals over the momenta, as we did before in the case of the quantity σ_0^2 related to the propagator, for each relevant value of d. Whenever it becomes necessary to use a minimum but non-vanishing value of the modulus of the momentum as a lower integration limit, use m_R as that value.

3.2.5. Use the result in equation (3.2.4) for the dimensionfull renormalized coupling constant Λ_R in $d = 3$, as well as the one-loop result for the renormalized mass m_R obtained in section 1.3, in order to exhibit explicitly flows $[\alpha(N), \lambda(N)]$ that approach the Gaussian point in the continuum limit and for which both m_R and Λ_R have finite and non-vanishing limits. Assume, if necessary, that Λ_R is small compared to m_R. The solutions should tend asymptotically to the line tangent to the critical curve at the Gaussian point, and both $\alpha(N)$ and $\lambda(N)$ should go to zero as $1/N$. Hint: try

$$\alpha = \frac{A}{N} + \frac{B}{N^2} \quad \text{and} \quad \lambda = \frac{C}{N}.$$

3.2.6. Write programs to calculate numerically the sum s_2 given in equation (3.2.1) in dimensions from $d = 1$ to $d = 5$ and confirm the asymptotic results obtained in problem 3.2.4. These sums should be calculated with the same numerical techniques that were used for the calculation of the sums that appear in the quantity σ_0^2, which were calculated in the section in reference [49]. In fact, it suffices to make small changes in the programs written for that case in order to produce the programs needed in this case.

3.2.7. ($\star\star$) Calculate v_R to order ε^2 and thus obtain the two-loop evaluation of the equation of the critical curve. Start by calculating v_R in the broken-symmetrical phase and obtain the result

$$\begin{aligned} 0 \;=\; & \left(\lambda v^2 + \alpha + 3\lambda\sigma_0^2\right)\left(3\lambda v^2 + \alpha - 2\alpha_0 + 3\lambda\sigma_0^2\right) \\ & + \lambda\left[9\lambda v^2 + 3(\alpha - \alpha_0) + 9\lambda\sigma_0^2\right]\alpha_0 \sum_{\vec{n}} g_0^2(\vec{0},\vec{n}) + 6\lambda^2\alpha_0 \sum_{\vec{n}} g_0^3(\vec{0},\vec{n}). \end{aligned}$$

Next evaluate the asymptotic behavior of the new sum that appears,

$$s_3(\alpha_0) = \sum_{\vec{v}} g_0^3(\vec{0},\vec{n}) = \frac{1}{N^{2d}} \sum_{\vec{k}_1,\vec{k}_2} \frac{1}{[\rho^2(\vec{k}_1) + \alpha_0][\rho^2(\vec{k}_2) + \alpha_0][\rho^2(\vec{k}_1 + \vec{k}_2) + \alpha_0]}.$$

[2]Note: the answer to this problem is currently unknown.

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

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

which, in momentum space, can be written as

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

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.

3.2.9. ($\star$) Calculate $\lambda_R(\vec{k})$ for equal non-vanishing momenta $\vec{k}$ on all the external legs, entering in two of them and going out in the other two, using for this purpose the expression for this quantity that results from problem 3.1.7, to second order, with the choice $\alpha_0 = \alpha_R$, in the symmetrical phase, obtaining the result quoted in the text,

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

where the sum s_2 is given by

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

Note that, since $\lambda_R(\vec{k})$ itself is already a first-order quantity in λ, in order to keep consistent orders of the expansion parameter the numerator of the equation that defines $\lambda_R(\vec{k})$ should be calculated to second order, while it is enough to calculate the denominator to first order.

3.2.10. Evaluate the asymptotic behavior of the difference of sums given in equation (3.2.5), for large values of N, approximating the sums by integrals over the momenta, as we did before in problem 3.2.4. Whenever it becomes necessary to use a minimum but non-vanishing value of the modulus of the momentum as a lower integration limit, use m_R as that value.

Bibliography

[1] **TODO**: fix all entries; reorganize the references to the first book; include HTML links for the web version.

[2] "The Gaussian Model – An Exploration into the Foundations of Quantum Field Theory", http://latt.if.usp.br/books/tgm/

[3] Gradshteyn, I. S., and Ryzhik, I. M., *Table of Integrals, Series and Products*, 4.$^{\underline{\text{th}}}$ ed., Academic Press, 1965.

[4] *Ibid.*[2], chapter 4, section 1.

[5] *Ibid.*[2], chapter 3, section 2.

[6] *Ibid.*[2], chapter 4, section 3.

[7] *Ibid.*[2], chapter 4, section 1.

[8] *Ibid.*[2], chapter 4, section 2.

[9] *Ibid.*[2], chapter 3, section 4, problem c3-s04-exr-posfactor.

[10] *Ibid.*[2], chapter 1, section 3.

[11] *Ibid.*[2], chapter 4, section 1.

[12] *Ibid.*[2], chapter 3, section 4.

[13] *Ibid.*[2], chapter 3, section 2.

[14] *Ibid.*[3], formula 3.461, equations 2 and 3, page 337.

[15] *Ibid.*[3], formula 3.462, equation 1, page 337.

[16] *Ibid.*[3], formula 3.621, equations 3 and 4, page 369.

[17] *Ibid.*[3], formula 8.250, equation 1, page 930.

[18] *Ibid.*[3], formula 8.335, equation 1, page 938.

[19] *Ibid.*[3], formula 8.431, equations 3 and 4, page 958.

[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.

[31] In one dimension, see: D. Rouelle, "Statistical Mechanics of a One-Dimensional Lattice Gas", Commun. Math. Phys. **9,** (1968), 267–278; F. Dyson, "Non-Existence of Spontaneous Magnetization in a One-Dimensional Ising Ferromagnet", Commun. Math. Phys. **12,** (1969), 212–215. In two dimensions, see: D. Mermin and H. Wagner, "Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models", Phys. Rev. Lett. **17,** (1966), 1133–1136; R. L. Dobrushin and S. B. Shlosman, "Absence of Breakdown of Continuous Symmetry in Two-Dimensional Models of Statistical Physics", Commun. Math. Phys. **42,** (1975), 31–40.

[32] *Ibid.*[2], chapter 2, section 2.

[33] *Ibid.*[2], chapter 2, section 8.

[34] *Ibid.*[2], chapter 2, section 11.

[35] *Ibid.*[2], chapter 6, section 2.

[36] The number quoted was obtained from M. N. Barber, R. B. Pearson, D. Toussaint and J. L. Richardson, Phys. Rev. **B32** (1985), 1720; it can also be found in N. Ito and M. Suzuki, J. of the Phys. Soc. of Japan **60** (1991), 1978–1987.

[37] The number quoted is $\ln\sqrt{1+\sqrt{2}}$ and was obtained from T. L. Hill, "Statistical Mechanics", McGraw-Hill, (1956, reprinted by Dover, 1987), page 329.

[38] "Numerical Solution for the Mean-Field Critical Curves of the $SO(\mathfrak{N})$ Polynomial $\lambda\varphi^4$ Models", http://latt.if.usp.br/scientific-pages/nsmfccsonpm/

[39] *Ibid.*[2], chapter 3, section 6.

[40] *Ibid.*[2], chapter 3, section 6, problem c3-s06-exr-threepnt.

[41] *Ibid.*[2], chapter 3, section 6, equation c3-s06-eqn-defZ.

[42] *Ibid.*[2], chapter 3, section 6, equation c3-s06-eqn-connecprop.

[43] *Ibid.*[2], chapter 3, section 6, problem c3-s06-exr-threepnt.

[44] *Ibid.*[2], chapter 3, section 4.

[45] *Ibid.*[2], chapter 3, section 7.

[46] *Ibid.*[2], chapter 3, section 7, equation c3-s07-eqn-invprop.

[47] *Ibid.*[2], chapter 3, section 7, equation c3-s07-eqn-invprop.

[48] *Ibid.*[2], chapter 3, section 7, equation c3-s07-eqn-defbox.

[49] *Ibid.*[2], chapter 4, section 1.