Received: *February 8, 2001*, Accepted: *June 25, 2001*

A model for gauge theories with Higgs fields

Frank Ferrari*

Joseph Henry Laboratories, Princeton University
Princeton, New Jersey 08544, USA
E-mail: fferrari@feynman.princeton.edu

Abstract: We discuss in details a simple, purely bosonic, quantum field theory belonging to larger class of models with the following properties:
(a) They are asymptotically free, with a dynamically generated mass scale.
(b) They have a space of parameters which gets quantum corrections drastically modifying the classical singularity structure. The quantum theory can have massless solitons, Argyres-Douglas-like CFTs, exhibit confinement, etc...
(c) The physics can, to a large extent, be worked out in models with a large number of supersymmetries as well as in purely bosonic ones. In the former case, exact BPS mass formulas can be derived, brane constructions and embedding in M theory do exist.
(d) The models have an interesting $1/N$ expansion, and it is possible to define a double scaling limit in the sense of the "old" matrix models when approaching the singularities in parameter space.
These properties make these theories very good toy models for four dimensional gauge theories with Higgs fields, and provide a framework where the effects of breaking supersymmetry can be explicitly studied. In our model, we work out in details the quantum space of parameters. We obtain the non-local lagrangian description of the Argyres-Douglas-like CFT, and show that it admits a strongly coupled fixed point. We also explicitly demonstrate property (d). The possibility of defining such double scaling limits was not anticipated on the gauge theory side, and could be of interest to understand the gauge theory/string theory correspondence.

Keywords: Sigma Models, Nonperturbative Effects, 1/N Expansion.

To Clélia.

*On leave from Centre National de la Recherche Scientifique, École Normale Supérieure, Paris.

JHEP06(2001)057

Contents

1. General presentation

1.1 Motivations

The present paper, which extends and provides full details on a previous work [1], grew up from the contradictory feelings one might have with regard to the dramatic progresses the theory of strongly coupled supersymmetric gauge fields and strings have undergone in the past six years (for reviews, see [2, 3, 4]). Though it is likely that the general intuition gained in studying supersymmetric examples will be useful to eventually understand more realistic theories, the specific methods and the analytic solutions of the supersymmetric theories will probably be irrelevant for solving genuinely non-supersymmetric models. This is due to the fact that those theories cannot be viewed as perturbation of supersymmetric theories (at least if the supersymmetric theories have at least eight supercharges), as early [5] as well as more recent [6] works tend to demonstrate. For example, a gravity approximation to the hypothetical string theory description of gauge theories, which can yield useful insights in supersymmetric or nearly supersymmetric models, cannot apply to QCD since the hadron spectrum is string-like. These limitations have led me to try to find a simple framework where the effects of breaking supersymmetry could be analyzed. The present paper is devoted to a detailed study of the simplest non-trivial, non-supersymmetric, model belonging to a large class of theories which are *simple enough to be tractable even in their strongly coupled, non-supersymmetric, regime, but complex enough so that many interesting questions about gauge theories have a counterpart in the simple models.* More precisely, and as will be explained in the following, our simple models have all of the following general properties:

(a) The models are tractable from the non-supersymmetric versions to the supersymmetric ones. In the latter case exact results can be obtained (BPS mass formulas in particular) in strict parallel with what is known for four dimensional supersymmetric gauge theories. Formulas can actually quantitatively coincide in these cases. Asymptotically free as well as conformal field theories can be studied.

(*b*) The models have an analogue of a moduli space, with generically both weakly coupled and strongly coupled regions. Strong quantum corrections then drastically modify the classical structure. At weak coupling we can have solitons playing the rôle of magnetic monopoles or dyons, and these can become massless at strong coupling singularities. Argyres-Douglas -like CFTs [7] can appear at strong coupling. All these phenomena can occur and be studied in both supersymmetric and non-supersymmetric theories.

(*c*) The supersymmetric versions of the models admit constructions in terms of branes, and they can be solved via M theory.

(*d*) The models have a non-trivial $1/N$ expansion à la 't Hooft [8]. The large-N limit can be unconventional, as for $N = 2$ super Yang-Mills [9].

We will illustrate (*a*), (*b*) and (*d*) in this article; (*c*) is already known, as we will review below. Two additional properties would also be desirable,

(*e*) The supersymmetric versions of the models can be geometrically engineered as in [10].

(*f*) The models are dual à la Maldacena to some kind of string theory.

Though I am not aware of any explicit construction, it is very likely that (*e*) is true, as explained later. As for (*f*), it is plausible that it could be true in view of (*c*) and (*d*), but we will have unfortunately nothing to say about this fascinating possibility in this paper.

A fundamental question of principle, that we would also like to address, is whether it is possible to prove, or at least to get a good general understanding, that gauge theories (supersymmetric or not) or other four dimensional field theories can have a description in terms of string theories. The modern starting point is a conjecture [11] motivated by the relationship between supersymmetric D-branes and solitons in closed string theories. It is not clear whether this intuitive understanding of the gauge theory/string theory correspondence makes sense in the general non-supersymmetric case. Interestingly, the results of the present paper suggest a way to understand the possible proliferation of dualities between four dimensional field theories and string theories. The idea is to show that double scaling limits [13] can be defined in the vicinity of the singularities of the moduli (or parameter) spaces of the gauge theories when a non-trivial interacting physics develops at low energies. The double scaled theory is then a string theory [12] that can be shown to describe the interacting low energy degrees of freedom. We will explicitly demonstrate that such double scaling limits can be defined for the model we consider in this paper. Remarkably, these ideas can be explicitly generalized to four dimensional theories as well [14].

1.2 A family of toy models for gauge theories with Higgs fields

The fact that good toy models exist for four dimensional gauge theories is of course not new. A quarter of a century ago, Polyakov showed in [15] that two dimensional non-linear σ models can be asymptotically free, undergo dimensional transmutation, and develop infrared slavery, the landmarks of interesting four dimensional gauge theories. Based on this idea, many interesting results were then obtained (see e.g. [16]). The models we will propose are only modest extensions of the original non-linear σ models, the main new input being a way to mimic the presence of Higgs fields, which are conspicuous in modern studies. It seems that either it was not known that the standard non-linear σ models could be modified in a way that would make them suitable to compare with modern gauge theory studies, or it was not known that with these modifications the models would still be tractable enough in the non-supersymmetric cases. It is of course disappointing that we have to restrict ourselves to two-dimensional models, but, even at the turn of the millenium, this is still the price to pay to discuss the effect of breaking supersymmetry in strongly coupled theories. I hope that the results of [1] and of this paper will convince the reader that the models we are proposing constitute a nice, if modest, playground to study this fundamental question.

We now turn to describe the basic idea which led to the construction of our models. One important peculiarity of supersymmetric gauge theories à la Seiberg-Witten [17] is that they have massless scalar fields and a continuous moduli space of vacua. This is heavily used in the discussion of the theories (as well as in many other supersymmetric systems), and may appear as being an obvious and impassable obstacle in trying to find non-supersymmetric analogues. In non-supersymmetric theories, any vacuum degeneracy that may be present classically is generically lifted quantum mechanically. In two dimensions, the situation is even worse: even supersymmetric theories cannot have a continuous moduli space, because the strong infrared fluctuations always make the vacuum wave functional to spread over the whole would-be moduli space (a discrete set may remain [18]). However, it turns out that this is only an outward problem, for the following reason. In four dimensions, the moduli space is parametrized by Higgs vacuum expectation values (or more precisely by the vevs of independent gauge invariant combinations of Higgs fields). The physics is interesting because the masses of the gauge bosons, which govern the low energy coupling, depend on the Higgs vevs via the usual Higgs mechanism. Moving on the moduli space is then equivalent to varying the low energy coupling, and a very interesting physics is associated to the transition from weak coupling to strong coupling:

appearance of strong coupling singularities [17], non-trivial CFTs [7], rearrangement of the spectrum of stable states [19], etc... All this physics is largely independent of the fact that the Higgs potential has flat directions, but strongly depends on our ability to vary the low energy coupling. We are thus led to the conclusion that a good non-supersymmetric analogue of the moduli space of supersymmetric theories could be a space of parameters on which the low energy coupling depends. This is too vague, since there are a priori many ways to vary the low energy coupling, for example by putting the theory on a sphere of varying radius or considering a finite temperature. To stick as close as possible to the supersymmetric gauge theory case, we will in general consider parameters that correspond to giving masses to the fields that contributes with a minus sign to the β function. In the case of the non-linear σ model with target space the $N-1$ sphere S^{N-1} that we will consider in the present paper, this simply amounts to giving a mass to the $N-1$ would-be Goldstone bosons. These mass parameters play the rôle of Higgs vevs, and span a space of parameters $\mathcal{M}$. One of our main goal is to compute the quantum corrections to $\mathcal{M}$, in the same sense as quantum corrections to the moduli space of supersymmetric gauge theories were computed in [17] or [20].

Though the very simple idea presented above certainly suggests that the space of mass parameters in non-linear σ models is an interesting object to consider, the reader may not be convinced that it is really a good analogue of the moduli spaces of supersymmetric gauge theories. The present author himself actually became convinced that it is the case only after he became aware of papers by Hanany and Hori [21] and Dorey and collaborators [22], where the $\mathcal{N}=2$ supersymmetric $\mathbb{CP}^{N-1}$ non-linear σ model with mass terms is discussed (in this context, the mass terms are called "twisted masses," and are in one to one correspondence with the holomorphic isometries of the target Kähler manifold [23]). In [21], a brane construction of this supersymmetric model is given, and M theory is used to obtain non-perturbative results, in strict parallel to what was done in the case of $\mathcal{N}=2$ super Yang-Mills [24]. The striking, in some sense quantitative, similarity with $\mathcal{N}=2$ super Yang-Mills was then further discussed in [22]. The property making brane constructions possible is that the mass parameters can be interpreted in the $\mathbb{CP}^{N-1}$ model as vector multiplets vevs in a gauged linear formulation of the model, and thus literally parametrize a Coulomb branch. The works [21, 22] thus clearly demonstrate that an excellent toy model for $\mathcal{N}=2$ super Yang-Mills in four dimensions is the $\mathcal{N}=2$ $\mathbb{CP}^{N-1}$ non-linear σ model with mass terms in two dimensions (in spite of the fact that $\mathcal{N}=2$ in four dimensions corresponds to eight supercharges, whereas $\mathcal{N}=2$ in two dimensions corresponds to four supercharges). In retrospect, this analogy is not too surprising, and there are actually many independent arguments in favor of this correspondence. For example, both types of theories are known to have similar non-renormalization theorems ([25, 26]), and both admit topological twists ([27, 28]). I think that this latter property is particularly significant, because the exact results

4 D gauge theories	2 D non-linear σ models
Gauge bosons	Goldstone bosons
Gauge group	Isometry group
Number of colors	Dimension of the compact target space
Gauge coupling constant	Inverse radius of the target manifold
Higgs fields	Mass terms
Moduli space $\mathcal{M}$	Space of mass parameters $\mathcal{M}$
Monopoles and dyons	Kinks
S duality	Kramers-Wannier duality
Argyres-Douglas CFT	Ginzburg-Landau CFT
Eight supercharges	Four supercharges
Effective prepotential $\mathcal{F}$	Effective superpotential W
String description	Branched polymer description

Table 1: A gauge theory/σ model dictionary

à la Seiberg-Witten ([17, 2]) are likely to have a semi-topological origin, very much like the exact results one can obtain in two dimensions [29]. Another important fact is that effective superpotentials in two dimensions are very similar to effective prepotentials in four dimensions. This was emphasized for example in [30] where two dimensional superpotentials were geometrically engineered using singular Calabi-Yau fourfolds and mirror symmetry, in the same spirit as four dimensional prepotentials can be engineered using singular Calabi-Yau threefolds and mirror symmetry [10]. Though the models considered in [30] were different from the one we are proposing in this paper, a similar construction is probably possible.

We believe that this analogy between supersymmetric theories in four and two dimensions can be fruitfully extended to the non-supersymmetric cases, and used to study the effects of breaking supersymmetry. We present below a short dictionary for the correspondence between four dimensional and two dimensional models. Some of the entries will be exemplified in later sections.

1.3 Plan of the paper

We give a general discussion of mass terms, including their renormalization properties, in section 2.1. We work out a few simple examples when the target space is the sphere S^{N-1}, including a model that we will call after C. Neumann and on which we will focus. We also briefly discuss the supersymmetric generalizations. In section 2.2, the semiclassical properties of the Neumann model are explored, including the singularity structure of the classical space of parameters $\mathcal{M}_{\rm cl}$, bound states and solitons. section 3 is devoted to the $1/N$ expansion. Emphasis is put on the peculiarities introduced by the fact that the dimension of $\mathcal{M}$ is of order N, and on general limitations of the $1/N$ expansion in our model. We compute the quantum corrections to the metric as well as the mass of some stable particles. We then show how to go beyond the standard $1/N$ expansion, and we obtain a provisional picture of the singularity structure of the quantum space of parameters $\mathcal{M}_{\mathrm{q}}$ (also simply denoted by $\mathcal{M}$ in the following). In section 4 we work out the full structure of $\mathcal{M}$, in the large-N limit. In section 5 we analyze the theory in the vicinity of singularities on $\mathcal{M}$, for any finite N. We show that the non-trivial infrared physics is most naturally described in terms of a non-local lagrangian analogous to a theory of electric and magnetic charges. This brings the analogy with an Argyres-Douglas theory to a climax. Due to the simplicity of two dimensional theories, we are able to show that the non-local theory has a strongly coupled fixed point, and we recover the results of sections 3 and 4 independently of a large-N approximation. In section 6, we briefly show that double scaling limits, in the sense of [13], can be defined as the singularities on $\mathcal{M}$ are approached. This has never been studied in the context of two dimensional non-linear σ models, and our results suggest very interesting possibilities for four dimensional gauge theories that have not been anticipated. Moreover, these double scaling limits have interesting properties, which we will discuss elsewhere [14].

We have also included three appendices. In appendix A, we find the most general static finite energy solutions to the field equations of our model at arbitrary N. A notable result is that we obtain *static* solutions describing several standard sine-Gordon solitons at arbitrary distances from each other. In appendix B we derive several simple formulas used in the main text to study the large-N limit of the theory. Finally in appendix C we compute the $1/N$ corrections to the equation of the critical hypersurface obtained in section 3. The calculation illustrates some generic properties of the $1/N$ expansion, as well as some subtleties associated with infrared divergences that can plague this expansion in our model. The result gives a consistency check of the simple calculations of section 3.

To the opposite of [1], we provide in this work detailed derivations and elementary discussions of various points. This explains in part the length of the paper.

2. The Neumann model

2.1 Mass terms in the non-linear sigma model

In the following, the target space of our non-linear σ models is the $N-1$ sphere S^{N-1}. The discussion could be easily generalized to sigma models on symmetric spaces, for example. At the technical level, the renormalization properties of the two dimensional non-linear sigma model are non-trivial [31], and their renormalizability can be proven with the help of a sort of Zinn-Justin equation, very much like the renormalizability of gauge theories is analyzed (see e.g. [32]). In the latter case, it is not straightforward to add mass terms for the gauge bosons that do not spoil the renormalizability of the model. To do so, it is necessary to introduce scalar fields, the Higgs bosons, and the gauge bosons masses are then largely determined by the transformation properties of the Higgs fields under the gauge group G through the Higgs mechanism. The situation is similar for mass terms in non-linear sigma models: they generally spoil the renormalizability of the theory, and from this point of view they are most naturally characterized by their transformation properties under the isometry group $G = \mathrm{O}(N)$ of the target space. The explicit form of the mass terms in the lagrangian is then determined unambiguously by G invariance and power counting.

2.1.1 Renormalization theory

Working in the euclidean, the lagrangian without mass terms is

$$L_{\mathrm{kin}} = \frac{1}{2} \sum_{i,j=1}^{N-1} g_{ij}\, \partial_\alpha \Phi_i \partial_\alpha \Phi_j \, , \tag{2.1}$$

where g_{ij} is the standard $\mathrm{O}(N)$ invariant metric on a sphere of radius $1/g^2$, g being the dimensionless coupling constant analogous to the gauge coupling constant. By taking the Φ_is to be the usual cartesian coordinates, we have

$$g_{ij} = \delta_{ij} + \frac{g^2 \Phi_i \Phi_j}{1 - g^2 \sum_{i=1}^{N-1} \Phi_i^2} \, . \tag{2.2}$$

The model is defined by a path integral over the fields Φ_i, $1 \leq i \leq N-1$, with an $\mathrm{O}(N)$ invariant measure. We have interactions of the form $g^{2n} \, (\partial\Phi)^2 \, \Phi^{2n}$ for all $n \geq 1$. A mass term is defined to be a couple $(\mathbf{\Gamma}, d)$ where $\mathbf{\Gamma}$ is an irreducible representation of $\mathrm{SO}(N)$ and d is the canonical dimension of the mass parameters. In two dimensions, it is consistent to have $d = 1$ or $d = 2$. In the latter case, the parameters actually correspond to mass squared. Different choices of $\mathbf{\Gamma}$ and d do not necessarily correspond to independent mass terms, as we will see. If $d = 1$ the mass parameters are denoted by $\mathbf{m} \equiv m_{i_1 \cdots i_p}$, and if $d = 2$ by $\mathbf{h} \equiv h_{i_1 \cdots i_p}$. The i_js are $\mathrm{SO}(N)$ indices, and the tensors $\mathbf{m}$ or $\mathbf{h}$ satisfies some constraints depending on $\mathbf{\Gamma}$.

Instead of working with the lagrangian (2.1), which is possible but awkward, we will introduce a Lagrange multiplier field α and work in a representation where the $\mathrm{O}(N)$ symmetry is linearly realized:

$$L_{\mathrm{kin}} = \frac{1}{2} \, \partial_\alpha \mathbf{\Phi} \partial_\alpha \mathbf{\Phi} + \frac{1}{2} \, \alpha \, \left(\mathbf{\Phi}^2 - \frac{1}{g^2} \right) , \tag{2.3}$$

with $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_N)$. Eliminating Φ_N from (2.3) by using the constraint enforced by α, one recovers the non-linear lagrangian (2.1). We will regulate the theory by a simple momentum cutoff Λ_0, which is manifestly $\mathrm{O}(N)$ invariant. Introducing $\mathbf{m}$ or $\mathbf{h}$ into the game, the action must then be taken to be the most general relativistic local functional invariant under O(N) (by varying both the fields and the parameters) and compatible with power counting. The canonical dimensions of the fields are

$$[\Phi] = 0\,, \qquad [\alpha] = 2\,. \tag{2.4}$$

As we will further discussed in section 3.1, these dimensions are relevant to study the UV finiteness of both ordinary perturbation theory and of the $1/N$ expansion because the model is asymptotically free. Due to its canonical dimension, α can only appear in a term of the form $F(\mathbf{\Phi}^2)\,\alpha$, where F is an arbitrary function. By integrating over α, we get the quantum constraint

$$F(\mathbf{\Phi}^2) = 0\,, \tag{2.5}$$

which can always be solved in the UV as

$$\mathbf{\Phi}^2 = \frac{1}{ZZ_g g^2}\,, \tag{2.6}$$

where ZZ_g is a renormalization constant. This shows that, as long as we integrate over α without introducing a source for this field, F can be taken without loss of generality to be proportional to $\mathbf{\Phi}^2 - 1/(ZZ_g g^2)$, and $\mathbf{\Phi}^2$ can be replaced by its constant value in the other terms of the lagrangian. The subtleties associated with the renormalization of α-dependent quantities will only show up in appendix C, and is discussed there. Without any mass term, the only non-zero dimension two operator that remains is proportional to $\partial_\alpha \mathbf{\Phi} \partial_\alpha \mathbf{\Phi}$, and from this we deduce the form of the renormalized lagrangian [31]

$$L_{\text{kin}} = \frac{Z}{2}\,\partial_\alpha \mathbf{\Phi} \partial_\alpha \mathbf{\Phi} + \frac{Z}{2}\,\alpha\left(\mathbf{\Phi}^2 - \frac{1}{ZZ_g g^2}\right). \tag{2.7}$$

$\mathbf{\Phi}$ and g are now the renormalized fields and coupling constant, related to the bare quantities by

$$\mathbf{\Phi} = \frac{\mathbf{\Phi}_0}{\sqrt{Z}}\,, \qquad g = \frac{g_0}{\sqrt{Z_g}}\,. \tag{2.8}$$

In the following, we investigate the simplest mass terms and find the scalar potential V they correspond to.

Singlet. Because of the constraint (2.6), singlet mass terms are trivial and simply correspond to adding a constant to the lagrangian.

Vector h_i of dimension two. This term corresponds to a standard magnetic field,

$$V_{(v,2)} = -Z_{(v,2)}\,\mathbf{h}\mathbf{\Phi} = -Z_{(v,2)}\,h\Phi_N\,, \tag{2.9}$$

where we have used the $\mathrm{O}(N)$ symmetry to align the magnetic field with the Nth direction. This is a linear symmetry breaking term, and thus $Z_{(v,2)} = 1$. By eliminating Φ_N we have

$$V_{(v,2)} = -Z_{(v,2)} \, \frac{h}{g} \sqrt{1 - g^2 \sum_{i=1}^{N-1} \Phi_i^2} \,. \tag{2.10}$$

In addition to giving a mass $\sqrt{Z_{(v,2)} gh}$ to the would-be Goldstone bosons, this term also produces an infinite sum of new interactions. This model can easily be studied in the large-N limit, but it is not particularly interesting.

Vector m_i of dimension one. The only new invariant operators of dimension 2 are $(\mathbf{m\Phi})^2$ and $\mathbf{m}^2$. One has also the dimension one operator $\mathbf{m\Phi}$. The potential is then

$$\begin{aligned} V_{(v,1)} &= -\frac{Z_{(v,1)}}{2} \, (\mathbf{m\Phi})^2 - Z'_{(v,1)} \, M \mathbf{m\Phi} - Z''_{(v,1)} \mathbf{m}^2 \\ &= -\frac{Z_{(v,1)}}{2} \, m^2 \Phi_N^2 - Z'_{(v,1)} \, M m \Phi_N - Z''_{(v,1)} m^2 \,. \end{aligned} \tag{2.11}$$

M is a new mass parameter that generically enters the problem. However, it can be consistently set to zero thanks to the symmetry $\Phi_N \mapsto -\Phi_N$. $Z''_{(v,1)}$ renormalizes the vacuum energy.

Antisymmetric tensor of dimension two. No new operator can be constructed,

$$V_{(a,2)} = 0 \,. \tag{2.12}$$

Traceless symmetric tensor h_{ij} of dimension two This corresponds to

$$V_{(s,2)} = -\frac{Z_{(s,2)}}{2} \, \mathbf{\Phi h \Phi} = -\frac{Z_{(s,2)}}{2} \sum_{i=1}^{N} h_i \Phi_i^2 \,. \tag{2.13}$$

By diagonalizing the matrix h_{ij}, we see that we have $N-1$ independent parameters, since the h_is satisfy $\sum_{i=1}^{N} h_i = 0$. This tracelessness condition can actually be waived without changing the physics, since the trace part of $\mathbf{h}$ corresponds to adding a constant to the lagrangian. This model, that we will call the Neumann model for reasons to become clear later in this section, is the one on which we want to focus in this paper. We will usually use the $N-1$ independent variables

$$v_i = h_N - h_i \,, \qquad 1 \leq i \leq N-1 \,. \tag{2.14}$$

Note that in the particular case $v_1 = \cdots = v_{N-1} = m^2$ we recover the model for a dimension one vector mass parameter with $M = 0$. By eliminating Φ_N we find, up to a constant term, the very simple potential

$$V = \frac{Z_{(s,2)}}{2} \sum_{i=1}^{N-1} v_i \Phi_i^2 \,. \tag{2.15}$$

Antisymmetric tensor a_{ij} of dimension one. Dimension two operators can be constructed by using $t_{ijkl} = a_{ij}a_{kl}$. This tensor decomposes into four irreducible representations of SO(N), amongst which only two can be used to construct invariants with the help of the Φ_is: the trivial representation $\operatorname{tr}\mathbf{a}^2$, and t_{ikjk}. We thus get a special case of the previous model, for which

$$h_{ij} = -\sum_{k=1}^{N} a_{ik}a_{jk}\,. \tag{2.16}$$

This special case is particularly significant, however, because it is on this form that the Neumann model can be supersymmetrized [33].

The Neumann model is singled out by the fact that it is the most general model with only quadratic terms in the fields Φ when the Lagrange multiplier α is introduced. This makes the analysis of the large-N limit very simple as we will see in section 3. More general models can nevertheless be studied in the large-N limit, with the ideas presented in the present paper, but at the expense of introducing additional auxiliary fields (the large-N limit can be subtle in some cases, though). In this sense the Neumann model is "minimal," since it can be studied by introducing the minimal number of auxiliary fields.

It is natural to consider another simple model.

Traceless symmetric tensor m_{ij} of dimension 1. To work out the potential term, we must decompose $t_{ijkl} = m_{ij}m_{kl}$ into irreducible representations of SO(N). Four such representations occur, three of which can be used to construct three independent invariants, including the trivial representation which renormalizes the vacuum energy. This implies that, in addition to the mass parameters themselves, the model depends on a new dimensionless coupling constant G. Generically, we will also have the dimension one operator $\mathbf{\Phi m \Phi}$ and a new mass parameter M. The potential is

$$\begin{aligned} V_{(s,1)} &= \frac{Z_{(s,1)}}{2}\,\mathbf{\Phi m}^2\mathbf{\Phi} - \frac{Z'_{(s,1)}G^2}{2}\,(\mathbf{\Phi m \Phi})^2 - Z''_{(s,1)}\,M\,\mathbf{\Phi m \Phi} - Z'''_{(s,1)}\operatorname{tr}\mathbf{m}^2 \\ &= \frac{Z_{(s,1)}}{2}\sum_{i=1}^{N} m_i^2\Phi_i^2 - \frac{Z'_{(s,1)}G^2}{2}\left(\sum_{i=1}^{N} m_i\Phi_i^2\right)^2 - Z''_{(s,1)}M\sum_{i=1}^{N} m_i\Phi_i^2 - Z'''_{(s,1)}\sum_{i=1}^{N} m_i^2\,. \end{aligned} \tag{2.17}$$

M can be set to zero thanks to the symmetry $m_i \mapsto -m_i$. G can also be consistently set to zero, since we then recover a special case of the Neumann model, for which

$$h_{ij} = -\sum_{k=1}^{N} m_{ik}m_{jk}\,. \tag{2.18}$$

Considering $G \neq 0$ is however interesting, since the model can be supersymmetrized for $G = g$ [33] (in the non-supersymmetric version, it is not consistent to set $G = g$, and G and g are then two independent coupling constants, see below). We thus have two natural supersymmetric versions of the Neumann model, with $\mathbf{h}$ taking the special form (2.16), or (2.18) with $G = g$. In the latter case, the scalar potential can be written most elegantly by introducing a new auxiliary field σ, which turns out to be in the same supersymmetry multiplet as the Lagrange multiplier α,

$$V_{(s,1)} = \frac{Z_{\mathrm{SUSY}}}{2} \sum_{i=1}^{N} (\sigma + m_i)^2\, \Phi_i^2\,. \tag{2.19}$$

This is the $\mathcal{N} = 1$ supersymmetric version of the $\mathcal{N} = 2$ potential considered in [21] and [22], which is

$$V_{\mathcal{N}=2} = \frac{1}{2} \sum_{i=1}^{N} |\sigma + m_i|^2\, |\Phi_i|^2, \tag{2.20}$$

where σ and the Φ_is are now complex fields. The $\mathcal{N} = 2$ theory with the potential (2.20) shows quantitative similarities with $\mathcal{N} = 2$ super Yang-Mills in four dimensions [21, 22], as we have already pointed out in section 1. We thus see that the Neumann model can be viewed as a bosonic version of the supersymmetric theories studied in [21, 22].

2.1.2 Renormalization constants

One-loop formulas for the various renormalization constants can be easily obtained with the background field method and using Riemann normal coordinates, which is very elementary in the case of the sphere. We recover the well known formulas

$$\begin{aligned} Z_g &= 1 + \frac{N-2}{2\pi}\, g^2 \ln \frac{\mu}{\Lambda_0} + \mathcal{O}(g^4)\,, \\ Z &= 1 + \frac{g^2}{2\pi} \ln \frac{\mu}{\Lambda_0} + \mathcal{O}(g^4)\,, \end{aligned} \tag{2.21}$$

and we also obtain

$$Z_{(s,2)} = 1 - \frac{g^2}{2\pi} \ln \frac{\mu}{\Lambda_0} + \mathcal{O}(g^4)\,. \tag{2.22}$$

Λ_0 is the UV cutoff and μ an arbitrary renormalization scale. The RG functions of the Neumann model are then (we have indicated β up to two loops [34], since this is needed later)

$$\beta(g^2) = \frac{\partial g^2}{\partial \ln \mu} = -\frac{N-2}{2\pi}\, g^4 - \frac{N-2}{(2\pi)^2}\, g^6 + \mathcal{O}(g^8)\,, \tag{2.23}$$

$$\gamma(g^2) = \frac{\partial \ln Z}{\partial \ln \mu} = \frac{g^2}{2\pi} + \mathcal{O}(g^4)\,, \tag{2.24}$$

$$\sigma(g^2) = \frac{\partial \ln Z_{(s,2)}}{\partial \ln \mu} = -\frac{g^2}{2\pi} + \mathcal{O}(g^4)\,. \tag{2.25}$$

JHEP06(2001)057

The theory is asymptotically free for $N \geq 3$ [15], with a dynamically generated mass scale Λ defined at one loop by the equation for the coupling at scale μ,

$$\frac{1}{g^2(\mu)} = \frac{N-2}{2\pi} \ln \frac{\mu}{\Lambda} \,. \tag{2.26}$$

We have also computed

$$Z_{(s,1)} = 1 - \frac{g^2 + 4G^2}{2\pi} \ln \frac{\mu}{\Lambda_0} + \mathcal{O}(g^4) \,, \tag{2.27}$$

$$Z'_{(s,1)} = 1 - \frac{3g^2}{\pi} + \mathcal{O}(g^4) \,, \tag{2.28}$$

from which we can deduce the β function for the coupling G,

$$\beta_G(G^2, g^2) = -\frac{G^2(2G^2 - 3g^2)}{\pi} + \mathcal{O}(g^4) \,. \tag{2.29}$$

This shows unambiguously that G is a new coupling, independent of g, and thus that introducing mass terms for the Goldstone bosons will generically introduce new dimensionless coupling constants into the theory. This is another common point with gauge theories, where in addition to the gauge coupling constant we have the couplings in the Higgs potential. However, in two dimensions, the new couplings cannot spoil asymptotic freedom: in spite of possible plus signs in β functions like (2.29), the physical coupling is really something like $G^2 m^2$ where m is some mass parameter, and is always irrelevant in the UV.

2.2 The Neumann model

From now on, we focus exclusively on the Neumann model, whose minkowskian classical lagrangian and field equations are

$$L = \frac{1}{2} \sum_{i=1}^{N} \Big(\partial_\mu \Phi_i \partial^\mu \Phi_i + h_i \Phi_i^2 \Big) - \frac{\alpha}{2} \left(\sum_{i=1}^{N} \Phi_i^2 - \frac{1}{g^2} \right), \tag{2.30}$$

$$\partial_\mu \partial^\mu \Phi_i = (h_i - \alpha)\, \Phi_i \,. \tag{2.31}$$

The space of parameters $\mathcal{M}$ is spanned by the $N-1$ real independent variables

$$v_i = h_N - h_i \,. \tag{2.32}$$

A singularity on $\mathcal{M}$ is a point where some of the degrees of freedom become massless. The set of all the singularities on $\mathcal{M}$ is called the critical hypersurface $\mathcal{H} \subset \mathcal{M}$. We will distinguish between the classical critical hypersurface $\mathcal{H}_{\rm cl}$ and the quantum critical hypersurface $\mathcal{H}_{\rm q}$. We will also speak loosely of a classical $\mathcal{M}_{\rm cl}$ and quantum $\mathcal{M}_q$ space of parameters, meaning $\mathcal{M}$ equipped with $\mathcal{H}_{\rm cl}$ or $\mathcal{H}_q$ respectively. For generic values of the parameters v_i, the model has a symmetry $\mathbb{Z}_{2(1)} \times \cdots \times \mathbb{Z}_{2(N)}$ with

$$\mathbb{Z}_{2(i)} : \Phi_j \longmapsto (-1)^{\delta_{ij}} \, \Phi_j . \tag{2.33}$$

When the v_is coincide for p disjoint subsets of indices $I_1, \ldots, I_p$, of respective cardinality $k_1, \ldots, k_p$, the symmetry group is enlarged to $\times_{i \notin I_1 \bigcup \cdots \bigcup I_p} \mathbb{Z}_{2(i)} \times \mathrm{O}(k_1) \times \cdots \times \mathrm{O}(k_p)$.

The Neumann model, in addition of being a natural extension of the standard non-linear σ model, is also a generalization of the sine-Gordon model. Indeed the classical potential

$$V_{\mathrm{cl}} = -\frac{1}{2} \sum_{i=1}^{N} h_i \Phi_i^2 \tag{2.34}$$

reduces for $N = 2$ to the sine-Gordon potential

$$V_{\mathrm{sG}} = -\frac{v_1}{4g^2} \cos 2\theta \tag{2.35}$$

by writing

$$\Phi_1 = \frac{\sin\theta}{g} , \qquad \Phi_2 = \frac{\cos\theta}{g} . \tag{2.36}$$

In the case $N = 3$, our model is also related, through Haldane's map [35], to the anisotropic XYZ quantum large integer spin chain in one dimension, or equivalently to the anisotropic classical Heisenberg model in two dimensions. This analogy is quite helpful to understand the physics of the model. The model for $N = 3$ has instantons, and we could introduce in this case a non-zero θ angle. The physics would then strongly depend on θ. We will not do that in the following, however, since we want to study the models for arbitrary N in a unified way. Considering a non-zero θ would be more natural in the context of a work on the $\mathbb{CP}^N$ model with mass terms.

The model defined by (2.30) is also the field theoretic generalization of a famous integrable system in classical mechanics corresponding to the motion of a particle on a sphere with a quadratic potential,

$$\frac{d^2\Phi_i}{dx^2} = (\alpha - h_i)\, \Phi_i, \quad \sum_{i=1}^{N} \Phi_i^2 = \frac{1}{g^2} . \tag{2.37}$$

This problem was first studied by C. Neumann 150 years ago (for $N = 3$), hence the name "Neumann" for the model (2.30). The time independent solutions to the field theory satisfy (2.37). In appendix A we use integrability to find the most general solitonic solutions of the Neumann model, generalizing the famous sine-Gordon solitons.

2.3 The weakly coupled theory

At weak coupling, the classical vacua of the theory correspond to the minima of the potential (2.34) defined on the sphere $\sum_{i=1}^{N}\Phi_i^2 = 1/g^2$. If we restrict ourselves, without loss of generality, to the region $\mathcal{M}_N$ of parameter space defined by $v_i \geq 0$, then we have generically two equivalent vacua related by the spontaneously broken $\mathbb{Z}_{2(N)}$ symmetry, such that

$$\langle\Phi_i\rangle = \pm\frac{\delta_{iN}}{g}\,. \tag{2.38}$$

By expanding around one of these vacua, we read from the lagrangian that the $N-1$ "Goldstone bosons" Φ_i, $1\leq i\leq N-1$, have masses

$$m_i = \sqrt{v_i}. \tag{2.39}$$

This relation together with (2.26) shows that the coupling is weak in the region

$$v_i \gg \Lambda^2 \qquad \text{for } 1\leq i\leq N-1\,. \tag{2.40}$$

It is actually sufficient to have $N-2$ heavy fields for the physics to be weakly interacting, for example

$$v_i \gg \Lambda^2 \qquad \text{for } 1\leq i\leq N-2\,. \tag{2.41}$$

The low energy theory is then a sine-Gordon model whose elementary field has a mass $\sqrt{v_{N-1}}$ and whose non-running coupling is small.

The classical hypersurface of singularities $\mathcal{H}_{\text{cl}}$ in $\mathcal{M}_N$ is trivially determined by (2.39) to be the union of the hyperplanes $v_i = 0$. In addition to $\mathcal{M}_N$, there are $N-1$ other regions $\mathcal{M}_j \subset \mathcal{M}$ defined by

$$\mathcal{M}_j = \Big\{(v_1,\ldots,v_{N-1})\mid v_i^{(j)} = v_i - v_j \geq 0\Big\}\,, \tag{2.42}$$

or equivalently by $h_j = \max_i h_i$ in $\mathcal{M}_j$. $\mathcal{H}_{\text{cl}}\cap\mathcal{M}_j$ is then the union of the hyperplanes $v_i^{(j)} = 0$ in $\mathcal{M}_j$. We have reprensented $\mathcal{M}_{\text{cl}}$ for the cases $N=3$ and $N=4$ in figure 1.

2.3.1 Integrability?

In view of the relation of our model with the Neumann integrable system, the $\mathrm{O}(N)$ non-linear σ model and the sine-Gordon model, the reader may wonder whether it could be integrable itself or not. However, the simple arguments that can be used to prove integrability when $v_1=\cdots=v_{N-1}=0$ or $N=2$ no longer work, because the model combines in general the complications due to both an explicit mass term and a non-zero β function. In other words, there are enough operators to spoil the higher conservation laws of the standard non-linear σ model. Integrability can actually be

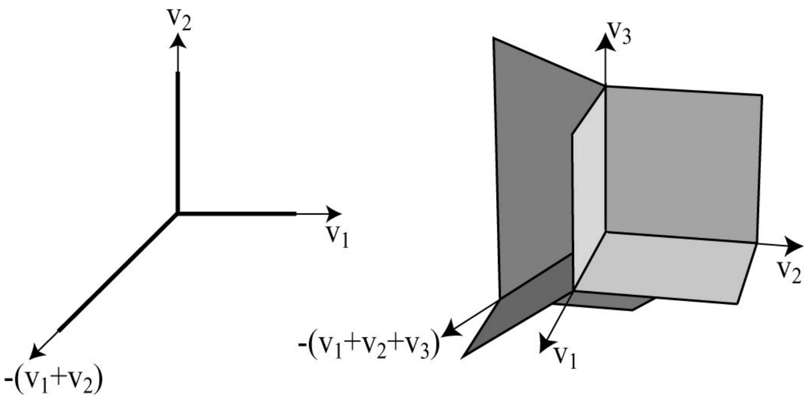

Figure 1: The classical space of parameters $\mathcal{M}_{\rm cl}$ in the cases $N = 3$ (left) and $N = 4$ (right).

easily disproved in the weakly coupled region of parameter space, for example by computing the S matrix element for the scattering of elementary particles,

$$(i, \mathbf{q}_1) + (j, \mathbf{q}_2) \longrightarrow (k, \mathbf{p}_1) + (l, \mathbf{p}_2)\,, \tag{2.43}$$

where $\mathbf{q}_1 = (\omega_{i,q_1}, q_1), \mathbf{q}_2$ and $\mathbf{p}_1, \mathbf{p}_2$ are the incoming and outgoing two-momenta respectively. In the Born approximation, which is valid at weak coupling, we have

$$\begin{aligned}\langle k, p_1; l, p_2 \mid S - 1 \mid i, q_1; j, q_2\rangle &= \delta^{(2)}(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{q}_1 - \mathbf{q}_2) \times \\ &\quad\times \frac{ig^2}{16\pi^2}\Big(s\delta_{ij}\delta_{kl} + t\,\delta_{ik}\delta_{jl} + u\,\delta_{il}\delta_{jk}\Big)\,, \end{aligned} \tag{2.44}$$

where $s = (\mathbf{q}_1 + \mathbf{q}_2)^2$, $t = (\mathbf{q}_1 - \mathbf{p}_1)^2$ and $u = (\mathbf{q}_1 - \mathbf{p}_2)^2$ are the usual Mandelstam variables, and with the normalization

$$\langle j, p \mid i, q\rangle = \omega_{i,q}\,\delta_{ij}\delta(q - p) = \sqrt{q^2 + m_i^2}\,\delta_{ij}\delta(q - p)\,. \tag{2.45}$$

It is clear from (2.44) that processes where the number of particles of a given mass changes are allowed, which proves that the model cannot be integrable in this regime [37]. In the O$(N-1)$ symmetric case $v_i = \cdots = v_{N-1}$ where all the elementary particles have the same mass, the S matrix element can be written

$$\begin{aligned}\langle k, p_1; l, p_2 \mid S - 1 \mid i, q_1; j, q_2\rangle &= \omega_{q_1}\delta(q_1 - p_1)\,\omega_{q_2}\delta(q_2 - p_2) \times \\ &\quad\times(\delta_{ij}\delta_{kl}S_1 + \delta_{ik}\delta_{jl}S_2 + \delta_{il}\delta_{jk}S_3) + (k \leftrightarrow l, p_1 \leftrightarrow p_2)\,, \end{aligned} \tag{2.46}$$

with

$$S_1 = \frac{ig^2}{8\pi^2}\,\frac{s}{\sqrt{-su}}\,, \qquad S_2 = 1\,, \qquad S_3 = \frac{ig^2}{8\pi^2}\,\frac{u}{\sqrt{-su}}\,. \tag{2.47}$$

On this form, it is clear that the Yang-Baxter equation is violated [37], and thus the theory cannot be integrable even in the most symmetric case. The above reasoning does not exclude that the model could be integrable for some special values of the parameters at strong coupling, but apart from the case $v_1 = \cdots = v_{N-1} = 0$ we think it is very unlikely. This is confirmed by the analysis in the large-N approximation, see section 3.

2.3.2 Bound states

From what is known for the sine-Gordon model, we can suspect that we have bound states of the elementary particles at weak coupling. This is a priori possible because the interactions between the Φ_is are attractive thanks to the derivatives in the interaction terms (see (2.1) and (2.2)). At weak coupling, we can use a non-relativistic approximation to investigate the bound state spectrum. The quantum mechanical attractive potential in the two-particle subspace can be straightforwardly deduced from the S matrix element (2.44) and is simply

$$V(X) = -\frac{1}{2}\, g^2\, \delta(X) \otimes J\,, \tag{2.48}$$

where J acts in internal space,

$$\langle kl \mid J \mid ij \rangle = \delta_{ij}\delta_{kl}\,. \tag{2.49}$$

We see that processes like $(ii) \to (jj)$ with $i \neq j$ are possible, as long as the kinematical non-relativistic constraint $m_i = m_j$ is satisfied. In general, stable bound states will exist between particles of the lowest mass. Let us study the two-particle bound states in the symmetric case $v_1 = \cdots = v_{N-1} = v$. This amounts to solving the Schrödinger equation for two particles of mass $\sqrt{v}$ interacting through (2.48). The wavefunctions $\psi_{ij}(x_1, x_2)$ must satisfy $\psi_{ii}(x_1, x_2) = \psi_{ii}(x_2, x_1)$ due to bose statistics. It is very elementary to see that there is a unique bound state which is a singlet of $\mathrm{O}(N-1)$ ($\psi_{ij} \propto \delta_{ij}$) and whose mass is

$$m_{\mathrm{b}} = 2\sqrt{v}\left(1 - \frac{1}{32}\,(N-1)^2 g^4\right). \tag{2.50}$$

The symmetric mixing between the $N-1$ Φ_i–Φ_i two-particle states making up the bound state dramatically stabilizes it. This is particularly significant for the physics of the model, as we will see later. In particular, in the large-N limit à la 't Hooft [8]

$$N \to \infty\,, \qquad g \to 0\,, \qquad g^2 N = \text{constant}\,, \tag{2.51}$$

the interactions between the fields Φ_i are of order $1/N$, but they can nevertheless form a bound state whose binding energy is of order N^0, as (2.50) shows. We will discuss the properties of this bound state further in section 3. It is also natural to look for multiparticle bound states, which we know must exist in the limit where the model approach the sine-Gordon model (when one of the v_is is much smaller than the others). This is non-trivial, because the J matrices corresponding to different pairs of particles don't always commute, and we have not tried to study the corresponding Schrödinger equation. Note that these multiparticle bound states will be unstable in the field theory, which is unlike the case of the sine-Gordon model, but their lifetime can be large. In the case where the v_is are not all equal, we will still have a stable bound state in which the pair Φ-Φ corresponding to the particle of lowest mass dominate. We will see how (2.50) generalizes in this case in section 3 in the large-N limit.

2.3.3 Solitons

The most general time independent, finite energy solutions to the field eqs. (2.31) are derived in appendix A. They satisfy (2.37) and must tend towards one of the vacua (2.38) of our theory. Let us assume for the moment that $v_1 > v_2 > \cdots > v_{N-1}$. The only solutions corresponding to stable particles are then found by restricting the fields to $\Phi_1 = \cdots = \Phi_{N-2} = 0$, and correspond to the standard sine-Gordon soliton and anti-soliton of mass

$$m_{\mathrm{sol}} = \frac{2\sqrt{v_{N-1}}}{g^2}\,. \tag{2.52}$$

For example the soliton solution joining the North pole of the sphere S^{N-1} at $x = -\infty$ to the South pole at $x = +\infty$ and centered at $x = x_c$ is given by

$$\Phi_{N-1} = \frac{\sin\theta}{g}\,, \qquad \Phi_N = \frac{\cos\theta}{g}\,, \tag{2.53}$$

with

$$\tan\frac{\theta}{2} = e^{\sqrt{v_{N-1}}(x-x_c)}\,. \tag{2.54}$$

There are also solutions corresponding to sine-Gordon solitons of masses

$$m_{\mathrm{i,sol}} = \frac{2\sqrt{v_i}}{g^2}\,, \tag{2.55}$$

when the only varying fields are taken to be Φ_N and Φ_i. The particles associated with these solutions are unstable and would decay to the stable solution (2.53), (2.54) by emitting elementary quanta.

There are also more general solutions, which are worked out in details in appendix A, describing a succession of sine-Gordon kinks and anti-kinks of different masses. The force between these kinks magically cancel at the classical level, a consequence of the integrability of the equations (2.37). It is very unlikely that this property is maintained at the quantum level. The static solution for the case $N = 3$ is depicted in figure 1.

When several of the v_is coincide, the soliton solutions have collective coordinates correponding to the enhanced $\mathrm{O}(p)$ symmetries. Let us discuss briefly the most

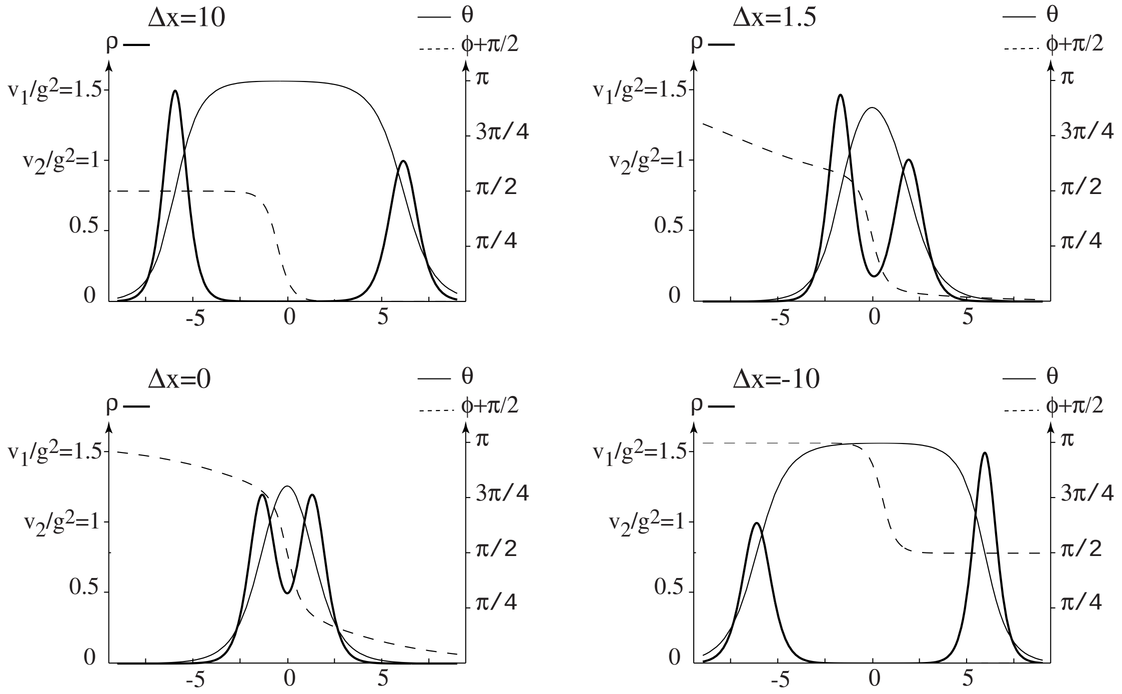

Figure 2: A static solution for the case $N = 3$, $v_1 = 1.5$, $v_2 = 1$, and $g = 1$, describing two kinks of respective masses $2\sqrt{v_1}/g^2$ and $2\sqrt{v_2}/g^2$, at different values of the distance $\Delta x_c = x_{c2} - x_{c1}$ between their centers. The energy density ρ is represented in thick solid line, and the spherical angles θ and ϕ in thin solid line and dashed line respectively. The maximum energy density of a given isolated kink is v_i/g^2. The formulas used to obtain the figure are given in appendix A.3.

symmetric case $v_1 = \cdots = v_{N-1} = v$. The general solution $\Phi_{i,\mathrm{sol}}(x - x_c; \xi_\alpha)$ is obtained from (2.53) and (2.54) by applying an arbitrary $\mathrm{SO}(N-1)$ rotation $R(\xi_\alpha)$. The angles ξ_α can be taken to parametrize the coset $\mathrm{SO}(N-1)/\,\mathrm{SO}(N-2) = S^{N-2}$. The quantization in the one-soliton sector, in the moduli space (low energy) approximation, then proceeds in the usual way, by restricting the fields to be of the form

$$\Phi_i(x,t) = \Phi_{i,\mathrm{sol}}(x - x_c(t); \xi^\alpha(t)), \tag{2.56}$$

and replacing this ansatz in the lagrangian (2.30) to find the quantum mechanics governing the collective coordinates $\xi^\alpha(t)$ and $x_c(t)$. The resulting lagrangian turns out to be

$$L_{\mathrm{moduli}} = -m_{\mathrm{sol}} + \frac{1}{2}\, m_{\mathrm{sol}}\, \dot{x}_c^2 + \frac{m_{\mathrm{sol}}}{2v}\, g_{\alpha\beta}\, \dot{\xi}^\alpha \dot{\xi}^\beta\,, \tag{2.57}$$

where $g_{\alpha\beta}$ is the $\mathrm{O(N-1)}$ invariant metric on the sphere S^{N-2} of radius unity. The Schrödinger equation corresponding to (2.57) involves the Beltrami laplacian on the sphere S^{N-1}, whose eigenvalues and eigenvectors are well-known. We get a tower of states of mass

$$m_{\mathrm{sol},J} = m_{\mathrm{sol}} + \frac{v}{2m_{\mathrm{sol}}}\, J(J+N-3) \tag{2.58}$$

filling multiplets of $\mathrm{SO}(N-1)$ corresponding to the completely symmetric and traceless tensors of rank J. The low-lying states must be stable at weak coupling, since then $v/m_{\rm sol} = \sqrt{v}\, g^2/2 \ll \sqrt{v}$, the mass of the elementary particles.

We also have soliton/anti-soliton bound states described by the breather solution of the sine-Gordon model. It could be interesting to investigate, for example in the case $v_1 = \cdots = v_{N-1}$, the possible relationship between these two-solitons bound states and the multiparticles bound states formed by the Φ_is. We have not performed this analysis, however; for our purposes, only the stable Φ–Φ singlet will be relevant.

3. The 1/N expansion

3.1 General properties and limitations of the large-N approximation

The large-N expansion à la 't Hooft (2.51) is one of the main non-perturbative tool at our disposal. The basic idea (for reviews see for example [38]) is to integrate *exactly* over the fields Φ_i in (2.30), which yields an effective action proportional to N. The $1/N$ expansion is then nothing but a loop expansion for this effective action, and it is non-perturbative with respect to the other parameters of the theory. For example, the connected vacuum amplitude can be written in the Neumann model as

$$W = \sum_{l=0}^{\infty} N^{1-l}\, W_l(v_i/\Lambda)\,. \tag{3.1}$$

A $2n$-point function would have an additional factor of $1/N^{n-1}$. Though it is very useful, the $1/N$ expansion has nevertheless two important limitations that will bother us in the following, and that are likely to show up in gauge theories as well.

The first limitation comes from the fact that our space of parameters $\mathcal{M}$ is $N-1$ dimensional. In the large-N limit, instead of working with a discrete set of dimensionless parameters $x_i = v_i/\Lambda$, one should really introduce a density

$$r(x) = \frac{1}{N-1}\sum_{i=1}^{N-1}\delta(x-x_i)\,, \tag{3.2}$$

and consider only smooth enough functions $r(x)$. What this means in practice is that we cannot study in this approximation scheme non-generic parameters where a small number of masses are adjusted to take particular values. To be concrete, let us suppose that we want to investigate the region in parameter space where one of the v_is, say v_1, goes to zero. In the $1/N$ expansion, the v_is typically appear through combinations of the form

$$\frac{1}{V} = \frac{1}{N-1}\sum_{i=1}^{N-1}\frac{1}{v_i}\,. \tag{3.3}$$

We see that, though V is generically of order N^0, it is of order N when $v_1 \to 0$. This means that the $1/N$ counting is completely modified in this limit, and the standard expansion can no longer be trusted. This problem is of course a serious drawback, since we know in particular from gauge theory studies that interesting phenomena are usually associated with fine tuning of parameters ("critical points"). Fortunately, there are two ways out of this problem. First, one can consider only p dimensional subspaces $\mathcal{S}_p \subset \mathcal{M}$, with $p \ll N$, for example by choosing $v_1 = \cdots = v_{[(N-1)/p]} = w_1$, $v_{[(N-1)/p]+1} = \cdots = v_{[2(N-1)/p]} = w_2$, etc... The physics on $\mathcal{S}_p$, including special points, can then be studied reliably in the large-N expansion. This trick is useful in the supersymmetric generalizations of the Neumann model [33]. In the present case, we need to do better, however, because interesting physics can be associated with only *one* parameter going to a special value, like $v_1 \to 0$. The idea is then to treat exactly the small number of degrees of freedom that play a special rôle in the particular limit we are interested in. For example, in the limit $v_1 \to 0$ with all the other v_is positive, we will integrate over the $N-2$ fields $\Phi_2, \ldots, \Phi_{N-1}$ in (2.30), while keeping explicitly the fields Φ_1 and Φ_N together with α. More generally, it is actually very convenient to always keep explicitly the field Φ_N when working on $\mathcal{M}_N = \{v_i \geq 0\}$ for example, since this field is singled out by the fact that $\langle \Phi_N \rangle \neq 0$ at weak coupling in this region.

The second limitation of the $1/N$ expansion that we will encounter comes from the fact that the terms in the expansion can behave badly for some values of the parameters. For example, it could be that some of the coefficients $W_l(v_i/\Lambda)$ in (3.1) are singular for some (generic) values of the v_is. Typically, these coefficients can be expanded in powers of the coupling constant at weak coupling, but when the coupling grows the series might become singular. Beyond that point, the coefficients might again be expandable in terms of some "dual" coupling. At the singular points, the standard $1/N$ expansion fails completely. We will see in section 3.3 that the physics can nevertheless be extracted in the large-N limit, by resumming in some sense the $1/N$ series. We will also see in section 6 that in our model the singular behaviour of the coefficients of the $1/N$ expansion is very specific, and that this has some very interesting consequences. The same kind of phenomena are likely to occur in gauge theories as well, for example in the vicinity of an Argyres-Douglas point.

Renormalization. At the technical level, the $1/N$ expansion needs to be renormalized. It is actually easy to find the divergent part of the renormalization constants that make the theory finite to all orders in $1/N$. This is possible because the theory is asymptotically free, and the $1/N$ expansion is non-perturbative in g. The divergent contributions can then be obtained by summing up ordinary perturbation theory by using the renormalization group. As is well known, these divergent terms are completely determined by the two-loop β function and the one-loop γ and σ functions, all given in (2.23), (2.24) and (2.25). More precisely, we have the exact formulas

$$\frac{1}{g_0^2} = \frac{N-2}{2\pi} \ln\frac{\Lambda_0}{\Lambda} + \frac{1}{2\pi} \ln\ln\frac{\Lambda_0}{\Lambda} + \text{finite terms}\,, \tag{3.4}$$

$$\ln Z = -\frac{1}{N-2}\ln\ln\frac{\Lambda_0}{\Lambda} + \text{finite terms}\,, \tag{3.5}$$

$$\ln Z_{(s,2)} = \frac{1}{N-2}\ln\ln\frac{\Lambda_0}{\Lambda} + \text{finite terms}\,, \tag{3.6}$$

which yield

$$Z = 1 - \frac{1}{N}\ln\ln\frac{\Lambda_0}{\Lambda} + \frac{k}{N} + \mathcal{O}(1/N^2)\,, \tag{3.7}$$

$$Z_{(s,2)} = 1 + \frac{1}{N}\ln\ln\frac{\Lambda_0}{\Lambda} + \frac{k'}{N} + \mathcal{O}(1/N^2)\,, \tag{3.8}$$

where k and k' are finite constants. The finite terms in $\ln Z$ and $\ln Z_{(s,2)}$ will contribute to the divergent terms in Z and $Z_{(s,2)}$ at order $1/N^2$ and higher, however.

3.2 Massless states

To keep the formulas as simple as possible, let us begin by restricting ourselves to the $\mathcal{O}(N-1)$ symmetric case $v_1 = \cdots = v_{N-1} = v \geq 0$. We can then take without loss of generality $h_N = v$ and $h_i = 0$ for $1 \leq i \leq N-1$. By integrating over the $N-1$ fields Φ_i, $1 \leq i \leq N-1$, in (2.30), and Wick rotating to the euclidean, we obtain, in the large-N limit, the non-local effective action

$$S_{\text{eff}}[\alpha, \Phi_N] = N\, s_{\text{eff}}[\alpha, \varphi] \tag{3.9}$$

with

$$\Phi_N = \sqrt{N}\,\varphi \tag{3.10}$$

and

$$s_{\text{eff}}[\alpha, \varphi] = \int d^2x \left(\frac{1}{2}\,\partial_\alpha\varphi\partial_\alpha\varphi + \frac{\alpha - v}{2}\,\varphi^2 - \frac{\alpha}{2Ng^2}\right) + s[\alpha]\,. \tag{3.11}$$

The functional $s[\alpha]$ is defined and studied in details in appendix B, to which we refer the reader, and g is the renormalized coupling (2.26). The effective potential corresponding to s_{eff} is

$$v_{\text{eff}}(\alpha, \varphi) = \frac{\alpha - v}{2}\,\varphi^2 - \frac{\alpha}{8\pi}\,\ln\frac{\alpha}{e\Lambda^2}\,, \tag{3.12}$$

and the saddle point equations governing the physics in the limit $N \to \infty$ are

$$\begin{cases} (\alpha_* - v)\,\varphi_* = 0 \\ 4\pi\varphi_*^2 = \ln\dfrac{\alpha_*}{\Lambda^2}\,. \end{cases} \tag{3.13}$$

At weak coupling $v \gg \Lambda^2$, the $\mathbb{Z}_{2(N)}$ symmetry is spontaneously broken and $\varphi_* \neq 0$ (2.38). We will loosely call the "weak coupling" region in parameter space the whole region where $\mathbb{Z}_{2(N)}$ is broken and where we have

$$\text{Weak coupling:} \quad \begin{cases} \alpha_* = v \\ \varphi_*^2 = \dfrac{1}{4\pi}\,\ln\dfrac{v}{\Lambda^2}\,. \end{cases} \tag{3.14}$$

Equations (3.14) give

$$\langle \Phi_N^2 \rangle = \frac{1}{g^2(\mu = \sqrt{v})} = \frac{1}{g_{\text{eff}}^2}\,, \tag{3.15}$$

which is simply the classical formula (2.38) where the bare coupling has been replaced by the effective low energy coupling g_{eff}, the running coupling (2.26) evaluated at the scale $\sqrt{v}$ corresponding to the mass of the elementary particles. On the other hand, when $v = 0$, we know that $\langle \Phi_N \rangle = 0$ as a consequence of the SO(N) symmetry. Thus, there must be a phase transition at strong coupling for some $v = v_{\text{c}} \sim \Lambda^2$, where the $\mathbb{Z}_{2(N)}$ is restored. We would like to understand the nature of this transition, and whether $v_{\text{c}} = 0$ or $v_{\text{c}} > 0$. Since it occurs at strong coupling in a non-abelian system, the standard intuition would suggest that a mass gap is formed and thus that any phase transition must be first order. This would be the case for example if $v_{\text{c}} = 0$. However, we are going to show that this is not what happens: the transition is second order, and we have massless states at strong coupling. To do so, we evaluate the stability of the saddle point (3.14). The effective potential for φ is obtained by integrating out α from (3.12),

$$v_{\text{eff}}(\varphi) = \frac{\Lambda^2}{8\pi}\left(e^{4\pi\varphi^2} - 4\pi\varphi^2\,\frac{v}{\Lambda^2}\right). \tag{3.16}$$

This shows that we have a second order transition at $v = v_{\text{c}} = \Lambda^2$, between a broken symmetry phase $v > \Lambda^2$ and a phase $v < \Lambda^2$ where the $\mathbb{Z}_{2(N)}$ is restored. This is the "strong coupling" region of parameter space, where the saddle point is

$$\text{Strong coupling:}\quad \begin{cases} \alpha_* = \Lambda^2 \\ \varphi_* = 0\,. \end{cases} \tag{3.17}$$

It is interesting to investigate in details what happens at the point $v = \Lambda^2$. In the $1/N$ expansion, the effective coupling constant g_{eff} goes to infinity at that point, which is directly related to the fact that we have massless degrees of freedom. This phenomenon is also known to occur in *supersymmetric* gauge theories when a magnetic monopole becomes massless. Actually we will see in section 4 and 5 that effects which are non-perturbative in $1/N$ make the effectiveq coupling constant finite at the transition point $v = \Lambda^2$. The critical theory will be argued to be very similar to an Argyres-Douglas CFT [7], where both magnetic monopoles (topologically stable solitons) and electrically charged particles (created by the elementary fields) can become massless. In our case, it is clear that the stable SO($N-1$) singlet soliton found at the end of section 2.3.3 must be massless when $v = \Lambda^2$. This is simply due to the fact that the two degenerate minima of the potential (3.16) merge at that point, and any solitonic solution to the effective action must then become trivial. To find out whether any perturbative state could become massless together with the soliton, it is natural to first look at the mass of the elementary fields Φ_i. One can calculate straightforwardly the two point functions in the leading $N \to \infty$ approximation. In momentum space they are simply

$$\langle \Phi_i \Phi_j \rangle(p) = \frac{\delta_{ij}}{p^2 + \alpha_*}\,, \tag{3.18}$$

which shows that the fields Φ_i, $1 \leq i \leq N-1$, create free particles of masses $m_i = \sqrt{\alpha_*}$. The m_is thus take their classical values $m_i = \sqrt{v}$ as long as $v > \Lambda^2$, and then take the constant value $m_i = \Lambda$ when $v < \Lambda^2$: at strong coupling, a mass gap is created for these particles, as the standard lore suggests. To find massless perturbative states at $v = \Lambda^2$, we thus have to look more carefully at the menu of stable particles described in section 2.3. As explained there, though the interactions between the elementary particles are of order $1/N$, a bound state whose binding energy is of order N^0 can be formed because of a dramatic stabilization due to the mixing between the different flavors. This bound state is created by the operator $\sum_{i=1}^{N-1} \Phi_i^2$, but can equivalently be associated with

$$\varphi = \frac{1}{\sqrt{Ng^2}} \sqrt{1 - g^2 \sum_{i=1}^{N-1} \Phi_i^2}\,, \tag{3.19}$$

and thus its mass must appear as a pole in the $\langle \varphi\varphi \rangle$ correlation function. By expanding

$$\varphi = \varphi_* + \frac{\chi}{\sqrt{N}}\,, \qquad \alpha = \alpha_* + \frac{\beta}{\sqrt{N}}\,, \tag{3.20}$$

we immediately obtain from (3.11) the leading term in the large-N expansion,

$$\langle \chi\chi \rangle(p) = \frac{\tilde{s}^{(2)}(p^2;\alpha_*)}{(p^2 + \alpha_* - v)\,\tilde{s}^{(2)}(p^2;\alpha_*) - \varphi_*^2}\,. \tag{3.21}$$

The function $\tilde{s}^{(2)}$ is defined in appendix B. The bound state mass m_{b} thus satisfies the equation

$$m_{\mathrm{b}}^2 = \alpha_* - v - \frac{\varphi_*^2}{\tilde{s}^{(2)}(p^2 = -m_{\mathrm{b}}^2;\alpha_*)}\,. \tag{3.22}$$

At strong coupling the solution is simply

$$m_{\mathrm{b}} = \sqrt{\Lambda^2 - v}\,, \qquad \text{for } v \leq \Lambda^2\,. \tag{3.23}$$

We see that when $v = 0$ and we have the full $\mathrm{O}(N)$ symmetry, $m_{\mathrm{b}} = \Lambda = m_i$, and the spectrum consists in a $\mathrm{O}(N)$ vector [37] made up by the $N-1$ elementary particles and the φ field. We also discover that when $v = \Lambda^2$ the mass of the bound state vanishes!

When $v > \Lambda^2$, we have to solve

$$\sqrt{\frac{4v}{m_{\mathrm{b}}^2} - 1}\, \ln \frac{v}{\Lambda^2} = \pi - 2 \arctan \sqrt{\frac{4v}{m_{\mathrm{b}}^2} - 1}\,. \tag{3.24}$$

JHEP06(2001)057

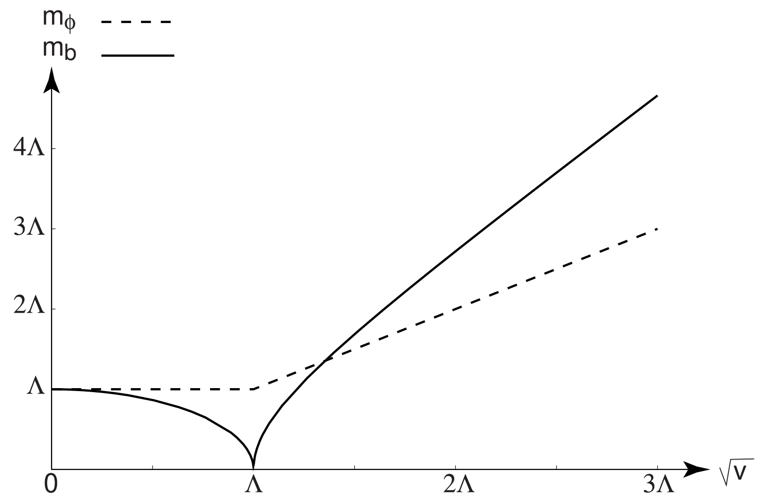

Figure 3: The mass of the elementary particles created by the Φ_is, $1 \leq i \leq N-1$ (dashed line) and of the Φ-Φ bound state (solid line), as a function of $\sqrt{v} = \sqrt{v_1} = \cdots = \sqrt{v_{N-1}}$, in the leading $1/N$ approximation.

When $v \gg \Lambda^2$, this yields

$$m_{\mathrm{b}} = 2\sqrt{v}\left(1 - \frac{1}{32}(Ng^2)^2 + \frac{1}{32\pi}(Ng^2)^3 + \mathcal{O}\left((Ng^2)^4, \frac{1}{N}\right)\right), \tag{3.25}$$

in agreement with (2.50). In the vicinity of $v = \Lambda^2$, (3.23) and (3.24) predicts

$$\begin{aligned} m_{\mathrm{b}}^2 &\simeq \Lambda^2\left(1 - \frac{v}{\Lambda^2}\right) \qquad \text{for } v \to \Lambda^2\,,\ v < \Lambda^2\,, \\ m_{\mathrm{b}}^2 &\simeq 2\Lambda^2\left(\frac{v}{\Lambda^2} - 1\right) \qquad \text{for } v \to \Lambda^2\,,\ v > \Lambda^2\,. \end{aligned} \tag{3.26}$$

We have depicted in figure 1 the masses of the elementary particles and of the bound state as a function of $\sqrt{v}$, as given by the leading term in the $1/N$ expansion. As we will see in section 3.4, this leading approximation is actually incorrect near $v = \Lambda^2$, and in particular the formulas (3.26) are wrong. However, the most important fact that the soliton and the bound state are massless will remain, and we will see that the correct form of (3.26) is simply

$$m_{\mathrm{b}}^2 \propto m_{\mathrm{sol}}^2 \propto \Lambda^2\left(1 - \frac{v}{\Lambda^2}\right)^2 \quad \text{for } v \to \Lambda^2. \tag{3.27}$$

3.3 Correlators and quantum corrected metric

We continue to restrict ourselves to the symmetric case $v_1 = \cdots = v_{N-1}$ in this subsection. This makes the formulas simpler, but do not affect the physics we want to discuss. Our aim is to obtain the large-N formulas for the S matrix, generalizing (2.47), and to compute the quantum corrections to the metric on the sphere

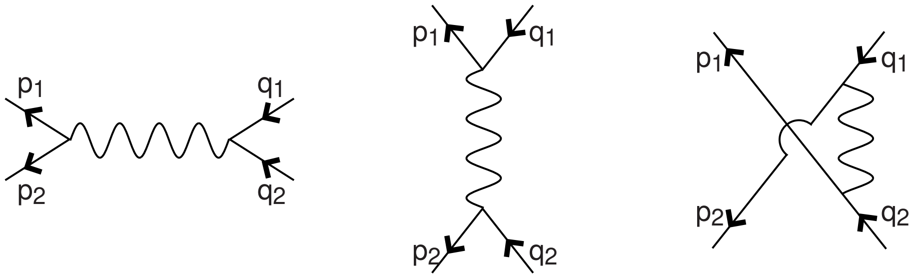

Figure 4: Large-N Feynman diagrams for the four-point function. The wavy line corresponds to the β field propagator, which is itself a sum of bubble diagrams in ordinary perturbation theory.

target space. The generating functional $Z[J]$ for the correlation functions is given by a path integral

$$\begin{aligned} Z[J] = \int \mathcal{D}\chi\mathcal{D}\beta \exp\Big[-(N-1)\, s_{\text{eff}} \Big[\alpha_* + \frac{\beta}{\sqrt{N}}, \varphi_* + \frac{\chi}{\sqrt{N}}\Big] + \\ + \frac{1}{2}\int d^2x \sum_{i=1}^{N-1} J_i \, \frac{1}{-\partial^2 + \alpha_* + \beta/\sqrt{N}} \, J_i \Big] \Big/ \\ \int \mathcal{D}\chi\mathcal{D}\beta \exp\Big[-(N-1)\, s_{\text{eff}} \Big[\alpha_* + \frac{\beta}{\sqrt{N}}, \varphi_* + \frac{\chi}{\sqrt{N}}\Big]\Big] . \end{aligned} \tag{3.28}$$

The large-N Feynman diagrams contributing to the four point function $\langle \Phi_{i_1}\Phi_{i_2}\Phi_{i_3}\Phi_{i_4}\rangle$ are indicated in figure 1. The wavy lines correspond to the β field propagator

$$\langle \beta\beta\rangle(p) = \frac{p^2 + \alpha_* - v}{(p^2 + \alpha_* - v)\tilde{s}^{(2)}(p^2, \alpha_*) - \varphi_*^2} . \tag{3.29}$$

The functions S_1, S_2 and S_3 defined by (2.46) can then be deduced from the four-point function. At weak coupling ($v > \Lambda^2$), we have

$$S_1 = \frac{ig_{\text{eff}}^2}{8\pi^2}\, \frac{s}{\sqrt{-su}}\, \frac{1}{1 + Ng_{\text{eff}}^2 s\, \tilde{s}^{(2)}(p^2 = -s; v)} , \qquad S_2 = 1 , \quad S_3 = S_1(s \mapsto u) , \tag{3.30}$$

or equivalently in terms of the rapidity $\theta > 0$ such that $s = 4v\cosh^2(\theta/2)$ and $u = -4v\sinh^2(\theta/2)$,

$$S_1(\theta) = \frac{ig_{\text{eff}}^2}{8\pi^2}\, \frac{1}{\tanh(\theta/2)}\, \frac{1}{1 + Ng_{\text{eff}}^2\, \frac{\theta - i\pi}{4\pi\tanh(\theta/2)}} , \qquad S_2(\theta) = 1 , \quad S_3(\theta) = S_1(i\pi - \theta) . \tag{3.31}$$

g_{eff}^2 is defined as usual to be the low energy coupling, i.e. the running coupling (2.26) evaluated at $\mu = \sqrt{v}$. These formulas illustrate nicely how the $1/N$ expansion works and produces non-perturbative results. The genuinely non-perturbative information lies on the full θ dependence of the denominators, not on the fact that g_{eff} appears in these denominators. In particular, when $\theta \to \infty$, the non-perturbative corrections become dominant, even if $v \gg \Lambda^2$, but due to asymptotic freedom, we know that in this regime the S matrix can also be deduced from perturbation theory by using the renormalization group. Indeed we have

$$S_1(\theta) \underset{\theta\to\infty}{=} \frac{1}{2\pi N\theta} = \frac{ig^2(\mu = \sqrt{s})}{8\pi^2} \tag{3.32}$$

which is the perturbative result (2.47) at high energy, but with the coupling g evaluated at the center of mass energy, which is different from the low energy coupling. In the strong coupling region $v < \Lambda^2$, we have

$$S_1 = \frac{i}{8\pi^2 N\sqrt{-su}}\,\frac{1}{\tilde{s}^{(2)}(p^2 = -s; \Lambda^2)}\,, \qquad S_2 = 1 - \frac{i\Lambda^2}{\pi N\sqrt{-su}}\,, \qquad S_3 = S_1(s \mapsto u), \tag{3.33}$$

or equivalently

$$S_1(\theta) = \frac{i}{2\pi N}\,\frac{1}{\theta - i\pi}\,, \qquad S_2(\theta) = 1 - \frac{i}{2\pi N\sinh\theta}\,, \qquad S_3(\theta) = -\frac{i}{2\pi N\theta}\,. \tag{3.34}$$

The functions S_is are independent of v in this regime, but this is true only in the leading $1/N$ approximation. The formulas (3.33) show that the Φ_is, $1 \le i \le N-1$, no longer form a two particles bound state for $v < \Lambda^2$, and thus that the field Φ_N must be considered as an independent degree of freedom in this region. Clearly the geometric interpretation of the Φ_is as living on a sphere cannot be valid for $v < \Lambda^2$. To better understand what is really going on, it is natural to compute the quantum corrections to the metric on the target space sphere as a function of v. This amounts to computing the coefficient of p^2 in a low energy expansion of all the connected, one-particle irreducible, $2n$-point functions $\langle \Phi_{i_1} \cdots \Phi_{i_{2n}} \rangle$. The $\mathrm{SO}(N-1)$ symmetry restricts a priori the metric to be of the form

$$g_{ij} = f_1\left(\sum_{k=1}^{N-1} \Phi_k^2\right)\delta_{ij} + f_2\left(\sum_{k=1}^{N-1} \Phi_k^2\right)\Phi_i\Phi_j \tag{3.35}$$

for unknown functions f_1 and f_2. Actually, we are going to show that in the leading $1/N$ approximation and for $v > \Lambda^2$

$$g_{ij} = \delta_{ij} + \frac{\Phi_i\Phi_j}{\frac{N}{4\pi}\ln\frac{v}{\Lambda^2} - \sum_{k=1}^{N-1}\Phi_k^2}\,, \tag{3.36}$$

which is simply the $\mathrm{SO}(N)$ invariant metric on a sphere of radius $1/g_{\mathrm{eff}}$. This very simple result is equivalent to the statement that the p^2 terms in the low energy expansion of the $2n$ point functions is simply given by the perturbative, tree level result, with the bare coupling being replaced by the low energy coupling g_{eff}. This simplification occurs because the low energy expansion of the β field propagator (3.29) is

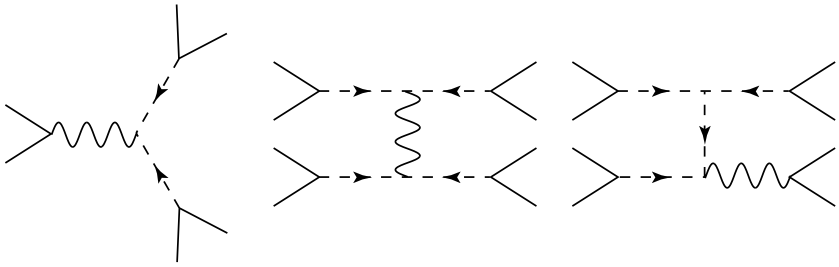

Figure 5: Typical diagrams relevant to the calculation of the metric and contributing to the 6-point (left) and 8-point (center and right) functions. The $\langle\beta\beta\rangle$ propagator is represented by a wavy line, and the $\langle\beta\chi\rangle$ contraction by a dashed line with an arrow pointing toward the χ insertion.

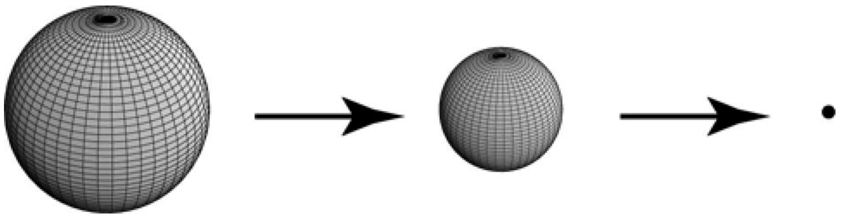

Figure 6: The effective target space sphere shrinks symmetrically to a point when $v \to \Lambda^2$ in the leading $1/N$ approximation. At strong coupling, the geometrical interpretation of the σ model is lost. This is reminiscent of the stringy geometry described in the context of supersymmetric non-linear σ model.

$\langle\beta\beta\rangle(p) = -Ng_{\text{eff}}^2\, p^2 + \mathcal{O}(p^4)$, and thus only one such propagator can appear in a diagram contributing to the metric. A connected diagram of this type for the $2n$ points function must then be entirely constructed from one $\langle\beta\beta\rangle$ propagator, $n-2$ vertices $\beta\chi^2/\sqrt{N}$ coming from the expansion of s_{eff} (3.11) and (3.20), and $2n-4$ contractions between β and χ which each give a factor of $\sqrt{Ng_{\text{eff}}^2}$, in addition to the n $\Phi\Phi\beta/\sqrt{N}$ vertices. Such diagrams at low energies depend on N and g_{eff} only through a multiplicative factor $Ng_{\text{eff}}^2\, N^{-(n-2)/2}\, (Ng_{\text{eff}}^2)^{(2n-4)/2}\, N^{-n/2} = N^{1-n}\, (Ng_{\text{eff}}^2)^{n-1} = g_{\text{eff}}^{2(n-1)}$. This dependence on the coupling is the same as the one found in a tree level perturbative calculation. The dependence in the momenta must then also coincide, since the large-N result must reduce to the tree level result when $g_{\text{eff}} \ll 1$. This proves (3.36).

We thus see that the effective target space sphere shrinks symmetrically to a point when $v \to \Lambda^2$, $v > \Lambda^2$. The geometrical interpretation of the σ model is thus lost in the strong coupling region $v < \Lambda^2$, which corresponds in some sense to a continuation of the sphere to negative radius. Similar continuations have already been seen in the context of supersymmetric models [39], and when the σ model is conformal they actually play an important rôle in understanding the topology changing transitions in string theory (see for example [40]). We will further discuss the consequences of this interpretation in section 5.

3.4 The critical hypersurface $\mathcal{H}_{\mathrm{q}}$

Let us now discuss the general case where the $v_i \geq 0$ can be distinct. Formulas (3.11)—(3.14) and (3.17) are replaced respectively by

$$s_{\mathrm{eff}}[\alpha,\varphi] = \int d^2x \left(\frac{1}{2}\partial_\alpha\varphi\partial_\alpha\varphi + \frac{\alpha - h_N}{2}\varphi^2 - \frac{\alpha}{2Ng^2}\right) + \frac{1}{N}\sum_{i=1}^{N-1} s[\alpha - h_i], \quad (3.37)$$

$$v_{\mathrm{eff}}(\alpha,\varphi) = \frac{\alpha - h_N}{2}\varphi^2 - \frac{1}{8\pi N}\sum_{i=1}^{N-1}(\alpha - h_i)\ln\frac{\alpha - h_i}{e\Lambda^2}\,, \quad (3.38)$$

$$\begin{cases} (\alpha_* - h_N)\,\varphi_* = 0 \\ 4\pi\varphi_*^2 = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N-1}\ln\frac{\alpha_* - h_i}{\Lambda^2}\,, \end{cases} \quad (3.39)$$

$$\text{Weak coupling}: \begin{cases} \alpha_* = h_N \\ \varphi_*^2 = \dfrac{1}{4\pi N}\displaystyle\sum_{i=1}^{N-1}\ln\frac{v_i}{\Lambda^2}\,, \end{cases} \quad (3.40)$$

$$\text{Strong coupling}: \begin{cases} \displaystyle\sum_{i=1}^{N-1}\ln\frac{\alpha_* - h_i}{\Lambda^2} = 0 \\ \varphi_* = 0\,. \end{cases} \quad (3.41)$$

Note that the various sums over i should be more rigorously replaced by integrals with the help of (3.2). This will always be understood in the following. The border between the weak coupling region $\langle\Phi_N\rangle \neq 0$ and the strong coupling region $\langle\Phi_N\rangle = 0$ is a codimension one hypersurface one which both the soliton and the bound state become massless. The equation for this quantum hypersurface of singularity $\mathcal{H}_{\mathrm{q}}$ in $\mathcal{M}_N$ is immediately deduced from (3.40) and (3.41),

$$\mathcal{H}_{\mathrm{q}} \cap \mathcal{M}_N : \quad \prod_{i=1}^{N-1} v_i = \Lambda^{2(N-1)}\,. \quad (3.42)$$

The critical point $v_1 = \cdots = v_{N-1} = v_{\mathrm{c}} = \Lambda^2$ discussed previously is of course on $\mathcal{H}_{\mathrm{q}}$. $\mathcal{H}_{\mathrm{q}}$ has also other sheets in the other, physically equivalent, regions $\mathcal{M}_i$ of parameter space (2.42), whose equations are simply

$$\mathcal{H}_{\mathrm{q}} \cap \mathcal{M}_i : \quad \prod_{j\neq i} v_j^{(i)} = \Lambda^{2(N-1)}\,. \quad (3.43)$$

We thus get a preliminary picture (it is only partially correct as will be discussed in section 4) of the quantum space of parameters, in the large-N limit.

It is very likely that the qualitative picture is still valid even for values of N as small as $N = 3$ (we will give strong evidence that this is the case in section 5), and we have used equations (3.42) and (3.43) in the cases $N = 3$ and $N = 4$ to make figure 7.

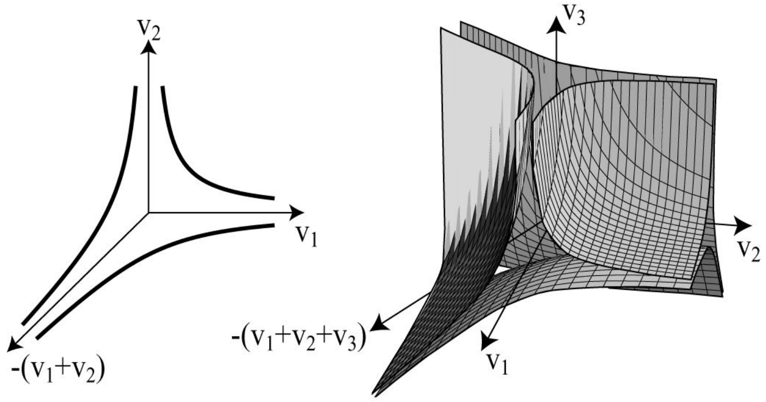

Figure 7: A preliminary picture of the quantum space of parameters $\mathcal{M}_{\mathrm{q}}$ in the cases $N=3$ (left) and $N=4$ (right).

This figure should be compared with the classical case given in figure 1. One might have expected that the singularities on parameter space would have disappeared in the quantum case, a mass gap being created, and the $\mathbb{Z}_{2(N)}$ symmetry being restored through a first order phase transition. We see that on the contrary the classical singular hyperplanes split into two in the quantum case, and the $\mathbb{Z}_{2(N)}$ symmetry is restored through a second order phase transition. We discuss in the next subsection how we can understand the nature (Ising) of this transition in the large-N limit, and we will also discuss it at finite N in section 5.

We can also write down the formulas generalizing (3.23) and (3.24) to the case of arbitrary positive v_is. At strong coupling, the mass m_{b} satisfies

$$m_{\mathrm{b}} = \sqrt{\alpha_* - h_N}\,, \tag{3.44}$$

with α_* determined by (3.41), and at weak coupling we have

$$\sum_{i=1}^{N-1} \ln \frac{v_i}{\Lambda^2} = 2 \sum_{i=1}^{N-1} \frac{1}{\sqrt{4v_i/m_{\mathrm{b}}^2 - 1}} \arctan \frac{1}{\sqrt{4v_i/m_{\mathrm{b}}^2 - 1}}\,. \tag{3.45}$$

3.5 Beyond the $1/N$ expansion: the Ginzburg-Landau description

In the vicinity of the critical hypersurface $\mathcal{H}_{\mathrm{q}}$, the usual $1/N$ expansion is inconsistent because of infrared divergences. These divergences are due to two different effects. First, the χ field becomes massless and thus the $\langle\chi\chi\rangle$ propagator (3.21) goes to $1/p^2$. Consequently the $1/N$ expansion cannot be defined because of the usual IR divergences associated with massless scalar propagators in two dimensions. Second, and independently of the fact that we are in two dimensions, the $1/N$ corrections are calculated by taking into account interaction terms corresponding to relevant operators. Such interactions always produce IR divergences in perturbing around a massless theory. This is strictly analogous to the textbook discussion of the corrections to the mean field approximation for the ϕ^4 theory near dimension 4, except that we are dealing presently with the $1/N$ perturbation theory instead of the ordinary perturbation theory. To analyse the divergences, one could study in details the large-N Feynman graphs and isolate the most divergent contributions. Instead, we will proceed in a slightly more heuristic, but actually equivalent, way. In a first step, we would like to integrate out the field $\alpha = \alpha_* + \beta/\sqrt{N}$, since the divergences are due to the masslessness of χ. At large N, this can be done by solving the variational equation

$$\frac{\delta s_{\text{eff}}[\alpha, \varphi]}{\delta \alpha(x)} = 0\,, \tag{3.46}$$

which at low energy reduces to

$$\frac{\partial v_{\text{eff}}}{\partial \alpha} = 0\,. \tag{3.47}$$

By using (3.38) and (3.37), we then obtain a low energy effective action valid in the vicinity of $\mathcal{H}_{\text{q}}$, and depending on $\chi = \Phi_N$ only,

$$S_{\text{eff}}[\chi] = \int d^2x \left(\frac{1}{2}\, \partial_\alpha \chi \partial_\alpha \chi - \frac{\delta v}{2}\, \chi^2 + \frac{\pi V}{N}\, \chi^4 + \mathcal{O}\left(\frac{\chi^6}{N^2}\right)\right), \tag{3.48}$$

where we have defined

$$\frac{1}{V} = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{1}{v_i}\,, \tag{3.49}$$

and

$$\delta v = \frac{V}{N-1} \sum_{i=1}^{N-1} \ln \frac{v_i}{\Lambda^2}\,. \tag{3.50}$$

The equation for $\mathcal{H}_{\text{q}}$ is simply $\delta v = 0$. Higher derivative corrections to (3.47) produce terms like $\chi^2 \partial_\alpha \chi \partial_\alpha \chi / N$ which are irrelevant in the IR. Similarly, we have dropped in (3.48) irrelevant terms or order χ^{2k}/N^{k-1}, $k \geq 3$. In a $1/N$ expansion at $\delta v = 0$, their contribution is subleading with respect to the χ^4/N term, to *all* orders in $1/N$. Indeed, IR divergences compensate for the $1/N$ factors coming from the χ^4/N interaction, which should thus be considered as giving leading N^0 contributions when $\delta v \to 0$. A way to understand this is to rescale the space-time and momentum variables

$$x = \sqrt{N}\, x'\,, \quad p = p'/\sqrt{N}, \tag{3.51}$$

which eliminates any N dependence in (3.48) on $\mathcal{H}_{\text{q}}$,

$$S_{\text{eff}}[\chi] = \int d^2x' \left(\frac{1}{2}\, \partial'_\alpha \chi \partial'_\alpha \chi + \pi V\, \chi^4 + \mathcal{O}\left(\frac{\chi^6}{N}\right)\right). \tag{3.52}$$

The IR properties $p \to 0$ can now be studied by working at fixed $p' = p\sqrt{N}$ and using a simple minded $1/N$ expansion in the p' variables. This amounts to perturbing around the interacting theory (3.52) where the χ^4 term is taken into account exactly, and contrary to the perturbation around the free massless theory this is perfectly well-defined.

JHEP06(2001)057

We recognize in (3.48) the Ginzburg-Landau description of an Ising critical point [41]. The low energy physics on $\mathcal{H}_{\rm q}$ ($\delta v = 0$) is thus governed by an Ising CFT. Formula (3.27) is then simply derived by noting that a deviation from the critical value $v = \Lambda^2$ corresponds to turning on the energy operator in the Ising CFT, a dimension one operator. Note that displacements along the critical surface correspond to operators that are irrelevant in the IR, since the Ising CFT does not have any marginal deformation. The appearance of a non-trivial interacting CFT at a point where both a soliton and a perturbative state become massless is of course very reminiscent of an Argyres-Douglas CFT [7], though in our case it appears to correspond to a very banal critical theory with a local description (3.48). Nevertheless, we will see in section 5 that a natural way to understand the appearance of this interacting CFT, independently of a large-N approximation, is through a description in terms of a more exotic non-local theory of electric and magnetic charges (spin waves and vortices). The Ising CFT is then nothing but a strong coupling fixed point of this non-local theory. This latter description is more akin to what one may expect in the context of gauge theories, where simple local description à la Ginzburg-Landau may not exist.

The reader may be concerned at this point by the fact that even though the standard $1/N$ expansion is not valid in the vicinity of $\mathcal{H}_{\rm q}$, the very existence of $\mathcal{H}_{\rm q}$ itself has been demonstrated in this simple framework. To investigate the consistency of the analysis, one should compute the $1/N$ corrections to the equation for $\mathcal{H}_{\rm q}$ (3.42), and see if they are small. This calculation, which illustrates several properties of the $1/N$ expansion, is presented in appendix C. The result is that though we do encounter IR divergences when computing the corrections to (3.42), they are mild logarithmic divergences. The only consequence is that the leading corrections, instead of being of order $1/N$, are of order $(\ln N)/N$. This is still a small correction when $N \to \infty$, as required.

4. The quantum space of parameters $\mathcal{M}_{\rm q}$

The equation (3.42) for $\mathcal{H}_{\rm q}$ is not reliable when a small number of the v_is goes to zero. Indeed, this equation is obtained in the leading $N \to \infty$ approximation in the form

$$\mathcal{H}_{\rm q} \cap \mathcal{M}_N : \quad \frac{1}{N} \sum_{i=1}^{N-1} \ln \frac{v_i}{\Lambda^2} = 0\,. \tag{4.1}$$

Then, when for example v_{N-1} is much smaller than the other v_is, the corresponding term in the sum (4.1), which is of order $1/N$, inconsistently becomes larger than the sum of the other terms, which is of order N^0. Naively, one might expect that this problem could be eliminated by taking into account the $1/N$ corrections to the equation for $\mathcal{H}_{\rm q}$ (C.30), but this is not true: the fact that the large-N counting is changed in the limit $v_{N-1} \to 0$ affects all orders of the $1/N$ expansion used in section 3.

4.1 Weak coupling

Actually, we can readily understand that (4.1) is not correct, even qualitatively, on the whole parameter space. Indeed, when $v_{N-1} = 0$ and all the other v_is are much larger than Λ^2, the low energy physics is governed by a weakly coupled O(2) non-linear σ model. This is a *massless* phase, and thus we see that at weak coupling the classical (figure 1) and quantum hypersurface of singularities must coincide. This is unlike the prediction based on (4.1), see figure 1. More precisely, for $v_{N-1} \ll v_i$ and $v_i \gg \Lambda^2$, $1 \leq i \leq N-2$, the low energy effective action is reliably determined by a one-loop perturbative calculation (2.23) to be

$$S_{\text{eff}} = \frac{1}{2g_{\text{eff}}^2} \int d^2x \left(\partial_\alpha \theta \partial_\alpha \theta - \frac{v_{N-1}}{2} \cos 2\theta \right) , \tag{4.2}$$

with

$$\frac{1}{g_{\text{eff}}^2} = \frac{1}{4\pi} \ln \frac{\prod_{i=1}^{N-2} v_i}{\Lambda^{2(N-2)}} \, . \tag{4.3}$$

As already discussed in section 2, we get a sine-Gordon model whose coupling in standard normalizations is

$$\beta_{\text{SG}} = 2g_{\text{eff}} \, . \tag{4.4}$$

It is important to note that, though $\mathcal{H}_{\text{q}} = \mathcal{H}_{\text{cl}}$ at weak coupling, the classical and quantum low energy physics are drastically different. Classically, we have a free massless theory independent of g_{eff}. Quantum mechanically, we have the non-trivial CFT of a boson compactified on a circle of radius $1/g_{\text{eff}}$. Even at very small coupling, the physics (anomalous dimensions, etc. . .) depends on g_{eff} in a non-trivial way. The fact that the classical and quantum singularity structure can coincide though the corresponding physics are different is also seen in supersymmetric gauge theories, for example in $\mathcal{N} = 1$ super Yang-Mills with $N_{\text{f}} = N_{\text{c}} + 1$ [2].

The description in terms of (4.2) is very useful at small g_{eff} (for a review on the sine-Gordon model including the original references, see e.g. [42]). We can for example use (4.2) to compute the mass of the two particle bound state,

$$m_{\text{b}} = \frac{4\sqrt{v_{N-1}}}{g_{\text{eff}}^2} \left(1 - \frac{g_{\text{eff}}^2}{2\pi} \right) \sin \frac{g_{\text{eff}}^2}{2(1 - g_{\text{eff}}^2/2\pi)} \, . \tag{4.5}$$

This formula should be compared to the analogous formulas (2.50) and (3.45) which were obtained in different regimes. The mass of the soliton (2.54) is also known,

$$m_{\text{sol}} = \frac{2\sqrt{v_{N-1}}}{g_{\text{eff}}^2} \left(1 - \frac{g_{\text{eff}}^2}{2\pi} \right) . \tag{4.6}$$

The action (4.2) also predicts that the bound state becomes unstable when $g_{\text{eff}}^2 \geq 2\pi/3$, and that we have a massless phase for $g_{\text{eff}}^2 > 2\pi$ even when $v_1 \neq 0$. This is of course incorrect, and (4.2) is not a good description of the physics at strong coupling. We know that the non-abelian degrees of freedom must come into the game and create a mass gap Λ. Moreover, the results of section 3 indicates that for $v_{N-1} \neq 0$, there should be a critical value of g_{eff} at which both the bound state and the soliton become massless and the low energy theory is an Ising CFT. This is consistent with the idea that (4.5) only gives an upper bound on the mass of the bound state, which can actually be significantly lower when the coupling increases due to the mixing with the other degrees of freedom (see section 2).

4.2 Strong coupling

To investigate the strongly coupled region, the idea is to modify the large N expansion of section 3 in order to treat exactly the fields Φ_N *and* Φ_{N-1} is the region of small v_{N-1}. This amounts to integrating out the $N-2$ fields $\Phi_1, \dots \Phi_{N-2}$ from (2.30), while keeping explicitly $\Phi_N = \sqrt{N}\varphi$ and $\Phi_{N-1} = \sqrt{N}\varphi'$. The effective action, which replaces (3.37), is then simply

$$s_{\text{eff}}[\alpha,\varphi,\varphi'] = \int d^2x \left(\frac{1}{2}\partial_\alpha\varphi\partial_\alpha\varphi + \frac{1}{2}\partial_\alpha\varphi'\partial_\alpha\varphi' + \frac{\alpha - h_N}{2}\varphi^2 + \frac{\alpha - h_{N-1}}{2}\varphi'^2 - \frac{\alpha}{2Ng^2}\right) + \\ + \frac{1}{N}\sum_{i=1}^{N-2} s[\alpha - h_i]\,. \tag{4.7}$$

We can then repeat straightforwardly the arguments of sections 3.4 and 3.5. The large-N low energy effective action at small v_{N-1} coincide with the one-loop result (4.2) except in a vanishingly small region in parameter space corresponding to $\tilde{\delta} v \ll \Lambda^2$ where

$$\tilde{\delta} v = \frac{\tilde{V}}{N-2}\sum_{i=1}^{N-2} \ln\frac{v_i}{\Lambda^2}\,, \tag{4.8}$$

with

$$\frac{1}{\tilde{V}} = \frac{1}{N-2}\sum_{i=1}^{N-2}\frac{1}{v_i}\,. \tag{4.9}$$

In this region, the fluctuations of the radius of the effective target space circle become important, and the low energy physics is governed by the Ginzburg-Landau action

$$S_{\text{eff}} = \int d^2x \left(\frac{1}{2}\partial_\alpha\chi\partial_\alpha\chi + \frac{1}{2}\partial_\alpha\chi'\partial_\alpha\chi' + \frac{v_{N-1}}{2}\chi'^2 - \frac{\tilde{\delta} v}{2}\left(\chi^2 + \chi'^2\right) + \frac{\pi\tilde{V}}{N}\left(\chi^2 + \chi'^2\right)^2\right). \tag{4.10}$$

This type of Ginzburg-Landau theory has already been studied [43], following the ideas of [41]. If $v_{N-1} = 0$ and $\tilde{\delta} v > 0$, it gives a description of the compactified boson CFT. The marginal operator corresponding to a change in the radius is $\chi^2 + \chi'^2$. We thus get a smooth interpolation with the physics described by (4.2). If $v_{N-1} > 0$, we have an Ising critical point corresponding to the restoration of the $\mathbb{Z}_{2(N)}$ symmetry at $\tilde{\delta v} = 0$, which means that the equation for $\mathcal{H}_{\mathrm{q}}$ is

$$\mathcal{H}_{\mathrm{q}} \cap \mathcal{M}_N : \quad \prod_{i=1}^{N-2} v_i = \Lambda^{2(N-2)} , \qquad \text{for } 0 < v_{N-1} \ll \Lambda^2 . \tag{4.11}$$

We thus also obtain a smooth interpolation of the physics described by (3.48), but we now see that $\mathcal{H}_{\mathrm{q}}$ must intersect the hyperplane $v_{N-1} = 0$. Similarly, if $v_{N-1} < 0$, we have an Ising critical point corresponding to the restoration of the $\mathbb{Z}_{2(N-1)}$ symmetry at $\tilde{\delta v} = v_{N-1}$, and thus

$$\mathcal{H}_{\mathrm{q}} \cap \mathcal{M}_{N-1} : \quad \prod_{i=1}^{N-2} v_i = \left(\Lambda^2 e^{v_{N-1}/\tilde{V}}\right)^{N-2} , \qquad \text{for } -\Lambda^2 \ll v_{N-1} < 0 . \tag{4.12}$$

For $v_{N-1} = 0$ and $\tilde{\delta v} = 0$, the two Ising CFT found for $v_{N-1} > 0$ and $v_{N-1} < 0$ respectively are coupled through an $\mathrm{O}(2)$ invariant interaction, giving an Ashkin-Teller critical point. This CFT is equivalent to the compactified boson CFT described by (4.2) for $v_{N-1} = 0$ at $g^2_{\text{eff}} = \pi/2$, which is itself equivalent to the $\mathbb{Z}_2$ orbifold CFT at the self-dual point [44] (for an elementary discussion, see [45]). The marginal operator corresponding to the variation of the orbifold radius is $\chi^2\chi'^2$ [43], but this is not associated to any microscopic operator in the Neumann model. For this reason, the equivalence with an orbifold theory does not seem to play a particular rôle in our model. What is extremely significant, however, is that the value $g^2_{\text{eff}} = \pi/2$ also corresponds to the critical coupling for a Kosterlitz-Thouless phase transition [46]. For $v_{N-1} = 0$ and $\tilde{\delta v} < 0$, (4.10) is indeed believed [43] to describe a massive phase, and no longer the compactified boson CFT. This is a very interesting physics: the creation of the mass gap in our model, which is fundamentally due to the non-abelian nature of the degrees of freedom, can be understood as coming from a condensation of vortices in the abelian description (4.2). This aspect will be discussed in great details in the next section.

The fact that the transition must occur at $g^2_{\text{eff}} = \pi/2$, and not at $g^2_{\text{eff}} = \infty$ as predicted by the $N \to \infty$ approximation (4.11), (4.3), is an exact result that cannot be deduced in any simple approximation scheme. Physically, it means that the effective target space does not literally shrinks to a point, but that the effective radius nevertheless becomes so small that its quantum fluctuations are important and any classical geometric interpretation of the target space is impossible. This is likely to be true anywhere on $\mathcal{H}_{\mathrm{q}}$, and thus for example the endpoint of figure 1 should rather be viewed as a very small, but non-vanishing, fluctuating quantum sphere.

4.3 The ansatz for the equation of $\mathcal{H}_{\mathrm{q}}$

We would like to present an equation for $\mathcal{H}_{\mathrm{q}}$ valid in the large-N limit on the whole parameter space $\mathcal{M}$, that interpolates smoothly between the different regimes (3.42),

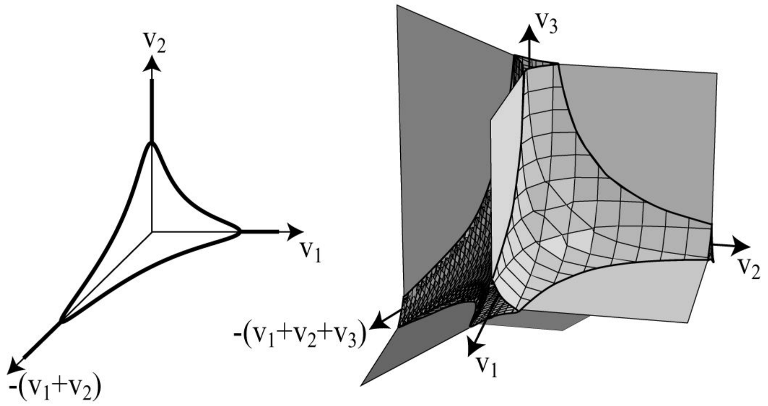

Figure 8: The quantum space of parameters $\mathcal{M}_{\mathrm{q}}$ in the cases $N = 3$ (left) and $N = 4$ (right). We have chosen arbitrarily $\tilde{\Lambda}/\Lambda = 3/2$.

(3.43), (4.11) and (4.12). To get such an equation, we clearly need to treat all the regions $\mathcal{M}_i$ (2.42) symmetrically. It does not seem to be possible to do this rigorously, for reasons explained at lenght above: we need different $1/N$ expansions in the different regions $\mathcal{M}_i$, and yet another expansion to understand the transition between the different regions. However, there is a field, the Lagrange multiplier α, that we can treat symmetrically on the whole parameter space. In particular, we can compute an $N \to \infty$ effective potential for α by integrating out the N fields Φ_i from (2.30),

$$v_{\mathrm{eff}}(\alpha) = -\frac{1}{8\pi N} \sum_{i=1}^{N} (\alpha - h_i) \ln \frac{\alpha - h_i}{e\Lambda^2} \,. \tag{4.13}$$

The saddle point α_* satisfies

$$\frac{dv_{\mathrm{eff}}(\alpha = \alpha_*)}{d\alpha} = -\frac{1}{8\pi N} \sum_{i=1}^{N-1} \ln \frac{\alpha_* - h_i}{\Lambda^2} = 0. \tag{4.14}$$

It is easy to see that (4.14) is consistent with our preceding formulas (3.40), (3.41). For example, if we consider the region $\mathcal{M}_N$ at weak coupling (Λ small), (4.14) implies an expansion $\alpha_* = h_N + \Lambda^{2(N-1)}/\prod_{i=1}^{N-1} v_i + \cdots$, and thus the corrections to (3.40) go to zero except on a vanishingly small region of $\mathcal{M}_N$ when $N \to \infty$. The advantage of (4.14) is that it is smooth over the entire $\mathcal{M}$. An improved equation for $\mathcal{H}_{\mathrm{q}}$ in $\mathcal{M}_N$ is then to use (3.39),

$$\prod_{i=1}^{N-1} (\alpha_* - h_i) = \Lambda^{2(N-1)}, \tag{4.15}$$

but with α_* now being determined by (4.14). A natural generalization of (4.15) on the whole of parameter space is

$$\mathcal{H}_{\mathrm{q}} \ : \quad \sum_{i=1}^{N} \prod_{\substack{j=1 \\ j \neq i}}^{N} (\alpha_* - h_j) = \kappa \Lambda^{2(N-1)} = \tilde{\Lambda}^{2(N-1)}, \tag{4.16}$$

where we have introduced a "phenomenological" constant $\kappa > 1$ which, though it cannot be seen in the $N \to \infty$ limit, is necessary to make $\mathcal{H}_{\mathrm{q}}$ intersect with the hyperplanes separating the different regions $\mathcal{M}_i$. κ can be considered as a non-perturbative quantity analogous to the Kosterlitz-Thouless critical coupling $g_{\mathrm{eff}} = \pi/2$ that cannot be calculated in a large-N expansion. The best justification we can find for (4.16) is by looking a posteriori to its consequences: by construction, is reproduces the correct result in the regions $\mathcal{M}_i$ separately (section 3.4), and it can be checked easily that it also gives a correct interpolation between the different regions (section 4.2). Though our results have been obtained at large N, it is very likely that they are qualitatively valid even for $N = 3$, and we have used (4.14) and (4.16) to obtain a sketch of the quantum space of parameters $\mathcal{M}_{\mathrm{q}}$ in the cases $N = 3$ and $N = 4$ in figure 1.

In our model, it is necessary to cross a surface of singularities to go from weak coupling to strong coupling. This is of course not true in general. For example, one may turn on a magnetic field (a $(v, 2)$ mass term in the terminology of section 2) in addition to the $(s, 2)$ term considered so far. This would make the phase transitions first order.

5. The non-local theory of electric and magnetic charges

5.1 General discussion

The picture of the quantum space of parameters in figure 1 clearly shows that we have two qualitatively different regimes in our model. Outside the hypersurface of singularities $\mathcal{H}_{\mathrm{q}}$, the usual geometrical interpretation of the non-linear σ model as a sum over maps from $\mathbb{R}^2$ to S^{N-1} is valid. The use of the coordinates on the sphere as the fundamental physical degrees of freedom is then justified. On the contrary, in the interior of $\mathcal{H}_{\mathrm{q}}$, the geometrical interpretation breaks down. It is of course still possible to define formally the model as a sum over maps from $\mathbb{R}^2$ to S^{N-1}, but the path integral will be dominated by highly singular maps from which the image of a smooth target space manifold cannot be reconstructed. A very interesting possibility is that there could be a (non-local) change of variables in the path integral to new degrees of freedom that would give a natural description of the physics at strong coupling. These new degrees of freedom may, or, more probably in our case, may not, have a geometrical interpretation in terms of coordinates on a new target space. In the case of $\mathcal{N} = 2$ supersymmetric non-linear σ models, this change of variables can be argued to exist in many cases and is called the "mirror map" (see e.g. [40, 47]). In the case of the integrable $\mathrm{O}(N)$ non-linear σ model, which corresponds to the origin of our space of parameters (figure 1), it is actually a long standing problem to find these "good" degrees of freedom (see e.g. [48]). When $N = 3$, a possibility is to use variables associated with instantons. In our model, instanton calculations can be done reliably at weak coupling since the v_is provide there an IR cutoff. However, I think that it is unlikely that instantons, which are smooth configurations of the original fields, can account for the physics near and in the interior of $\mathcal{H}_{\mathrm{q}}$ (it would be interesting to investigate this point, though). Moreover, one would like to have a unified description for all $N \geq 3$, and not only for $N = 3$, since the physics is qualitatively the same for all those cases. It is more natural to suspect that the relevant degrees of freedom at strong coupling are more singular configurations. Figure 1 suggests that configurations of the fields where

$$\mathbf{\Phi} = 0 \tag{5.1}$$

at some space-time points will be conspicuous. Around such points, the angular coordinates on the sphere can vary abruptly. These configurations are suppressed classically by the kinetic energy term in the action, but at strong coupling we see that they can actually contribute significantly to the path integral. For example, in the case of the two-sphere, we have two coordinates θ and ϕ with the identifications

$$(\theta, \phi) \equiv (\theta + 2\pi, \phi)\,, \qquad (\theta, \phi) \equiv (\theta, \phi + 2\pi)\,, \qquad (\theta, \phi) \equiv (-\theta, \phi + \pi)\,, \tag{5.2}$$

which suggest that relevant degrees of freedom could be vortices in the θ and ϕ variables as well as configurations where θ has square root branch cuts, as in a $\mathbb{Z}_2$ orbifold. Unfortunately, I do not know how to take into account this three types of defects together in an $\mathrm{O}(3)$ invariant way. Nevertheless, we can more modestly look at the regime $v_{N-1} \to 0$ studied in section 4 where the low energy effective theory is abelian. When $v_{N-1} = 0$, we have demonstrated at large N that the transition from weak (massless phase) to strong (massive phase) coupling is of the Kosterlitz-Thouless type. It is of course well-known [46] that such a transition can be understood as being triggered by a condensation of vortices $\theta \equiv \theta + 2\pi$. Let us stress an important non-trivial point of principle here. The low energy action at $v_{N-1} = 0$,

$$S_{\mathrm{eff}} = \frac{1}{2g_{\mathrm{eff}}^2} \int d^2x\, \partial_\alpha\theta\partial_\alpha\theta\,, \tag{5.3}$$

does *not* by itself predicts a Kosterlitz-Thouless transition at $g_{\mathrm{eff}}^2 = \pi/2$, but rather a massless phase at any g_{eff}. The action (5.3) is often used as a continuum version of the $\mathrm{O}(2)$ spin model defined on a lattice, g_{eff}^2 playing the rôle of a temperature. The lattice model is the object of interest in condensed matter physics, and was studied by Kosterlitz and Thouless [46]. In the lattice hamiltonian, configurations for which $\theta \to \theta + 2\pi$ abruptly are not suppressed; the suppression is an artefact of the continuum description given by (5.3). By introducing the vortices by hand in the continuum formulation (which means that we choose a non-zero fugacity for the singular field configurations corresponding to vortices in the definition of the path integral), one can get a good modeling of the lattice model [46], and in particular predict the mass gap at high temperature and the correct behaviour at the transition. In our case, however, we are not dealing with lattice models, but with a bona fide field theory. What has been proven in the large-N limit in section 4 is that a non-zero fugacity for the vortices is *dynamically* generated in the model because of the non-abelian (continuous) degrees of freedom. The physical origin of the condensation of vortices is thus completely different in our case and in the well-known abelian lattice model. Note that the vortices are not instantons of the non-abelian model and can exist for all N. This suggests that, more generally, the $\mathrm{O}(N)$ model may be naturally described in terms of "non-abelian" generalized vortices (5.2).

In the following, we are going to study how the standard picture of vortex condensation is modified when we go to the regime where v_{N-1} is small but non-zero. In sections 3 and 4 we obtained in the large-N limit a simple Ginzburg-Landau description of the physics on $\mathcal{H}_{\mathrm{q}}$. We do not expect to get as simple a description in the present context. In general, both weak coupling and strong coupling degrees of freedom can be relevant in the vicinity of $\mathcal{H}_{\mathrm{q}}$, where the transition between the weak coupling and strong coupling behaviours takes place, giving an unusual non-local description of the low energy physics. The same kind on physics has of course already been studied in condensed matter works on abelian lattice models. A $\cos p\theta$ potential in the lagrangian is interpreted in this context as a p-fold symmetry breaking term modeling crystal anisotropy. When $p \geq 4$, a perturbative analysis à la Kosterlitz is possible ([49, 50]), but when $p \leq 3$, and in particular in the case $p = 2$ we are interested in (4.2), this is not possible as we will review. Even in those cases, the nature of the physics is rather clear from the lattice point of view in particular (see the appendix of [51] or [52]). The general ideas presented later in this section are thus certainly not new, though a comprehensive analysis of the kind we offer does not seem to have appeared. Our main goal is really to emphasize the striking similarity between our non-supersymmetric system and the physics in the vicinity of an Argyres-Douglas point in supersymmetric gauge theories. Whether similar things can happen in non-supersymmetric four dimensional theories is of course still an open problem.

Before we turn to these points, let us make a simple heuristic remark. In our model, we have a clear interpretation of the fact that a compactified boson can be equivalent to the O(2) symmetric Ashkin-Teller model, and that this happens precisely at the Kosterlitz-Thouless coupling. One may then wonder whether a natural explanation for the fact that these CFTs are also equivalent to the $\mathbb{Z}_2$ orbifold at the self-dual radius could emerge. A naive argument in our context could be the following: since the Kosterlitz-Thouless transition is triggered by the non-abelian degrees of freedom, the angle θ should know somehow that is lives on a sphere, and an orbifold-like identification $\theta \equiv -\theta$ should then be implemented (5.2). However, the orbifold of the circle theory at the Kosterlitz-Thouless radius *is not* the orbifold at the self-dual radius. The naive interpretation thus does no make sense as it stands. It would be interesting to try to include the effect of the other Φ coordinates on the sphere to see if this picture might be made consistent.

5.2 The non-local description in the abelian regime

A general configuration of the field θ, including vortices, can be written

$$\theta(x) = \theta_0(x) + \theta_v(x)\,, \tag{5.4}$$

where θ_0 is the non-singular part and

$$\theta_v(x) = \sum_i n_i \arg(x - x_i) \tag{5.5}$$

corresponds to vortices ("magnetic charges") of charge n_i centered at the points x_i. When the total charge $\sum n_i$ is zero, the action (4.2) can be finite and is given by

$$\begin{aligned} S_{\text{eff}}(n_i; x_i) = {} & \frac{1}{2g_{\text{eff}}^2}\int d^2x \partial_\alpha\theta_0\partial_\alpha\theta_0 - \frac{\pi}{g_{\text{eff}}^2}\sum_{i,j} n_i n_j \ln|x_i - x_j| - \\ & - \frac{v_{N-1}}{8g_{\text{eff}}^2}\int d^2x \left[\prod_i \left(\frac{z - z_i}{\bar{z} - \bar{z}_i}\right)^{n_i} e^{2i\theta_0} + \prod_i \left(\frac{\bar{z} - \bar{z}_i}{z - z_i}\right)^{n_i} e^{-2i\theta_0}\right]. \end{aligned} \tag{5.6}$$

We recognize the standard Coulomb interaction between magnetic charges in two dimensions, but we have also a complicated interaction term between the magnetic charges and the fundamental spin waves ("electric charges"). We have noted $z_i = x_i^1 + ix_i^2$ and $\bar{z}_i = x_i^1 - ix_i^2$. The charges $n_i = \pm 1$ are the most relevant, as can be checked straightforwardly along the lines of the calculation presented below. We will thus write the partition function in term of the dynamically generated fugacity $F = f\Lambda^2$ (f is a dimensionless constant) for the $n_i = \pm 1$ vortices as

$$Z = \sum_{n=1}^{\infty} \frac{f^{2n}\Lambda^{4n}}{(n!)^2}\int\prod_{i=1}^{n}(d^2x_i d^2y_i)\int \mathcal{D}\theta_0\, e^{-S_{\text{eff}}[\theta_0; x_i, y_i]}\,, \tag{5.7}$$

where the x_is and y_is parametrize the positions of the charge $+1$ and charge -1 vortices respectively, and $S_{\text{eff}}[\theta_0; x_i, y_i]$ is the action (5.6) for such a configuration. By expanding the exponential in powers of v_{N-1} and calculating the path integral over θ_0, we get

$$Z = \sum_{m,n=0}^{\infty} \frac{(v_{N-1}/2g_{\text{eff}}^2)^{2m}}{(m!)^2} \frac{(f\Lambda^2)^{2n}}{(n!)^2} \int \prod_{p=1}^{m} (d^2u_p d^2w_p) \prod_{i=1}^{n} (d^2x_i d^2y_i) \times$$
$$\times \prod_{1\le p<q\le m} (|u_p - u_q||w_p - w_q|)^{2g_{\text{eff}}^2/\pi} \prod_{1\le p,q\le m} |u_p - w_q|^{-2g_{\text{eff}}^2/\pi} \times$$
$$\times \prod_{1\le i<j\le n} (|x_i - x_j||y_i - y_j|)^{2\pi/g_{\text{eff}}^2} \prod_{1\le i,j\le n} |x_i - y_j|^{-2\pi/g_{\text{eff}}^2} \times$$
$$\times \prod_{\substack{1\le i\le n\\ 1\le p\le m}} \frac{(u_p - x_i)(w_p - y_i)(\bar u_p - \bar y_i)(\bar w_p - \bar x_i)}{(\bar u_p - \bar x_i)(\bar w_p - \bar y_i)(u_p - y_i)(w_p - x_i)} . \tag{5.8}$$

The second line in (5.8) corresponds to the standard mass expansion of the original sine-Gordon model (4.2), the third line corresponds to the interaction between vortices, which is also described by a sine-Gordon model of the type (4.2) but with a coupling $g'_{\text{eff}} = \pi/g_{\text{eff}}$, and the fourth line corresponds to the interactions between the electric and magnetic charges. This partition function is reproduced, including the electric/magnetic interactions, by the perturbative formula

$$Z = \langle e^{-S_{\text{int}}} \rangle_{\text{free}} \tag{5.9}$$

where the average is computed with the free field weight (5.3) and

$$S_{\text{int}} = -\int d^2x \left(\frac{v_{N-1}}{4g_{\text{eff}}^2} \cos 2\theta + 2f\Lambda^2 \cos \frac{2\pi}{g_{\text{eff}}^2} \tilde\theta \right) . \tag{5.10}$$

The dual field $\tilde\theta$ is defined by

$$d\tilde\theta = -i * d\theta . \tag{5.11}$$

The formula (5.9) for Z is perturbative, since it involves an average over free fields. Interestingly, an off-shell, non-perturbative, formulation of this type of theories exists [53]. The basic idea is to treat the fields θ and $\tilde\theta$ as independent variables, and try to find an action whose equations of motions give (5.11) when the interactions are turned off. By writing (5.11) in the minkowskian,

$$\partial_0 \tilde\theta = -\partial_1 \theta , \qquad \partial_1 \tilde\theta = -\partial_0 \theta , \tag{5.12}$$

we see that $\partial_1 \tilde\theta$ is in some sense the canonical momentum associated with θ. The phase space path integral then suggests to try the action [53]

$$S_{\text{em}} = -\frac{1}{2g_{\text{eff}}^2} \int d^2x \left(\partial_0\theta\partial_1\tilde\theta + \partial_0\tilde\theta\partial_1\theta + (\partial_1\theta)^2 + (\partial_1\tilde\theta)^2 \right) + S_{\text{int}}[\theta, \tilde\theta] , \tag{5.13}$$

which indeed yields the correct equations of motion. If S_{int} depends on θ only, then by integrating out $\tilde\theta$ we would recover the standard Lorentz invariant action for a scalar, but in general it is impossible to write down a manifestly Lorentz invariant action for both θ and its dual. Operators associated with a dyon of electric charge n_e and magnetic charge n_m are

$$\mathcal{O}_{(n_e,n_m)} = e^{\pm i(n_e\theta + 2\pi n_m\tilde{\theta}/g_{\text{eff}})} . \tag{5.14}$$

We thus see that our model (5.10) corresponds to the case where both a magnetic charge (the vortex) ($n_e = 0, n_m = 1$) and an electric charge ($n_e = 2, n_m = 0$) are included. By rescaling the fields

$$\theta = g_{\text{eff}}\,\phi\,, \qquad \tilde{\theta} = g_{\text{eff}}\,\tilde{\phi} \tag{5.15}$$

and defining

$$F_e = \frac{v_{N-1}}{2g_{\text{eff}}^2}\,, \quad F_m = 4f\Lambda^2 \tag{5.16}$$

we can bring the action in a form

$$S_{\text{em}} = -\frac{1}{2}\int d^2x \Big(\partial_0\phi\partial_1\tilde{\phi} + \partial_0\tilde{\phi}\partial_1\phi + (\partial_1\phi)^2 + (\partial_1\tilde{\phi})^2 + F_e\cos(2g_{\text{eff}}\phi) + F_m\cos(2\pi\tilde{\phi}/g_{\text{eff}})\Big) \tag{5.17}$$

which is manifestly invariant under a strong/weak coupling S duality,

$$g_{\text{eff}} \overset{S}{\longleftrightarrow} \frac{\pi}{g_{\text{eff}}}\,, \qquad F_e \overset{S}{\longleftrightarrow} F_m\,, \qquad \phi \overset{S}{\longleftrightarrow} \tilde{\phi}\,. \tag{5.18}$$

Note that S is not the usual T duality that would exchange the unit magnetic charge with the unit electric charge and g_{eff} with $2\pi/g_{\text{eff}}$, in the same way as the monodromy at an Argyres-Douglas point is not the full duality group.

The quantization of (5.17) can be done in the standard way. We have two second class constraints,

$$\Pi = -\frac{1}{2}\,\partial_1\tilde{\phi}\,, \qquad \tilde{\Pi} = -\frac{1}{2}\,\partial_1\phi\,, \tag{5.19}$$

and the Dirac bracket yields the equal time commutation relation

$$\Big[\phi(x^0,x^1), \tilde{\phi}(x^0,x'^1)\Big] = i\,\Theta(x^1 - x'^1) \tag{5.20}$$

where Θ is the Heavyside step function. We see that the non-locality of our theory, which describes the interaction of electric and magnetic degrees of freedom, is simply encoded in this commutation relation. Let us note finally that actions describing both electric and magnetic charges in four dimensions, and that should play a rôle in understanding Argyres-Douglas CFTs, have been constructed along the lines of (5.17) [54].

5.3 The strongly coupled fixed point à la Argyres-Douglas

The standard way to treat (5.17) would be to use perturbation theory in F_e and F_m ([49, 55]). This makes sense when the corresponding operators are nearly marginal. What are we expecting in our case? From the results of section 4, or from the intuition gained in lattice model, we would like to show that (5.17) has an Ising fixed point. If we have only one fixed point, which is likely, it must occur at the S-invariant point

$$g_{\mathrm{eff}} = g_* = \sqrt{\pi}\,, \qquad F_e = F_m\,. \tag{5.21}$$

On the other hand, the perturbative dimensions of the "electric" and "magnetic" operators are respectively

$$\Delta_{\mathrm{elec}} = \frac{g_{\mathrm{eff}}^2}{\pi}\,, \qquad \Delta_{\mathrm{mag}} = \frac{\pi}{g_{\mathrm{eff}}^2}\,, \tag{5.22}$$

which would correspond at the fixed point to

$$\Delta_{\mathrm{elec}*} = \Delta_{\mathrm{mag}*} = 1\,, \tag{5.23}$$

values far below the threshold of marginality $\Delta = 2$ where perturbation theory is valid. This means that the fixed point we are looking for must be at strong coupling. This is also what is expected in four dimensions, and in general it would imply, even in two dimensions, that the theory in not tractable. Luckily, in the very particular case at hand, we have a powerful tool at our disposal that can solve the problem: we can fermionize and replace the scalar ϕ by a Dirac spinor ψ. Fermionizing is of course a very natural thing to do since we are looking for an Ising critical point. Moreover, the self-dual dimensions (5.23) strongly suggest that the electric and magnetic operators can be viewed as fermion mass terms near the fixed point. Fermionization is a bit unusual in our case, however, because the "topological" symmetry $\tilde{\phi} \to \tilde{\phi} + \text{constant}$ is broken by the interaction term in (5.17) and thus we do not have a conserved fermion current $j \propto *d\phi$ (in other words, the relations (5.11) or (5.12) are no longer correct due to the interactions between ϕ and $\tilde{\phi}$). The perturbative reasoning of Coleman [56] can nevertheless be reproduced straightforwardly, starting from (5.9). In the same way as the sine-Gordon potential introduces Dirac mass terms, the $\cos(2\pi\tilde{\phi}/g_{\mathrm{eff}})$ term introduces Majorana mass terms. More precisely, we can identify

$$\mu\,\cos(2g_{\mathrm{eff}}\phi) = -\pi\,\bar{\psi}\psi\,, \qquad \mu\,\cos(2\pi\tilde{\phi}/g_{\mathrm{eff}}) = -\pi\,\bar{\psi}^C\psi\,, \tag{5.24}$$

where we have introduced an arbitrary renormalization scale μ. The theory (5.17) is then equivalent to the fermionic theory with lagrangian

$$S_{\mathrm{ferm}} = \int d^2x\,\Big(i\bar{\psi}\gamma^\mu\partial_\mu\psi + \frac{\pi F_e}{\mu}\,\bar{\psi}\psi + \frac{\pi F_m}{\mu}\,\bar{\psi}^C\psi - \frac{G}{2}\,\bar{\psi}\gamma^\mu\psi\,\bar{\psi}\gamma_\mu\psi\Big) \tag{5.25}$$

with

$$\frac{G}{\pi} = \frac{\pi - g_{\rm eff}^2}{g_{\rm eff}^2}\,. \tag{5.26}$$

It is useful to decompose the Dirac fermion ψ in terms of two Majorana fermions $\lambda = -\lambda^C$ and $\chi = -\chi^C$ such that

$$\begin{aligned}\psi &= \begin{pmatrix}\psi_- \\ \psi_+\end{pmatrix} = \frac{\lambda + i\chi}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}\lambda_- + i\chi_- \\ \lambda_+ + i\chi_+\end{pmatrix}\\ \bar\psi &= (\bar\psi_+, \bar\psi_-) = \frac{\bar\lambda - i\bar\chi}{\sqrt{2}} = \frac{1}{\sqrt{2}}(\lambda_+ - i\chi_+, \lambda_- - i\chi_-)\,.\end{aligned} \tag{5.27}$$

In terms of these new variables, and by introducing the light-cone coordinates $x^\pm = x^0 \pm x^1$ and

$$M = \frac{\pi(F_e - F_m)}{\mu}\,, \qquad M' = \frac{\pi(F_e + F_m)}{\mu}\,, \tag{5.28}$$

we have

$$\begin{aligned}S_{\rm ferm} = \int d^2x\, \Bigl(&\lambda_+\partial_-\lambda_+ + \chi_+\partial_-\chi_+ - \lambda_-\partial_+\lambda_- - \chi_-\partial_+\chi_- - \\ &- iM'\,\chi_-\lambda_+ - iM\,\chi_+\lambda_- - \frac{\pi(\pi - g_{\rm eff}^2)}{2g_{\rm eff}^2}\,\lambda_-\chi_-\lambda_+\chi_+\Bigr)\,.\end{aligned} \tag{5.29}$$

We thus see that at the self-dual point (5.21), for which $M = 0$, we have at low energy a free Majorana fermion (λ_-, χ_+), which indeed describes an Ising critical point as was to be shown. The S duality (5.18) acts as

$$M \overset{S}{\longleftrightarrow} -M\,, \qquad M' \overset{S}{\longleftrightarrow} M'\,, \qquad \chi_\pm \overset{S}{\longleftrightarrow} \mp\chi_\pm\,, \qquad \lambda_\pm \overset{S}{\longleftrightarrow} \lambda_\pm\,, \tag{5.30}$$

and thus reduces at low energy to the Kramers-Wannier duality of the Ising model.

I would like to conclude this section by giving a more rigorous derivation of (5.25) and explaining at the same time a point that might puzzle the reader. The puzzle is the following. The duality (5.18) should be valid not only at the fixed point, but also for any coupling $g_{\rm eff} \neq g_*$. Using (5.18) and (5.30), the four-fermions interaction term in (5.25) naively transforms under S as

$$G\,\bar\psi\gamma^\mu\psi\,\bar\psi\gamma_\mu\psi \overset{S}{\longleftrightarrow} \frac{G}{1 + G/\pi}\,\bar\psi\gamma^\mu\psi\,\bar\psi\gamma_\mu\psi\,. \tag{5.31}$$

This is clearly inconsistent, since by repeating the transformation (5.31), we could make the coupling arbitrarily small. The subtlety comes from the definition of the fermion current from which the four-fermions interaction is constructed [57]. The canonical definition, that must be used for example in a hamiltonian approach, uses a point-splitting regularization *at equal time*. However, such a definition does not produce a Lorentz vector. A correct definition of the current must actually involve both a space-like and a time-like splitting. There is a one-parameter family of Lorentz covariant current that can be defined in that way as we will review below. Products of fields taken at time-like intervals depend on the dynamics of the theory. The Lorentz invariant four-fermions interaction must then depend in some hidden way on the coupling constant G, and this invalidates (5.31). Since we don't want to repeat the tedious arguments of [57], we are going to show much more straightforwardly how this comes about by giving at the same time a rigorous non-perturbative derivation of (5.25) starting from (5.17). We will consider the hamiltonian in the Schrödinger picture,

$$H = \frac{1}{2}\left[\left(\frac{d\phi}{dx}\right)^2 + \left(\frac{d\tilde{\phi}}{dx}\right)^2 + F_e \cos(2g_{\rm eff}\phi) + F_m \cos\left(\frac{2\pi\tilde{\phi}}{g_{\rm eff}}\right)\right], \tag{5.32}$$

with the quantization condition (5.20), and apply the bosonization/fermionization formulas on H. In the Schrödinger picture, all operators are of course regularized by point splitting at equal time. Fermions can be defined by using any scalar fields X and $\tilde{X}$ satisfying the same quantization condition as ϕ and $\tilde{\phi}$ (5.20) by

$$\psi_- = \sqrt{\frac{\mu}{2\pi}}\, e^{i\sqrt{\pi}(\tilde{X}+X)}, \qquad \psi_+ = \sqrt{\frac{\mu}{2\pi}}\, e^{i\sqrt{\pi}(\tilde{X}-X)}. \tag{5.33}$$

For our purposes, we need

$$X = \frac{g_{\rm eff}}{\sqrt{\pi}}\,\phi, \qquad \tilde{X} = \frac{\sqrt{\pi}}{g_{\rm eff}}\,\tilde{\phi}. \tag{5.34}$$

The current $j^\mu = \bar{\psi}\gamma^\mu\psi$ is

$$j^0 = i(\psi_+\bar{\psi}_+ - \psi_-\bar{\psi}_-) = \frac{g_{\rm eff}}{\pi}\frac{d\phi}{dx}, \qquad j^1 = -i(\psi_+\bar{\psi}_+ + \psi_-\bar{\psi}_-) = \frac{1}{g_{\rm eff}}\frac{d\tilde{\phi}}{dx}, \tag{5.35}$$

and this is not a Lorentz vector. The four-fermions interaction is then defined by

$$\bar{\psi}\gamma^\mu\psi\,\bar{\psi}\gamma_\mu\psi \equiv q_0 (j^0)^2 - q_1 (j^1)^2, \tag{5.36}$$

where q_0 and q_1 are parameters that we will adjust later to make the interaction a Lorentz scalar. It is straightforward to check that the fermion hamiltonian

$$\begin{aligned} H_{\rm ferm} = {}& \bar{\psi}_-\frac{d\psi_-}{dx} + \bar{\psi}_+\frac{d\psi_+}{dx} + \frac{\pi F_e}{\mu}(\psi_-\bar{\psi}_+ + \psi_+\bar{\psi}_-) + \\ & + \frac{\pi F_m}{\mu}(\psi_-\psi_+ + \psi_+\psi_-) + \frac{G}{2}\left(q_0(j^0)^2 - q_1(j^1)^2\right) \end{aligned} \tag{5.37}$$

is equivalent to the bosonic hamiltonian

$$\begin{aligned} H_{\rm bos} = {}& \frac{1}{2}\left[\frac{g_{\rm eff}^2}{\pi}\left(1 + \frac{Gq_0}{\pi}\right)\left(\frac{d\phi}{dx}\right)^2 + \frac{\pi}{g_{\rm eff}^2}\left(1 - \frac{Gq_1}{\pi}\right)\left(\frac{d\tilde{\phi}}{dx}\right)^2\right] + \\ & + F_e \cos(2g_{\rm eff}\phi) + F_m \cos\left(\frac{2\pi\tilde{\phi}}{g_{\rm eff}}\right). \end{aligned} \tag{5.38}$$

The four-fermions interaction will be a Lorentz scalar if the $G \neq 0$ theory is Lorentz invariant with the same speed of light $c = 1$ as the free $G = 0$ fermionic theory. This in turn is equivalent to the condition that H_{bos} coincide with H (5.32), which happens for

$$q_1 = \frac{g_{\text{eff}}^2}{\pi}\, q_0 \tag{5.39}$$

and

$$\frac{G}{\pi} = \frac{1}{q_0}\,\frac{\pi - g_{\text{eff}}^2}{g_{\text{eff}}^2}\,. \tag{5.40}$$

Equation (5.39) shows that we have one free parameter q_0 in defining the scalar four-fermions interaction, and also that this interaction depends implicitly on g_{eff} through q_1. Equation (5.40) generalizes the standard Coleman's formula (5.26) which corresponds to the choice $q_0 = 1$. Of course the physics does not depend on q_0. We can now deduce the correct duality transformations of the four-fermions term. Equations (5.36), (5.39) and (5.40) implies

$$G\,\bar\psi\gamma^\mu\psi\,\bar\psi\gamma_\mu\psi = \frac{\pi(\pi - g_{\text{eff}}^2)}{g_{\text{eff}}^2}\left((j^0)^2 - \frac{g_{\text{eff}}^2}{\pi}(j^1)^2\right). \tag{5.41}$$

Equations (5.35), (5.27) and (5.30) show that

$$j^0 \overset{S}{\longleftrightarrow} j^1\,, \tag{5.42}$$

and thus by using (5.18) we finally deduce that the four-fermions term is actually *invariant* under the S duality.

It would be interesting to study possible strongly coupled fixed points in theories of the same kind as (5.17), but including other electric and magnetic charges, multi-components fields, and θ angles [58]. The fact that the Argyres-Douglas CFTs appearing in four dimensional supersymmetric gauge theories admit an ADE classification [59] suggests that all the minimal unitary CFTs, that also follow the ADE pattern [60], could be obtained in this way. A very simple example is the case where the electric charge $n_e = 2$ of our model is replaced by an electric charge $n_e = 3$. This yields a strongly coupled fixed point of the Potts type.

6. The double scaling limits

We have discussed three different descriptions of the non-trivial low energy physics appearing in our model near the singularities of parameter space. The Ginzburg-Landau description (sections 3.5 and 4.2) and the fermionic description (section 5.3) are very simple, and can exist because of peculiarities of two-dimensional physics. The third description (section 5.2), in terms of a theory of electric and magnetic charges, is a priori more generic, and I have emphasized the strong similarity with

Argyres-Douglas CFTs in four dimensions in particular. We would like now to derive a fourth description, whose counterpart in four dimensions would be a string theory description of the vicinity of an Argyres-Douglas CFT (or more generally of the vicinity of generic non-trivial critical points that can be found in the context of gauge theories with Higgs fields). We will be rather brief on this interesting aspect of our model here, since we hope to provide more details in [14]. We are simply going to show explicitly that a double scaling limit in the sense of the "old" matrix models [13] can be defined when we approach the Ising or Ashkin-Teller singularities on $\mathcal{H}_{\mathrm{q}}$, and that the double scaled theory coincide with the low energy theory in the vicinity of $\mathcal{H}_{\mathrm{q}}$. This provides a duality between the interacting theory describing the low energy degrees of freedom and a theory of randomly branched polymers, as explained in [61] and as will be reviewed in [14]. The discussion is similar to the case of the ordinary vector models, as reviewed for example by J. Zinn-Justin in [38].

There are three different types of critical behaviour occuring in our model (Ising, Ashkin-Teller and compactified boson), and thus a priori three different types of double scaling limits can be considered. Let us first discuss the vicinity of the Ising critical point. As discussed in section 3.5, when we approach the critical hypersurface $\mathcal{H}_{\mathrm{q}}$, $\delta v \to 0$, the standard $1/N$ expansion in plagued by IR divergences. This means that in the large-N expansion of physical quantities, like (3.1), the coefficients W_l will diverge as $\delta v \to 0$. Proving that a double scaling limit can be defined when $\delta v \to 0$ amounts to proving that these divergences have a specific form for any l such that they can be compensated for by taking $N \to \infty$ and $\delta v \to 0$ in a correlated way. The double scaled theory then gets contributions from all orders in $1/N$. This means in particular that we must take into account the renormalization of the original non-linear σ model to all orders. Remarkably enough, this can be done very simply. The crucial point is of course that the divergences in W_l are due to IR effects.

We have shown in section 3.5 that at momenta $p \ll \Lambda$, the physics is correctly described by the action (3.48). By using the rescaled variables (3.51), we deduce that physical quantities can be correctly described by the action

$$S[\chi] = \int d^2x' \left(\frac{1}{2}\, \partial'_\alpha \chi \partial'_\alpha \chi - \frac{N\delta v}{2}\, \chi^2 + \pi V\, \chi^4 \right) \tag{6.1}$$

for momenta

$$p' \ll \Lambda\sqrt{N}\,. \tag{6.2}$$

If we now take the limit $N \to \infty$, the effective UV cut-off of (6.1), which is of order

$$\Lambda_{0,\mathrm{eff}} \sim \Lambda\sqrt{N}\,, \tag{6.3}$$

goes to infinity. The action (6.1) thus needs to be renormalized accordingly, in order to get a finite limit. This is done with the help of the standard normal ordering,

$$\chi^4 = :\chi^4: + 6\, \frac{1}{2\pi} \ln \frac{\Lambda_{0,\mathrm{eff}}}{\Lambda}\, \chi^2 + 3 \left(\frac{1}{2\pi} \ln \frac{\Lambda_{0,\mathrm{eff}}}{\Lambda} \right)^2 . \tag{6.4}$$

By replacing into (6.1), we see that, up to some trivial factors, we obtain a finite double scaled partition function in the limit

$$N \to \infty\,, \qquad \delta v \to 0\,, \qquad N\,\frac{\delta v}{V} - 3\ln N = \text{constant}\,, \tag{6.5}$$

as was to be shown. In momentum dependent quantities, $p' = p\sqrt{N}$ must be held fixed. The same reasoning show that the correct double scaling limit in the vicinity of the Ashkin-Teller critical point (4.10) is defined by

$$\begin{aligned}
&N \to \infty\,, \qquad \tilde{\delta} v \to 0\,, \qquad v_{N-1} \to 0\,,\\
&N\,\frac{\tilde{\delta} v}{\tilde{V}} - 4\ln N = \text{constant}\,, \qquad N\,\frac{\tilde{\delta} v - v_{N-1}}{\tilde{V}} - 4\ln N = \text{constant}'\,.
\end{aligned} \tag{6.6}$$

The last case that we have to consider is the vicinity of the compactified boson CFT, for example $v_{\mathrm{N}-1} \to 0$ with $\prod_{i=1}^{N-2} v_i > \Lambda^{2(N-2)}$. We cannot define a double scaling limit in this case. Indeed, in the $N \to \infty$ limit, the physics is well described by the action (4.2), since $g_{\text{eff}}^2 \sim 1/N$ is then very small. We see that the $1/N$ corrections are simply corrections to the radius of the compactified boson, which correspond to a marginal operator. This is very different to the case of the Ising or Ashkin-Teller CFTs, and does not produce the wild IR divergences that are necessary to be able to define a double scaling limit.

Acknowledgments

I would like to acknowledge interesting discussions with Duncan Haldane.

A. General solitary waves solutions

As discussed in section 2.3.3, the most general static, finite energy solutions to the field equations (2.31) satisfy

$$\frac{d^2\Phi_i}{dx^2} = (\alpha - h_i)\,\Phi_i \tag{A.1}$$

with

$$\lim_{x\to\pm\infty} \Phi_i^2 = \frac{\delta_{i,N}}{g^2} \tag{A.2}$$

when $h_N = \max_{1\le i\le N} h_i$, which we shall assume in the following. α is a Lagrange multiplier implementing the constraint

$$\sum_{i=1}^{N} \Phi_i^2 = \frac{1}{g^2}\,. \tag{A.3}$$

The set of equations (A.1) and (A.3) can be interpreted as the Newton equations for the motion of a particle constrained to move on the sphere S^{N-1} of radius $1/g$ in the quadratic potential

$$U(\Phi) = +\frac{1}{2}\sum_{i=1}^{N} h_i \Phi_i^2 = -V(\Phi). \tag{A.4}$$

The rôle of the time in the mechanical problem is played by x, and $V(\Phi) = -U(\Phi)$ is the potential of the corresponding field theory. That such a mechanical analogy is possible for solitons in two dimensions is of course banal (see e.g. [42]). The peculiarity of the soliton problem is that we are interested only in the motions satisfying (A.2).

We show in this appendix that all the solitonic solutions to our model can be found, and are expressed in terms of elementary functions only. This generalizes the similar result for the sine-Gordon theory. This is possible because the classical mechanical problem (A.1), (A.3) is integrable in the Liouville sense. It was first studied by C. Neumann in the case $N = 3$, who shows that one can separate the variables in the Hamilton-Jacobi equation [36], and discussed much later in the general case [62]. In the following, we present a self-contained elementary analysis in the case where all the h_is are distinct. The more symmetric cases where some of the h_is would coincide can be easily obtained by taking suitable limits, but we will not spell out these details. In section A.1 we explain how to solve the equations of motion in general, and then we specialize to the soliton problem.

A.1 General analysis

The starting point are the formulas for the constants of the motion [62],

$$\mathcal{I}_i = \frac{\Phi_i^2}{g^2} + \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{(\Phi_j \Phi_i' - \Phi_j' \Phi_i)^2}{h_i - h_j}, \qquad 1 \leq i \leq N, \tag{A.5}$$

where $\Phi_i' = d\Phi_i/dx$. Only $N-1$ of the $\mathcal{I}_i$s are independent because of the identity

$$\sum_{i=1}^{N} \mathcal{I}_i = \frac{1}{g^4}. \tag{A.6}$$

It is straightforward to check that $d\mathcal{I}_i/dx = 0$ by using (A.1) and (A.3). These $N-1$ independent integrals of the motion actually form an involutive system with respect to the Poisson bracket (by identifying Φ_i' with the conjugate momentum), which prove Liouville integrability. The energy is simply expressed in terms of the $\mathcal{I}_i$s,

$$\mathcal{E} = \frac{1}{2}\sum_{i=1}^{N}\left(\Phi_i'^2 + h_i\Phi_i^2\right) = \frac{g^2}{2}\sum_{i=1}^{N} h_i \mathcal{I}_i. \tag{A.7}$$

We are thus left with $N-1$ independent *first order* differential equations, instead of the second order equations (A.1). In order to solve these equations, it is very useful to rewrite the definition of the $\mathcal{I}_i$s in a more compact, elegant way. For this purpose, we introduce an arbitrary complex parameter z and define

$$q(x,y;z) = \sum_{i=1}^{N} \frac{x_i y_i}{z-h_i}\,, \qquad q(x;z) = q(x,x;z)\,, \tag{A.8}$$

$$\mathcal{I}(\Phi,\Phi';z) = \sum_{i=1}^{N} \frac{\mathcal{I}_i(\Phi,\Phi')}{z-h_i}\,. \tag{A.9}$$

The formulas (A.5) are equivalent to the equation

$$\mathcal{I}(\Phi,\Phi';z) = q(\Phi;z)\left(q(\Phi';z)+\frac{1}{g^2}\right) - q(\Phi,\Phi';z)^2\,, \tag{A.10}$$

and the differential equations we want to solve can be written

$$\mathcal{I}(\Phi,\Phi';z) = \sum_{i=1}^{N} \frac{\mathcal{I}_i}{z-h_i} = \frac{1}{g^4}\frac{b(z)}{a(z)}\,, \tag{A.11}$$

where

$$a(z) = \prod_{i=1}^{N}(z-h_i)\,, \tag{A.12}$$

and $b(z)$ is defined by (A.11). The arbitrary parameter z is then chosen to make the equations (A.11) as simple as possible. Equation (A.10) suggests that a good choice is

$$q(\Phi; z=\mu) = 0\,. \tag{A.13}$$

This equation has generically $N-1$ solutions $z=\mu_1,\ldots,\mu_{N-1}$ that can be expressed in terms of the $N-1$ independent coordinates Φ_i on the sphere, and vice-versa. The μ_is are usually called elliptic coordinates. Explicitly, by introducing

$$m(z;\mu) = \prod_{i=1}^{N-1}(z-\mu_i), \tag{A.14}$$

we have

$$q(\Phi;z) = \frac{1}{g^2}\frac{m(z;\mu)}{a(z)} \tag{A.15}$$

and

$$g^2\,\Phi_i^2 = \frac{m(z=h_i;\mu)}{a'(z=h_i)} = \frac{\prod_{j=1}^{N-1}(h_i-\mu_j)}{\prod_{j=1,j\neq i}^{N}(h_i-h_j)}\,. \tag{A.16}$$

We now make the change of variable $\Phi \mapsto \mu$ in (A.11). By taking the derivative of (A.13) with respect to x we get

$$q(\Phi, \Phi'; \mu_i) = \frac{1}{2}\mu_i' \sum_{j=1}^{N} \frac{\Phi_j^2}{(\mu_i - h_k)^2}\,, \tag{A.17}$$

and by taking the derivative of (A.15) with respect to z we get

$$\sum_{j=1}^{N} \frac{\Phi_j^2}{(\mu_i - h_j)^2} = -\frac{1}{g^2}\,\frac{\partial_z m(z=\mu_i;\mu)}{a(z=\mu_i)}\,. \tag{A.18}$$

From these relations and (A.10), (A.11) and (A.13) we finally obtain

$$\mu_i' = 2\epsilon_i \frac{\sqrt{-\,a(\mu_i)b(\mu_i)}}{\prod_{j=1,j\neq i}^{N-1}(\mu_i - \mu_j)}\,, \tag{A.19}$$

with $\epsilon_i = \pm 1$. We can put (A.19) in a more suggestive form by using the

Lemma 1. *If $P(z) = \prod_{i=1}^{n}(z - r_i)$ is an arbitrary polynomial, then*

$$\sum_{i=1}^{n} \frac{r_i^{n-j}}{P'(r_i)} = \delta_{j,1}\,, \qquad j \geq 1. \tag{A.20}$$

This is proven by calculating

$$\oint_C \frac{dz}{2i\pi}\,\frac{z^{n-j}}{P(z)} \tag{A.21}$$

with the help of the residue theorem and by taking C to be a circle centered at $z = 0$ of radius $R \to \infty$.

The lemma implies that (A.19) is equivalent to

$$\sum_{i=1}^{N-1} \epsilon_i \frac{\mu_i^{N-1-j}\mu_i'}{2\sqrt{-a(\mu_i)b(\mu_i)}} = \delta_{j,1}\,, \qquad 1 \leq j \leq N-1\,, \qquad \epsilon_i = \pm 1\,. \tag{A.22}$$

The general solution of the equations of motion are then given by elliptic integrals associated with the genus $N-1$ hyperelliptic curve

$$y^2 = -4a(x)b(x)\,. \tag{A.23}$$

However, as already stressed, we are not interested with the most general solutions, but with the ones corresponding to the solitons of our field theory model. We now focus on this special case.

JHEP06(2001)057

A.2 The solitons for general N

A.2.1 Basic equations

The boundary conditions (A.2) were imposed in order to have a finite mass M for the solitons. M is simply given by the action of the mechanical problem,

$$M = \frac{1}{2}\int_{-\infty}^{+\infty} dx\left(\sum_{i=1}^{N}\Phi_i'^2 + v_i\Phi_i^2\right), \tag{A.24}$$

where $v_i = h_N - h_i$ as usual. For M to be finite, it is necessary to have

$$\mathcal{I}_i = \frac{\delta_{i,N}}{g^4}, \tag{A.25}$$

that is to say,

$$b(z) = \prod_{i=1}^{N-1}(z-h_i)\,. \tag{A.26}$$

We will see below that (A.25) is actually sufficient to have a finite M. The conditions (A.25) thus characterize the particular motions that can be interpreted as solitary waves in our original theory.

With this particular choice for $b(z)$, (A.22) yields

$$\sum_{i=1}^{N-1}\epsilon_i\,\frac{\mu_i^{N-1-j}\,d\mu_i}{2\sqrt{h_N-\mu_i}\,\prod_{j=1}^{N-1}(\mu_i-h_j)} = \delta_{j,1}\,dx\,,\qquad 1\leq j\leq N-1\,,\qquad \epsilon_i=\pm1\,. \tag{A.27}$$

We introduce the new coordinates

$$y_i = \epsilon_i\sqrt{h_N-\mu_i}\,, \tag{A.28}$$

and we finally get our basic equation that can be written in two equivalent forms corresponding respectively to (A.19) and (A.22),

$$\frac{\prod_{j=1,j\neq i}^{N-1}(y_j^2-y_i^2)}{\prod_{j=1}^{N-1}(v_j-y_i^2)}\,dy_i = -dx\,,\qquad 1\leq i\leq N-1\,, \tag{A.29}$$

$$\sum_{i=1}^{N-1}\frac{(h_N-y_i^2)^{N-1-j}\,dy_i}{\prod_{j=1}^{N-1}(v_j-y_i^2)} = -\delta_{j,1}\,dx\,,\qquad 1\leq j\leq N-1\,. \tag{A.30}$$

Equation (A.30) can be integrated immediately in terms of elementary functions and $N-1$ constants of integration, which gives the full solution of the problem. The fact that only elemetary functions are involved is due to the fact that the solitonic solutions correspond to a highly degenerate hyperelliptic curve (A.23). Explicit formulas will be given for the case $N=3$ in the next subsection. Note that though (A.30) is best suited to an explicit integration, a qualitative analysis of the solution, to which we now turn, is most easily performed by using (A.29).

A.2.2 Description of the solutions

We choose without loss of generality $h_N > h_{N-1} > \cdots > h_1$. Equation (A.16) shows that $\Phi_j^2 \geq 0$ implies that each interval $[h_i, h_{i+1}]$ contains one and only one of the μ_js, or equivalently that each interval $[v_i, v_{i-1}]$ contains one and only one of the y_js. We consider a solution such that

$$\lim_{x\to-\infty} \Phi_i = +\frac{\delta_{i,N}}{g}, \tag{A.31}$$

and we conventionally choose the ordering of the coordinates such that this boundary condition corresponds to

$$\lim_{x\to-\infty} y_i = \sqrt{v_i}\,. \tag{A.32}$$

By looking at (A.29), we see that $dy_1 < 0$ when $x \to -\infty$. More generally, as long as $y_i \in [\sqrt{v_{i+1}}, \sqrt{v_i}]$, which is true at least up to some finite value of x because one can have only one y in each of these intervals, a simple recursive reasoning using (A.29) shows that $dy_i < 0$ for all i. Two y_i^2 can eventually swap their intervals. Equation (A.29) shows that it is possible to have $y_i^2 = v_j$ at some finite x as long as some other $y_k^2 = v_j$ at the same x, the pole in the denominator being then compensated by a zero in the numerator. With our particular boundary conditions (A.32), the only consistent way this can happen is by having first $y_{N-1} \to -\sqrt{v_{N-1}}$ and $y_{N-2} \to \sqrt{v_{N-1}}$ (step 1), then $y_{N-1} \to -\sqrt{v_{N-2}}$ and $y_{N-3} \to \sqrt{v_{N-2}}$ (step 2), and so on up to step $N-2$ when $y_{N-1} \to -\sqrt{v_2}$ and $y_1 \to \sqrt{v_2}$. This is followed by $N-3$ additional steps during which $y_1 \to \sqrt{v_3}$ and $y_{N-2} \to -\sqrt{v_3}$ (step $N-1$), $y_1 \to \sqrt{v_4}$ and $y_{N-3} \to -\sqrt{v_4}$ (step N), and so on up to step $2N-5$ when $y_1 \to \sqrt{v_{N-1}}$ and $y_2 \to -\sqrt{v_{N-1}}$. At each intermediate step it can be checked using (A.30) that $dy_i < 0$ for all i, so the y_is are monotonic decreasing functions of x. At the end of the day we have

$$\lim_{x\to+\infty} y_i = -\sqrt{v_{N-i}}\,. \tag{A.33}$$

What we have been describing is the generic solution. By restricting oneself to motions for which some of the Φ_is are equal to zero, one can obtain other solutions, corresponding to the generic solution at a lower value of N.

To get a better physical understanding of the nature of the solution, let us consider the coordinates Φ_i. First, let us note that the transition between the different steps correspond to the successive crossing of the hyperplanes $\Phi_{N-1} = 0, \Phi_{N-2} = 0, \ldots, \Phi_2 = 0, \Phi_3 = 0, \ldots, \Phi_{N-1} = 0$. This shows that the trajectory lies entirely on the half sphere $\Phi_1 > 0$ or $\Phi_1 < 0$, the two cases being related by the symmetry $Z_{2(1)}$. Second, and more importantly, as x increases we have successively $y_{N-1} = 0$, $y_{N-2} = 0$, and so on up to $y_1 = 0$. This shows that

$$\Phi_N^2 = \frac{1}{g^2}\frac{y_1^2 \cdots y_{N-1}^2}{v_1 \cdots v_{N-1}} \tag{A.34}$$

has $N-1$ zeros when x goes from $-\infty$ to $+\infty$. From this we can conclude that when N is odd the solution is topologically trivial, while when N is even, it belongs to the topologically non-trivial sector of the theory. Equivalently, we have

$$\lim_{x\to+\infty} \Phi_i = \frac{(-1)^{N-1}\delta_{i,N}}{g}\,. \tag{A.35}$$

The fact that we cross the equator $\Phi_N = 0$ $N-1$ times suggests that the solution actually corresponds to the succession of $N-1$ kinks and anti-kinks. The $N-1$ parameters (constants of integration) would be in this interpretation the global center of mass and the relative separations of the individual kinks. To prove that this interpretation is indeed correct, we are going to show that the total mass (A.24) of our solution is the sum of the masses of the individual sine-Gordon solitons of the model (2.55),

$$M = \frac{2}{g^2}\sum_{i=1}^{N-1}\sqrt{v_i}\,. \tag{A.36}$$

Note that M does not depend of the $N-1$ parameters of the solution, which is a direct consequence of the fact that M coincide with the action of the mechanical problem, and that the boundary conditions in the action are independent of those parameters. We have thus obtained a *static* solution consisting of a succession of sine-Gordon kinks and anti-kinks! As pointed out before, we actually have solutions for any number $k \le N-1$ kinks and anti-kinks. Note that all the kinks or anti-kinks appearing in the solution have different masses. A way to prove (A.36) is to choose the constants of integration in some limit that simplify the general solution (kinks and anti-kinks in a limit of very large separation). We will present instead a direct derivation which turns out to be simpler.

Our starting point is

$$M = \int_{-\infty}^{+\infty} dx \sum_{i=1}^{N-1} v_i\Phi_i^2\,, \tag{A.37}$$

which is derived from (A.24), (A.25) and (A.7). We then use the following algebraic identities:

$$\sum_{i=1}^{N-1} v_i\Phi_i^2 = \frac{1}{g^2}\sum_{i=1}^{N-1}(v_i - y_i^2)\,, \tag{A.38}$$

$$\sum_{i=1}^{N-1}(v_i - y_i^2) = \sum_{i=1}^{N-1}\frac{\prod_{j=1}^{N-1}(v_j - y_i^2)}{\prod_{j=1,j\neq i}^{N-1}(y_j^2 - y_i^2)}\,. \tag{A.39}$$

Equation (A.38) is proven by writing (A.15) as

$$\frac{1}{g^2}\,m(z;\mu) = a(z)q(\Phi;z) \tag{A.40}$$

and identifying on both sides the coefficient of z^{N-2}. Equation (A.39) is proven by using the fact that the right hand side can be viewed as a rational function of y_1^2 of degree $d \leq 1$, and that it is actually a linear function of y_1^2 since it has a finite limit when $y_1^2 \to y_i^2$, for all $i = 2, \ldots N-1$. Taking into account the permutation symmetry amongst the y_is, this leaves us with two unknown coefficients which are determined easily by looking at the $y_1^2 \to \infty$ limit up to terms $\mathcal{O}(1/y_1^2)$. By using successively (A.37), (A.38), (A.39) and (A.29), we get

$$M = -\frac{2}{g^2} \sum_{i=1}^{N-1} \int dy_i \,. \tag{A.41}$$

By using the fact that the y_is are monotonic functions of x, and (A.32) and (A.33), we finally obtain the desired result (A.36). From our understanding of the solution, we deduce in particular that the mass density

$$\rho(x) = \sum_{i=1}^{N-1} v_i \Phi_i^2 \tag{A.42}$$

will have $N-1$ maxima, located at the positions of the kinks and anti-kinks, whose values approach the corresponding values for sine-Gordon solitons, v_i/g^2, at large separations.

A.3 Formulas for the case $N = 3$

We give below the explicit formulas used in section 2.3.3 to make the figure 1. By integrating (A.30) for $N = 3$ we get

$$\frac{(\sqrt{v_1} - y_1)(\sqrt{v_1} - y_2)}{(\sqrt{v_1} + y_1)(\sqrt{v_1} + y_2)} = e^{2\sqrt{v_1}(x - x_1)} \,, \qquad \frac{(y_1 - \sqrt{v_2})(\sqrt{v_2} - y_2)}{(\sqrt{v_2} + y_1)(\sqrt{v_2} + y_2)} = e^{2\sqrt{v_2}(x - x_2)} \,. \tag{A.43}$$

When $|x_2 - x_1| \gg 1/\sqrt{v_2}$, we have two well separated kinks centered at $x = x_1$ and $x = x_2$. Solving explicitly for y_1 and y_2 amounts to solving a degree 2 polynomial equation. The relevant root is picked up by using (A.32); the resulting formulas are very complicated and will not be listed here. The spherical coordinates (θ, ϕ) are then given by

$$\theta = \arccos \frac{y_1 y_2}{\sqrt{v_1 v_2}} \,, \qquad \phi = \mathrm{sign}(y_1 - \sqrt{v_2}) \arctan \sqrt{\frac{v_1}{v_2} \frac{(y_1^2 - v_2)(v_2 - y_2^2)}{(v_1 - y_1^2)(v_1 - y_2^2)}} \,. \tag{A.44}$$

B. The functional $s[f]$

In this appendix, we study the functional $s[f]$ defined by

$$s[f] = \frac{i}{2} \operatorname{tr} \ln(-\partial_\mu \partial^\mu - f + i\epsilon) + \frac{1}{4\pi} \ln \frac{\Lambda_0}{\mu} \int d^2x \, f(x) \,, \tag{B.1}$$

or equivalently in the euclidean by

$$s[f] = \frac{1}{2}\,\mathrm{tr}\ln(-\partial^2 + f) - \frac{1}{4\pi}\ln\frac{\Lambda_0}{\mu}\int d^2x\, f(x)\,. \tag{B.2}$$

This functional is ubiquitous in the study of the large-N limit of our model. $s[f]$ depends on an arbitrary renormalization scale μ.

All UV divergent integrals are regulated by a momentum cutoff $|p| \leq \Lambda_0$. In finite quantities, we will take the limit $\Lambda_0 \to \infty$.

B.1 $s[f]$ for constant f

For constant f, $s[f]$ can be written in terms of a potential $v(f)$ as

$$s[f = \mathrm{cst}] = \int d^2x\, v(f)\,. \tag{B.3}$$

When f is a *positive* constant, the euclidean formula (B.2) yields

$$v(f) = \frac{1}{2}\int \frac{d^2p}{(2\pi)^2}\ln(p^2 + f) - \frac{f}{4\pi}\ln\frac{\Lambda_0}{\mu}\,, \tag{B.4}$$

which is, up to f-independent terms,

$$v(f) = -\frac{f}{8\pi}\ln\frac{f}{e\mu^2}\,,\qquad \text{for } f > 0\,. \tag{B.5}$$

The correct formula for $f < 0$ is obtained by using Feynman's $i\epsilon$ prescription (B.1),

$$v(f) = -\frac{f}{8\pi}\left(\ln\frac{-f}{e\mu^2} - i\pi\right),\qquad \text{for } f < 0\,. \tag{B.6}$$

Equation (B.6) is the analytic continuation of (B.5) by going to $\mathrm{Im}\, f < 0$.

B.2 $s[f]$ for arbitrary f

When f is arbitrary, one can expand $s[f]$ in terms of ordinary one-loop Feynman diagrams. Introducing an arbitrary mass scale m and defining $\phi(x) = f(x) - m^2$, we have

$$s[f] = \sum_{n=0}^{\infty} s_n[f; m^2]\,, \tag{B.7}$$

where

$$\begin{aligned} s_n[f; m^2] &= s_n[m^2 + \phi; m^2] \\ &= \frac{1}{n!}\int\prod_{i=1}^{n} d^2x_i s^{(n)}(x_1, \ldots, x_n; m^2)\phi(x_1)\cdots\phi(x_n) \\ &= \frac{1}{n!}\int\prod_{i=1}^{n}\frac{d^2p_i}{(2\pi)^2}(2\pi)^2\delta^{(2)}\left(\sum_{i=1}^{n} p_i\right)\tilde{s}^{(n)}(p_1, \ldots, p_n; m^2)\tilde{\phi}(p_1)\cdots\tilde{\phi}(p_n)\,, \end{aligned} \tag{B.8}$$

with

$$\tilde{\phi}(p) = \int d^2x\, e^{-ipx} \phi(x), \tag{B.9}$$

$$(2\pi)^2 \delta^{(2)} \left(\sum_{i=1}^{n} p_i \right) \tilde{s}^{(n)}(p_1, \dots, p_n; m^2) = \int \prod_{i=1}^{n} d^2x_i e^{i \sum_{j=1}^{n} p_j x_j} s^{(n)}(x_1, \dots, x_n; m^2)\,. \tag{B.10}$$

The euclidean formula for $\tilde{s}^{(n)}$ is

$$\tilde{s}^{(n)}(p_1, \dots p_n; m^2) = (-1)^{n-1} \frac{(n-1)!}{8\pi^2} \int \frac{d^2k}{\prod_{i=0}^{n-1}((k + \sum_{j=1}^{i} p_j)^2 + m^2)} - \frac{\delta_{n,1}}{4\pi} \ln \frac{\Lambda_0}{\mu}\,. \tag{B.11}$$

Explicitly, the linear term is given by

$$\tilde{s}^{(1)} = \frac{1}{8\pi} \ln \frac{\mu^2}{m^2} \tag{B.12}$$

or equivalently

$$s_1[f; m^2] = \frac{1}{8\pi} \ln \frac{\mu^2}{m^2} \int d^2x\, \phi(x)\,. \tag{B.13}$$

In the euclidean regime $p^2 > 0$, the quadratic term is given by

$$\begin{aligned} \tilde{s}^{(2)}(p, -p; m^2) &= \tilde{s}^{(2)}(p^2, m^2) \qquad\qquad (B.14)\\ &= -\frac{1}{4\pi} \frac{1}{p^2\sqrt{1 + 4m^2/p^2}} \ln \frac{\sqrt{1 + 4m^2/p^2} + 1}{\sqrt{1 + 4m^2/p^2} - 1}\,, \qquad \text{for } p^2 > 0\,. \end{aligned}$$

At low momentum we can use the expansion

$$\tilde{s}^{(2)}(p^2, m^2) = -\frac{1}{8\pi m^2} \left(1 - \frac{p^2}{6m^2} + \mathcal{O}(p^4)\right) \tag{B.15}$$

or equivalently

$$s_2[f; m^2] = -\frac{1}{16\pi m^2} \int d^2x\, \phi^2(x) - \frac{1}{96\pi m^4} \int d^2x\, \phi(x) \partial^2 \phi(x) + \mathcal{O}(\partial^4). \tag{B.16}$$

For physical momenta below the pair production threshold, $0 < -p^2 < 4m^2$, the correct analytic continuation is given by

$$\tilde{s}^{(2)}(p^2; m^2) = \frac{1}{2\pi} \frac{1}{p^2\sqrt{-1 - 4m^2/p^2}} \arctan \frac{1}{\sqrt{-1 - 4m^2/p^2}}\,, \quad \text{for } 0 < -p^2 < 4m^2\,. \tag{B.17}$$

Finally, for $-p^2 > 4m^2$, which is the region relevant for two-particle scattering, we have

$$\tilde{s}^{(2)}(p^2, m^2) = -\frac{1}{4\pi} \frac{1}{p^2\sqrt{1 + 4m^2/p^2}} \left(\ln \frac{1 + \sqrt{1 + 4m^2/p^2}}{1 - \sqrt{1 + 4m^2/p^2}} - i\pi \right), \quad \text{for } -p^2 > 4m^2\,. \tag{B.18}$$

In the latter case, it is convenient to introduce the rapidity parameter $\theta > 0$ such that

$$-p^2 = 4m^2 \cosh^2(\theta/2) \tag{B.19}$$

and in terms of which

$$\tilde{s}^{(2)}(p^2, m^2) = \frac{1}{8\pi m^2} \frac{\theta - i\pi}{\sinh\theta}\,. \tag{B.20}$$

Let us quote to finish the asymptotic behaviour of $\tilde{s}^{(2)}$ at large euclidean p^2, which is useful to investigate the UV properties of the $1/N$ expansion,

$$\tilde{s}^{(2)}(p^2; m^2) = -\frac{1}{4\pi p^2}\left(\ln\frac{p^2}{m^2} + \frac{2m^2}{p^2}\left(1 - \ln\frac{p^2}{m^2}\right) + \mathcal{O}\left(\frac{m^4}{p^4}\ln\frac{p^2}{m^2}\right)\right). \tag{B.21}$$

C. The 1/N corrections to the critical hypersurface

We discuss in this appendix the $1/N$ corrections to the equation

$$\frac{1}{N}\sum_{i=1}^{N-1} \ln\frac{v_i}{\Lambda^2} = 0 \tag{C.1}$$

for the hypersurface of singularities $\mathcal{H}_{\mathrm{q}}$ derived in section 3.4. The main goal of this calculation is to understand the rôle of the infrared divergences, which generically plague the $1/N$ expansion near $\mathcal{H}_{\mathrm{q}}$. As the very existence of $\mathcal{H}_{\mathrm{q}}$ was derived within the large-N expansion, the consistency of the analysis requires that the equation for $\mathcal{H}_{\mathrm{q}}$ itself gets only small corrections when we go beyond the leading approximation. We show below that the first corrections are of order $N^{-1}\ln N$, larger than the naive N^{-1}, but nevertheless smaller that the leading N^0 term. The calculation also illustrates nicely how renormalization works within the $1/N$ expansion, consistently with the discussion of section 3.1.

C.1 The effective potential

The starting point is the effective action

$$S_{\mathrm{eff}}[\alpha, \Phi_N] = (N-1)\, s_{\mathrm{eff}}[\alpha, \varphi]\,, \tag{C.2}$$

where

$$s_{\mathrm{eff}}[\alpha, \varphi] = \int d^2x \left[\frac{1}{2}\,\partial_\alpha\varphi\partial_\alpha\varphi + \frac{\alpha - h_{N,0}}{2}\,\varphi^2 - \frac{\alpha}{4\pi}\ln\frac{\mu}{\lambda}\right] + \frac{1}{N-1}\sum_{i=1}^{N-1} s[\alpha - h_{i,0}]\,. \tag{C.3}$$

We were careful in keeping subleading terms in (C.3), since we are now willing to take them into account. We have defined

$$\varphi = \frac{\Phi_N}{\sqrt{N-1}} \tag{C.4}$$

and λ by the equation

$$\frac{1}{g_0^2} = \frac{N-1}{2\pi} \ln \frac{\Lambda_0}{\lambda} \,. \tag{C.5}$$

To order N^0, we have $\lambda = \Lambda$, but to order N^{-1}, λ must pick an infinite, Λ_0-dependent contribution, coming from the $1/N$ corrections to the β function. Equations (3.4), (3.7) and (3.8) of section 3.1 gives the form of this contribution, as well as the renormalization of $h_{i,0} = h_i Z_{(s,2)}/Z$:

$$\lambda^2 = \Lambda^2 \left(1 + \frac{1}{N} \ln \frac{\Lambda_0^2}{\Lambda^2} - \frac{2}{N} \ln\ln \frac{\Lambda_0}{\Lambda} + \frac{c_1}{N} \right), \tag{C.6}$$

$$h_{i,0} = h_i \left(1 + \frac{2}{N} \ln\ln \frac{\Lambda_0}{\Lambda} + \frac{c_2}{N} \right), \tag{C.7}$$

where c_1 and c_2 are finite constants to be determined by some renormalization conditions.

The $1/N$ corrections correspond to the one-loop diagrams derived from the non-local action s_{eff}. In particular, the effective potential v_{eff} to order N^{-1} is such that

$$\begin{aligned} \int d^2x v_{\mathrm{eff}}(\alpha,\varphi) = \int d^2x \Bigg(& \frac{1}{2}(\alpha - h_{N,0})\,\varphi^2 - \frac{\alpha}{4\pi} \ln \frac{\mu}{\lambda} - \\ & - \frac{1}{8\pi(N-1)} \sum_{i=1}^{N-1} (\alpha - h_{i,0}) \ln \frac{\alpha - h_{i,0}}{e\mu^2} \Bigg) + \frac{1}{2N} \operatorname{tr} \ln \mathcal{J}[\alpha,\varphi] \,. \end{aligned} \tag{C.8}$$

$\mathcal{J}$ is the functional hessian of s_{eff}, for which

$$\frac{1}{2} \operatorname{tr} \ln \mathcal{J}[\alpha,\varphi] = \frac{1}{8\pi^2} \int d^2x \int d^2p \, \operatorname{tr} \ln \begin{pmatrix} p^2 + \alpha - h_{N,0} & \varphi \\ \varphi & \tilde{S}(p^2;\alpha,h_{1,0},\ldots,h_{N-1,0}) \end{pmatrix} \tag{C.9}$$

where we have defined

$$\tilde{S}(p^2;\alpha,h_1,\ldots,h_{N-1}) = \frac{1}{N-1} \sum_{i=1}^{N-1} \tilde{s}^{(2)}(p^2;\alpha - h_i) \,. \tag{C.10}$$

From (C.8), (C.9) we get the basic formulas, valid up to terms of order $1/N^2$,

$$\frac{\partial v_{\mathrm{eff}}}{\partial \varphi} = (\alpha - h_{N,0})\varphi - \frac{1}{N} \frac{\varphi}{4\pi^2} \int d^2p \frac{1}{(p^2 + \alpha - h_N)\tilde{S}(p^2;\alpha,h_1,\ldots,h_{N-1}) - \varphi^2}, \tag{C.11}$$

$$\frac{\partial v_{\mathrm{eff}}}{\partial \alpha} = \frac{\varphi^2}{2} - \frac{1}{8\pi(N-1)} \sum_{i=1}^{N-1} \ln \frac{\alpha - h_{i,0}}{\lambda^2} + \frac{1}{8\pi^2 N} \int d^2p \frac{\tilde{S} + (p^2 + \alpha - h_N)\partial_\alpha \tilde{S}}{(p^2 + \alpha - h_N)\tilde{S} - \varphi^2} \,. \tag{C.12}$$

C.2 The equation for $\mathcal{H}_{\mathrm{q}}$

The naive strategy to get the equation for $\mathcal{H}_{\mathrm{q}}$ up to order $1/N$ would be to solve

$$\frac{\partial v_{\mathrm{eff}}}{\partial \varphi} = \frac{\partial v_{\mathrm{eff}}}{\partial \alpha} = 0 \tag{C.13}$$

by substituting α and φ in the terms of order $1/N$ by their respective values at order N^0, that is $\alpha = h_N$ and $\varphi = 0$ (see section 3). However, this yields IR divergent integrals, and thus is not correct. Instead, we will identify the terms responsible for the IR divergences, and set $\alpha = h_N$ and $\varphi = 0$ in the other terms only. The IR divergences can be analysed easily by using

$$\tilde{S}(p^2; \alpha = h_N, h_1, \ldots, h_{N-1}) = -\frac{1}{8\pi V} + \mathcal{O}(p^2)\,, \tag{C.14}$$

$$\partial_\alpha \tilde{S}(p^2; \alpha = h_N, h_1, \ldots, h_{N-1}) = \frac{1}{8\pi(N-1)} \sum_{i=1}^{N-1} \frac{1}{v_i^2} + \mathcal{O}(p^2)\,, \tag{C.15}$$

where V is defined in (3.49). The term containing $\partial_\alpha \tilde{S}$ in $\partial v_{\text{eff}}/\partial\alpha$ is not IR divergent, while the other term can be written

$$\begin{aligned} &\int \frac{d^2p}{8\pi^2} \left[\frac{\tilde{S}(p^2; \alpha, h_1, \ldots, h_{N-1})}{(p^2 + \alpha - h_N)\tilde{S}(p^2; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2} - \right. \\ &\qquad \left. - \frac{\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1})}{(p^2 + \alpha - h_N)\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2} \right] + \\ &+ \int \frac{d^2p}{8\pi^2} \frac{\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1})}{(p^2 + \alpha - h_N)\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2}\,. \end{aligned} \tag{C.16}$$

The first term is now IR convergent, and goes actually to zero when $\alpha \to h_N$ and $\varphi \to 0$, while the second term is easily calculated. The condition $\partial v_{\text{eff}}/\partial\alpha = 0$ is thus equivalent to

$$0 = \frac{\varphi^2}{2} - \frac{1}{8\pi(N-1)} \sum_{i=1}^{N-1} \ln \frac{\alpha - h_{i,0}}{\lambda^2} + \frac{1}{8\pi N} \ln \frac{\Lambda_0^2}{\alpha - h_N + 8\pi V \varphi^2} + \frac{I_\alpha(v_1, \ldots, v_{N-1})}{8\pi N}\,, \tag{C.17}$$

where

$$I_\alpha(v_1, \ldots, v_{N-1}) = \frac{1}{\pi} \int d^2p\, \frac{\partial_\alpha \tilde{S}(p^2; \alpha = h_N, h_1, \ldots, h_{N-1})}{\tilde{S}(p^2; \alpha = h_N, h_1, \ldots, h_{N-1})}\,. \tag{C.18}$$

We treat $\partial v_{\text{eff}}/\partial\varphi$ is a similar way. We write the $1/N$ contribution as

$$\begin{aligned} &-\int \frac{d^2p}{4\pi^2} \left[\frac{1}{(p^2 + \alpha - h_N)\tilde{S}(p^2; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2} - \right. \\ &\qquad \left. - \frac{1}{(p^2 + \alpha - h_N)\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2} \right] - \\ &- \int \frac{d^2p}{4\pi^2} \frac{1}{(p^2 + \alpha - h_N)\tilde{S}(p^2 = 0; \alpha, h_1, \ldots, h_{N-1}) - \varphi^2}\,, \end{aligned} \tag{C.19}$$

check that the first term is now IR convergent while the second term is easy to compute, and conclude that the condition $\partial v_{\text{eff}}/\partial\varphi = 0$ is equivalent to

$$0 = \alpha - h_{N,0} + \frac{V}{N} \left(2 \ln \frac{\Lambda_0^2}{\alpha - h_N + 8\pi V \varphi^2} - I_\varphi(v_1, \ldots, v_{N-1}) \right), \tag{C.20}$$

where

$$I_\varphi(v_1,\ldots,v_{N-1}) = \frac{1}{4\pi^2 V}\int\frac{d^2p}{p^2}\left(\frac{1}{\tilde S(p^2;\alpha=h_N,h_1,\ldots,h_{N-1})} - \right.$$
$$\left. - \frac{1}{\tilde S(p^2=0;\alpha=h_N,h_1,\ldots,h_{N-1})}\right). \quad \text{(C.21)}$$

We have chosen to approach $\mathcal{H}_\mathrm{q}$ from weak coupling, and thus we have discarded the solution $\varphi=0$ to the equation $\partial v_\mathrm{eff}/\partial\varphi=0$. $\mathcal{H}_\mathrm{q}$ is precisely the locus in parameter space where $\varphi=0$ becomes the true minimum of the effective potential for φ, see section 3.4.

The equations (C.17) and (C.20) are our new starting point. They are badly UV divergent, as discussed further in the next subsection, and the renormalizations (C.6) and (C.7) are not enough to make them finite. This is not surprising: as explained in section 2.1, formulas containing the field α need an infinite number of counterterms. To obtain a finite, physically sensible, formula, we must eliminate α using (C.17), and then check that the result is finite. To make the presentation as clear as possible, let us introduce

$$\tilde\alpha = \alpha - h_{N,0}\,. \quad \text{(C.22)}$$

We know from section 3.4 that $\tilde\alpha=0$ on $\mathcal{H}_\mathrm{q}$ to order N^0 (this is a direct consequence of (C.20). We will thus always set $\tilde\alpha=0$ in $1/N$ corrections, as long as this does not lead to an IR divergence. Moreover, to leading order, (C.17) shows that $\tilde\alpha = \tilde\alpha_{N^0}(v_{1,0},\ldots,v_{N-1,0};\lambda)$ with

$$\sum_{i=1}^{N-1}\ln\frac{\tilde\alpha_{N^0}+v_{i,0}}{\lambda^2} = 0\,. \quad \text{(C.23)}$$

In order to solve (C.17), we then substitute $\tilde\alpha = \tilde\alpha_{N^0}+\mathcal{O}(1/N)$ and use (C.23) to obtain

$$\tilde\alpha = \tilde\alpha_{N^0} + \frac{V}{N}\left(I_\alpha - \ln\frac{\tilde\alpha_{N^0}}{\Lambda_0^2}\right). \quad \text{(C.24)}$$

Finally, using (C.20), we obtain the α-independent equation

$$\tilde\alpha_{N^0}(v_{1,0},\ldots,v_{N-1,0};\lambda) = \frac{V}{N}\left(I_\varphi - I_\alpha + 3\ln\frac{\tilde\alpha_{N^0}(v_{1,0},\ldots,v_{N-1,0};\lambda)}{\Lambda_0^2}\right). \quad \text{(C.25)}$$

This is the equation for $\mathcal{H}_\mathrm{q}$ at order $1/N$. A highly non-trivial check is that it is actually Λ_0-independent. A crucial ingredient for this to be possible is that only the combination $I_\varphi - I_\alpha$ appears, so that most of the UV divergences in I_φ and I_α cancel each other (see (C.41) for an explicit formula). Moreover, by using (C.6) and (C.7), it is straightforward to show that up to terms of order $1/N^2$,

$$\tilde\alpha_{N^0}(v_{1,0},\ldots,v_{N-1,0};\lambda) = \tilde\alpha_{N^0}(v_1,\ldots,v_{N-1};\Lambda) + \frac{V}{N}\left(\ln\frac{\Lambda_0^2}{\Lambda^2} - 4\ln\ln\frac{\Lambda_0}{\Lambda} + c_1 - c_2\right). \quad \text{(C.26)}$$

This turns out to be exactly what is required to cancel the divergences in (C.25). In terms of finite quantities only, the equation for $\mathcal{H}_{\rm q}$ can then be written

$$\tilde{\alpha}_{N^0} = \frac{V}{N}\Big(I - 4\ln 2 + 3\ln\frac{\tilde{\alpha}_{N^0}}{\Lambda^2} + c_2 - c_1\Big). \tag{C.27}$$

The finite integral $I(v_1, \dots v_{N-1})$ is defined by (C.42). It remains to enforce a renormalization condition in order to determine $c_2 - c_1$. The study of this seemingly secondary point is actually important, because the correction in (C.27) are small only if we can choose $c_2 - c_1$ to be much smaller than N. In particular, the naive guess that $c_2 - c_1$ can be an arbitrary finite constant independent of N is not correct. Our renormalization condition will be written at $v_1 = \cdots = v_{N-1} = v$, values for which $I = 0$ and $\tilde{\alpha}_{N^0} = -v + \Lambda^2$. We impose that

$$v = \Lambda^2\Big(1 - \frac{\Delta}{N}\Big), \tag{C.28}$$

Δ being an N-independent, strictly positive, constant. It is impossible to choose $\Delta = 0$, due to the IR divergences, very much like it is impossible to use renormalization conditions at zero momentum in massless theories. Equation (C.28) implies

$$c_2 - c_1 = 3\ln N + \Delta - 3\ln\Delta + 4\ln 2\,, \tag{C.29}$$

and the final equation for $\mathcal{H}_{\rm q}$ is

$$\tilde{\alpha}_{N^0} = \frac{V}{N}\left(3\ln N + I + \Delta + 3\ln\frac{\tilde{\alpha}_{N^0}}{\Delta\Lambda^2}\right). \tag{C.30}$$

We recall that $\tilde{\alpha}_{N^0}$ is the unique solution to

$$\sum_{i=1}^{N-1}\ln\frac{\tilde{\alpha}_{N^0} + v_i}{\Lambda^2} = 0\,, \tag{C.31}$$

I is defined by (C.42), and $\Delta > 0$ is a renormalization constant independent of N. We thus see that the IR instability is responsible for a correction of order $(\ln N)/N$, larger than the expected $1/N$ but much smaller than the leading N^0 term, as required.

C.3 Formulas for I_φ, I_α and $I_\varphi - I_\alpha$

In view of (B.14), it is natural to introduce the variables

$$x_i(p^2) = \frac{\sqrt{1 + 4v_i/p^2} + 1}{\sqrt{1 + 4v_i/p^2} - 1}\,. \tag{C.32}$$

This set of variables is particularly well suited to study the integrals I_φ and I_α when the v_is are equal or nearly equal. The UV behaviour is also easily studied in this representation. One has

$$\frac{x_j - 1}{x_j + 1} = \frac{x_i - 1}{x_i + 1} g_{ji}(x_i)\,, \tag{C.33}$$

with

$$g_{ji}(x_i) = \frac{1}{\sqrt{1 + \frac{4(v_j - v_i)}{v_i} \frac{x_i}{(x_i+1)^2}}}\,, \tag{C.34}$$

so that

$$g_{ij}(x_k) g_{ji}(x_i) = g_{jk}(x_k). \tag{C.35}$$

We have the explicit formulas

$$I_\alpha = -\sum_{i=1}^{N-1} \int_1^{x_i(\Lambda_0^2)} dx_i \frac{\frac{(x_i-1)^3}{x_i^2(x_i+1)} + \frac{2}{x_i}\left(\frac{x_i-1}{x_i+1}\right)^2 \ln x_i}{\sum_{j=1}^{N-1} \frac{x_j(x_i)-1}{x_j(x_i)+1} \ln x_j(x_i)}\,, \tag{C.36}$$

$$I_\varphi = -\sum_{i=1}^{N-1} \int_1^{x_i(\Lambda_0^2)} dx_i \left[\frac{\frac{x_i^2-1}{x_i^2}}{\sum_{j=1}^{N-1} \frac{x_j(x_i)-1}{x_j(x_i)+1} \ln x_j(x_i)} - \frac{2}{N-1} \frac{x_i+1}{x_i(x_i-1)} \right]. \tag{C.37}$$

I_φ and I_α have separately untamable UV divergencies. These divergencies can be studied for example for $v_1 = \cdots = v_{N-1}$. In that case one has

$$I_\varphi(v_1 = \cdots = v_{N-1} = v) = 2 \ln \frac{\Lambda_0^2}{v} - 2 \ln \ln \frac{\Lambda_0^2}{v} + 2\gamma - \mathrm{li}\, x(\Lambda_0^2), \tag{C.38}$$

with

$$x(\Lambda_0^2) = \frac{\Lambda_0^2}{v} \left(1 + \frac{v}{\Lambda_0^2} + \mathcal{O}(\frac{v^2}{\Lambda_0^4}) \right) \tag{C.39}$$

and li is the logarithmic integral, whose expansion is

$$\mathrm{li}\, x = \mathrm{P} \int_0^x \frac{dx}{\ln x} = \gamma + \ln \ln x + \sum_{k=1}^{\infty} \frac{(\ln x)^k}{k!\, k} \qquad \text{for } x > 1\,. \tag{C.40}$$

On the other hand, $I_\varphi - I_\alpha$ has simple UV divergencies that are renormalizable with a finite number of counterterms. The general formula is

$$I_\varphi - I_\alpha = I - 4 \ln \ln \frac{\Lambda_0}{\Lambda} + 4 \ln \frac{\Lambda_0^2}{\Lambda^2} - 4 \ln 2 + \frac{4}{N-1} \sum_{i=1}^{N-1} \ln \frac{\Lambda^2}{v_i}\,, \tag{C.41}$$

where I is a finite integral (which vanishes when $v_1 = \cdots = v_{N-1}$),

$$\begin{aligned} I(v_1, \ldots, v_{N-1}) = \sum_{i=1}^{N-1} \int_1^{\infty} dx_i \Bigg[& \frac{4}{N-1} \frac{1}{x_i \ln x_i} - \frac{2}{N-1} \frac{x_i - 1}{x_i(x_i+1)} - \\ & - \frac{\frac{4(x_i-1)}{x_i(x_i+1)}(1 - \frac{1}{2} \frac{x_i-1}{x_i+1} \ln x_i)}{\sum_{j=1}^{N-1} \frac{x_j(x_i)-1}{x_j(x_i)+1} \ln x_j(x_i)} \Bigg]. \end{aligned} \tag{C.42}$$

On the critical hypersurface $\mathcal{H}_\mathrm{q}$, the last term in (C.41) is of order $1/N$ and can be neglected.

References

[1] F. Ferrari, *Non-supersymmetric cousins of supersymmetric gauge theories: quantum space of parameters and double scaling limits*, *Phys. Lett.* **B 496** (2000) 212 [hep-th/0003142].

[2] A. Bilal, *Duality in $N = 2$ SUSY* SU(2) *Yang-Mills theory: a pedagogical introduction to the work of Seiberg and Witten*, hep-th/9601007;
L. Álvarez-Gaumé and S.F. Hassan, *Introduction to S-duality in $N = 2$ supersymmetric gauge theories: a pedagogical review of the work of Seiberg and Witten*, *Fortsch. Phys.* **45** (1997) 159 [hep-th/9701069];
K. Intriligator and N. Seiberg, *Lectures on supersymmetric gauge theories and electric-magnetic dualities*, *Nucl. Phys.* **55** *(Proc. Suppl.)* (1997) 157.

[3] J. Polchinski, *Tasi lectures on D-branes*, hep-th/9611050;
A. Giveon and D. Kutasov, *Brane dynamics and gauge theory*, *Rev. Mod. Phys.* **71** (1999) 983 [hep-th/9802067];
C.V. Johnson, *D-brane primer*, hep-th/0007170.

[4] I.R. Klebanov, *Tasi lectures: introduction to the AdS/CFT correspondence*, hep-th/0009139;
O. Aharony, S.S. Gubser, J. Maldacena, H. Ooguri and Y. Oz, *Large-N field theories, string theory and gravity*, *Phys. Rep.* **323** (2000) 183 [hep-th/9905111].

[5] L. Álvarez-Gaumé, J. Distler, C. Kounnas, and M. Mariño, *Softly broken $N = 2$ QCD*, *Int. J. Mod. Phys.* **A 11** (1996) 4745 [hep-th/9604004];
L. Álvarez-Gaumé and M. Mariño, *More on softly broken $N = 2$ QCD*, *Int. J. Mod. Phys.* **A 12** (1997) 975 [hep-th/9606191];
L. Álvarez-Gaumé, M. Mariño, and F. Zamora, *Softly broken $N = 2$ QCD with massive quark hypermultiplets, 1*, *Int. J. Mod. Phys.* **A 13** (1998) 403 [hep-th/9703072]; *Softly broken $N = 2$ QCD with massive quark hypermultiplets, 2*, *Int. J. Mod. Phys.* **A 13** (1998) 1847 [hep-th/9707017].

[6] J. Polchinski and M.J. Strassler, *The string dual of a confining four-dimensional gauge theory*, hep-th/0003136.

[7] P.C. Argyres and M.R. Douglas, *New phenomena in* SU(3) *supersymmetric gauge theory*, *Nucl. Phys.* **B 448** (1995) 93 [hep-th/9505062];
P.C. Argyres, M.R. Plesser, N. Seiberg and E. Witten, *New $N = 2$ superconformal field theories in four dimensions*, *Nucl. Phys.* **B 461** (1996) 71 [hep-th/9511154].

[8] G. 't Hooft, *A planar diagram theory for strong interactions*, *Nucl. Phys.* **B 72** (1974) 461.

[9] M.R. Douglas and S.H. Shenker, *Dynamics of* SU(N) *supersymmetric gauge theory*, *Nucl. Phys.* **B 447** (1995) 271 [hep-th/9503163].

JHEP06(2001)057

[10] A. Klemm, W. Lerche, P. Mayr, C. Vafa and N.P. Warner, *Self-dual strings and $N = 2$ supersymmetric field theory*, *Nucl. Phys.* **B 477** (1996) 746 [hep-th/9604034];
S. Katz, A. Klemm and C. Vafa, *Geometric engineering of quantum field theories*, *Nucl. Phys.* **B 497** (1997) 173 [hep-th/9609239].

[11] J. Maldacena, *The large-N limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231 [hep-th/9711200];
S.S. Gubser, I.R. Klebanov and A.M. Polyakov, *Gauge theory correlators from non-critical string theory*, *Phys. Lett.* **B 428** (1998) 105 [hep-th/9802109];
E. Witten, *Anti-de sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [hep-th/9802150].

[12] F. David, *Planar diagrams, two-dimensional lattice gravity and surface models*, *Nucl. Phys.* **B 257** (1985) 45;
V.A. Kazakov, *Phys. Lett.* **B 150** (1985) 28;
J. Ambjörn, B. Durhuus and J. Fröhlich, *Diseases of triangulated random surface models and possible cures*, *Nucl. Phys.* **B 257** (1985) 433.

[13] É. Brézin and V.A. Kazakov, *Exactly solvable field theories of closed strings*, *Phys. Lett.* **B 236** (1990) 144;
M.R. Douglas and S. Shenker, *Strings in less than one dimension*, *Nucl. Phys.* **B 355** (1990) 635;
D.J. Gross and A.A. Migdal, *Nonperturbative two-dimensional quantum gravity*, *Phys. Rev. Lett.* **64** (1990) 127.

[14] F. Ferrari, *The large-N limit of $\mathcal{N} = 2$ super Yang-Mills, fractional instantons, and infrared divergences*, hep-th/0106192; *Exact amplitudes in four dimensional non-critical string theories,* PUPT-1993, LPTENS-01/10, in preparation.

[15] A.M. Polyakov, *Interaction of goldstone particles in two-dimensions. applications to ferromagnets and massive Yang-Mills fields*, *Phys. Lett.* **B 59** (1975) 79.

[16] A. D'Adda, M. Luscher and P.D. Vecchia, *A $1/N$ expandable series of nonlinear sigma models with instantons*, *Nucl. Phys.* **B 146** (1978) 63.

[17] N. Seiberg and E. Witten, *Electric-magnetic duality, monopole condensation and confinement in $N = 2$ supersymmetric Yang-Mills theory*, *Nucl. Phys.* **B 426** (1994) 19 [hep-th/9407087], erratum *ibid.* **B430** (1994) 485;
Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD, *Nucl. Phys.* **B 431** (1994) 484 [hep-th/9408099].

[18] E. Witten, *Some comments on string dynamics*, in *Future perspectives in string theory*, Strings '95, I. Bars et al. eds. World Scientific Publishing, Singapore 1996.

[19] F. Ferrari and A. Bilal, *The strong-coupling spectrum of the Seiberg-Witten theory*, *Nucl. Phys.* **B 469** (1996) 387 [hep-th/9602082];
A. Bilal and F. Ferrari, *Curves of marginal stability and weak and strong-coupling*

JHEP06(2001)057

BPS spectra in $N = 2$ supersymmetric QCD, *Nucl. Phys.* **B 480** (1996) 589 [hep-th/9605101];
F. Ferrari, *The dyon spectra of finite gauge theories*, *Nucl. Phys.* **B 501** (1997) 53 [hep-th/9702166];
A. Bilal and F. Ferrari, *The BPS spectra and superconformal points in massive $N = 2$ supersymmetric QCD*, *Nucl. Phys.* **B 516** (1998) 175 [hep-th/9706145].

[20] N. Seiberg, *Exact results on the space of vacua of four-dimensional susy gauge theories*, *Phys. Rev.* **D 49** (1994) 6857 [hep-th/9402044];
N. Seiberg, *Electric-magnetic duality in supersymmetric nonabelian gauge theories*, *Nucl. Phys.* **B 435** (1995) 129 [hep-th/9411149].

[21] A. Hanany and K. Hori, *Branes and $N = 2$ theories in two dimensions*, *Nucl. Phys.* **B 513** (1998) 119 [hep-th/9707192].

[22] N. Dorey, *The BPS spectra of two-dimensional supersymmetric gauge theories with twisted mass terms*, *J. High Energy Phys.* **11** (1998) 005 [hep-th/9806056];
N. Dorey, T.J. Hollowood and D. Tong, *The BPS spectra of gauge theories in two and four dimensions*, *J. High Energy Phys.* **05** (1999) 006 [hep-th/9902134].

[23] L. Álvarez-Gaumé and D.Z. Freedman, *Potentials for the supersymmetric nonlinear sigma model*, *Comm. Math. Phys.* **91** (1983) 87.

[24] E. Witten, *Solutions of four-dimensional field theories via M-theory*, *Nucl. Phys.* **B 500** (1997) 3 [hep-th/9703166].

[25] N. Seiberg, *Supersymmetry and nonperturbative beta functions*, *Phys. Lett.* **B 206** (1988) 75.

[26] L. Alvarez-Gaume and D.Z. Freedman, *Kähler geometry and the renormalization of supersymmetric sigma models*, *Phys. Rev.* **D 22** (1980) 846; *Geometrical structure and ultraviolet finiteness in the supersymmetric sigma model*, *Comm. Math. Phys.* **80** (1981) 443.

[27] E. Witten, *Topological quantum field theory*, *Comm. Math. Phys.* **117** (1988) 353.

[28] E. Witten, *Topological sigma models*, *Comm. Math. Phys.* **118** (1988) 411.

[29] S. Cecotti and C. Vafa, *Topological antitopological fusion*, *Nucl. Phys.* **B 367** (1991) 359; *Comm. Math. Phys.* **158** (1993) 569.

[30] W. Lerche, *Fayet-Iliopoulos potentials from four-folds*, *J. High Energy Phys.* **11** (1997) 004 [hep-th/9709146].

[31] É. Brézin, J. Zinn-Justin and J.C. Le Guillou, *Renormalization of the nonlinear sigma model in $(2 + \epsilon)$ dimension*, *Phys. Rev.* **D 14** (1976) 2615.

[32] J. Zinn-Justin, *Quantum field theory and critical phenomena*, Oxford Science Publications.

[33] F. Ferrari, *The supersymmetric non-linear sigma model with mass terms*, in preparation.

[34] É. Brézin and J. Zinn-Justin, *Renormalization of the nonlinear sigma model in $2+\epsilon$ dimensions: application to the Heisenberg ferromagnets*, *Phys. Rev. Lett.* **36** (1976) 691; *Spontaneous breakdown of continuous symmetries near two dimensions*, *Phys. Rev.* **B 14** (1976) 3110.

[35] F.D.M. Haldane, *Continuum dynamics of the 1-d Heisenberg antiferromagnetic identification with the $O(3)$ nonlinear sigma model*, *Phys. Lett.* **A 93** (1983) 464; *Nonlinear field theory of large spin Heisenberg antiferromagnets: semiclassically quantized solitons of the one-dimensional easy axis Néel state*, *Phys. Rev. Lett.* **50** (1983) 1153.

[36] C. Neumann, *De problemate quodam mechanics, quod ad primam integralium ultraellipticorum classem revocatur*, *Journal für die Reine und Angewandte Mathematik* **56** (1859) 46.

[37] A.M. Polyakov, *Hidden symmetry of the two-dimensional chiral fields*, *Phys. Lett.* **B 72** (1977) 224;
A.B. Zamolodchikov and Al.B. Zamolodchikov, *Ann. Phys. (NY)* **120** (1979) 253.

[38] S. Coleman, *1/N* in *Aspects of symmetry*, Cambridge University Press 1985;
J. Zinn-Justin, *Vector models in the large-N limit: a few applications*, lectures at the 11th Taiwan Spring School on Particles and Fields 1997, hep-th/9810198.

[39] S. Cecotti and C. Vafa, *Exact results for supersymmetric sigma models*, *Phys. Rev. Lett.* **68** (1992) 903 [hep-th/9111016].

[40] B.R. Greene, *String theory on Calabi-Yau manifolds*, Les Houches '95 and TASI '96 lectures, hep-th/9702155.

[41] A.B. Zamolodchikov, *Conformal symmetry and multicritical points in two-dimensional quantum field theory*, *Sov. J. Nucl. Phys.* **44** (1987) 529.

[42] R. Rajaraman, *Solitons and instantons*, North Holland 1989.

[43] K. Li, *Some Landau-Ginzburg models from conformal field theory*, *Int. J. Mod. Phys.* **A 5** (1990) 2343.

[44] H. Saleur, *Partition function of the two-dimensional Ashkin-Teller model on the critical line*, *J. Phys.* **A 20** (1987) L1127.

[45] P. Ginsparg, *Applied conformal field theory*, Les Houches '88 lectures, in *Champs, Cordes, et Phénomènes critiques*, É. Brézin and J. Zinn-Justin eds., Elsevier Science Publishers 1989.

[46] J.M. Kosterlitz and D.J. Thouless, *Ordering, metastability and phase transitions in two dimensional systems*, *J. Phys.* **C 6** (1973) 1181;
J.M. Kosterlitz, *The critical properties of the two dimensional XY model*, *J. Phys.* **C7** (1974) 1046.

[47] K. Hori and C. Vafa, *Mirror symmetry*, `hep-th/0002222`.

[48] A.M. Polyakov, *Gauge fields and strings*, Harwood Academic Publishers 1987.

[49] J.V. José, L.P. Kadanoff, and S. Kirkpatrick, *Renormalization, vortices, and symmetry breaking perturbations on the two-dimensional planar model*, *Phys. Rev.* **B 16** (1977) 1217;
L.P. Kadanoff, *Multicritical behaviour at the Kosterlitz-Thouless critical point*, *Ann. Phys. (NY)* **120** (1979) 39;
L.P. Kadanoff and A.C. Brown, *Correlation functions on the critical lines of the Baxter and Ashkin-Teller models*, *Ann. Phys. (NY)* **121** (1979) 318.

[50] D. Boyanovsky, *Field theoretical renormalization and fixed point structure of a generalized Coulomb gas*, *J. Phys.* **A22** (1989) 2601;
D. Boyanovsky and R. Holman, *The Coulomb gas, vertex operators and the c theorem*, *Nucl. Phys.* **B 332** (1990) 641.

[51] J.L. Cardy, *Duality and the theta parameter in abelian lattice models*, *Nucl. Phys.* **B 205** (1982) 17.

[52] M.C. Ogilvie, *Spin waves, vortices, fermions, and duality in the Ising and Baxter models*, *Ann. Phys. (NY)* **136** (1981) 273.

[53] R. Floreanini and R. Jackiw, *Selfdual fields as charge density solitons*, *Phys. Rev. Lett.* **59** (1987) 1873;
A.A. Tseytlin, *Duality symmetric formulation of string world sheet dynamics*, *Phys. Lett.* **B 242** (1990) 163; *Duality symmetric closed string theory and interacting chiral scalars*, *Nucl. Phys.* **B 350** (1991) 395.

[54] J.H. Schwarz and A. Sen, *Duality symmetric actions*, *Nucl. Phys.* **B 411** (1994) 35 [`hep-th/9304154`].

[55] B. Nienhuis, *Coulomb gas formulation of two-dimensional phase transitions*, in *Phase transitions and critical phenomena*, vol. 11, C. Domb and J.L. Lebowitz eds., Academic Press 1987.

[56] S. Coleman, *Quantum sine-gordon equation as the massive thirring model*, *Phys. Rev.* **D 11** (1975) 2088.

[57] K. Johnson, *Solution of the equations for the Green's functions of a two dimensional relativistic field theory*, *Nuovo Cim.* **20** (1961) 2085;
C.M. Sommerfield, *On the definition of currents and the action principle in field theories of one spacial dimension*, *Ann. Phys. (NY)* **26** (1963) 1.

[58] J.L. Cardy, *Universal properties of* U(1) *gauge theories*, *Nucl. Phys.* **B 170** (1980) 369;
J.L. Cardy and E. Rabinovici, *Phase structure of Z_p models in the presence of a theta parameter*, *Nucl. Phys.* **B 205** (1982) 1;
A. Shapere and F. Wilczek, *Selfdual models with theta terms*, *Nucl. Phys.* **B 320** (1989) 669.

[59] T. Eguchi, K. Hori, K. Ito and S.-K. Yang, *Study of $N = 2$ superconformal field theories in* 4 *dimensions*, *Nucl. Phys.* **B 471** (1996) 430 [hep-th/9603002];
T. Eguchi and K. Hori, *$N = 2$ superconformal field theories in 4 dimensions and ADE classification*, talk given at *Mathematical beauty of physics* in memory of Claude Itzykson, Saclay, hep-th/9607125.

[60] A. Cappelli, C. Itzykson, and J.B. Zuber, *The ADE classification of minimal and $A_1^{(1)}$ conformal invariant theories*, *Comm. Math. Phys.* **113** (1987) 1.

[61] J. Ambjörn, B. Durhuus and T. Jonsson, *Summing over all genera for $d > 1$: a toy model*, *Phys. Lett.* **B 244** (1990) 403;
S. Nishigaki and T. Yoneya, *A nonperturbative theory of randomly branching chains*, *Nucl. Phys.* **B 348** (1991) 787;
P.D. Vecchia, M. Kato and N. Ohta, *Double scaling limit in o(n) vector models*, *Nucl. Phys.* **B 357** (1991) 495.

[62] K.K. Uhlenbeck, *Equivariant harmonic maps into spheres*, in *Harmonic maps*, Proceedings, New Orleans 1980, Lecture Notes in Mathematics, Springer-Verlag 1982, p. 146;
J. Moser, *Integrable hamiltonian systems and spectral theory*, Scuola Normale Superiore, Pisa 1981.

JHEP06(2001)057