Phase Transitions in Matrix Models

Louk Rademaker, Leiden University,
under supervision of Koenraad Schalm

September 15, 2008

Foreword

Matrix models are "toy models", in a sense that they do not contain physical reality in themselves, but can be used to approximate or understand other "real" models. It is possible that aspects of the abstract "unreal" matrix models can tell us something about "real" Nature. Here I write "real" between parentheses, since the search for an essence of reality in mathematical equations might the most difficult task for theoretical physicists today. Many times we have to plough through tedious calculations before we see even a small glimpse of Nature.

Anyway, I was allowed to play a little with these matrix toys. And 'playing around' means that I tried some known concepts to these matrix models, and just watched what happened. Thereby I was able to learn many aspects of theoretical physics - which still is, of course, the main goal of doing that study.

This thesis reports on my research into matrix models. Roughly the first half contains already known concepts and definitions. With those backgrounds I try to tackle the matrix model in the second half of this thesis.

For me, the most important result is that I somehow managed to translate the bizarre mathematics of a matrix model into a language that physicists would understand. Instead of Hermitian matrices and their eigenvalues, I can now talk about phase transitions and the free energy.

To conclude this foreword, I would like to thank my supervisor Koenraad Schalm. Working on matrix models was difficult for me personally, and many times when I exclaimed that "these matrix models are just plain nonsense!" he got me back on track by showing the "real" physics behind all the equations. Thanks, Koenraad!

Louk Rademaker
September 2008

Contents

Chapter 1

Basic concepts of renormalization

Matrix models have many degrees of freedom - a feature that will become clear in the rest of this thesis. Therefore we are obliged to introduce in this chapter 'renormalization', which is a method of dealing with large numbers of degrees of freedom. Normally one would associate a certain length scale with a system (for example, atomic length scale for many-atom systems). By averaging the interactions over larger scales, we can understand how the effective interactions between the degrees of freedom change whenever we change the scale.

However, the effective interaction between relevant degrees of freedom can suddenly change. Such phase transitions play a crucial role in theories with a large number of degrees of freedom.

In this chapter, we explain the concepts of renormalization and phase transitions at the hand of two specific models: the planar model (also known as XY model, or sine-Gordon model, or Thirring model, or periodic Gaussian model) and the 2d Ising model (which has the opposite nomenclature problem - there are many different models that bear that name, but no worries, we will only use one type).

It is the framework introduced in this chapter that will be used to investigate the properties of the matrix models. Those models will be introduced in the next chapter.

1.1 Theories with many degrees of freedom

The most difficult problems in physics arise when there are large numbers of degrees of freedom involved. For example, a gas or a liquid can contain more than 10^{23} atoms, each of them having several degrees of freedom. It is impossible to calculate the momenta and positions of all these molecules.

Of course no physicist will calculate all possible interactions between all degrees of freedom. Most of the time we can deal with those huge quantities by only considering collective, statistical properties. The aforementioned gas can effectively be described by means of only temperature, pressure and volume.

In this section we will review two physical models with large numbers of degrees of freedom. At first we will discuss the branch of quantum field theories, all of them having

initially infinitely many degrees of freedom. Thereafter we will discuss a - at first glance - completely different field of physics: the critical phenomena in statistical mechanics.

Of both theories we will mention their major methods and concepts. Next to that we will, of course, emphasize why they have many degrees of freedom.

1.1.1 Quantum field theories

Over the last fifty years, quantum field theories have proven to give enormous insight into many previously unintelligible problems, such as the beta decay or the anomalies in the electron charge. A good introduction of the theory is given by Peskin and Schroeder [1] .

But even more important than these practical results is the beautiful conceptual framework on which quantum field theories are based.

QFT is based on classical field theory. The elementary particles or *quanta* are then small perturbations from the ground state of the field (the vacuum). The imposed *local gauge symmetry* between the quanta gives rise to interactions.

It is expected - or it is at least hoped for - that the concept of perturbative quanta with symmetries can be used to describe all forces of nature. Currently we still lack a field theory of gravity (see the epilogue of [1]).

Even though the concepts and results of QFT are beautiful, the amount of degrees of freedom is extremely large. Every field $\phi(x)$ has a each spacetime point x one degree of freedom. Assuming that spacetime is continuous, we have to conclude that the theory contains uncountable many degrees of freedom [2].

The uncountable many degrees of freedom do not only constitute an unpractical amount of work. In fact, the quantities that we want to compute - such as scattering cross sections - become infinite as well.

These *ultraviolet divergences* arise because in the path integral formalism we need to integrate over quanta with every possible momentum. For high momenta, or equivalently: small spatial distances, the integrand diverges.

1.1.2 Critical phenomena in statistical mechanics

Onto another branch of physics: statistical mechanics. This theory describes the mechanics of many-body systems. That is done by summing over all the possible configurations of all the microscopic degrees of freedom in the so-called partition function Z. Out of this partition function, a few general macroscopic degrees of freedom can be derived.

Take, for example, a glass of water. The glass contains more than 10^{23} molecules, all having many degrees of freedom. Still everything you want to know about your glass of water is given by it's temperature, pressure and volume. These are the relevant degrees of freedom of the system.

Why is this possible? At what size did my physical system change from 'just a number of molecules' - with quantities such as momenta and dipole moment - to 'a glass of water' - with quantities such as temperature and pressure?

The minimum size of a collection of water molecules must have in order to speak of 'a glass of water' is called the *correlation length* $\xi(T)$. An extended physical picture is given by [2]. In favorable circumstances, ξ is only one or two molecular spacings.

When a region the size of the correlation length ξ contains only a few degrees of freedom, it is possible to effectively describe the system with macroscopic quantities - such as pressure and temperature. We can also state this the other way around: as long as the scale under study is much larger than the correlation length ξ, it is possible to introduce macroscopic variables.

In general, this correlation length ξ is of molecular or atomic scale. However, at the critical point marking the onset of a phase transition the correlation length diverges. We will clarify this divergence in section 1.3.1. So if $\xi \to \infty$, it will at some point become as big as our system under study. All degrees of freedom appear to be relevant, and there is no simple way of finding effective large-scale degrees of freedom.

1.2 Renormalization

We just mentioned two tough physical problems with an extremely large number of degrees of freedom. And by what other means should we handle them than by reducing the number of degrees of freedom?

Throughout this section we will show two different, complementary techniques to reduce that number of freedom: the lattice formulation and the renormalization group. And meanwhile we notice that quantum field theory and statistical mechanics are conceptually equivalent.

1.2.1 Lattice gauge theories

It's quite impossible to deal with uncountable many degrees of freedom. Bringing that number back to countable infinite is already a great achievement.

Before we remarked that the problem of uncountable degrees of freedom manifests itself through ultraviolet divergences. In quantum field theory we therefore introduce a high momentum cut-off to get rid of these divergences. This is equivalent to a small-distance cut-off. In other words, spacetime is no longer continuous but discrete. It has become a lattice.

We will see that a lattice theory can be related to a quantum field theory by sending the lattice spacing to zero - the continuum limit.

A thorough introduction to lattice gauge theories is given by [3].

Ising model

The simplest lattice theory is the 2d Ising spin model. Consider a two-dimensional $N \times N$ square lattice with at each site $n = (n_1, n_2)$ a "spin" variable $s(n)$. This variable can take

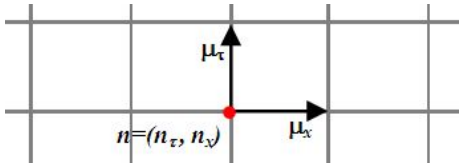


Figure 1.1: *The square lattice of the Ising model. Each site* $n = (n_\tau, n_x)$ *has two nearest neighbors; at* $n + \mu_\tau$ *and* $n + \mu_x$.

the values "up" or "down",

$$s(n) = \begin{cases} \text{up}, & +1 \\ \text{down}, & -1 \end{cases} . \tag{1.1}$$

The energy, or "action" of the Ising model is given by a nearest-neighbor interaction

$$S = -J \sum_n \sum_\mu s(n)s(n+\mu). \tag{1.2}$$

The summation over μ denotes the two nearest-neighbors of the site n in the positive direction, so μ is either $(0,1)$ or $(1,0)$, see figure 1.1. We demand that the coupling constant J is positive, so that the action is minimal if all the spins are aligned.

The Ising model just introduced can be characterized by two main properties:

1. The action is local, as far as that is possible with discrete spacings.

2. The action has a global $\mathbf{Z}_2$-symmetry. That is a mathematical way of saying that the action does not change if we flip all spins.

The Ising model can be treated as a statistical mechanics model. When the system is put at finite temperature by contact with a heatbath, the probability that a particular spins configuration occurs is proportional to

$$P = \exp(-\beta S) \tag{1.3}$$

where β equals the famous Boltzmann factor $1/kT$. The collective properties can be obtained from the partition function

$$Z = \sum_{\{\text{configs}\}} \exp(-\beta S) \tag{1.4}$$

where we sum runs over all the possible configuration. For example, the free energy is given by $F = -kT \log Z$. The statistical correlation between two spin sites at distance n is defined by the spin-spin correlation function

$$\Gamma(n) \equiv \langle s(n)s(0)\rangle \equiv \frac{1}{Z} \sum_{\{\text{configs}\}} s(n)s(0)e^{-\beta S}. \tag{1.5}$$

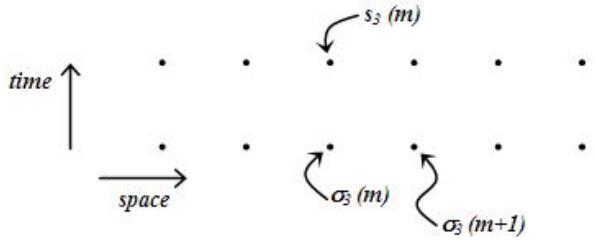


Figure 1.2: *The square lattice of the Ising model, now with the appropriate names used in the τ-continuum limit.*

This spin-spin correlation allows us to explicitly define the *correlation length* ξ that we introduced earlier,

$$\Gamma(n) \sim \exp^{-|n|/\xi(T)}. \tag{1.6}$$

τ-continuum limit

The nomenclature 'μ_τ' and 'μ_x' already suggests that the two axes of the Ising model can represent time and space. Define τ as the lattice spacing in the time direction. We want to make the time evolution continuous ($\tau \to 0$), without qualitatively changing the physics. This is the so-called τ-continuum limit.

In quantum mechanics, the time step operator $\hat{T}$ is related to the (infinitesimal time step) Hamiltonian by

$$\hat{T}(\tau) = e^{-\hat{H}\tau}. \tag{1.7}$$

The probability that if time $t = t_i$ the physical system is in state $|\alpha\rangle$ at the next time instance $t = t_i + \tau$ the physical state is $|\beta\rangle$, is equal to

$$P = \langle\beta|\hat{T}(\tau)|\alpha\rangle. \tag{1.8}$$

And fortunately enough, we already knew all statistical weights in the Ising model! We just have to sum all configurations that equal $|\alpha\rangle$ at the row $t = t_i$ and equal $|\beta\rangle$ at the row $t = t_i + \tau$. We will call the operator $\hat{T}$ the *transfer matrix*.

The first step therefore is to rewrite the action of the Ising model so that we isolate a 'Lagrangian' at each moment of constant time. We will call all sites with at the same time a 'row'. The spin variables in a row will be called $\sigma_3(m)$. The spin variables at one time instance later will be called $s_3(m)$. This is visualized in figure 1.2.

The action of the Ising model now equals

$$S = \sum_{\text{rows } n_\tau} L(n_\tau) \tag{1.9}$$

where the 'Lagrangian' of each row is given by

$$L = \frac{1}{2}J_\tau \sum_m \left[s_3(m) - \sigma_3(m)\right]^2 - \frac{1}{2}J_x \sum_m \left[\sigma_3(m+1)\sigma_3(m) + s_3(m+1)s_3(m)\right]. \qquad (1.10)$$

Note that we allow that the coupling in the time direction is different than in the spatial direction - we now have J_τ and J_x. This is, of course, because we want to take the limit $\tau \to 0$, but we want to keep the spatial lattice spacing x constant.

Using the Lagrangian (1.10) we can derive the transfer matrix. We know that the statistical weight of any configuration if proportional to e^{-S}. So the probability that if at some row the physical system is in state $|\alpha\rangle$ at the next row the physical state is $|\beta\rangle$, is equal to

$$P \sim \langle\beta|e^{-\beta L}|\alpha\rangle. \qquad (1.11)$$

Compare this result with equation (1.8). So if we investigate all possible state α and β in the low τ limit, we can find the Hamiltonian $\hat{H}$.

Define $K_x = J_x/kT$ and $K_\tau = J_\tau/kT$. Suppose at first that $\alpha = \beta$ (so $\sigma_3 = s_3$), which implies that the two adjacent rows are the same. Physically, this means that in the time step τ there have been no spin flips.

$$\begin{aligned} \hat{T}(0 \text{ flips}) &= e^{-\beta L} = e^{K_x \sum_m \sigma_3(m+1)\sigma_3(m)} \\ &= 1 - \tau\hat{H}\Big|_{(0 \text{ flips})} + \mathcal{O}(\tau^2), \end{aligned} \qquad (1.12)$$

This result implies that K_x is proportional to τ. Now if two successive rows only differ one spin, the Lagrangian obtains an extra term $2J_\tau$ (namely at the point where $\sigma_3(m) - s_3(m) = \pm 2$). The transfer matrix $\hat{T}$ for one spin flip equals

$$\begin{aligned} \hat{T}(1 \text{ flip}) &= e^{-2K_\tau} e^{\frac{1}{2}K_x \sum_m [\sigma_3(m+1)\sigma_3(m) + s_3(m+1)s_3(m)]} \\ &= -\tau\hat{H}\Big|_{(1 \text{ flip})} + \mathcal{O}(\tau^2). \end{aligned} \qquad (1.13)$$

We now see that $\exp[-2K_\tau] \propto \tau$. A fixed ratio between $\exp[-2K_\tau]$ and K_x can be defined, we call this constant λ. Let us suppose also that $K_x = \lambda\tau$,

$$K_x = \lambda e^{-2K_\tau} = \lambda\tau. \qquad (1.14)$$

Now let us turn our attention to the form of the Hamiltonian. We can write the Hamiltonian operator in terms of matrices. The diagonal elements then correspond to time evolution that doesn't change the physical state. From (1.12) and (1.14) we deduce

$$\hat{H}_{(\text{Diagonal})} = -\lambda \sum_m \hat{\sigma_3}(m+1)\hat{\sigma_3}(m) + \mathcal{O}(\tau). \qquad (1.15)$$

Similarly the off-diagonal elements of $\hat{H}$ can be derived from (1.13). Now remember that the first Pauli matrix $\hat{\sigma}_1(m) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ flips the spin at mth site. Hence the off-diagonal Hamiltonian can be stated in terms of $\hat{\sigma_1}(m)$,

$$\hat{H}_{(\text{Off}-\text{Diagonal})} = -\sum_m \hat{\sigma_1}(m) + \mathcal{O}(\tau). \qquad (1.16)$$

Statistical mechanics		*Field theory*
Free energy density	$\leftrightarrow$	Vacuum energy density
Correlation function	$\leftrightarrow$	Propagator
Reciprocal of the correlation length	$\leftrightarrow$	Mass gap

Figure 1.3: *Related concepts in Field Theory and Statistical Mechanics.*

In the continuum limit we set $\tau \to 0$, which corresponds to $K_x \to 0$ and $K_\tau \to \infty$. However, we found that their ratio λ remains constant, so that we are left with the combined Hamiltonian

$$\hat{H} = -\sum_m \hat{\sigma}_1(m) - \lambda \sum_m \hat{\sigma}_3(m+1)\hat{\sigma}_3(m). \tag{1.17}$$

Hence we can conclude that

- the physics of the 2d Ising lattice model and the one-dimensional quantum Hamiltonian (1.17) are the same.

Statistical mechanics meets quantum field theory

We have just related a 2d lattice model with a 1d quantum model. Both theories have a global $\mathbf{Z}_2$-symmetry. We can use the formalism of the transfer matrix to generalize the observed relation between statistical mechanics and quantum mechanics. We then find that we can relate the path integral of a quantum field theory to the partition function of a statistical mechanics model,

$$Z = \sum_{\text{configs}} e^{-\beta S} = \int_{\text{paths}} e^{\frac{i}{\hbar}S}. \tag{1.18}$$

The exact equivalence can be established by

- a Wick rotation $t \to -i\tau$;
- relating the temperature with Planck's constant $kT = \hbar$;
- impose a spatial cut-off at the size of the lattice spacing.

This implies that some concepts in Quantum Field Theory and Statistical Mechanics are related to one another. An overview of the most important concepts and their relations is given in figure 1.3.

1.2.2 Wilsons renormalization group

We have dealt with a QFT by introducing a finite, arbitrary, lattice spacing. Introduction of a cut-off is called *regularization*. However, the physics behind the model should not depend on that arbitrary choice of cut-off. We should be able to describe the same physical problem with different lattice spacings. The process in which we change the lattice spacing but do not change the physics behind the model, is called *renormalization*.

The physics of the model is encoded in the path integral Z as a function of the coupling constant J and the cut-off a. Since the physics should be cut-off independent, we find

$$\frac{dZ}{da} = 0 = \frac{\partial Z}{\partial a} + \frac{\partial Z}{\partial J}\frac{\partial J}{\partial a}. \tag{1.19}$$

In other words: we need to change the coupling J whenever we change the cut-off a. The equations $\frac{\partial J}{\partial a}$ for all coupling constants J form together the *renormalization group*.

Renormalization group of the sine-Gordon model

As an example we will derive the discrete renormalization group of the *sine-Gordon* model. Kogut [3] has shown that the *planar model*, a 2d lattice model with action

$$S = J \sum_{\langle mn \rangle} \cos(\theta_n - \theta_m) \tag{1.20}$$

where $\langle mn \rangle$ sums over all neighboring sites, can be rewritten into the 2d sine-Gordon model. That model has action

$$S[\phi_\Lambda, \mu, \beta] = \int d^2x \left\{ -\frac{1}{2}\phi_\Lambda(x)\partial^2\phi_\Lambda(x) - \mu \cos\left[2\pi\sqrt{\beta}\phi_\Lambda(x)\right] \right\}, \tag{1.21}$$

and the momentum cut-off $\Lambda = a^{-1}$ is imposed via the Fourier transform of the field $\phi(x)$,

$$\phi_\Lambda(x) = \int_{0<p<\Lambda} \frac{d^2p}{(2\pi)^2} e^{ipx}\phi(p). \tag{1.22}$$

The path integral of the theory is

$$Z_\Lambda(\mu, \beta) = \int_{0<p<\Lambda} \mathcal{D}\phi(p) \exp\left\{-S[\phi_\Lambda]\right\}. \tag{1.23}$$

Our aim is to find out how the coupling constants μ and β should depend on the scale Λ. This is obtained by finding a relation between the path integral at two different cut-offs,

$$Z_\Lambda(\mu, \beta) \propto Z_{\Lambda'}(\mu', \beta'), \tag{1.24}$$

with $\Lambda > \Lambda'$. We then can introduce the 'high-momentum' part of the field $\phi(x)$ by

$$h(x) = \phi_\Lambda(x) - \phi_{\Lambda'}(x) = \int_{\Lambda'<p<\Lambda} \frac{d^2p}{(2\pi)^2} e^{ipx}\phi(p). \tag{1.25}$$

The kinetic part of the action then easily splits

$$\int d^2x \frac{1}{2}\phi_\Lambda(x)\partial^2\phi_\Lambda(x) = \int_{0<p<\Lambda} \frac{d^2p}{(2\pi)^2}(-\frac{1}{2}p^2)|\phi(p)|^2 \quad (1.26)$$

$$= \frac{1}{2}\int d^2x \left\{\phi_{\Lambda'}\partial^2\phi_{\Lambda'} + h\partial^2 h\right\}. \quad (1.27)$$

The path integral can be written as

$$Z = \int_{0<p<\Lambda'} \mathcal{D}\phi\ e^{\frac{1}{2}\int \phi_{\Lambda'}\partial^2\phi_{\Lambda'}} \int_{\Lambda'<p<\Lambda} \mathcal{D}\phi\ e^{\frac{1}{2}\int h\partial^2 h + \mu\cos\left[2\pi\sqrt{\beta}(\phi_{\Lambda'}+h)\right]}. \quad (1.28)$$

It requires some difficult mathematics to split the cosine interaction term. We will do this in a familiar way: suppose that μ is small and make a perturbation series. That is,

$$e^{\mu\int d^2x\cos(..x)} = 1 + \mu\int d^2x\ \cos(..x) + \frac{1}{2}\mu^2\int d^2x d^2y\ \cos(..x)\cos(..y) + .. \quad (1.29)$$

Since we still have to take the path integral over these terms, we introduce a 'high-momentum average' defined by

$$\langle\theta(\phi)\rangle_h = \frac{\left[\int_{\Lambda'<p<\Lambda}\mathcal{D}\phi\ e^{\frac{1}{2}\int h\partial^2 h}\theta(\phi)\right]}{\left[\int_{\Lambda'<p<\Lambda}\mathcal{D}\phi\ e^{\frac{1}{2}\int h\partial^2 h}\right]}, \quad (1.30)$$

so that the 'high-momentum' part of the full path integral (1.28) becomes

$$Z_h = \left[\int \mathcal{D}\phi\ e^{\frac{1}{2}\int h\partial^2 h}\right] \times \left[1 + \mu\int d^2x\ \langle\cos 2\pi\sqrt{\beta}(\phi_{\Lambda'}+h)\rangle_h + \mathcal{O}(\mu^2)\right]. \quad (1.31)$$

The free path integral over h is just a multiplicative factor, which is of no further interest. Now we will calculate these 'high-momentum averages', starting with the first order term. We split the cosine term so we can get the $\phi_{\Lambda'}$ part outside the averaging brackets,

$$\langle\cos 2\pi\sqrt{\beta}(\phi_{\Lambda'}+h)\rangle_h = \frac{1}{2}e^{i2\pi\sqrt{\beta}\phi_{\Lambda'}}\langle e^{i2\pi\sqrt{\beta}h}\rangle_h + \text{h.c.} \quad (1.32)$$

Here h.c. stands for the Hermitian conjugate of the first term. We are left with the task of computing

$$\langle e^{i2\pi\sqrt{\beta}h}\rangle_h = \frac{\left[\int_{\Lambda'<p<\Lambda}\mathcal{D}\phi\ e^{(\frac{1}{2}\int h\partial^2 h) + i2\pi\sqrt{\beta}h}\right]}{\left[\int \mathcal{D}\phi\ e^{\frac{1}{2}\int h\partial^2 h}\right]}. \quad (1.33)$$

In momentum space, the exponent in this equation becomes

$$\int_{\Lambda'<p<\Lambda}\frac{d^2p}{(2\pi)^2}\left\{-\frac{1}{2}p^2|\phi(p)|^2 + 2\pi i\sqrt{\beta}e^{ipx}\phi(p)\right\}. \quad (1.34)$$

With the transformation

$$\phi'(p) = \phi(p) - \frac{2\pi i\sqrt{\beta}}{p^2} e^{-ipx} \tag{1.35}$$

we can rewrite this as

$$\int_{\Lambda'<p<\Lambda} \frac{d^2p}{(2\pi)^2} \left\{ -\frac{1}{2} p^2 \phi'(p)\phi'(-p) - \frac{2\pi^2\beta}{p^2} \right\}. \tag{1.36}$$

The transformation (1.35) does not change the measure of the path integral, $\mathcal{D}\phi = \mathcal{D}\phi'$. Now we can get rid of the path integrals in (1.33). Define

$$G_h(x) = \int_{\Lambda'<p<\Lambda} \frac{d^2p}{(2\pi)^2} e^{ipx} \frac{1}{p^2} \tag{1.37}$$

and

$$A_x = e^{-2\pi^2\beta G_h(x)}. \tag{1.38}$$

These definitions now lead to the result

$$\langle \cos 2\pi\sqrt{\beta}(\phi_{\Lambda'} + h)\rangle_h = A_0 \cos 2\pi\sqrt{\beta}\phi_{\Lambda'} \equiv A_0 \cos\alpha\phi_{\Lambda'}. \tag{1.39}$$

Using the same technique, we can derive that the second order term is, where $\phi_x \equiv \phi_{\Lambda'}(x)$ and $h_x \equiv h(x)$,

$$\begin{aligned} \langle \cos\alpha(\phi_x + h_x)\cos\alpha(\phi_y + h_y)\rangle_h &= \frac{1}{2} A^2_{(x-y)} A_0^2 \cos\alpha(\phi_x + \phi_y) + \\ &\quad \frac{1}{2} A^{-2}_{(x-y)} A_0^2 \cos\alpha(\phi_x - \phi_y). \end{aligned} \tag{1.40}$$

Recall that we wanted to find back the original partition function but with new coupling constants. Therefore we need to exponentiate the perturbation series again. Just like we did with the Feynman expansion in quantum field theory,

$$\{\text{all diagrams}\} = e^{\{\text{connected diagrams}\}}. \tag{1.41}$$

In this particular expansion, we can see that the 'connected' term in second order is

$$\langle \cos_x \cos_y\rangle - \langle \cos_x\rangle\langle \cos_y\rangle. \tag{1.42}$$

Hence the high-momentum part of the partition function (1.31) becomes

$$\begin{aligned} Z_h \;\propto\; \exp\Big\{ & \mu A_0 \int d^2x \, \cos\alpha\phi_x \\ & + \frac{1}{4} A_0^2\mu^2 \int d^2x d^2y \, [A^2_{(x-y)} - 1] \cos\alpha(\phi_x + \phi_y) \\ & + \frac{1}{4} A_0^2\mu^2 \int d^2x d^2y \, [A^{-2}_{(x-y)} - 1] \cos\alpha(\phi_x - \phi_y) + .. \Big\}. \end{aligned} \tag{1.43}$$

We are now faced with a difficult integration over x and y. This is most easily done via the transformation

$$\begin{cases} \xi = x - y \\ z = \frac{1}{2}(x+y) \end{cases} . \tag{1.44}$$

Because the propagator $G_h(x)$ is expected to fall of very fast, the presence of the $A_{(x-y)}$-factors allows us to approximate $\phi_x - \phi_y \approx \partial\phi(z)$. Define

$$a_2 = \int \xi^2 d\xi \, \left[A_\xi^2 - 1\right] \tag{1.45}$$

so that the only interesting part of Z_h becomes

$$Z_h \propto e^{\int d^2z \mu A_0 \cos 2\pi\sqrt{\beta}\phi_z - \frac{1}{2}\mu^2 A_0^2 \pi^2 \beta a_2 (\partial\phi_z)^2}. \tag{1.46}$$

The full partition function is now

$$Z_\Lambda(\mu,\beta) \propto \int_{0<p<\Lambda'} \mathcal{D}\phi' e^{\int d^2z \left(-1-\mu A_0^2\pi^2\beta a_2\right)(\partial\phi_z)^2 + \mu A_0 \cos 2\pi\sqrt{\beta}\phi_z}. \tag{1.47}$$

Now this up to a multiplicative constant equal to the original path integral (1.23), but now with different couplings,

$$Z_\Lambda(\mu,\beta) \propto Z_{\Lambda'}(\mu',\beta'), \tag{1.48}$$

namely

$$\phi' = \sqrt{1+\mu^2\pi^2 A_0^2 \beta a_2}\, \phi, \tag{1.49}$$

$$\mu' = \mu A_0, \tag{1.50}$$

$$\beta' = \beta / \left[1 + \mu^2\pi^2 A_0^2 \beta a_2\right]. \tag{1.51}$$

The three scaling relations (1.49)-(1.51) form the *renormalization group* equations. They describe how the coupling constants μ and β change whenever the momentum cut-off changes. Note that the cut-off dependence is implicitly included via the factors A_0 and a_2.

To summarize, we followed these steps to derive the renormalization group:

1. Split the regularized field $\phi_\Lambda(x)$ in a high momentum part $h(x)$ and a low momentum field $\phi_{\Lambda'}(x)$.
2. Split the kinetic terms in the action. One part is dependent on $\phi_{\Lambda'}(x)$, the other part is dependent on $h(x)$ only.
3. Make a perturbation series in the coupling constant μ.
4. Use the technique of a 'high-momentum average' to compute all the terms in perturbation series.
5. Exponentiate the series again. This is done by only considering the 'connected' terms, just like for Feynman diagrams.
6. Compute the integrals under the assumption that all distances are small, $x \approx y$.
7. Relate Z_Λ with $Z_{\Lambda'}$ to find the scaling relations.

Infinitesimal renormalization group equations

It is interesting to find the continuous dependence of the coupling constants on the scale. We can find that dependence by integrating out a infinitesimal small momentum 'slice'

$$\Lambda' = \Lambda - d\Lambda < p < \Lambda. \tag{1.52}$$

That only effects the 'high-momentum propagator' $G_h(x)$. For the planar model, Kosterlitz [4] has proven that the renormalization group equations are

$$dy = -xy\frac{da}{a}, \tag{1.53}$$

$$dx = -y^2\frac{da}{a}. \tag{1.54}$$

The relation between the planar model and the sine-Gordon model is $y = \frac{1}{2}\mu a^2$ and $x = \pi\beta - 2$. The momentum cut-off Λ is the reciprocal of the lattice spacing a.[1]

Note that the renormalization group of the planar model has a fixed point at $x = y = 0$. That is, if $y = 0$ the coupling constants are independent of cut-off. In other words: the model with $x = y = 0$ is scale invariant.

The model with $x = y = 0$ corresponds, in the language of statistical mechanics, to a model with critical temperature

$$T_c \approx \frac{\pi J}{2k}. \tag{1.55}$$

We see that the planar model is independent of scale at the critical temperature T_c.

1.3 Phase transitions and dualities

In the course of the previous section we derived the renormalization group because we wanted to find a relation between the coupling constants and the cut-off size. With this technique we could find the real (cut-off independent) physics for any arbitrary choice of cut-off. And there we found a critical point of specific coupling constants, where the theory was cut-off independent!

1.3.1 Critical points in the Ising model

As mentioned in section 1.1.2, the minimum size that a system can have without changing it's physical properties is the correlation length. But at what happens at the critical point is that the correlation length will become infinitely large.

What causes this divergence?

[1] However, the equations (1.53)-(1.54) are not infinitesimal extensions of (1.49)-(1.51). The calculation involves a special technique of choosing a suitable "smooth momentum slice $d\Lambda$". See for more information [3] and its references.

Let's study that in the case of the 2d Ising model. Even though the the model seems so simple, in the presence of an external magnetic field with interaction $-B\sum_n s(n)$ the model acquires a critical point at a critical temperature T_c.

We expect that for high temperatures ($T > T_c$) the thermal fluctuations dominate and that there are no long range correlations between the spins. At the other hand, for $T < T_c$ the system is simply magnetized, and we do not find infinitely correlated spins as well. The typical correlation length $\xi(T)$ is in both cases negligible .

The spin-spin correlation function $\Gamma(n)$ can be calculated, and for $T > T_c$ we can use the definition of the correlation length (1.6),

$$\Gamma(n) \sim e^{-|n|/\xi(T)},\ (T > T_c). \tag{1.56}$$

Precisely at the critical point, we expect the correlation function to be a power law of the distance n,

$$\Gamma(n) \sim |n|^{-(d-2+\eta)},\ (T = T_c) \tag{1.57}$$

since that implies longe-range correlations. In order for (1.57) and (1.56) to be compatible, the correlation length $\xi(T)$ must become infinite at $T = T_c$. Its behavior is described by the critical exponent ν via

$$\xi(T) = (T - T_c)^{-\nu}. \tag{1.58}$$

For the 2d Ising model, $\nu = 1$.

Critical exponents like ν and η can be calculated whenever critical phenomena are involved. Luckily, the critical exponents of a wide scope of different physical system are identical.

This has lead to the *correlation length scaling hypothesis*, which states that the only relevant information of a system near $T \sim T_c$ is given by the correlation length $\xi(T)$. Precisely at the critical point, the system has become invariant of scale. All physical behavior can be related to the only relevant critical exponent ν.

Still, the hypothesis categorizes physical models according to their critical behavior. It is an interesting but open question what exactly determines these *universality classes* of physical models. We do know that the symmetry in the action and the dimensionality of the system are relevant.

1.3.2 Perturbative and topological solutions

Now that we have unravelled the first mysteries of a critical point, it is worthwhile to note that a critical point divides a physical model in two phases. We have different physical behavior for $T > T_c$ and for $T < T_c$. It is the planar model that gives a fundamental insight in the differences between these two thermal regions.

Recall that the planar model consists of a 2d square lattice with at each site n a spin variable $s(n) = \begin{pmatrix} \cos\theta_n \\ \sin\theta_n \end{pmatrix}$. The action is given by

$$S = J \sum_{n,\mu} \cos \Delta_\mu \theta_n, \tag{1.59}$$

where $\Delta_\mu \theta_n = \theta_n - \theta_{n+\mu}$. The summation over μ is over nearest neighbors. Let us calculate the correlation function,

$$\Gamma(n) = \langle s(n)s(0)\rangle = \langle e^{i(\theta_0 - \theta_n)}\rangle + \text{h.c.}. \tag{1.60}$$

The part that we need to calculate explicitly is

$$\langle e^{i(\theta_0-\theta_n)}\rangle = \frac{1}{Z}\int d\theta e^{i(\theta_0-\theta_n)} e^{-\frac{J}{kT}\sum_{n,\mu}\cos\Delta_\mu\theta_n} \tag{1.61}$$

For $T > T_c$, we can expand the e^{-S} term in powers of $\left(\frac{J}{kT}\right)$. Since integrals over odd powers of θ_n vanish, we are only left with the $\left(\frac{J}{kT}\right)^{|n|}$ term. That implies, when we bring the correlation function in the same form as (1.56),

$$\Gamma(n) \sim e^{-|n|\log\frac{kT}{J}},\ T > T_c. \tag{1.62}$$

Now calculate the correlation function for the phase $T < T_c$. The relative absence of heat implies that there are not many thermal fluctuations. We can therefore assume that $\Delta_\mu\theta_n$ is small, so

$$S \approx J \sum_{n,\mu} \frac{1}{2}(\Delta_\mu \theta_n)^2 \tag{1.63}$$

which turns the integral (1.61) into a Gaussian. This leads to a correlation function similar to (1.57),

$$\Gamma(n) \sim |n|^{-kT/2\pi J},\ T < T_c. \tag{1.64}$$

Clearly, the system behaves different for the phases below and above the critical temperature T_c. We have a phase transition at $T = T_c$. What causes this transition?

An intriguing insight was given through the work of J.M. Kosterlitz and D.J. Thouless [5]. They suggested that due to the periodicity of the variable θ_n there could exist singular spin configurations. Below the critical temperature, the system is dominated by small perturbation from a ground state: waves. Above the critical temperature the system is dominated by the singular configurations.

A singular spin configuration is defined is the same way as singular point in an analytic function. We can find it by considering a contour integral over the configuration,

$$\Sigma\ \Delta_\mu \theta_n = 2\pi q, \tag{1.65}$$

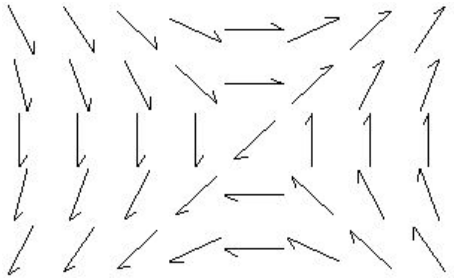


Figure 1.4: *A vortex configuration with winding number* $q = -1$.

where the *winding number* q is an integer

$$q = 0, \pm 1, \pm 2, ... \tag{1.66}$$

A configuration with $q \neq 0$ is called a *vortex*. A typical vortex with winding number $q = -1$ can be seen in figure 1.4.

Consider now a spin system that contains a vortex. The energy (action) of the vortex is

$$S = \pi J \log R/a, \tag{1.67}$$

where R is the size of the vortex system and a is the lattice spacing. (Or equivalently: R/a is the number of sites that the vortex system spans.) Note that if we take the continuum limit $a \to 0$, the vortex energy becomes infinite. The vortex is really a singularity in the system, and it is therefore called a *topological solution*.

To estimate the incidence of vortices, we compute the free energy. The free energy of any system is

$$F = \text{Action} - \text{Temperature} \times \text{Entropy}. \tag{1.68}$$

The entropy is equal to the logarithm of the multiplicity of the configuration, which implies that the entropy of a vortex is $k \log(R/a)^2$. Hence we conclude that a vortex has free energy

$$F = (\pi J - 2kT) \log R/a. \tag{1.69}$$

So we expect vortices to occur whenever

$$T > T_c = \frac{\pi J}{2k}. \tag{1.70}$$

The phase $T > T_c$ with it's singular configurations stands in great contrast with $T < T_c$, where the spin configurations consist of small deviations from a mean spin value. The latter can be seen as waves or, in quantum mechanical terms, as perturbative solutions. To conclude, the two phases can be classified as shown in figure 1.5.

	$T < T_c$	$T > T_c$
Correlation function	Power law, $\Gamma \sim \lvert n\rvert^{-kT/2\pi J}$	Falls off exponentially, $\Gamma \sim e^{-\lvert n\rvert \log \frac{kT}{J}}$
Configurations	Small deviations, waves	Vortices
Type of solutions	Perturbative	Topological

Figure 1.5: *The two phases of the planar model, with a Kosterlitz-Thouless phase transition at* $T_c = \pi J/2k$.

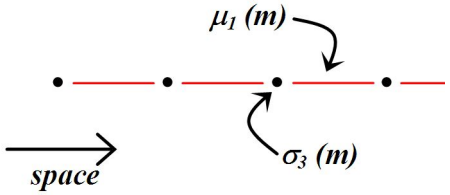


Figure 1.6: *The dual of the Ising model. The original lattice consists of spin variables* $\sigma_3(n)$ *at each site, the dual lattice consists of their links* $\mu_1(n) = \sigma_3(n+1)\sigma_3(n)$.

1.3.3 Dualities

Now that we have classified the difference between the two phases of the planar model, it is instructive to look for similarities. This can be easily done with the beautiful concept of duality.

The self-dual Ising model

In section 1.2.1 we saw that the τ-continuum limit of the Ising model leads to the Hamiltonian

$$\hat{H}[\sigma, \lambda] = -\sum_m \hat{\sigma}_1(m) - \lambda \sum_m \hat{\sigma}_3(m+1)\hat{\sigma}_3(m). \tag{1.71}$$

Now introduce the *dual* lattice that consists of the links between the sites n (figure 1.6). With each link we associate a variable

$$\mu_1(n) = \sigma_3(n+1)\sigma_3(n). \tag{1.72}$$

A corresponding 'flipping operator' can easily be defined as

$$\mu_3(n) = \prod_{m<n} \sigma_1(n). \tag{1.73}$$

Note that the group $\langle \mu_1, \mu_3 \rangle$ has the same symmetry as the group $\langle \sigma_1, \sigma_3 \rangle$. When we rewrite the Hamiltonian (1.71) into the new operators μ we find

$$\hat{H}[\sigma, \lambda] = \lambda \left[-\sum_m \hat{\mu}_1 - \lambda^{-1} \sum_m \hat{\mu}_3(m+1)\hat{\mu}_3(m) \right] \equiv \lambda \hat{H}[\mu, \lambda^{-1}]. \tag{1.74}$$

This means that the Ising model is *self-dual*; the Hamiltonian remains the same when we switch to the dual lattice. However, what did change was the coupling constant $\lambda \to \lambda^{-1}$.

Since $\lambda \propto 1/T$, the self-duality of the Ising model implies that the physics are the same for low T and high T.

Stop! This is strange. Wasn't it that we showed in the previous section that the low T and high T phases are very different? Perturbative solutions for $T < T_c$ and topological solutions when $T > T_c$?

The sine-Gordon model

Let us therefore get back to the planar model. Which was, as mentioned before, equivalent to the sine-Gordon model. It is now time to introduce this sine-Gordon model in a little more detail. (Not as detailed as [6]!)

The name 'sine-Gordon' is a joke that refers to the fact that it is a field theory

$$S_{SG} = \int d^2x \left[\frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{m^2}{\beta^2} (\cos \beta\phi - 1) \right], \tag{1.75}$$

with a corresponding equation of motion that looks just like the Klein-Gordon equation, but then with a sine function:

$$\frac{\beta^2}{m^2} \left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) \phi + \sin \beta\phi = 0. \tag{1.76}$$

The sine-Gordon model allows for static ($\partial_t \phi = 0$) and constant solutions ($\partial_x \phi = 0$), when $\sin \beta\phi = 0$. This is when

$$\beta\phi = \pi N, \quad N \in \mathbf{Z}. \tag{1.77}$$

The energy of these solutions is given by

$$H = \int dx \left\{ \frac{1}{2} (\partial_t \phi)^2 + \frac{1}{2} (\partial_x \phi)^2 - \frac{m^2}{\beta^2} (\cos \beta\phi - 1) \right\}, \tag{1.78}$$

so the energy of the odd-numbered solutions in (1.77) equals $\int dx 2m^2/\beta^2$ which is infinite. Hence the only static constant solutions with finite (even zero) energy are

$$\beta\phi_N = 2\pi N, \quad N \in \mathbf{Z}, \tag{1.79}$$

whose energy is vanishing: they represent vacuum solutions. We call all classical solutions of finite energy *solitons*. The Hamiltonian density of any soliton ϕ must vanish at spatial infinity, for the total energy to be finite,

$$\lim_{x \to \infty} \mathcal{H}[\phi] = 0. \tag{1.80}$$

Since the only constant solutions with zero energy are given by ϕ_N in (1.79), all solitons must satisfy

$$\lim_{x\to\infty} \phi = \phi_N. \tag{1.81}$$

We can therefore associate a *topological charge* with each soliton, which is simply the difference between ϕ_N at plus infinity and ϕ_N at minus infinity.

$$Q = N_{+\infty} - N_{-\infty} = \frac{\beta}{2\pi}\phi(x)|_{-\infty}^{x=+\infty} = \frac{\beta}{2\pi}\int_{-\infty}^{+\infty} dx \frac{\partial\phi}{dx}. \tag{1.82}$$

The charge density must be the zeroth component of a corresponding topological current

$$J^\mu = \frac{\beta}{2\pi}\epsilon^{\mu\nu}\partial_\nu\phi = \frac{\beta}{2\pi}\begin{pmatrix} \partial_x\phi \\ -\partial_t\phi \end{pmatrix}. \tag{1.83}$$

We see that this current is conserved, even without imposing the equation of motion. On pure topological grounds we find

$$\partial_\mu J^\mu = \frac{\beta}{2\pi}\left(\partial_t\partial_x - \partial_x\partial_t\right)\phi = 0. \tag{1.84}$$

By explicit calculus we can find the static soliton solutions of the sine-Gordon equation that have topological charge $Q = \pm 1$.

$$\phi_\pm = \pm\frac{4}{\beta}\arctan\left[e^{m(x-x_0)}\right]. \tag{1.85}$$

Via a Lorentz boost $x \to \gamma(x - vt)$ we obtain the moving solutions. The energy, computed via the Hamiltonian, of the static one-soliton solution (1.85) equals

$$M = \frac{8m}{\beta^2}. \tag{1.86}$$

Hence solitons can be viewed as particles with topological charge $Q = +1$ and rest-mass M.

The sine-Gordon model has next to this soliton solutions also quantum solutions, which are small perturbations from any ground state ϕ_N. The mass of these *meson excitations* is given by the quadratic term in the potential energy, which is m.

These mesons have a ϕ^4 interaction in the expansion of the cosine, with coupling $\lambda = m^2\beta^2$. In the weak coupling limit $\lambda, \beta \to 0$, the mass of the solitons becomes very large while the meson mass remains constant.

Depending on the strength of the coupling constant, we could have a system that is dominated by the perturbative mesons or by the topological solitons. We have rediscovered the two phases of the planar model, now within the quantum field theoretical framework of the sine-Gordon model.

Duality between sine-Gordon and Thirring

In 1975 Sidney Coleman [7] established an equivalence between the aforementioned sine-Gordon model and the Thirring model. The Thirring model is a two-dimensional theory of a Dirac field with action

$$S_T = \int d^2x \left[\overline{\psi} i \gamma^\mu \partial_\mu \psi - m_F \overline{\psi} \psi - \frac{g}{2} \left(\overline{\psi} \gamma^\mu \psi \right)^2 \right]. \tag{1.87}$$

When writing out the perturbation series of both theory explicitly, he found that the theories describe the same physics if we relate their coupling constants in the following way,

$$\frac{4\pi}{\beta^2} = 1 + \frac{g}{\pi}. \tag{1.88}$$

So weak coupling in the Thirring model ($g \to 0$) corresponds to strong coupling in the sine-Gordon model. Via this duality we have related the perturbative phase of the Thirring model with the topological phase of the sine-Gordon model, and vice versa.

Note that the topological current of a soliton (1.83) is equivalent to the Noether current of the Dirac field.

1.3.4 The continuum limit and the renormalization group

In short, all results of this section can be summarized in the following two statements:

- Models with a Kosterlitz-Thouless phase transition have one phase where the topological solutions dominate, and one phase where the perturbative solutions dominate.
- Duality exchanges topological solutions with perturbative solutions, and correspondingly topological currents with Noether currents;

We started this section by noting that the planar model had a critical point, which we discovered using the renormalization group approach. This gives us now the opportunity to relate the concept of duality with the renormalization group.

The renormalization group of the planar model was given by equations (1.53) and (1.54). The group flow is shown in figure 1.7. Making the lattice spacing a smaller is equivalent to following the arrows in the group flow diagram. Recall that the action was proportional to

$$S \propto -\frac{1}{2} \phi \partial^2 \phi - \mu \cos\left(2\pi \sqrt{\beta}\, \phi \right), \tag{1.89}$$

where $\mu = 2y/a^2$ and $x = \pi\beta - 2$.

We see that in the continuum limit $a \to 0$, in the phase $T < T_c$, the renormalization yields $y \to 0$. This is equivalent to setting the coupling constant $\mu \to 0$. Hence we conclude from the renormalization group that in the phase $T < T_c$ we can neglige the cosine interaction. The theory can be described by free waves with action $S \propto \phi \partial^2 \phi$.

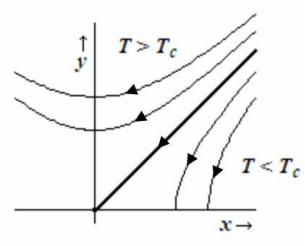


Figure 1.7: *The renormalization group flow of the planar model according to equations (1.53) and (1.54).*

However, in the phase $T > T_c$ the continuum limit gives $y \approx -x$, both going up to $+\infty$. The cosine interaction term dominates the kinetic term, and the model can effectively be described by the periodic action $S \propto \cos\phi'$.

Hence the relation between the two phases of the planar model and it's renormalization group is as follows.

- The fixed points of the renormalization group correspond to either the perturbative phase ($T < T_c$) or to the topological phase ($T > T_c$).

Chapter 2

Introducing matrix models

Now that we are familiar with some basic concepts, it is time to introduce the matrix models themselves. As mentioned in the introduction,these models are just "toy models" that can be used to study various physical theories. An extensive overview of application of matrix models is given by [8].

In this chapter we will define matrix models. As an example of their applications, we derive their relation to two-dimensional quantum gravity.

2.1 Matrix models

A matrix model is described by a partition function Z that depends on a Hermitian $N \times N$ matrix M, with energy/action $S[M]$. The partition function of the M^3-interacting theory is for example

$$Z_N(g) = \int dM \; e^{-\frac{1}{2}\mathrm{Tr}M^2 - g\mathrm{Tr}M^3}. \tag{2.1}$$

Because the integral over dM is not Gaussian, it is very difficult to evaluate the partition function. In quantum field theory it is then usual to expand the interaction part

$$e^{-g\mathrm{Tr}M^3} = 1 - g\mathrm{Tr}M^3 + \frac{1}{2}g^2\left[\mathrm{Tr}M^3\right]^2 + \mathcal{O}(M^9), \tag{2.2}$$

so that the partition function becomes

$$Z_N(g) = \sum_{k=0}^{\infty} \int dM \; \frac{1}{k!}\left[-g\mathrm{Tr}M^3\right]^k \; e^{-\frac{1}{2}\mathrm{Tr}M^2}. \tag{2.3}$$

Each term in the summation over k can be evaluated by introducing a source term in the action

$$-\sum_{i,j} J_i{}^j M_j{}^i = -\mathrm{Tr}JM, \tag{2.4}$$

since

$$\frac{\delta}{\delta J_i{}^j}\, e^{-\mathrm{Tr}JM} = \left(-M_j{}^i\right)\, e^{-\mathrm{Tr}JM}. \tag{2.5}$$

Now with the notation

$$\mathrm{Tr}\left(\frac{-\delta}{\delta J}\right)^3 = \sum_{i,j,k}\left(\frac{-\delta}{\delta J_i{}^j}\right)\left(\frac{-\delta}{\delta J_j{}^k}\right)\left(\frac{-\delta}{\delta J_k{}^i}\right) \tag{2.6}$$

the g-expansion of the interacting partition function will be

$$Z_N(g, J=0) = \int dM\, e^{-\frac{1}{2}\mathrm{Tr}M^2 - g\mathrm{Tr}M^3} \tag{2.7}$$

$$= \sum_{k=0}^{\infty}\frac{1}{k!}\left[g\mathrm{Tr}\left(\frac{\delta}{\delta J}\right)^3\right]^k Z_N(g=0,J)\Bigg|_{J=0}. \tag{2.8}$$

This is a simplification due to the fact that we can explicitly integrate out M in $Z_N(g=0,J)$, which yields

$$Z_N(g=0,J) = \pi^{\frac{1}{2}N^2}2^{\frac{1}{2}N}\, e^{\frac{1}{2}\mathrm{Tr}J^2}. \tag{2.9}$$

Again, we can steal the ideas from quantum field theory. This time we use the concept of the Feynman diagrams. 't Hooft [9] proposed to picture a matrix $M_i{}^j$ with

- an ingoing arrow for the upper index and
- an outgoing arrow for the lower index.

Whenever we take the trace of the identity matrix we form a closed index loop, which gives us a factor $\mathrm{Tr}I = N$. Each vertex is represented by a factor g. This leads to the Feynman rules shown in figure 2.1. Each diagram can now be associated with a factor

$$\text{diagram} \propto g^V N^L, \tag{2.10}$$

where V is the number of vertices and L the number of closed index loops.

Note that every diagram in the expansion of Z satisfies a topological relation between the number of propagators P, the number of vertices V and the number of closed index loops L. First, if we put a dot on both ends of a propagator we find that

$$2P = 3V. \tag{2.11}$$

Secondly, if we look upon the diagram as being a polyhedra with L faces, P edges and V corners, we can use the Euler characteristic formula

$$\chi = 2 - 2H = L - P + V. \tag{2.12}$$

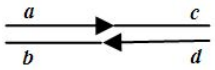
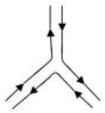
Name	*Mathematical representation*	*Feynman representation*
Source	$J_a{}^b$	
Matrix Propagator	$\delta_a^c \delta_d^b$	
$3 \cdot$ Vertex	$-g = -\frac{\lambda}{\sqrt{N}}$	
Index Loop	N	

Figure 2.1: *The Feynman rules for the M^3 matrix model.*

The Euler characteristic χ of the diagram is determined only by the integer H, which represents the number of handles of the surface. (So a sphere has $H = 0$, a torus $H = 1$.)

Using this topology relations, the factor of a diagram can be rewritten as follows.

$$\begin{aligned} \text{diagram} &\propto g^V N^L && (2.13)\\ &= g^{2P-2V} N^L && (2.14)\\ &= \left(g^2 N\right)^{\frac{1}{2}V} N^\chi. && (2.15) \end{aligned}$$

This suggests that we can introduce a renormalized coupling constant λ that satisfies

$$g^2 N = \lambda^2, \tag{2.16}$$

so as to split the topology-dependent part from the expansion in the number of vertices,

$$\text{diagram} \propto \lambda^V N^\chi. \tag{2.17}$$

The Euler characteristic χ is either 2, 0 or negative. Since we are considering $N \geq 2$ models, it is obvious that for constant V the diagram with $\chi = 2$ contains the largest factor N. The $\chi = 2$ diagrams are called *planar diagrams*.

We see that in the limit of $N \to \infty$ only the planar diagrams remain.

2.2 Relation to 2d quantum gravity

Because the concept of a Quantum Field is strikingly beautiful, several attempts have been done to describe the force of gravitation within this framework. The first few paragraphs of this section are explained in further detail by [10].

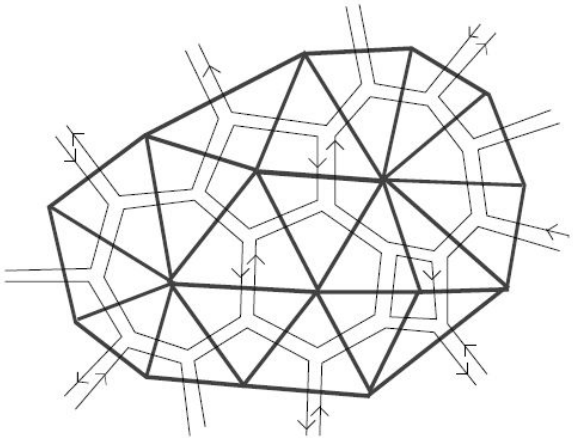


Figure 2.2: *A random triangulation of a surface. Each of the triangles is dual to a M^3-vertex of a matrix theory. Each triangulation vertex is dual to a closed index loop in the matrix model. Picture copied from [10].*

2.2.1 Topology of general relativity

The degrees of freedom of the theory of gravitation (general relativity) are represented by the metric $g_{\mu\nu}$. A quantum field theory needs an action with a varying field, and the only logical choice for a gravitational theory is the Einstein-Hilbert action.

$$S_{EH} = \int d^n x \sqrt{g} \frac{R}{2\kappa^2} \tag{2.18}$$

The famous Gauss-Bonnet theorem relates this curvature-dependent integral to topological concepts. In two dimensions, it relates R to the Euler characteristic,

$$\int d^2 x \sqrt{g} R = 4\pi\chi. \tag{2.19}$$

The starting point for a 2d quantum theory of gravitation is, when we also add an area term $-\beta A = -\beta \int \sqrt{g}$, the partition function

$$Z_{QG} = \sum_{\text{Topologies}} \int \mathcal{D}g \; e^{-\beta A + \gamma\chi} \tag{2.20}$$

for some parameters β, γ.

The difficult aspect of the partition function (2.20) is the path integral over all possible metrics. To simplify this integral we discretize the 2d metrical surface by equilateral triangles of unit area. One can read off the curvature at the points where the triangles touch (the vertices). The curvature is zero ($R = 0$, Euclidean space) when six triangles come together. The curvature of any vertex is given by

$$R_i = 2\pi(6/N_i - 1), \tag{2.21}$$

where N_i is the number of triangles that come together at the given vertex. If we then identify the volume element $d^2x\sqrt{g}$ with $N_i/3$, we can replace the path integral by a discrete summation over all possible triangulations.

$$Z = \sum_{\text{Random triangulations}} e^{-\beta A + \gamma \chi} \tag{2.22}$$

2.2.2 The dual of the triangulation

In figure 2.2 we have a drawing of such a random triangulation. If we place inside each triangle a M^3 matrix model vertex, we obtain the dual lattice of the triangulation.

This dual lattice is a planar diagram of the g expansion of the M^3 matrix model. We see that the number of vertices V in the dual lattice equals the number of triangles, which equals the area A of the surface (remember that the triangles had unit area!).

$$V = A \tag{2.23}$$

Now make the following identification,

$$e^{-\beta} = \lambda, \tag{2.24}$$

$$e^{\gamma} = N. \tag{2.25}$$

Hence the 2d quantum gravity partition function becomes

$$Z_{QG} = \sum_{\text{connected planar diagrams}} \lambda^V N^\chi. \tag{2.26}$$

The sum over connected diagrams is equal to the logarithm of the sum over all diagrams. Hence we have related the 2d quantum gravity partition function to the M^3 matrix model,

$$\lim_{N\to\infty} \log Z_N[\lambda] = Z_{QG}. \tag{2.27}$$

Chapter 3

Phase transitions in matrix models

We defined matrix models in the previous chapter with a partition function depending on Hermitian matrices. In this chapter we will try to understand the basic properties of matrix models in terms of their eigenvalues. We will strip the matrix models from any context and consider them as full, independent theories. These theories then exhibit a phase transition.

3.1 Interpretation of the matrix model

The model introduced in chapter 2 is a statistical model, defined by it's partition function

$$Z = \int dM\ e^{-S[M]}. \tag{3.1}$$

Similar to any statistical model, we can speak of the N^2 degrees of freedom (namely the matrix elements $M_{\mathtt{a}}^{\mathtt{b}}$) and a specific configuration of these degrees of freedom given by the Hermitian matrix M with a certain energy $S[M]$.

Note that a matrix model is static: it does not contain any movement in time or space and/or derivatives of any kind.

The partition function, in turn, defines the *free energy* [11, 12] of the system,

$$e^{-F} = Z. \tag{3.2}$$

In statistical models we are not interested in the precise configurations of the degrees of freedom, but in general properties of the system as a whole. In a matrix model the only macroscopic variable of the system is the free energy.

3.1.1 The use of solitons

In classical Lagrangian mechanics a physical system of many bodies must satisfy the principle of least action. That is, any real physical system equals a configuration that minimizes the

action / energy $S[M]$. In yet other words, the configuration M is a solution of the Euler-Lagrange equation

$$\frac{\delta S[M]}{\delta M_{\mathtt{a}}^{\mathtt{b}}} = 0. \tag{3.3}$$

Configurations M that satisfy the Euler-Lagrange equations and are stable[1] are called *solitons*[2] or *classical solutions*.

In the partition function Z, the exponential 'favors' configuration with a low energy. The contribution from matrices with higher-than-minimal action $S[M]$ is small compared to the minimal energy. The *method of steepest descent* [13] indeed proves that any partition function can be approximated by a Gaussian integral around the minimum of the action M_0,

$$\int dM\ e^{-S[M]} = e^{-S[M_0]} \int dM\ e^{-\frac{1}{2}|S''[M_0]|\ \mathrm{Tr}[M-M_0]^2+\mathcal{O}(M^3)} = e^{-F}. \tag{3.4}$$

Given this equation one would expect that the free energy F is therefore closely related to the energy of the minimal solution, namely the lowest soliton. Similar to the terminology of field theories we would call this lowest energy soliton the *vacuum state* or *ground state* [1, 14]. The system as a whole can be approximated using perturbations from this ground state.

The use of the ground state soliton is that its energy resembles, in lowest order, the free energy. From there we can compute the exact free energy with perturbation series.

3.2 The matrix gM^4 model

Let us explicitly investigate the gM^4 matrix model with action

$$S[M] = \frac{m}{2}\mathrm{Tr}M^2 + \frac{g}{4}\mathrm{Tr}M^4. \tag{3.5}$$

The Euler-Lagrange equation for this model is

$$mM + gM^3 = 0. \tag{3.6}$$

If m is positive, there exists only one stable solution with energy $S[M_0] = 0$. This is in line with the perturbative approach that we advocated in the previous section.

However, if m is negative, there exists $(N+1)$ continuous sets of stable solutions.[3] They consist of matrices with eigenvalues $\pm\sqrt{-m/g}$, and the energy of each configuration is $S = -\mu_C N$. The factor μ_C equals the 'depth of the well' and equals

$$\mu_C = \frac{m^2}{4g}. \tag{3.7}$$

[1] Stable implies that the second derivative of the action $S[M]$ with respect to M is positive.

[2] Hereby we use the definition as put forward by Marciano and Pagels [25]. Their section 6.1.1 ("What are solitons?") starts with: "*Solitons are stable finite energy solutions to the classical equations of motion of Lagrangian field theories.*"

[3] We can rotate each diagonal matrix by $U(N)$ transformations, which doesn't affect the Euler-Lagrange equation. If we then count the number of positive eigenvalues, we can identify $(N+1)$ distinct continuous sets of solitons.

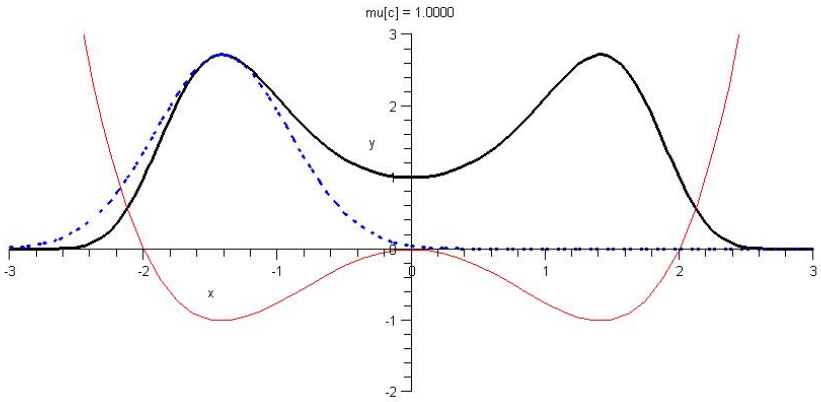


Figure 3.1: *The thin red line corresponds to the action of the $N = 1$ matrix model for $\mu_C = 1$. The integrand e^{-S} in the partition function Z is the thick black line. The dotted blue line represents the Gaussian approximation by the method of steepest descent.*

How can we still perform perturbation theory if there are so many candidates for a ground state? A first step is to integrate out an obvious degree of freedom, namely the rotational symmetry of M. We are then left with a partition function that depends on the real eigenvalues λ_i of the Hermitian matrix M [15],

$$Z = \int \prod_{i=1}^{N} d\lambda_i \prod_{i<j}^{N} (\lambda_i - \lambda_j)^2 \, e^{-\sum_{i=1}^{N} \left(\frac{m}{2}\lambda_i^2 + \frac{g}{4}\lambda_i^4 \right)}, \tag{3.8}$$

where we should impose an ordering on the eigenvalues,

$$\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n. \tag{3.9}$$

If we exponentiate the Vandermonde determinant $\prod_{i<j}^{N}(\lambda_i - \lambda_j)^2$ we can describe our physical system in terms of an effective action in terms of the eigenvalues

$$S[\lambda_i, N, m, g] = \sum_{i=1}^{N} \left(\frac{m}{2}\lambda_i^2 + \frac{g}{4}\lambda_i^4 \right) - \sum_{i<j}^{N} \log(\lambda_i - \lambda_j)^2. \tag{3.10}$$

Since the calculation above is quite abstract, it is an instructive exercise to investigate the model for the simple $N = 1$ and $N = 2$.

3.2.1 Soliton superposition in the $N = 1$ model

We noticed that there exist multiple solitons if m is negative. What soliton could then possibly be the ground state? In figure 3.1 we see intuitively that a Gaussian approximation by the method of steepest descent (3.4) from one of the minima does not match the 'real' integrand e^{-S}.

We can make that explicit by the following perturbative series. Let us make a Taylor expansion around the minimum at $\lambda_0 = -\sqrt{-m/g}$, which yields an action

$$S = -\mu_C - m\lambda^2 - \sqrt{-mg}\lambda^3 + \frac{g}{4}\lambda^4. \tag{3.11}$$

The partition function $Z = \int d\lambda\, e^{-S}$ clearly converges as long as g is positive. Now we write the partition function as a perturbation series

$$Z = e^{-\mu_C} \int d\lambda \sum_{k,l=0}^{\infty} \frac{(-mg)^{k/2}\,(-g/4)^l}{k!\,l!}\, \lambda^{(3k+4l)}\, e^{m\lambda^2} \equiv e^{-\mu_C}\, I_{(k,l)}. \tag{3.12}$$

The zeroeth order term in the expansion, $I_{(k,l)}$, equals of course the Gaussian integral. Higher order can be computed exactly using the Gamma function. If k is odd, $I_{(k,l)} = 0$. In all other cases, it equals

$$I_{(k,l)} = |m|^{-\frac{1}{2}} \left(\frac{-1}{4}\right)^l \left(\frac{m^2}{g}\right)^{-\frac{1}{2}k-l} \frac{\Gamma(\frac{3}{2}k + 2l + \frac{1}{2})}{k!\,l!} \tag{3.13}$$

The norm of this expression diverges due to the last Gamma function. We cannot speak of a "perturbation series" since every next term is bigger than the previous!

Again we find that *the multiplicity of solitons leads to a breakdown of perturbation theory.*

How can we make a sensible approximation to the free energy of our matrix model in this case? Remember that we introduced a breakdown of perturbative calculations in other models. In the nonperturbative phase of the XY model, and we cite [3], "*we expect vortices to blend into the ground state - vortex condensation.*" In other words: the ground state is now composed of a superposition of the soliton solutions.

The $N = 1$ model allows us to visualize this in figure 3.2. Instead of just a Gaussian approximation around one of the minima, we state that the partition function is the sum of two Gaussians around the two minima,

$$Z \approx e^{-\mu_C} \int d\lambda\, e^{m(\lambda-\lambda_0)^2} + e^{-\mu_C} \int d\lambda\, e^{m(\lambda+\lambda_0)^2}. \tag{3.14}$$

The ground state, or the zeroth order approximation to the free energy, is hence given by a superposition of all the solitons. We can speak of a *soliton condensate.*

3.2.2 Phase transition in the $N = 2$ model

Let us pass on to the first non-trivial model, the $N = 2$ model. Now the action[4] does contain a logarithmic term, and we have the ordering (3.9). Still the Euler-Lagrange equations can be stated in one simple line,

$$m\lambda_1 + g\lambda_1^3 = \frac{2}{\lambda_1 - \lambda_2} = -m\lambda_2 - g\lambda_2^3. \tag{3.15}$$

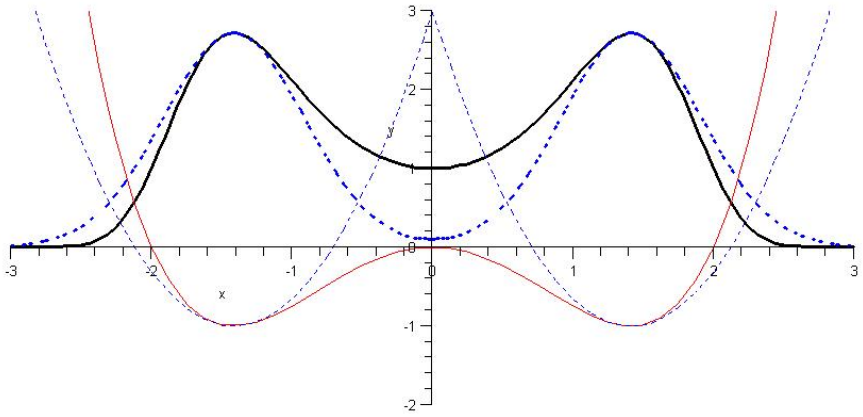


Figure 3.2: *Similar to figure 3.1 we display the $N = 1$ matrix model action S and the exponent e^{-S}. Now the dashed thin line represent the two Gaussian (quadratic) approximations around the two minima. The thick dotted line shows the superposition of the Gaussian approximations of the two minima, see equation (3.14).*

What are the solitons? Before we start performing an exact calculation, let us visualize the model. In figure 3.3 we see two contour plots and three side views of the action of the $N = 2$ model. We see that for low μ_C, there is only one (symmetric) soliton. For high μ_C there can be three solitons, among them are asymmetric ones!

Exact derivation

The upper left and upper right hand side of the Euler-Lagrange equations (3.15) together imply that either $\lambda_1 = -\lambda_2$ or that we have an asymmetric solution. Let us calculate first the symmetric soliton.

Symmetric solution

The fact that $\lambda_1 = -\lambda_2$ reduces the Euler-Lagrange equations to

$$m\lambda^2 + g\lambda^4 = 1, \tag{3.16}$$

which has one very simple solution, namely

$$\lambda^2 = \frac{-m}{2g} + \frac{1}{2g}\sqrt{m^2 + 4g} = \frac{1}{2}\lambda_0 + \frac{1}{2}\sqrt{\lambda_0^2 + 4/g} \tag{3.17}$$

where we have introduced the parameter $\lambda_0 = \sqrt{-m/g}$, the position of the minimum of the potential. If $m >> g$, we see that $\lambda \approx \lambda_0$, which means that the two eigenvalues are indeed in the two minima of the potential. In that case the energy of the configuration is $S \approx -2\mu_C$, where $\mu_C = m^2/4g$ was defined as the depth of the potential well.

[4]Unless otherwise stated, we use the action in terms of the eigenvalues (3.10).

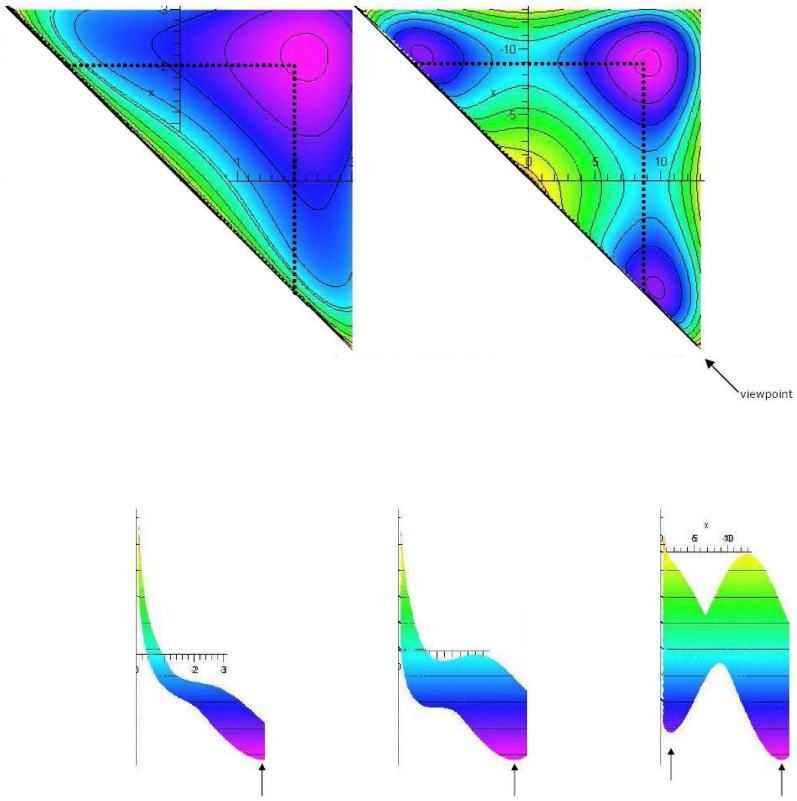


Figure 3.3: ***Above:*** *Two colored contour plots of the $N = 2$ matrix model action $S[M]$ as a function of the two eigenvalues λ_1 and λ_2, where $\lambda_1 < \lambda_2$. The dotted lines indicate the position of the minimum of the potential $V(\lambda)$. Left we see the model in the symmetric phase, with well depth $\mu_C = 1$. Clearly, there is one minimum of the action that is symmetric (so $\lambda_2 = -\lambda_1$). Right we see the model for $\mu_C = 20$. There now exist three minimal solutions of the action: one symmetric and two asymmetric solutions.* ***Below:*** *We can also compute the view of the upper plot along the direction $\lambda_1 = \lambda_2$ (see the arrow next to the upper plot). Below are these views for $\mu_C = 1, 2$ (critical) and 20. It is a nice view on the development of the asymmetric minima! All minima are pointed out with an arrow below the plot.*

Asymmetric solutions

The asymmetric solutions are obtained from rewriting the Euler-Lagrange equations (3.15) into

$$(\lambda_1 + \lambda_2)\left(m + g\lambda_1^2 + g\lambda_2^2 - g\lambda_1\lambda_2\right) = 0. \tag{3.18}$$

Hence we know that that solution must satisfy

$$\lambda_2 = \frac{1}{2}\lambda_1 + \frac{1}{2g}\sqrt{-4mg - 3g^2\lambda_1^2}. \tag{3.19}$$

The second Euler-Lagrange equation $m\lambda_1 + g\lambda^3 = 2/(\lambda_1 - \lambda_2)$ then becomes a quartic equation for λ_1^2,

$$g^2\lambda_1^8 + 3mg\lambda_1^6 + (3m^2 - 2g)\lambda_1^4 + (\frac{m^2}{g} - 2)m\lambda_1^2 + 4 = 0. \tag{3.20}$$

The substitution $x = g^{\frac{1}{2}}\lambda_1^2$ allows us to rewrite this into an equation that only depends on x and the depth of the well μ_C,

$$x^4 - 6\sqrt{\mu_C}x^3 + (12\mu_C - 2)x^2 - 2\sqrt{\mu_C}(4\mu_C - 2)x + 4 = 0. \tag{3.21}$$

A quartic equation has an exact solution by the method of Ferrari. Because this is a very complicated procedure, we will only investigate whether the system has a solution at all. The discriminant of the quartic polynomial,

$$\Delta_4 = 8\mu_C^3 - 23\mu_C^2 - 4\mu_C + 36 \tag{3.22}$$

is zero if and only if the quartic equation has multiple roots. The discriminant $\Delta_4 = 0$ with positive μ_C only at $\mu_C = 2$. After examination of the quartic equation, we find that it only has real solutions if $\mu_C \geq 2$. This implies that the asymmetric solutions only exist if $\mu_C \geq 2$.

To conclude, a numerical approximation of the energy of the asymmetric solitons suggests that their energy converges, for large μ_C, towards $S \approx -2\mu_C$.

Interpretation of the $N = 2$ model

If $\mu_C < 2$ we have only one solution so it is possible to construct a perturbation series. We can speak of a ground state (namely the symmetric soliton) and the free energy consists of the ground state plus small variations.

However, if $\mu_C \geq 2$ there are three solutions. We even found that for large μ_C the energies of the three solutions become the same! The ground state is now given by a superposition of these three states.

So we find, for this model, a *phase transition* at the critical $\mu_C^* = 2$. A rather straightforward calculation, which can be done by Maple, shows that the free energy is analytic at $\mu_C = 2$. Hence it is a *continuous phase transition*. Note that this phase transition looks similar to a Kosterlitz-Thouless phase transition: below μ_C^* we have a perturbative phase, above μ_C^* we have a nonperturbative phase where the ground state is a soliton condensate.

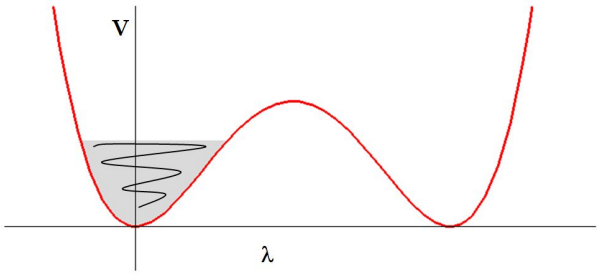


Figure 3.4: *A visualization of the complete asymmetric solution. The shifted potential energy V for the eigenvalues λ is filled up to the Fermi level μ_F around one of the two minima.*

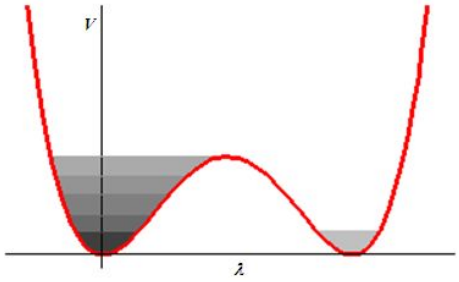


Figure 3.5: *An asymmetric soliton, though not completely asymmetric.*

3.3 Large N models

Does a continuous phase transition also arise in the case of large N matrix models? We expect it does, but it is quite impossible to solve large N models exactly.

However, we can search for the existence of asymmetric solitons. Imagine that each eigenvalue λ_i has a potential energy $V(\lambda_i)$. In this picture, we can draw a large N soliton into a graph of the potential (see figure 3.4). Suppose there exists a soliton solution where all eigenvalues are negative, the complete asymmetric soliton. Then the left potential well is 'filled up' to a certain energy "Fermi" level μ_F, whereby μ_F is of course lower than the depth of the well μ_C.

So for given model parameters (N, μ_C) we can find a Fermi level. The complete asymmetric soliton can only exist if $\mu_F(N, \mu_C) \leq \mu_C$. However, there are also other asymmetric solitons such as in figure 3.5. The soliton with one eigenvalue in the right well and $N-1$ eigenvalues in the left well could only exist if $\mu_F(N-1, \mu_C) \leq \mu_C$. If we continue with this reasoning, then there exists only one symmetric soliton if

$$\mu_F(N/2, \mu_C) > \mu_C. \tag{3.23}$$

Our computational task is thus to find the Fermi level μ_F as a function of N and μ_C. Then the phase transition occurs at $\mu_C^* = \mu_F(N/2, \mu_C^*)$.

Approximate calculation

We will compute that as follows. Take a fixed N and assume that all eigenvalues λ_i, $i = 1..N$, are in one of the potential wells. It is then convenient to shift the potential energy such that the minimum lies at $\lambda = 0$ and that $V(0) = 0$. This yields

$$\widetilde{V}(\lambda) = -m\lambda^2 + \frac{g}{4}\lambda^4 - \sqrt{-mg}\lambda^3. \tag{3.24}$$

The eigenvalues are distributed around the $\lambda = 0$ minimum such that the energy of each eigenvalue is at most the Fermi Level μ_F. Let us now state without loss of generality that $N = 2M + 1$ and that each eigenvalue is given by $\lambda_i = i\Delta + \epsilon_i$. The Euler-Lagrange equation becomes

$$\frac{\partial S}{\partial \lambda_i} = \frac{1}{i}\frac{\partial S}{\partial \Delta} + \frac{\partial S}{\partial \epsilon_i} = 0 \tag{3.25}$$

The derivative with respect to Δ involves the complete energy, and the derivative with respect to ϵ_i only the energy of one eigenvalue. We can therefore reasonably assume that the ϵ_i-dependent terms are of order $1/N$ smaller than the Δ dependent terms. Hence we neglect ϵ_i and the action becomes

$$S = \sum_{i=-M}^{M} V(i\Delta) - \sum_{i=1}^{N}(N-i)\log(i\Delta)^2, \tag{3.26}$$

and thence Euler-Lagrange equation becomes

$$\frac{\partial S}{\partial \Delta} = -\frac{1}{6}m(N^3 - N)\Delta + g\left(\frac{N^5}{80} - \frac{N^3}{24} + \frac{7N}{240}\right)\Delta^3 - \frac{N^2 - N}{\Delta} = 0. \tag{3.27}$$

We can rewrite this with implicit definition of $f_1(N), f_2(N)$ and $f_3(N)$ into

$$mf_1\Delta^2 + gf_2\Delta^4 = f_3. \tag{3.28}$$

This biquadratic equation has a simple solution that can be rewritten in terms of μ_C, g and the functions f_i.

$$\Delta^2 = \sqrt{\frac{1}{g}}\left|\frac{f_1}{f_2}\right|\left(\sqrt{\mu_C} + \sqrt{\mu_C + \frac{f_2 f_3}{f_1^2}}\right) = \sqrt{\frac{1}{g}}\,\widetilde{\Delta}^2. \tag{3.29}$$

The Fermi level is given by the energy level of the highest eigenvalue $\lambda = M\Delta = \frac{1}{2}(N-1)\Delta$. This yields, using the potential (3.24),

$$\mu_F(N, \mu_C) = V(M\Delta) = 2M^2\sqrt{\mu_C}\widetilde{\Delta}^2 + \frac{1}{4}M^4\,\widetilde{\Delta}^4 - M^3(4\mu_c)^{1/4}\,\widetilde{\Delta}^3. \tag{3.30}$$

Phase transition for large N

As explained above, the phase transition occurs whenever (3.23) becomes an equality. We can find numerical solutions using Maple[5]. In large N limit, this yields three solutions, but all of them are linear in N and all of them give $\mu_C^* = 0$ if $N = 2$. We choose the solution that is most compatible with our exact computation of the $N = 2$ model. This yields the critical value for μ_C,

$$\mu_C^* \approx 0.453N + \mathcal{O}(N^0). \tag{3.31}$$

[5]The code that we used is:

```
> f1 := N -> -1/6 * N ^ 3 + 1/6 * N:
> f2 := N -> 1/80 * N ^ 5 - 1/24 * N ^ 3 + 7/240 * N:
> f3 := N -> N ^ 2 - N:
> Delta := sqrt( -1/2 * ( m * f1(N) - sqrt( m ^ 2 * ( f1(N) ) ^ 2 + 4 * ( f2(N) ) * g * (
f3(N) ) ) ) / ( ( f2(N) ) * g ) ):
> V := lambda -> -1/4 * lambda ^ 2 * ( 4 * m - g * lambda ^ 2 + 4 * g * lambda * sqrt( -
m / g ) ):
> muF := N -> simplify( limit( subs( m = -sqrt( 4 * g * mu[c] ), V( (N-1)/2 * Delta) ),
g=1) ):
> muSol := [ solve( muF( N/2 ) = mu[c], mu[c] ) ]:
```

Chapter 4

Large N renormalization

Is the matrix dimension N an unphysical cut-off introduced in the matrix model?

We can answer that question using the techniques of the renormalization group. As Wilson and Kogut [2] stated it, renormalization is reducing the "density of degrees of freedom", without changing the overall properties of the physical system. This implies that we can integrate out degrees of freedom as to find new "effective" degrees of freedom. The free energy (the overall property of the system) will not change, due to a shift in the coupling constants m and g,

$$F(N, m, g) = F(N', m', g'). \tag{4.1}$$

So the renormalization group then consists of differential equations $\frac{\partial m}{\partial N} = f_m(N, m, g)$ and $\frac{\partial g}{\partial N} = f_g(N, m, g)$.

Besides answering the more fundamental question stated at the very beginning of this chapter, the renormalization group also forms a powerful mathematical tool to compute the free energy. After all, calculating the free energy for low N is easier that for large N. Renormalization allows us to relate the free energy of small and large N models. Note that the renormalization group flow also equals the lines of constant F.

In chapter 2 we saw that the matrix model for $N \to \infty$ can be related to 2d quantum gravity. In this case it is useful to look for fixed points of the renormalization group. If the model is at a fixed point (m^*, g^*), the free energy does not depend on the renormalization parameter N. This implies that if $N \to \infty$, the model converges to a fixed point if it converges at all. Hence fixed points of a matrix model coincide with a quantum gravity theory.

In this chapter we will first cite the results of Higuchi et al. [16] who derived a renormalization group of the matrix model to find the fixed points. We will then use their techniques to find the renormalization group flow of the gM^4 model. The central question there is: do the levels of constant free energy (the RG flow) cross the line of the phase transition that we found in the previous chapter?

4.1 Higuchi's renormalization group

In 1994 Higuchi et al. [16] derived a renormalization group equation, based on the work of Brézin and Zinn-Justin [17], for the general matrix model

$$Z_N(g_j) = \int dM \ \exp\left[-N\mathrm{Tr}\sum_{k\geq 1}\frac{g_k}{k}M^k\right] \tag{4.2}$$

and free energy defined by

$$F(N, g_j) = -\frac{1}{N^2}\log\left[\frac{Z_N(g_j)}{Z_N(g_2 = 1, \text{others} = 0)}\right]. \tag{4.3}$$

The renormalization group equation took the form of a Callan-Symanzik equation that relates the change in N with the change in all the coupling constant g_j,

$$\left[N\frac{\partial}{\partial N} + 2\right]F(N, g_j) = G\left(g_3, \ldots, g_m; \frac{\partial F}{\partial g_3}, \ldots, \frac{\partial F}{\partial g_m}\right) + \mathcal{O}\left(\frac{1}{N}\right). \tag{4.4}$$

The function G was found by integrating over the $(N+1)$-th eigenvalue in the Z_{N+1} partition function using the saddle point method. Of course, this only yields a solution for high N; the contributions of order $\mathcal{O}\left(\frac{1}{N}\right)$ are neglected. We will use this approach in the next chapter to find the running of the M^4 model coupling constants.

To conclude, Higuchi et al. also found three non-trivial fixed points in the model with M^3 and M^4 interactions and a fixed Gaussian term.[1]

4.2 Large N renormalization

We will now perform a renormalization based on the saddle-point method introduced in the previous section. We integrate out the $(N+1)$th eigenvalue, $\lambda_{N+1} \equiv \lambda$, which yields

$$Z_{N+1} = \int d\lambda_i \prod_{i<j}^{N}(\lambda_i - \lambda_j)^2 \ e^{-\sum_i^N V(\lambda_i)} \int d\lambda \ e^{-V(\lambda)+\sum_i^N \log(\lambda-\lambda_i)^2}. \tag{4.5}$$

The 'high part' of the partition function equals the renormalization of the potential

$$e^{-\delta S} = \int d\lambda \ e^{-V(\lambda)+\sum_i^N \log(\lambda-\lambda_i)^2} = \int d\lambda \ e^{-\widetilde{S}(\lambda)}. \tag{4.6}$$

We will use the method of steepest descent to evaluate this integral, where the saddle point λ_s is given by the equation

$$m\lambda_s + g\lambda_s^3 = \frac{2N}{\lambda_s} + \sum_{k=1}^{\infty}\frac{2}{\lambda_s^k}\mathrm{Tr}\Lambda^k = \frac{2N}{\lambda_s} + \mathrm{Tr}\frac{2\Lambda}{\lambda_s I - \Lambda}, \tag{4.7}$$

[1] Higuchi relates the fixed points to theories of (p, q)-minimal conformal matter coupled to two-dimensional quantum gravity. They found $(p = 2, q = 5)$ and $(p = 2, q = 3)$ fixed points.

where Λ is the $N \times N$ matrix of eigenvalues λ_i. This equation causes some trouble since it has, to quote Higuchi, "several branches of solutions". owever, large N suggests that the model is in the perturbative phase. This allows us to assume $\lambda_s > \Lambda$ and so we can neglect higher order terms. The saddle-point equation can be simplified to a biquadratic

$$m\lambda_s^2 + g\lambda_s^4 = 2N. \tag{4.8}$$

The solutions are, in terms of μ_C, g and N,

$$\lambda_s^2 = 2\sqrt{1/g}\left(\sqrt{\mu_C} + \sqrt{\mu_C + 2N}\right). \tag{4.9}$$

Now note that the whole system is invariant if we send all eigenvalues to it's negative, $\lambda_i \to -\lambda_i$. If we choose only one of the saddle-points given by (4.9), we break that symmetry which will give us odd powered terms in the effective action. A better approximation is to sum over both minima of the exponent, such that

$$e^{-\delta S} \approx e^{-\widetilde{S}(\lambda_s)} + e^{-\widetilde{S}(-\lambda_s)}, \tag{4.10}$$

and the change in the effective action equals $\delta S \approx \widetilde{S}(\lambda_s) + \widetilde{S}(-\lambda_s)$. The only terms that contribute are the logarithmic part since (4.8) implies that the potential equals $2N$. Hence

$$\begin{aligned} \delta S &= -\mathrm{Tr}\log(\lambda_s - \Lambda)^2 - \mathrm{Tr}\log(\lambda_s + \Lambda)^2 \\ &= -\log\lambda_s^2 + \frac{2}{\lambda_s^2}\mathrm{Tr}\Lambda^2 + \frac{1}{\lambda_s^4}\mathrm{Tr}\Lambda^4 + \mathcal{O}\left(\Lambda^6\right). \end{aligned} \tag{4.11}$$

The effective potential becomes

$$V_{\text{eff}}(\lambda_i) = \frac{1}{2}\left(m_{(N+1)} + \frac{4}{\lambda_s^2}\right)\lambda_i^2 + \frac{1}{4}\left(g_{(N+1)} + \frac{4}{\lambda_s^4}\right)\lambda_i^4 + \mathcal{O}(\lambda_i^6). \tag{4.12}$$

The coupling constants m and g change according to

$$m_N - m_{(N+1)} = \delta m = +\frac{4}{\lambda_s^2}, \tag{4.13}$$

$$g_N - g_{(N+1)} = \delta g = +\frac{4}{\lambda_s^4}. \tag{4.14}$$

Note that λ_s depends on the parameters $m_{(N+1)}$ and $g_{(N+1)}$. Also remark that λ_s^2 is positive so $m_N > m_{(N+1)}$. The corresponding renormalization of the top of the potential μ_C can be derived from

$$\mu_C(N) = \frac{(m_{(N+1)} + 4/\lambda_s^2)^2}{4(g_{(N+1)} + 4/\lambda_s^4)} \tag{4.15}$$

$$= \frac{(\lambda_s^2 m_{(N+1)} + 4)^2}{4(g_{(N+1)}\lambda_s^4 + 4)}. \tag{4.16}$$

We can now fill in λ_s. Define for the moment $\mu'_C \equiv \mu_C(N+1)$ so that we can write

$$\lambda_s^2 m_{N+1} = -4(\mu'_C + \sqrt{\mu'^2_C + 2N\mu'_C}), \tag{4.17}$$

$$\lambda_s^4 g_{N+1} = 8\mu'_C + 8\sqrt{\mu'^2_C + 2N\mu'_C} + 8N, \tag{4.18}$$

and

$$\mu_C(N) = \frac{(1 - \mu'_C - \sqrt{\mu'^2_C + 2N\mu'_C})^2}{1 + 2\mu'_C + 2\sqrt{\mu'^2_C + 2N\mu'_C} + 2N}. \tag{4.19}$$

We can write $\frac{\partial \mu_C}{\partial N} = \mu_C(N+1) - \mu_C(N)$, so

$$\frac{\partial \mu_C}{\partial N} = \frac{3\mu_C - 1 + 2\sqrt{\mu_C^2 + 2N\mu_C}}{2\mu_C + 1 + 2\sqrt{\mu_C^2 + 2N\mu_C} + 2N}. \tag{4.20}$$

A numerical approximation is given by Maple, which yields

$$\mu_C(N) \approx 0.825N + \mu_C(0) + \mathcal{O}\left(\frac{1}{N}\right). \tag{4.21}$$

4.3 A continuous phase transition

In figure 4.1 are the lines of constant F drawn together with the phase transition found in equation (3.31). First we note that the the lines of constant F lead in the large N limit most of the times to a system in the Fermi phase, how counterintuitive it might seem. Renormalization towards low N will eventually lead to positive m and thus a simple perturbative theory, or m will remain negative and the system will remain in the nonperturbative phase.

Secondly we note that the renormalization of the gM^4 model does not lead to fixed points that can be related to $N \to \infty$.

The third remarkable feature of the phase diagram 4.1 is the fact that the lines of constant F do cross the phase transition at all. This is again a confirmation of our result from chapter 3 that the phase transition is analytic in the macroscopic variable Z.

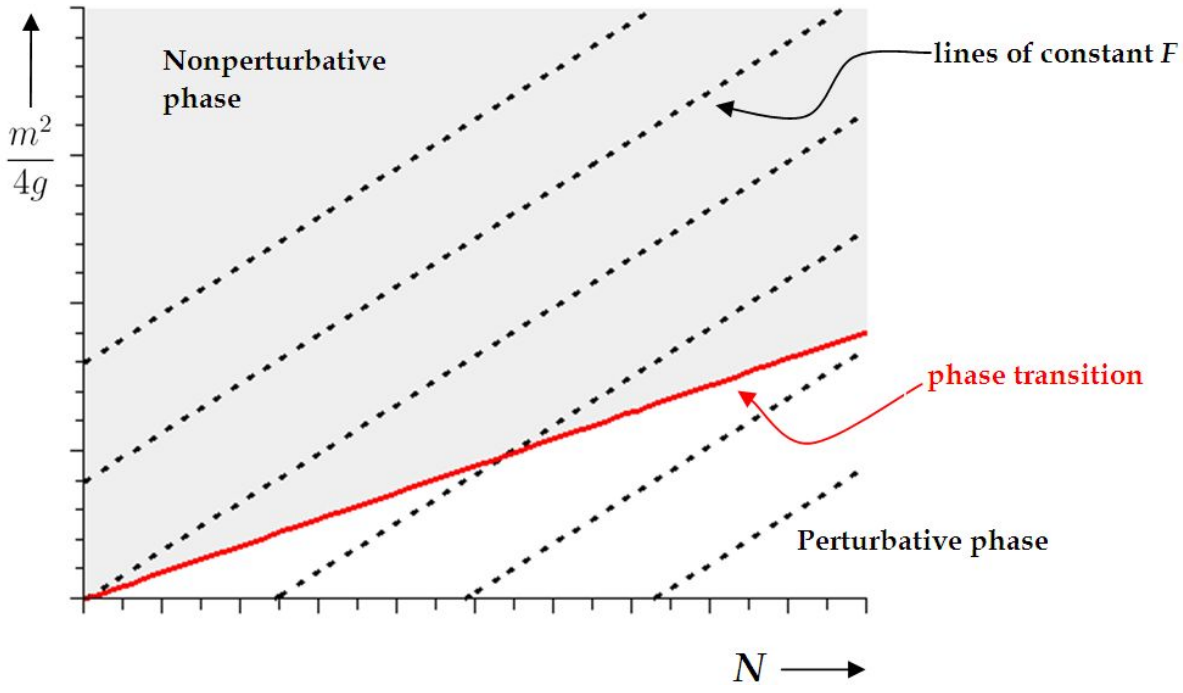


Figure 4.1: *The renormalization group flow of μ_C solved numerically with (4.21), which are the lines of equal free energy $F(N, m, g)$. Drawn together with the phases of the gM^4 matrix model given by (3.31).*

Chapter 5

Nonperturbative solutions

For small N and large μ_C, the gM^4 matrix model in the eigenvalue representation has many classical solutions with similar energy. In this nonperturbative phase the free energy could not be related to a perturbation series around the absolute minimum.

But what is it that makes the nonperturbative phase different from the perturbative phase? In chapter 1 we found that the occurrence of vortices drove the XY model into a Kosterlitz-Thouless phase transition. That phase transition looks similar to the one we found in the matrix model. So a logical question is: What are the vortices of the matrix models?

However, the only vortices that are yet described require the existence of a target space. The matrix models that we study here do not have a target space. Is it then possible to integrate out degrees of freedom to remove the nonperturbative phase of the matrix model?

5.1 Matrix models on a target space

In 1990 there was a series of articles published on the topic of nonperturbative solutions of matrix models [18, 19, 20, 21]. Although the matrix models that these publications discussed are not the same as ours, it is educational to review their results. The model was introduced by [22] and was defined by the path integral

$$Z_N(\beta) = \int \mathcal{D}^{N^2} M(t)\; e^{-\beta \int dt\, \mathrm{Tr}\left[\frac{1}{2}(\partial_t M)^2 - U(M)\right]}. \tag{5.1}$$

Note that in this model the matrix $M(t)$ depends on a time parameter t. The time parameter runs over the so-called target space.[1] A standard solution of this model is given by a transformation to the eigenvalues of $M(t)$,

$$M(t) = \Omega(t)\; \Lambda(t)\; \Omega^\dagger(t). \tag{5.2}$$

[1] The name 'target space' originally comes from string theory, where the D-dimensional spacetime is called the target space [23]. The matrix models mentioned here have a 1-dimensional spacetime, so the parameter t can also be called 'time'.

The integration measure changes with a factor equal to the Vandermonde determinant squared [15],

$$\mathcal{D}M = \mathcal{D}\Omega\mathcal{D}\Lambda \prod_{i<j}(\lambda_i - \lambda_j)^2. \tag{5.3}$$

Since the Vandermonde determinant is completely antisymmetric, the model is reduced to the problem of N non-interacting fermions λ_i with Hamiltonian

$$\hat{h} = -\frac{1}{2\beta^2}\frac{\partial^2}{\partial\lambda_i^2} + U(\lambda_i). \tag{5.4}$$

In the ground state of a fermionic system the energy levels of this Hamiltonian are filled up to it's Fermi level μ_F. As long as the Fermi level is lower than the top of the potential μ_C, the system can be described by means of perturbation theory. The nonperturbative behavior pops up when $\mu = \mu_C - \mu_F \approx 0$. The way this phase transition arises is similar to the phase transition of our matrix model.

The Fermi level μ_F can be found via the density of eigenstates e_n of the system

$$\rho(e) = \frac{1}{\beta}\sum_n \delta(e - e_n). \tag{5.5}$$

Gross and Miljkovic [18] found an exact differential equation that describes the density of states near it's Fermi level,

$$\frac{\partial\rho}{\partial\mu}(\mu_F) = \frac{1}{2\pi\beta\mu}\mathrm{Im}\int_0^\infty dt\; e^{-it}\;\frac{t/\beta\mu}{\sinh(t/\beta\mu)}. \tag{5.6}$$

The free energy contributions from the singlet sector (that is, the energy that only depends on the eigenvalues) is proportional to

$$F_s \sim N^2\mu^2|\ln\mu|. \tag{5.7}$$

However, due to the kinetic part $(\partial_t M)^2$ the path integral still depends on the angular part of the matrix M. Gross and Klebanov [21] transformed the path integral into a Hamiltonian for the eigenvalues and the angular momenta

$$\hat{H} = -\frac{1}{2\beta^2\Delta(\lambda)}\sum_i \frac{\partial^2}{\partial\lambda_i^2}\Delta(\lambda) + \sum_i U(\lambda_i) + \sum_{i<j}\frac{\Pi_{ij}^2 + \widetilde{\Pi}_{ij}^2}{(\lambda_i - \lambda_j)^2}, \tag{5.8}$$

where λ_i are the eigenvalues, $\Delta(\lambda)$ is the Vandermonde determinant and Π_{ij} and $\widetilde{\Pi}_{ij}$ are the generators of left rotations on the angular part $\Omega(t)$. This angular part of the Hamiltonian contributes to the ground state energy whenever the target space is a finite circle with radius R. The free energy contribution of the non-singlet sector is proportional to

$$F_{ns} \sim N^2\mu^R. \tag{5.9}$$

If $R < 2$ the non-singlet sector would dominate the free energy when $\mu \to 0$. If $R > 2$ then the singlet sector dominates the energy. This suggests the existence of a Kosterlitz-Thouless phase transition at $R = 2$.

The duality $R \to 1/R$ is destroyed by the KT phase transition. Below $R = 2$ the vortices, which are actually windings of the angular part around the target space, become dominant.

5.1.1 Derrick's theorem

In the search for nonperturbative solutions within the framework of quantum field theories, Derrick [24] found a theorem that limits the occurrence of these soliton solutions. We can look whether Derrick's theorem holds for matrix theories with a target space.

A soliton solution is defined as a stable solution of the classical Euler-Lagrange equation, see section 3.1.1. Derrick showed that a scalar field theory with D spatial dimensions,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - U[\phi], \tag{5.10}$$

could only have soliton solutions if $D = 1$. [25]

Let us generalize this to matrix models. Assume that we have a matrix model with a D-dimensional Euclidian target space

$$S = \int d^D x \, \mathrm{Tr} \left[\frac{1}{2}(\partial_\mu M)^2 + U(M) \right] \tag{5.11}$$

and assume that $U(M) \geq 0$. We have equivalence $U(M) = 0$ when M is a vacuum solution. Now suppose $M_s(x)$ is a soliton solution with energy $S = K + V$, K is the kinetic energy term and V the potential energy term. Now the scaled solution $M_s(x/a)$ has energy

$$S(a) = a^{D-2} K + a^D V. \tag{5.12}$$

Because we demanded that $M_s(x)$ was a stable solution, $S(a)$ must satisfy the differential equation

$$\left. \frac{\partial S(a)}{\partial a} \right|_{a=1} = (D-2)K + DV = 0. \tag{5.13}$$

Because K and V are both positive, this could only make sense if either $K = V = 0$ or if $D < 2$. The solitons obtained by Gross et al in the previous section were an example of the theories solitons in the $D = 1$ case.

5.1.2 Analysis of the target space models

Gross et al attribute the existence of vortices to the fact the the target space is a finite circle. This is compatible with Derrick's argument that vortices could only exist if the target space is one-dimensional.

However, the matrix model that we study has no target space. Hence the conclusions of the above discussion could not help us with the search for the nonperturbative solutions of the matrix model. The method of analysis nevertheless proved useful. In the previous sections we used the eigenvalue method and the concept of a Fermi level μ_F to find the phase transition.

It is interesting to note that the both the phase transition and the "fermionic" behavior occur in matrix models with or without target space. But if we remove the target space, the vortices are removed too. It interesting to ask ourselves: What is it then that drives the phase transition in the absence of a target space?

5.2 Integrating out the nonperturbative phase

As we noticed in chapter 3, the nonperturbative phase is characterized by $(N+1)$ distinct eigenvalue configurations. These solitons can be classified by the number of eigenvalues in one of the wells. Another parameter to order the solitons is the mean (or equivalently: sum) of all the eigenvalues

$$s = \sum_i \lambda_i. \tag{5.14}$$

We can test the 'reality' of the nonperturbative phase transition by integrating out the parameter s, which would presumably reduce the number of solitons. Such an integral can be evaluated if we make a linear transformation to coordinates s and $\Delta_i = \lambda_i - \lambda_{i+1}$, with $i = 1..(N-1)$. We set $\Delta_N \equiv s$. The transformation matrix defined by $\Delta_{\mathtt{i}} = (T_N)^{\mathtt{j}}_{\mathtt{i}}\lambda_{\mathtt{j}}$ is

$$T_N = \begin{bmatrix} 1 & -1 & 0 & \dots & & \\ 0 & 1 & -1 & & & \\ \vdots & & \ddots & \ddots & & \\ & & & & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{5.15}$$

The determinant of T_N can be easily found via the property that $\det T_N = \det T_{(N-1)} + 1$, hence $\det T_N = 1$. The inverse transformation matrix is given by

$$T_N^{-1} = \frac{1}{N}\begin{bmatrix} (N-1) & (N-2) & (N-3) & \dots & 1 & 1 \\ -1 & (N-2) & & & 1 & 1 \\ -1 & -2 & \ddots & & & \vdots \\ \vdots & & & & 1 & 1 \\ -1 & -2 & \dots & -(N-2) & -(N-1) & 1 \end{bmatrix}. \tag{5.16}$$

The complexity of the matrix T_N^{-1} shows that the resulting action $S[\Delta_i, s]$ is quite complicated. However, notice that the action has three main features.

1. The logarithmic terms $\sum_{i<j} \log(\lambda_i - \lambda_j)^2$ do not depend on the sum s.

2. The action is symmetric under interchange of Δ_i for $i = 1..(N-1)$.

3. The action is invariant under a change of sign of all Δ_i and s.

Especially the last property is important: it implies that the s-dependent terms of the action are just a quartic function. As long as g is positive, that part can be integrated out exactly. This is unfeasible for large N, we will therefore perform this calculation for the smallest non-trivial $N = 2$.

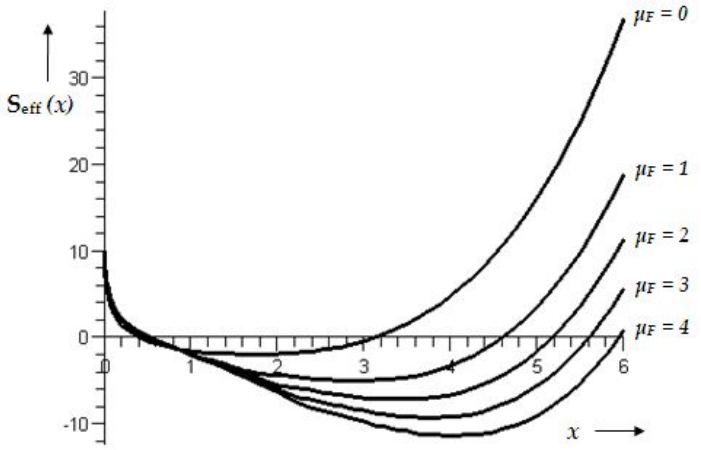


Figure 5.1: *The effective action of the $N = 2$ model as a function of $x = g^{1/4}\Delta = g^{1/4}(\lambda_1 - \lambda_2)$ for $\mu_C = 0, 1, 2, 3$ and 4.*

$N = 2$ nonperturbative phase

The partition function is, after transformation to $\Delta = \Delta_1$ and s given by

$$Z_2 = \frac{1}{2}\int d\Delta\, ds\; e^{-\frac{1}{4}m\Delta^2 - \frac{g}{32}\Delta^4 - \left(\frac{1}{4}m + \frac{3}{16}g\Delta^2\right)s^2 - \frac{g}{32}s^4 + \log\Delta^2}. \tag{5.17}$$

The parameter $s \in (-\infty, \infty)$ can be integrated out exactly yielding

$$\begin{aligned} Z_2 &= \frac{1}{(2g)^{1/4}}\int d\Delta\, \left(\frac{(4m + 3g\Delta^2)^2}{g}\right)^{1/4}\; K_{(1/4)}\left[\frac{1}{64g}(4m + 3g\Delta^2)^2\right] \\ &\quad \times\Delta^2 \exp\left\{\frac{1}{64g}(16m^2 + 8mg\Delta^2 + 7g^2\Delta^4)\right\}, \end{aligned} \tag{5.18}$$

where $K_\alpha[x]$ is the modified Bessel K function. Let us first rewrite this in terms of μ_C,

$$\begin{aligned} Z_2 &= \frac{1}{(4g)^{1/4}}\int d\Delta\, \sqrt{\left|3g^{\frac{1}{2}}\Delta^2 - 8\mu_C^{\frac{1}{2}}\right|}\; K_{(1/4)}\left[\left(\frac{3}{8}g^{\frac{1}{2}}\Delta^2 - \mu_C^{\frac{1}{2}}\right)^2\right] \\ &\quad \times\Delta^2 \exp\left\{\mu_C - \frac{1}{4}\mu_C^{\frac{1}{2}}g^{\frac{1}{2}}\Delta^2 + \frac{7}{64}g\Delta^4\right\}. \end{aligned} \tag{5.19}$$

We see that our parameter Δ always occurs with a factor $g^{1/4}$. That suggests a shift to the parameter $x = g^{1/4}\Delta$ which yields

$$\begin{aligned} Z_2 &= \frac{1}{\sqrt{2}g}\int dx\, \sqrt{\left|3x^2 - 8\mu_C^{1/2}\right|}\; K_{(1/4)}\left[\left(\frac{3}{8}x^2 - \mu_C^{1/2}\right)^2\right] \\ &\quad \times x^2 \exp\left\{\mu_C - \frac{1}{4}\mu_C^{1/2}x^2 + \frac{7}{64}x^4\right\} \end{aligned} \tag{5.20}$$

$$\equiv \frac{1}{\sqrt{2}g}\int dx\; e^{-S_{\rm eff}[x;\mu_C]}. \tag{5.21}$$

The effective action $S_{\text{eff}}[x;\mu_C]$ can be calculated for various μ_C. It is not defined at $x^2 = \frac{8}{3}\sqrt{\mu_C}$ but is continuous at that point. A plot of the effective action for several μ_C including the critical $\mu_C = 2$ is shown in figure 5.1.

Apart from the $x \to -x$ symmetry (which can be removed easily) the system clearly has one absolute minimum which allows perturbation theory, without constraint on μ_C. This implies that after integrating out $s = \lambda_1 + \lambda_2$, the continuous phase transition disappears.

5.3 Does the correlation function change?

In the previous section we found that we could integrate out the phase transition in the $N = 2$ model. This leads of course to speculations on the reality of that phase transition. Since that phase transition resembles a Kosterlitz-Thouless transition, we can now use the same techniques that were applied to the KT transition 1.3.2: by calculating the correlation function Γ.

The correlation function expresses relations between different microscopic degrees of freedom. So in the eigenvalue representation, the correlation function is defined by

$$\Gamma(m,n) =< \lambda_m \lambda_n > . \tag{5.22}$$

which means: it is the expectation value of $\lambda_m\lambda_n$. If we properly normalize it, then $\Gamma \approx 1$ if $\lambda_m \approx \lambda_n$. If $\Gamma \approx 0$, this implies that for any fixed λ_m, the probability for all values of λ_n is equal. However, we find that the meaning of a correlation function in a matrix model is somewhat distorted:

- The action couples each pair of eigenvalues via the Vandermonde term $\log(\lambda_m - \lambda_n)^2$. Hence "distance" has no intrinsic meaning in the matrix model, which makes differences between "short- and long-range" correlations also meaningless.

- Secondly, we had to impose an ordering onto the eigenvalues.[2] This in turn automatically induces a correlation: $\lambda_m \approx \lambda_n$ if $m \approx n$, and $\lambda_m \approx -\lambda_n$ if $m \approx N - n$.

It might still be useful to estimate the correlation function in the low and high μ_C regime. Even though it has no *meaning*, we might still see *mathematical* differences in the two phases of the matrix model.

So let us calculate the correlation function for the $N = 2$ model. Instead of computing $\langle\lambda_1\lambda_2\rangle$ we might as well calculate $\Gamma_2 \equiv \langle s^2\rangle$ in the basis introduced in the previous section. The correlation function now explicitly equals

$$\Gamma_2 = \frac{1}{4}\int_{-\infty}^{\infty} d\Delta \int_{-\infty}^{\infty} ds\ s^2\ e^{-\frac{1}{4}m\Delta^2 - \frac{g}{32}\Delta^4 - \left(\frac{1}{4}m + \frac{3}{16}g\Delta^2\right)s^2 - \frac{g}{32}s^4 + \log\Delta^2}. \tag{5.23}$$

[2] Recall that $\lambda_1 \le \lambda_2 \le ... \le \lambda_N$.

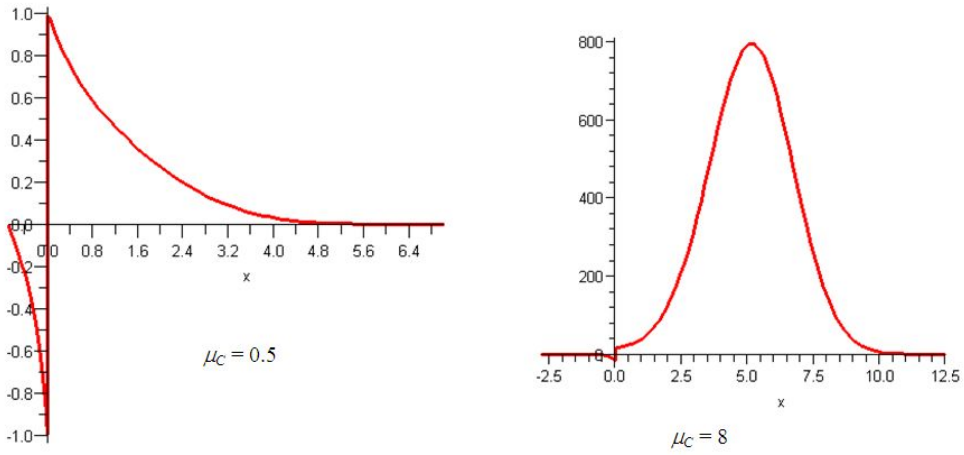


Figure 5.2: *The integrand G_{int} as a function of ξ for low $\mu_C = 0.5$ and 8. For low μ_C the integrand falls of exponentially. For high μ_C we see a Gaussian.*

First, we can perform the integral over Δ which yields

$$\Gamma_2 = -\frac{\sqrt{2}}{16g^{5/4}} \int ds\ s^2 \left(4m + 3gs^2\right) \tag{5.24}$$

$$\times \left[K_{(1/4)} \left(\frac{1}{64g}(4m + 3gs^2)^2 \right) - K_{(3/4)} \left(\frac{1}{64g}(4m + 3gs^2)^2 \right) \right] \tag{5.25}$$

$$\times\ \exp\left\{ \frac{1}{64g}(16m^2 + 8mgs^2 + 7g^2s^4) \right\} \left[\frac{1}{g}(4m + 3gs^2)^2 \right]^{1/4}. \tag{5.26}$$

By symmetry arguments we can change the integration domain to $(0, \infty)$. Then we are able to substitute $8\xi\sqrt{g} = (4m+3gs^2)$ such that the integration domain becomes $\xi \in (-\sqrt{\mu_C}, \infty)$. The correlation function equals

$$\Gamma_2 = \frac{4\sqrt{6}}{9g^2} \int_{-\sqrt{\mu_C}}^{\infty} d\xi\ \xi\ |\xi|^{1/2} \left[K_{(3/4)}(\xi^2) - K_{(1/4)}(\xi^2) \right]$$

$$\times e^{\frac{5}{18}\mu_C + \frac{8}{9}\sqrt{\mu_C}\xi + \frac{7}{9}\xi^2} \sqrt{\xi + \sqrt{\mu_C}} \tag{5.27}$$

$$\equiv \frac{4\sqrt{6}}{9g^2} \int_{-\sqrt{\mu_C}}^{\infty} d\xi\ G_{int}(\xi, \mu_C). \tag{5.28}$$

This integrand G_{int} has significant different behavior for low and high μ_C, see figure 5.2. We can therefore approximate the correlation function Γ_2 for those two phases.

Approximation for low μ_C

If μ_C is low, we can approximate the integrand with an exponential $G_{int} \sim a \exp -bx$. To find the parameter a and b we only need to find a linear expansion of G_{int} around $\xi = 0$.

First, we expand each term of G_{int},

$$\left[K_{(3/4)}(\xi^2) - K_{(1/4)}(\xi^2)\right] = \frac{1}{2}2^{3/4}\Gamma(3/4)\xi^{-3/2} - \frac{\pi 2^{3/4}}{2\Gamma(3/4)}\xi^{-1/2} + \mathcal{O}(\xi^{1/2}), \tag{5.29}$$

$$\xi\,|\xi|^{1/2}\,\sqrt{\xi+\sqrt{\mu_C}} = \mu_C^{1/4}\xi^{3/2} + \frac{\xi^{5/2}}{2\mu_C^{1/4}} + \mathcal{O}(\xi^{7/2}), \tag{5.30}$$

$$e^{\frac{5}{18}\mu_C + \frac{8}{9}\sqrt{\mu_C}\xi + \frac{7}{9}\xi^2} = e^{\frac{5}{18}\mu_C}\left[1 + \frac{8}{9}\sqrt{\mu_C}\xi + \mathcal{O}(\xi^2)\right]. \tag{5.31}$$

This yields all together a linear approximation of (5.27):

$$G_{int} = e^{\frac{5}{18}\mu_C}2^{3/4}\left\{\frac{1}{2}\Gamma\left(\frac{3}{4}\right)\mu_C^{1/4} + \left(\frac{4}{9}\Gamma\left(\frac{3}{4}\right)\mu_C^{3/4} - \frac{\pi\mu_C^{1/4}}{2\Gamma\left(\frac{3}{4}\right)} + \frac{\Gamma\left(\frac{3}{4}\right)}{4\mu_C^{1/4}}\right)\xi + \mathcal{O}(\xi^2)\right\}. \tag{5.32}$$

From this linear function we can read of the parameters a and b that define the exponential approximation. The integral is now simple, and this yields in leading order to the following dependence of Γ_2 on μ_C,

$$\Gamma_2(\text{low } \mu_C) \sim \mu_C^{3/4}e^{\frac{5}{18}\mu_C}. \tag{5.33}$$

Approximation for high μ_C

For large μ_C, the integrand μ_C resembles a Gaussian function. The top of the Gaussian is approximately at $\xi_0 \approx 2\sqrt{\mu_C - 2} \approx 2\mu_C^{1/2}$. Our approximation is given by

$$G_{int} \sim G_{int}(\xi_0)\ \exp\left\{-\frac{1}{2}\left|\frac{\partial^2 \log G_{int}}{\partial\xi^2}\right|_{\xi=\xi_0}(\xi-\xi_0)^2\right\}. \tag{5.34}$$

This equals, in leading order,

$$G_{int} \approx 2\sqrt{6}\left[K_{(3/4)}(4\mu_C) - K_{(1/4)}(\mu_C)\right]\mu_C\, e^{\frac{31}{6}\mu_C}\ \exp\left\{-\frac{2}{9}\left(\xi - 2\sqrt{\mu_C}\right)^2\right\}. \tag{5.35}$$

The correlation function Γ_2 now depends on μ_C in the following fashion,

$$\Gamma_2(\text{high } \mu_C) \sim \mu_C^{-1/2}\, e^{\frac{7}{6}\mu_C}. \tag{5.36}$$

Non-analyticity in the sources

We indeed see different mathematical behavior of the correlation for low μ_C (5.33) and for high μ_C (5.36). The approximate calculations suggest that there is a non-analyticity at the phase transition $\mu_C^* = 2$ in the correlation function. That, in turn, would imply that the partition function is non-analytic with respect to the source term coupled to the eigenvalues.

Let us make this picture somewhat clearer. A source term coupled to the eigenvalues is $\sum_i J_i \lambda_i$, so that the correlation function equals

$$\Gamma(i,j) \equiv \langle \lambda_1 \lambda_2 \rangle = \frac{\partial}{\partial J_i} \frac{\partial}{\partial J_j} Z(J) \bigg|_{J=0} . \tag{5.37}$$

So a non-analytic correlation function implies that Z is non-analytic with respect to J. Remark that the sources J preselect one specific set of microscopic variables. The appearance of the phase transition is therefore directly linked to a specific choice of microscopic variables.

Chapter 6

Conclusion & Discussion

Matrix models are toy models applicable in various fields of physics. The overall properties of such a matrix model are defined by its partition function, which is an integral over $N \times N$ Hermitian matrices M with energy/action $S[M]$ invariant under similarity transformations. Upon integrating over the rotational degrees of freedom, we described the action in terms of the eigenvalues of M. In that eigenvalue representation of the gM^4 model we found a phase transition. We calculated the critical parameter μ_C^*, and we found that renormalization could lead to a phase transition. It is that phase transition that remained intriguing. Therefore we will try to give an interpretation of the phase transition in the gM^4 matrix model.

6.1 An interpretation of the gM^4 phase transition

We studied, throughout this whole thesis, various exact classical solutions of the gM^4 matrix model. Thereby we used three bases of microscopic degrees of freedom. Let us summarize these three:

1. Equation (3.5) states the action of the matrix model in terms of the matrix elements $M_{\mathtt{a}}^{\mathtt{b}}$. That action had several continuous sets of solutions for all μ_C. Perturbation theory was not applicable.

2. Upon integrating of the $U(N)$ symmetry we described the model in terms of the eigenvalues λ_i, in equation (3.10). The model can be calculated perturbatively whenever $\mu_C < \mu_C^*$, where the critical μ_C is given by (3.31). For $\mu_C \geq \mu_C^*$ there are at most $(N+1)$ distinct solitons. At the critical point there exists a phase transition. We will call this a gM^4 *phase transition*. The correlation function Γ behaves different in the two phases (section 5.3).

3. The degeneracy in the eigenvalue solitons can be removed when we integrate out $s = \sum_i \lambda_i$ via the transformation (5.15). Now perturbation theory is always possible and there exists no phase transition. (This is only confirmed explicitly for $N = 2$.)

Thus on *mathematical grounds* we can conclude that the appearance of the gM^4 phase transition depends wholly on the *choice* of our microscopic degrees of freedom.

On *physical grounds* however, this does not make sense. How can a (physical) phase transition disappear if we only change the description of the system? The answer may be found in the relation between the microscopic degrees of freedom and the collective properties of the system as a whole.

Mathematically speaking all collective (macroscopic) properties can be expressed in terms of the partition function Z. The partition function Z is just a weighted sum of all configuration of some set of microscopic degrees of freedom. In mathematical terms, we can change the set of microscopic variables via either renormalization, transformation or just plain integration.

Now let us compare the gM^4 transition to the physics of a Kosterlitz-Thouless (KT) phase transition. In original derivations, a KT transition was found whenever the *correlation function* had different behavior for low and high T.[1] The definition of the correlation function is coupled to an explicit choice of microscopic variables. For example $\Gamma(n) = \langle s(n)s(0)\rangle$ in a lattice spin system, or equivalently the propagator $\Gamma \sim \frac{\partial^2}{\partial J^2} Z(J)\Big|_{J=0}$ in quantum field theories.[2] In all the models we encountered, a KT transition didn't change the macroscopic properties of the system. This is similar to the gM^4 matrix model, where the phase transition showed analyticity in the macroscopic variables but non-analyticity in the microscopic variables.

So remember what changed at the offset of a KT phase transition. The standard explanation of Kosterlitz and Thouless themselves was that the phase transition was driven by blending vortices into the ground state. This explanation only holds for the planar XY model, since the gM^4 model doesn't have vortices. A generalization could be that we interpret the gM^4 phase transition as a *shift in symmetry of a specific set of microscopic variables.*

Hence the gM^4 phase transition can be characterized by

- analyticity in the macroscopic variables;
- non-analyticity in a specific set of microscopic variables (when coupled to a source term in the action).

With this characterization the gM^4 resembles a Kosterlitz-Thouless phase transition.

It is, in this context, worthwhile to note the philosophical remarks by Fisher [26]. In his introduction on renormalization group theory he asks himself: How far is some concept only instrumental? And to what extent embody useful ideas and pictures an essence of reality?

[1] For the Ising model, see section 1.3.1. For the planar model, see section 1.3.2. In both cases the correlation function depended via a power law (low T) or exponentially (high T) on the spin site distance $|n|$.

[2] The source J is directly coupled to the microscopic degrees of freedom $\phi(x)$ since a quantum field theory action includes the term $-J\phi$.

Similarly we can ponder on the reality of the microscopic variables. Mathematically, we are free to change our description to other microscopic variables. Physically however, given the above discussion, the "reality" of the gM^4 phase transition is put *at the same level of reality* as the specific choice of microscopic variables.

6.2 Discussion

There are a few aspects of our interpretation of matrix models that need further attention.

Strict $N \to \infty$ limit

In this research we only focussed on finite N matrix models: from $N = 1$ to large N. It is interesting to compare our finite N results with earlier works on infinite N matrix models. Shimamune [27] derived the $N \to \infty$ limit of the M^4 model with parameter m negative. For $\mu_C \leq 4$, there exists only one (symmetric) solution to the Euler-Lagrange equation. For $\mu_C > 4$, there exists a symmetric solution ánd an asymmetric solution. Shimamune finds that at $\mu_C^* = 4$ there exists a third order phase transition.[3] So a matrix model for $N \to \infty$ shows behavior similar to the finite N models. There are two main differences, though. The analytic phase transition for finite N becomes third order in the infinite N limit. This can be easily explained as a feature of the limit: the third order critical point is an endpoint in the phase space. Secondly, the critical value for μ_C is not the large N limit of our result (3.31). This suggests that a more precise calculation of the critical value for μ_C should be performed.

Relation to Kosterlitz-Thouless phase transitions

In the introductory chapter we encountered KT phase transitions in the XY planar model, in the Ising model and in the sine-Gordon/Thirring model. These models all have collective properties depending on a partition function (or path integral), that where analytic at the KT phase transition. The coupling to a specific set of microscopic variables induced a non-analyticity. However, there does not exist a precise definition of the Kosterlitz-Thouless phase transition yet.

Before we concluded that the gM^4 model phase transition resembled a KT phase transition. It might be instructive to take our characterization of the gM^4 phase transition and investigate whether that definition makes sense for KT transitions in other models.

[3] Which means that the second derivative of the free energy is discontinuous.

Appendix A

Summary

Matrix models are toy models applicable in various fields of physics. The overall properties of such a matrix model are defined by its partition function, which is an integral over $N \times N$ Hermitian matrices M with energy/action $S[M]$ invariant under similarity transformations. Upon integrating over the rotational degrees of freedom, the action can be described in terms of the eigenvalues of M.

If the action has one unique absolute minimum, then the free energy $F = -\log Z_N$ can be approximated via a perturbation series around that minimum. Generically, however, the matrix model action will have multiple extrema, e.g. in the gM^4 model. Using the eigenvalue representation, we show that the gM^4 model exhibits a phase transition for a specific range of coupling constants. For high $\mu_C = m^2/4g$ (the depth of the potential well) the ground state consists of a superposition of multiple solitons. For low μ_C there exists one single minimum of the action, which allows a perturbation expansion of the free energy.

We find that the phase transition of the gM^4 model is analytic in the macroscopic parameters, but is non-analytic when the action is coupled to external sources for eigenvalues. This can be verified by computing the correlation function in both phases. Physically the source term preselects one specific set of microscopic variables. The non-analyticity in this microscopic parameter while analytic in all macroscopic parameters suggests that we are dealing with a Kosterlitz-Thouless phase transition.

Finally we construct a renormalization group flow of the theory with respect to changes in the matrix dimension N and show that the lines of constant Z (aka the renormalization group flows) do cross the line of critical μ_C in the phase diagram.

Bibliography

[1] M. E. Peskin and D. V. Schroeder, "An Introduction To Quantum Field Theory," *Reading, USA: Addison-Wesley (1995) 842 p*

[2] K. G. Wilson and J. B. Kogut, Phys. Rept. **12**, 75 (1974).

[3] J. B. Kogut, Rev. Mod. Phys. **51** (1979) 659.

[4] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).

[5] J. M. Kosterlitz and D. J. Thouless, J. Phys. C C **6**, 1181 (1973).

[6] S. V. Ketov, Fortsch. Phys. **45**, 237 (1997) [arXiv:hep-th/9611209].

[7] S. R. Coleman, Phys. Rev. D **11**, 2088 (1975).

[8] E. Brézin and S. R. Wadia, "The large N expansion in quantum field theory and statistical physics : from spin systems to 2-dimensional gravity," *Singapore: World Scientific (1993) 1130 p*

[9] G. 't Hooft, Nucl. Phys. B **72**, 461 (1974).

[10] P. H. Ginsparg, arXiv:hep-th/9112013.

[11] K. Huang, "Statistical Mechanics," *New York, USA: Wiley (1987) 493 p*

[12] L. D. Landau and E. M. Lifshitz, "Statistical Physics," *Oxford, UK: Pergamon (1980), 544 p*

[13] G. B. Arfken, "Mathematical methods for physicists," *New York, USA: Academic Press (1970) 815 p*

[14] P. van Baal, "A Course in Field Theory," *Leiden, The Netherlands: Lorentz Institute (2001) 109 p*

[15] M. L. Mehta, "Random Matrices," *San Diego, USA: Elsevier (2004) 688 p*

[16] S. Higuchi, C. Itoi, S. Nishigaki and N. Sakai, arXiv:hep-th/9409157.

[17] E. Brezin and J. Zinn-Justin, Phys. Lett. B **288**, 54 (1992) [arXiv:hep-th/9206035].

[18] D. J. Gross and N. Miljkovic, Phys. Lett. B **238**, 217 (1990).

[19] D. J. Gross and I. R. Klebanov, Nucl. Phys. B **344** (1990) 475.

[20] J. Ambjorn, J. Jurkiewicz and A. Krzywicki, Phys. Lett. B **243**, 373 (1990).

[21] D. J. Gross and I. R. Klebanov, Nucl. Phys. B **354** (1991) 459.

[22] E. Brezin, C. Itzykson, G. Parisi and J. B. Zuber, Commun. Math. Phys. **59**, 35 (1978).

[23] G. 't Hooft, "Introduction to String Theory," *Utrecht, The Netherlands: Institute for Theoretical Physics (2004) 79 p*

[24] G. H. Derrick, J. Math. Phys. **5**, 1252 (1964).

[25] W. J. Marciano and H. Pagels, Phys. Rept. **36**, 137 (1978).

[26] M. E. Fisher, Rev. Mod. Phys. **70**, 653 (1998).

[27] Y. Shimamune, Phys. Lett. B **108**, 407 (1982).

Index