17

The $1/N$ expansions

As we saw in chapter 16, perturbation theory, even when applicable, only describes a regime of the theory. In the case of the non-linear sigma model in $D = 2$ dimensions, and in Yang-Mills theory in $D = 4$ dimensions, the running coupling constant is weak only at short distances but the long-distance behavior is inaccessible to perturbation theory since the coupling constant runs to large values, outside the perturbative regime. So we have theories that at low energies have a vacuum state that is essentially different than the free field ground state. To understand the physics of the actual ground state requires the use of non-perturbative methods. A key tool in this respect is the study of the generalizations of the theories of interest in their "large-N" limits. Here N can mean the rank of the symmetry group or the rank of the representation. The behavior is different in each case. We begin by considering first the simpler, and more tractable, case of scalar fields.

There is a long history of studying theories in this limit, both in Statistical Physics and in Quantum Field Theory. In Statistical Physics it goes back to the classic work by Berlin and Kac on what they called a "spherical model" of a phase transition (Berlin and Kac, 1952). This solvable model was later on shown to be equivalent to the large-N limit of the classical Heisenberg model (with an N-component order parameter) by Stanley (Stanley, 1968). With the advent of the renormalization group, this limit was studied by S.-K. Ma in the context of a ϕ^4 theory with a global $O(N)$ symmetry (Ma, 1973). In Quantum Field Theory, large-N limits became a mainstay tool to study non-perturbative the behavior of asymptotically free theories such as non-linear sigma models (Brézin and Zinn-Justin, 1976b), the $\mathbb{CP}^{N-1}$ models, the Gross-Neveu models (Gross and Neveu, 1974) and, particularly, Yang-Mills gauge theory ('t Hooft, 1974; Witten, 1979b). These methods have also been extensively used in theories of the Kondo problem (Read and Newns, 1983), quantum antiferromagnetism (Sachdev and Read, 1991), and in the study of quantum phase transitions (Sachdev, 1999). It has also played an important role in the theory of random matrices (Mehta, 2004).

17.1 The ϕ^4 scalar field theory with $O(N)$ global symmetry

Let us begin by considering the $O(N)$ ϕ^4 theory. This theory has a scalar field $\boldsymbol{\phi}(x)$ which is an N-component vector that transforms in the fundamental (vector) representation of the global symmetry group $O(N)$. The (Euclidean) Lagrangian is

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}(x)\right)^2 + \frac{m_0^2}{2}\boldsymbol{\phi}^2(x) + \frac{g}{4!N}\left(\boldsymbol{\phi}^2(x)\right)^2 \tag{17.1}$$

where, as usual, repeated indices are summed over. Notice that we have made the replacement of the conventional coupling constant $\lambda \mapsto \frac{g}{N}$. Shortly we will see the necessity of this replacement. The interaction vertex of the $O(N)$ theory is shown in Fig.17.1, where, for clarity, we have formally split the contact interaction.

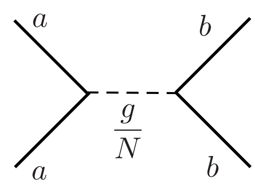


Figure 17.1 The interaction vertex of $O(N)$ scalar field theory with coupling constant $\lambda = \frac{g}{N}$.

17.1.1 *Diagrammatic approach to the large-N limit*

We will now ask what is the dependence on N of the Feynman diagrams of the 1PI two-point function, i.e. the ϕ field self-energy. The diagrams up to one-loop order are shown in Fig.17.2. It is easy to see that since each one-loop diagram contributes with a factor of the coupling constant, the rainbow diagram of Fig.17.2a contributes with a factor of $\frac{g}{N}$, while the tadpole diagram of Fig.17.2b contributes with a factor of $\frac{g}{N}N = g$, where the factor of N comes from the independent sum over the index b running inside the loop. Thus, in the limit $N \to \infty$, the leading term is Fig.17.2b and Fig.17.2a is a $1/N$ correction.

To see the emerging pattern in the large N limit, we will look at the

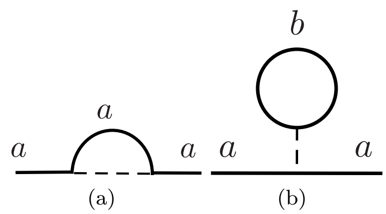


Figure 17.2 One-loop contributions to the 1PI two-point function of the ϕ field in the $O(N)$ theory; a) the rainbow diagram and b) the tadpole diagram.

two-loop diagrams of Fig.17.3. By counting powers of N, we see that the

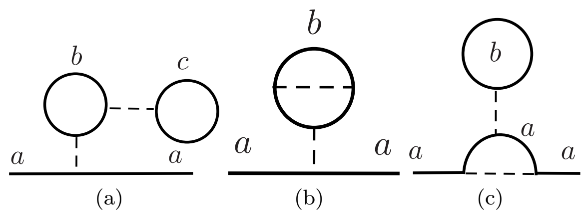


Figure 17.3 two-loop contributions to the 1PI two-point function of the ϕ field in the $O(N)$ theory; a) the two-loop tadpole diagram, b) the rainbow-tadpole two-loop diagram, and c) the two-loop rainbow diagram.

diagram of Fig.17.3a contributes with a factor of $\frac{g^2}{N^2}N^2 = g^2$, that the diagram of Fig.17.3b contributes with a factor of $\frac{g^2}{N^2}N = \frac{g^2}{N}$, and the diagram of Fig.17.3c contributes with a factor of $\frac{g^2}{N^2}$. Hence, only the diagram of Fig.17.3a survives in the $N \to \infty$ limit.

We now see the pattern: diagrams in which the number of independent sums over the internal indices is equal to the order in perturbation theory have a finite limit as $N \to \infty$, whereas the contributions of the other diagrams can be organized as a formal expansion in powers in $1/N$. A typical diagram that contributes to the two-point function in the $N \to \infty$ limit is shown in Fig.17.4. In fact, we have already reached this conclusion already in Section 11.6 where we showed that this sum of diagrams (also known as the Hartree approximation) yields the self-consistent expression for the one-loop self-energy of Eq.(11.60).

Moving on to the four-point function, it is easy to see that as $N \to \infty$ the only surviving contributions is the sum of the bubble diagrams shown in

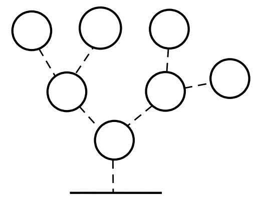


Figure 17.4 A typical contribution to the two-point function of the $O(N)$ scalar field in the $N \to \infty$ limit.

Fig.17.5 since each new diagram contributes with an extra factor of $\frac{g}{N}$ and a sum over an internal index that yields a factor of N.

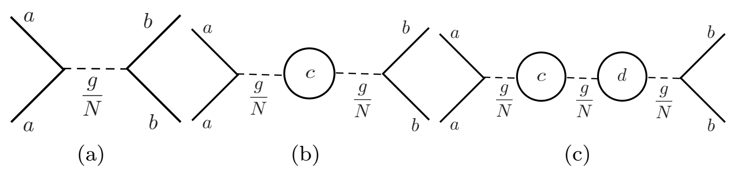


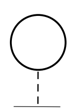
Figure 17.5 two-loop contributions to the 1PI four-point function of the ϕ field in the $O(N)$ theory; a) the bare vertex, b) the one-loop diagram, and c) a two-loop bubble diagram.

We can make these arguments explicit by writing the Dyson equation for the two-point function which, symbolically, becomes

$$G = G_0 + G_0\, \Sigma\, G \tag{17.2}$$

In the $N \to \infty$ limit, the self-energy Σ is the sum of all tree-like diagrams

$$\Sigma = \tag{17.3}$$



where the internal loop is the full two-point function in the $N \to \infty$ limit. More explicitly we can write

$$\Sigma(p) = -\frac{1}{6}\left(\frac{g}{N}\right) N \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m_0^2 - \Sigma(p)} \tag{17.4}$$

which does not depend on the external momentum p and the N dependence cancels out. This expression can be recast in terms of an effective mass

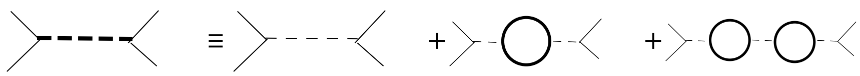
$$m^2 = m_0^2 - \Sigma(p) = m_0^2 + \frac{g}{6} \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m^2} \tag{17.5}$$

We recognize that this is just the renormalized mass (squared) at one-loop level, c.f. Eq.(11.62), discussed in Chapter 11. The only difference is that in the $N \to \infty$ limit this expression is exact. Thus, in this theory, the large N limit is equivalent to a self-consistent one-loop approximation.

We now have to consider the 1PI four-point vertex function $\Gamma^{(4)}_{ijkl}(p_1, \ldots, p_4)$, which by symmetry can be written is

$$\Gamma^{(4)}_{ijkl}(p_1, \ldots, p_4) = \\ = \frac{1}{3N}\left[\delta_{ij}\delta_{kl}F_4(p_1 + p_2) + \delta_{ik}\delta_{jl}F_4(P_1 + p_3) + \delta_{il}\delta_{jk}F_4(p_1 + p_4)\right] \quad (17.6)$$

The quantity $F_4(p)$ in the large-N limit is the sum of the bubble diagrams



$+ \ldots$ (17.7)

More explicitly, we can write the exact expression (in the $N \to \infty$ limit) for the sum of bubble diagrams $F_4(p)$

$$F_4(p) = \frac{g}{1 + \frac{g}{6}I(p)} \quad (17.8)$$

where $I(p)$ is the one-loop bubble diagram

$$I(p) = \int \frac{d^D q}{(2\pi)^D}\frac{1}{(q^2 + m^2)((p-q)^2 + m^2)} \quad (17.9)$$

Therefore, the sum of bubble diagrams yields the form of the effective interaction in the large-N limit. The sum of bubble diagrams is reminiscent of the Random Phase Approximation of Bohm and Pines (Pines and Bohm, 1952; Pines and Nozières, 1966) widely used in the theory of screening in electron fluids. In this context, the sum of tadpole diagrams leading to Eq.(17.3) is analogous to the Hartree approximation of the electron self-energy in the theory of the electron fluids.

The expression for $I(p)$ is UV finite for $D < 4$ and has a logarithmic divergence as $D \to 4^-$. The methods that we introduced to compute these integrals yield the result (in D Euclidean dimensions)

$$I(p) = \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{D/2}} 2^{4-D} \int_0^1 du((1-u^2)p^2 + 4m^2)^{(D-4)/2} \quad (17.10)$$

The important observation here is that if the renormalized mass vanishes, the IR behavior of the bubble leads to an IR singularity of the four point function $\Gamma^{(4)}(p) \propto p^{4-D}$. As we will see below, this result means that the effective coupling constant must have a non-trivial scale dependence, as dictated by the RG flow.

17.1.2 Path-integral approach

The diagrammatic analysis of the large-N limit of ϕ^4 theory is useful but it has the drawback that it seemingly applies only to the symmetric phase of the theory in which the $O(N)$ symmetry is not spontaneously broken. Although a similar diagrammatic analysis can be done in the broken symmetry phase, there is a more general approach that uses the large-N limit of the path integral for the partition function. To this end we introduce an auxiliary (Hubbard-Stratonovich) field $\alpha(x)$ to rewrite the partition function for the theory with the Lagrangian of Eq.(17.1)

$$
\begin{aligned}
Z &= \int \mathcal{D}\boldsymbol{\phi} \exp\left(-\int d^D x \left[\frac{1}{2}(\partial_\mu \boldsymbol{\phi})^2 + \frac{m_0^2}{2}\boldsymbol{\phi}^2 + \frac{g}{2N}(\boldsymbol{\phi}^2)^2\right]\right) \\
&= \int \mathcal{D}\alpha \int \mathcal{D}\boldsymbol{\phi} \exp\left(-\int d^D x \left[\frac{1}{2}(\partial_\mu \boldsymbol{\phi})^2 + \frac{m_0^2}{2}\boldsymbol{\phi}^2 - \frac{N}{2g}\alpha^2 + \alpha\boldsymbol{\phi}^2\right]\right) \\
&= \int \mathcal{D}\alpha \left(\mathrm{Det}\left[-\partial^2 + m_0^2 + \alpha\right]\right)^{-N/2} \exp\left(+\int d^D x \frac{N}{2g}\alpha^2\right)
\end{aligned}
\tag{17.11}
$$

After integrating-out the $\boldsymbol{\phi}$ fields, we can rewrite the partition function as a path-integral over the field $\alpha(x)$,

$$
Z = \int \mathcal{D}\alpha \exp\left(-N S_{\text{eff}}[\alpha]\right)
\tag{17.12}
$$

where the effective action $S_{\text{eff}}[\alpha]$ is

$$
S_{\text{eff}}[\alpha] = \frac{1}{2}\mathrm{tr}\ln\left[-\partial^2 + m_0^2 + 2\alpha\right] - \frac{1}{2g}\int d^D x\, \alpha^2(x)
\tag{17.13}
$$

Therefore, the large-N limit of ϕ^4 theory is the semiclassical limit of the theory of the field α, whose effective action is given by Eq.(17.13). Hence, in the $N \to \infty$ limit the path integral for the partition function is determined by the configurations $\alpha_c(x)$ that make the effective action stationary, i.e. the classical field $\alpha_c(x)$ that satisfies the saddle-point equation

$$
\frac{\delta S_{\text{eff}}}{\delta \alpha} = 0
\tag{17.14}
$$

which implies that the gap equation

$$
G(x, x; m^2) = \frac{\alpha_c}{g}
\tag{17.15}
$$

must be satisfied. Here we have set $m^2 = m_0^2 + 2\alpha_c$ and

$$G(x, y; m^2) = \langle x| \frac{1}{-\partial^2 + m^2} |y\rangle \tag{17.16}$$

is the propagator of a massive scalar field of mass squared m^2. Hence, the classical field α_c obeys the equation

$$\frac{\alpha_c}{g} = \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + m_0^2 + 2\alpha_c} \tag{17.17}$$

By analogy with the BCS theory of superconductivity, equations of this type are generally called "gap equations". It should be apparent that the self-energy we computed using the Dyson equation and the classical field are related by $\Sigma = -2\alpha_c$.

Returning to the effective action of the field $\alpha(x)$, we see that the 1PI two-point function of the field $\tilde{\alpha}(x) = \alpha(x) - \alpha_c$ is

$$\Gamma^{(2)}_{\alpha\alpha}(x, y) = 2G(x, y; m^2)G(y, x; m^2) + \frac{1}{g}\delta(x - y) \tag{17.18}$$

which, in momentum space, is easily seen to be the inverse of $F_4(p)/g$. We recognize the first term of the right hand side of this equation as the bubble diagram.

This analysis require a renormalization prescription. It should also be supplemented by an analysis of the broken symmetry state. For brevity we will do that only in the next section devoted to the large-N limit of the non-linear sigma-model.

17.2 The large-N limit of the $O(N)$ non-linear sigma model

We will now turn to the case of the $O(N)$ non-linear sigma model and its large-N limit. We have already discussed that, in the Euclidean metric, the non-linear sigma model is the formal continuum limit of the classical Heisenberg model. It has been know since the 1960s (Stanley, 1968) that the large-N limit of this model ie equivalent to the "spherical model" proposed in 1952 by Berlin and Kac (Berlin and Kac, 1952). Here we will work in D Euclidean dimensions.

As discussed in earlier chapters, at the classical level the $O(N)$ non-linear sigma model can be regarded as a limit of a ϕ^4 theory in its broken symmetry state. Indeed, we can rewrite the Euclidean Lagrangian of and N-component ϕ^4 theory as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \boldsymbol{\phi})^2 + \frac{\lambda}{4!}(\boldsymbol{\phi}^2 - \phi_0^2)^2 \tag{17.19}$$

which, in the $\lambda \to \infty$ limit, becomes the Lagrangian of a non-linear sigma model with and N component unit-vector field $\boldsymbol{n}(x) = (\sigma(x), \boldsymbol{\pi}(x))$, such that $\boldsymbol{n} = \sigma^2 + \boldsymbol{\pi}^2 = 1$ as a constraint everywhere in D-dimensional Euclidean space-time. The partition function of the $O(N)$ non-linear sigma model is the path integral

$$Z[H, \boldsymbol{J}] = \int \mathcal{D}\sigma \mathcal{D}\boldsymbol{\pi}\ \delta(\sigma^2 + \boldsymbol{\pi}^2 - 1)\ \exp\left(-\frac{S}{g}\right) \tag{17.20}$$

where $1/g = \phi_0^2$. Here, the δ-function acts at all points of D-dimensional Euclidean space-time. The action S is

$$S = \int d^D x \left[\frac{1}{2}(\partial_\mu \sigma)^2 + \frac{1}{2}(\partial_\mu \boldsymbol{\pi})^2 - H\sigma - \boldsymbol{J} \cdot \boldsymbol{\pi}\right] \tag{17.21}$$

where $(H, \boldsymbol{J})$ is a symmetry-breaking field. We will now see that even though classically (and in perturbation theory for $D > 2$) this is a theory of a broken symmetry state, this theory has (again for $D > 2$) both a broken symmetry phase and a symmetric phase separated by a critical (and non universal) value of the dimensionful coupling constant g.

17.2.1 Large-N limit

That the non-linear sigma model has a finite large-N limit can be gleaned from the perturbative expression for the 1PI two-point function of the $\boldsymbol{\pi}$ field, presented last chapter in Eq.(16.158) and in Fig.16.3. There we see that as N becomes large, we can have a finite limit provided the coupling constant g and the field H must be scaled by N and the field $\boldsymbol{\pi}$ must also be scaled by an appropriate power of N.

Here we will use a functional approach to study the large-N limit. We begin by implementing the constraint by means of an integral representation of the δ-function in terms of a Lagrange multiplier field $\alpha(x)$, in terms of which the partition function now reads

$$Z = \int \mathcal{D}\sigma \mathcal{D}\boldsymbol{\pi} \mathcal{D}\alpha \exp\left(-\frac{S}{g} + \int d^D x \frac{\alpha(x)}{2g}(1 - \sigma^2(x) - \boldsymbol{\pi}^2(x))\right) \tag{17.22}$$

Let us rescale the $N-1$-component fields $\boldsymbol{\pi} = \sqrt{g}\boldsymbol{\varphi}$ which in turn can be integrated-out to yield

$$Z = \int \mathcal{D}\sigma\mathcal{D}\alpha \left(\mathrm{Det}[-\partial^2 + \alpha]\right)^{-(N-1)/2}$$
$$\times \exp\left(\frac{1}{2}\int d^D x \int d^D y \; G(x-y;\alpha) \; \boldsymbol{J}(x) \cdot \boldsymbol{J}(y)\right)$$
$$\times \exp\left(-\frac{1}{g}\int d^D x \frac{1}{2}(\partial_\mu \sigma)^2 - H\sigma - \frac{1}{2}\alpha(\sigma^2 - 1)\right) \tag{17.23}$$

where the kernel $G(x-y;\alpha)$ satisfies

$$(-\partial^2 + \alpha(x))G(x-y;\alpha(x)) = \delta(x-y) \tag{17.24}$$

for an arbitrary configuration of the Lagrange multiplier field. Hereafter we will drop the explicit x-dependence of the field α. For the time being, we will set the sources $\boldsymbol{J}(x) = 0$ (although later they will be restored).

It will be convenient to rescale the bare coupling constant g and the field σ as follows

$$g = \frac{g_0}{N-1}, \quad \sigma(x) = \sqrt{g(N-1)}m(x) \tag{17.25}$$

The effective action of the rescaled field $m(x)$ and of the Lagrange multiplier field α becomes

$$S_{\text{eff}}(m,\alpha,H) = \int d^D x \left[\frac{1}{2}(\partial_\mu m)^2 + \frac{1}{2}\alpha m^2 - \frac{\alpha}{2g_0} - \frac{Hm}{\sqrt{g_0}}\right] + \frac{1}{2}\mathrm{tr}\ln\left(-\partial^2 + \alpha\right) \tag{17.26}$$

and the partition function is

$$Z[H] = \int \mathcal{D}m\mathcal{D}\alpha \exp\left(-(N-1)S_{\text{eff}}(m,\alpha,H)\right) \tag{17.27}$$

Hence, once again, the large-N limit is the semiclassical limit, in this case of the effective action for the fields m and α coupled to the source H. Therefore, in the large-N limit, the partition function of Eq.(17.27) is dominated by the configurations that leave the effective action of Eq.(17.26) stationary. By varying S_{eff} with respect to the field $m(x)$ we get

$$\frac{\delta S_{\text{eff}}}{\delta m(x)} = -\partial^2 m(x) + \alpha(x)m(x) - \frac{H}{\sqrt{g_0}} = 0 \tag{17.28}$$

Likewise, by varying S_{eff} with respect to $\alpha(x)$ we obtain

$$\frac{\delta S_{\text{eff}}}{\delta \alpha(x)} = -\frac{1}{2g_0} + \frac{1}{2}m^2(x) + \frac{1}{2}\frac{\delta}{\delta\alpha(x)}\mathrm{tr}\ln\left(-\partial^2 + \alpha\right) = 0 \tag{17.29}$$

where

$$\frac{\delta}{\delta\alpha(x)}\mathrm{tr}\ln\left(-\partial^2 + \alpha\right) = \langle x|\frac{1}{-\partial^2 + \alpha}|x\rangle \tag{17.30}$$

Hence, we obtain a second saddle point equation

$$\langle x| \frac{1}{-\partial^2 + \alpha} |x\rangle + m^2(x) - \frac{1}{g_0} = 0 \tag{17.31}$$

We will now seek a uniform solution of Eqs. (17.28) and (17.31) of the form

$$\sigma(x) = M, \quad \alpha(x) = \bar{\alpha} \tag{17.32}$$

with H constant, and $M^2 = g_0 m^2$. Thus, the solution of Eq.(17.28) is

$$\bar{\alpha} = \frac{H}{M} \tag{17.33}$$

while Eq.(17.31) becomes

$$g_0 \langle x| \frac{1}{-\partial^2 + \frac{H}{M}} |x\rangle = 1 - M^2 \tag{17.34}$$

or, what is the same

$$g_0 \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \frac{H}{M}} = 1 - M^2 \tag{17.35}$$

17.2.2 Renormalization

Provided $H/M \neq 0$, Eq.(17.35) is IR finite. However, for $D \geq 2$ the integral is UV divergent. We will absorb the strong dependence in the UV in a set of renormalization constants. Thus, we define a dimensionless renormalized coupling constant t, a renormalized M_R and a renormalized H_R through the relations (here κ is an arbitrary momentum scale) (see Eqs. (16.161) and (16.162))

$$\begin{aligned} g_0 &= t\kappa^{-\epsilon} Z_1 \\ M &= Z^{1/2} M_R \\ H &= Z_1 Z^{-1/2} H_R \end{aligned} \tag{17.36}$$

where we have set $\epsilon = D - 2$, and Z is the wave function renormalization. Notice that

$$\frac{H}{M} = \frac{Z_1}{Z} \frac{H_R}{M_R} \tag{17.37}$$

With these definitions Eq.(17.35) becomes

$$t\kappa^{-\epsilon} \frac{Z_1}{Z} \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \frac{H_R}{M_R} \frac{Z_1}{Z}} = \frac{1}{Z} - M_R^2 \tag{17.38}$$

We have encountered the integral in this equation several times before. It is given by

$$\int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \mu^2} = \frac{1}{(4\pi)^{D/2}} \Gamma\left(-\frac{\epsilon}{2}\right) \mu^{\epsilon} \tag{17.39}$$

Using this result we can recast Eq.(17.38) as

$$t\kappa^{-\epsilon} \left(\frac{Z_1}{Z}\right)^{1+\epsilon/2} \left(\frac{H_R}{M_R}\right)^{\epsilon/2} \frac{1}{(4\pi)^{D/2}} \Gamma\left(-\frac{\epsilon}{2}\right) = \frac{1}{Z} - M_R^2 \tag{17.40}$$

We now use dimensional regularization with (quasi) minimal subtraction by defining Z and Z_1 in such a way that the singular dependence in ϵ is cancelled. Thus, we choose

$$Z = Z_1 \tag{17.41}$$

and

$$\frac{1}{Z} = 1 + t \frac{1}{(4\pi)^{D/2}} \Gamma\left(-\frac{\epsilon}{2}\right) \tag{17.42}$$

With these choices Eq.(17.40) become

$$1 - M_R^2 = \frac{t}{t_c} \left(1 - \left(\frac{H_R}{\kappa^2 M_R}\right)^{\epsilon/2}\right) \tag{17.43}$$

which is finite as $\epsilon \to 0$. Here we introduced the quantity t_c,

$$t_c = \left(\frac{D-2}{2}\right) \frac{(4\pi)^{D/2}}{\Gamma\left(2 - \frac{D}{2}\right)} \tag{17.44}$$

which we will shortly identify with the value of the critical coupling constant. This equation relates M_R, the (renormalized) expectation value of the sigma field, to the renormalized dimensionless coupling constant t and the renormalized symmetry breaking field H_R. In the statistical mechanical interpretation, this is the equation of state (in the $N \to \infty$ limit).

17.2.3 Phase diagram and spectrum

We will now show that, in the $N \to \infty$ limit and for $D > 2$, the non-linear sigma model has a phase transition between a broken symmetry phase and a symmetric phase. To this end, let us seek a solution to Eq.(17.43) for $H_R = 0$ and $M_R \neq 0$, i.e.

$$1 - M_R^2 = \frac{t}{2\pi\epsilon} \tag{17.45}$$

and notice that, as $M_R \to 0$, t approaches (from below) the value t_c given in Eq.(17.44) which defines the critical coupling constant. It is worth to note that, as expected, $t_c \to 0$ as $\epsilon \to 0$. Using this expression for t_c, we can write Eq.(17.45) as

$$1 - M_R^2 = \frac{t}{t_c} \tag{17.46}$$

or, what is equivalent,

$$M_R = \left(1 - \frac{t}{t_c}\right)^{\beta} \tag{17.47}$$

where the exponent is $\beta = 1/2$. We already found this result in our study of the perturbative renormalization group where we noted that $\beta \to 1/2$ as $n \to \infty$. Hence, we expect that there will be corrections to this result if we go beyond the $N \to \infty$ limit. Thus, we see that for $t < t_c$ the $O(N)$ symmetry is indeed spontaneously broken.

Let us now examine the behavior for $t > t_c$. We will now use the definition of t_c of Eq.(17.44), to write Eq.(17.45) in the simpler form

$$1 - M_R^2 = \frac{t}{t_c}\left(1 - \left(\frac{H_R}{M_R \kappa^2}\right)^{\epsilon/2}\right) \tag{17.48}$$

We now recall that, by dimensional analysis and the Ward identity, the susceptibility χ_R is

$$\chi_R = \kappa^2 \frac{M_R}{H_R} \tag{17.49}$$

in terms of which we can write

$$1 - M_R^2 = \frac{t}{t_c}\left(1 - \left[\chi_R(t, H_R)\right]^{-\epsilon/2}\right) \tag{17.50}$$

We now take the limit $H_R \to 0$ and $M_R \to 0$ holding χ_R fixed, and obtain

$$\chi_R = \left(1 - \frac{t}{t_c}\right)^{-\gamma} \tag{17.51}$$

where the exponent is $\gamma = \frac{2}{\epsilon}$ (again in the $N \to \infty$ limit).

Having understood the question of spontaneous symmetry breaking, we will now inquire the behavior of the $\boldsymbol{\pi}$ fields. We expect that the $\boldsymbol{\pi}$ fields should be the Goldstone bosons of the broken symmetry state and, hence, should be massless in the broken symmetry phase, and massive in the unbroken phase. Moreover, we also expect the σ field to be massive in the broken phase and that all the masses become equal as $H \to 0$.

To this end we now restore the fields $\boldsymbol{J} \neq 0$, the sources of the $\boldsymbol{\pi}$ fields.

Although, in principle, we obtain more complicated saddle point equations if $\boldsymbol{J} \neq 0$, in practice we can still use our solutions obtained for $\boldsymbol{J} = 0$ in the regime in which $\boldsymbol{J}$ is infinitesimally small (i.e. "linear response"). Hence, we can approximate the kernel $G(x, y; \alpha)$ that enters in Eq.(17.23), by its approximate form with $\alpha = \bar{\alpha} = \frac{H}{M} = \frac{H_R}{M_R}$. We readily see that, in this limit, the two-point function of the $\boldsymbol{\pi}$ fields is

$$\langle \boldsymbol{\pi}(x) \cdot \boldsymbol{\pi}(y) \rangle = G\left(x, y; \frac{H_R}{M_R}\right) \tag{17.52}$$

In other words, provided $\frac{H_R}{M_R} \neq 0$, the $\boldsymbol{\pi}$ fields are massive and their mass (squared) is

$$m_{\boldsymbol{\pi}}^2 = \frac{1}{\xi_{\boldsymbol{\pi}}^2} = \frac{H_R}{M_R} \tag{17.53}$$

where $\xi_{\boldsymbol{\pi}}$ is the correlation length of the $\boldsymbol{\pi}$ fields, which, as we see, is given by

$$\xi_{\boldsymbol{\pi}} = \left(\chi_R \kappa^{-2}\right)^{1/2} = \kappa^{-1} \left|1 - \frac{t_c}{t}\right|^{-\nu} \tag{17.54}$$

where $\nu = \gamma/2 = 1/\epsilon$ (again, as $N \to \infty$).

Thus, in the symmetric phase, we find that at $H_R \to 0$,

$$m_{\boldsymbol{\pi}}^2 = \chi_R^{-1} = m_\sigma^2 \tag{17.55}$$

where m_σ^2 is the mass (squared) of the σ field. Hence, for $t > t_c$ we have a spectrum of N massive bosons forming a multiplet of $O(N)$ (as we should!). On the other hand, for $t < t_c$, the symmetry is spontaneously broken and $M_R \neq 0$ as $H_R \to 0$. Hence, in the broken symmetry state $m_{\boldsymbol{\pi}}^2 = 0$ and we have (as we should!) a spectrum of $N - 1$ massless (Goldstone) bosons.

In the special case of $D = 2$, the correlation length $\xi_\pi = \xi_\sigma = \xi$, becomes

$$\xi = m^{-1} = \kappa^{-1} \exp\left(\frac{2\pi}{t}\right) \tag{17.56}$$

This result shows the dependence of the correlation length on the dimensionless coupling constant t is an essential singularity. This implies that the $N \to \infty$ is truly non-perturbative. We will encounter this behavior in all asymptotically free theories.

17.2.4 Renormalization group

We now return to the definitions of Eq.(17.36) to obtain the renormalization group functions in the large-N limit. We begin with the beta function for the dimensionless coupling constant t,

$$\beta(t) = \kappa \frac{\partial t}{\partial \kappa}\Big|_B \tag{17.57}$$

were we hold the bare theory fixed. From the definition $g_0 = t\kappa^{-\epsilon} Z_1$ we find

$$\beta(t)\left(1 + t\frac{\partial \ln Z_1}{\partial t}\right) = \epsilon t \tag{17.58}$$

Using the expression for Z_1 (and of Z) of Eq.(17.42), and the definition of t_c of Eq.(17.44), we find that the beta function is exactly given by

$$\beta(t) = \epsilon t - \epsilon \frac{t^2}{t_c} \tag{17.59}$$

Hence, in the $N \to \infty$ limit the beta function terminates at the quadratic order in t, where it agrees with the one-loop result, c.f. Eq.(16.177), in the large-N limit and for $D = 2 + \epsilon$. We see that, for $D > 2$, the beta function has two zeros (or fixed points): a) a trivial fixed point at $t = 0$, and b) a non-trivial fixed point at $t = t_c$. Furthermore, the slope of the beta function is negative at the non-trivial fixed point. Hence, this fixed point us unstable in the IR (and hence stable in the UV). Conversely, the trivial fixed point is stable in the IR and unstable in the UV.

On the other hand, for $D = 2$ the beta function simply becomes

$$\beta(t) = \kappa \frac{\partial t}{\partial \kappa} = -\frac{t^2}{2\pi} \tag{17.60}$$

which says that the $O(N)$ non-linear sigma model is asymptotically free in $D = 2$, and the trivial fixed point at $t = 0$ is IR unstable. As we saw, the $N \to \infty$ limit predicts that the theory is in a massive phase with an unbroken global symmetry for all values of the coupling constant.

Similarly, we can compute the renormalization group function $\gamma(t)$,

$$\gamma(t) = \beta(t)\frac{\partial \ln Z}{\partial t} \tag{17.61}$$

and find

$$\gamma(t) = \frac{t}{t_c}\epsilon \tag{17.62}$$

Hence, the anomalous dimension η at the non-trivial fixed point at t_c is

$$\eta = \gamma(t_c) - \epsilon = 0 \tag{17.63}$$

In other terms, the anomalous dimension $\eta = O(1/N)$, and to determine it requires a computation of the leading correction in the $1/N$ expansion.

This is consistent with the result we obtained in the $2 + \epsilon$ expansion, c.f. Eq.(16.197), where we found that $\eta \propto 1/N$.

An important moral of this discussion is that, while for $D > 2$ the perturbative definition of the theory, based on an expansion in powers of the coupling constant, is non-renormalizable, here we find that the theory defined around the non-trivial fixed point is renormalizable. Notice that we were able to access this UV fixed point only by means of the use of the large-N limit. In principle, the $2 + \epsilon$ expansion could be used to the same end . However, this presents technical difficulties associated with the behavior of the expansion that we will not discuss here.

17.3 The $\mathbb{CP}^{N-1}$ model

We will now turn to the $\mathbb{CP}^{N-1}$ model (D'Adda et al., 1978; Witten, 1979b; Coleman, 1985), introduced in section 16.5.1. As we showed there, these models can be described by an N-component complex field $z(x)$ of unit norm, i.e. $\boldsymbol{z}(x) = (z_1(x), \dots, z_N(x))$ with $z_i(x) \in \mathbb{C}$ and $\boldsymbol{z}^\dagger(x) \cdot \boldsymbol{z}(x) = \sum_{i=1}^N |z_i(x)|^2 = 1$, minimally coupled to a $U(1)$ gauge field $A_\mu(x)$. The action of this model is (c.f. Eq.(16.103))

$$S[z_\alpha, z_\alpha^*, \mathcal{A}_\mu] = \frac{1}{g} \int d^D x \left| \left(\partial_\mu - i\mathcal{A}_\mu(x)\right) \boldsymbol{z}(x) \right|^2, \tag{17.64}$$

Here g is the coupling constant, which has units of $[L]^{D-2}$, as in the case of the $O(N)$ model. As we saw, this model is invariant under the local $U(1)$ gauge transformations,

$$\boldsymbol{z}(x) \mapsto \exp(i\phi(x))\boldsymbol{z}(x), \qquad \mathcal{A}_\mu(x) \mapsto \mathcal{A}_\mu(x) + \partial_\mu \phi(x) \tag{17.65}$$

Notice that the gauge field only enters in the action through the covariant derivative, and that we do not have a separate term for the gauge field, as we would in quantum electrodynamics.

In the absence of the $U(1)$ gauge field, this would be a theory of an N-component complex field with unit norm or, what is equivalent, a $2N$-component real vector of unit length. This would be a non-linear sigma model with $SU(N)$ global symmetry or, equivalently $O(2N)$. Hence, for $D > 2$, we would expect to have a broken symmetry state with $2N - 1$ Goldstone bosons. However, in the $\mathbb{CP}^{N-1}$ model a $U(1)$ subgroup of $SU(N)$ has been gauged. In this situation, we expect that in the broken symmetry state there is a Higgs mechanism (see section 18.11) and, consequently, one of the otherwise $2N-1$ massless Goldstone bosons should be absent, "eaten" by the gauge field which should be massive in this phase. On the other hand, in the symmetric phase, the $\boldsymbol{z}$ fields should be massive, and the gauge field should be strongly fluctuating.

Here we will use a functional approach very similar to what we did in the case of the $O(N)$ non-linear sigma model. Thus, we implement the local constraint, $\boldsymbol{z}^\dagger \boldsymbol{z} = 1$ with an integral representation of the delta function using a (real) Lagrange multiplier field $\alpha(x)$. The partition function of the $\mathbb{CP}^{N-1}$ model is

$$\mathcal{Z} = \int \mathcal{D}\boldsymbol{z}\mathcal{D}\boldsymbol{z}^\dagger \mathcal{D}\mathcal{A}_\mu \mathcal{D}\alpha \exp\left(-\frac{1}{g}\int d^D x \left[|(\partial_\mu - i\mathcal{A}_\mu)\boldsymbol{z}|^2 - \alpha(\boldsymbol{z}^\dagger \boldsymbol{z} - 1)\right]\right) \tag{17.66}$$

In order to have a well defined large-N limit we will rescale the coupling constant as $g = g_0/N$.

We can now integrate-out the N-component complex field $\boldsymbol{z}$ and obtain, after rescaling the $\boldsymbol{z}$ fields by $\sqrt{g_0}$, the partition function in terms of the effective action for the gauge field $\mathcal{A}_\mu$ and the Lagrange multiplier field α,

$$S_{\text{eff}}[\mathcal{A}_\mu, \alpha] = \operatorname{tr}\ln\left(-D_\mu[\mathcal{A}_\mu]^2 + \alpha\right) - \frac{1}{g_0}\int d^D x\alpha(x) \tag{17.67}$$

where $D_\mu[\mathcal{A}] = \partial_\mu - i\mathcal{A}_\mu$ is the covariant derivative. The partition function now has the form

$$\mathcal{Z} = \int \mathcal{D}\mathcal{A}_\mu \mathcal{D}\alpha \exp\left(-N S_{\text{eff}}[\mathcal{A}_\mu, \alpha]\right) \tag{17.68}$$

Notice that, unlike what we did in the $O(N)$ non-linear sigma model, we are treating symmetrically all the components of the $\boldsymbol{z}$ field. Although we will find a phase transition, the analysis of the broken symmetry state will be easier in a somewhat less symmetric formulation.

Clearly, as $N \to \infty$ this partition function will be dominated by the classical configurations that leave the effective action stationary. The difference between the $O(N)$ non-linear sigma model and the $\mathbb{CP}^{N-1}$ model is that, in addition to the Lagrange multiplier field α, we now have the gauge field $\mathcal{A}_\mu$. Thus we will have two saddle point equations, each requiring that the effective action be stationary under separate variations of the gauge field and of the Lagrange multiplier field.

Thus we find two conditions. The first one, obtained by varying S_{eff} with respect the Lagrange multiplier field $\alpha(x)$,

$$\frac{\delta S_{\text{eff}}}{\delta \alpha(x)} = 0 \tag{17.69}$$

implies that

$$\frac{1}{g_0} = \langle x| \frac{1}{-D[\mathcal{A}_\mu]^2 + \alpha} |x\rangle \tag{17.70}$$

The second saddle point equation, obtained by varying with respect to the gauge field $\mathcal{A}_\mu(x)$,

$$\frac{\delta S_{\text{eff}}}{\delta \mathcal{A}_\mu(x)} = 0 \tag{17.71}$$

is the condition that, at the classical level, the $U(1)$ gauge current vanishes,

$$\frac{\delta S}{\delta \mathcal{A}_\mu(x)} = j_\mu[\mathcal{A}] = 0 \tag{17.72}$$

where S is the action of the $\mathbb{CP}^{N-1}$ model. Thus, in the absence of external sources, we can set the classical configuration of the gauge field $\langle \mathcal{A}_\mu \rangle = 0$, which trivially satisfies the condition of a vanishing $U(1)$ gauge current.

Thus, the only equation to be solved in the large-N limit, Eq.(17.70), is the gap equation

$$\frac{1}{g_0} = \langle x| \frac{1}{-\partial^2 + \alpha} |x\rangle = \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \alpha_c} \tag{17.73}$$

Notice that the classical value of the Lagrange multiplier, α_c, plays the role of a mass term for the $\boldsymbol{z}$ fields, and

$$G(x-y|\alpha_c) = \langle x| \frac{1}{-\partial^2 + \alpha_c} |y\rangle \tag{17.74}$$

is the propagator for the $\boldsymbol{z}$ field.

We have encountered the integral on the right hand side of Eq.(17.73) several times, most recently in Eq.(17.39). This integral is UV divergent for $D > 2$ and logarithmically divergent for $D = 2$. Thus, as before, we need to define renormalized quantities. Here it will suffice to renormalize the coupling constant g_0^2. Using dimensional analysis we define a renormalized dimensionless coupling constant t,

$$g_0 = t\kappa^{-\epsilon} Z \tag{17.75}$$

where Z is a renormalization constant and κ is the renormalization scale. Using the expression of Eq.(17.39) for the integral, the saddle-point equation of Eq.(17.73) becomes

$$\frac{1}{Z} = t\kappa^{-\epsilon} \frac{1}{(4\pi)^{D/2}} \Gamma\left(-\frac{\epsilon}{2}\right) \alpha_c^{\epsilon/2} \tag{17.76}$$

Once again, we will use the minimal subtraction procedure to cancel the singular dependence of ϵ. To this end we choose Z to have the same form as in the non-linear sigma model, Eq.(17.42). This choice results in the saddle-point equation to become

$$1 = \frac{t}{t_c}\left(1 - (\alpha_c \kappa^{-2})^{\epsilon/2}\right) \tag{17.77}$$

where t_c is the same value for the critical coupling that we found for the non-linear sigma model, c.f. Eq.(17.44). We can now readily solve for the value α_c and find

$$\alpha_c = \kappa^2 \left(1 - \frac{t_c}{t}\right)^{2/\epsilon} \tag{17.78}$$

This solution is only allowed if $t > t_c$. For $t < t_c$, the only possible solution is $\alpha_c = 0$.

In other words, for $D > 2$ and in the limit $N \to \infty$, the $\mathbb{CP}^{N-1}$ model has two phases separated by a phase transition at t_c. For $t < t_c$ the $\boldsymbol{z}$ fields are massless, and for $t > t_c$ they are massive with mass $m_{\boldsymbol{z}}^2 = \alpha_c$. On the other hand in $D = 2$, $t_c \to 0$ and there is only one phase and the $\boldsymbol{z}$ fields are massive for all values of the coupling constant. It is easy to see that, in $D = 2$, the mass (squared) is

$$m_{\boldsymbol{z}}^2 = \alpha_c = \kappa^2 \exp\left(-\frac{\pi}{t}\right) \tag{17.79}$$

In other words, in $D = 2$ the $\mathbb{CP}^{N-1}$ model is be asymptotically free and exhibits dynamical mass generation.

That the $\mathbb{CP}^{N-1}$ model is asymptotically free in $D = 2$ can be seen by computing the beta function (in the large-N limit)

$$\beta(t) = \kappa \frac{\partial t}{\partial \kappa}\bigg|_B = \epsilon t - \frac{\epsilon}{t_c} t^2 \tag{17.80}$$

which is the same as the $O(N)$ non-linear sigma model in the $N \to \infty$ limit. Thus, here too, for $D > 2$ we have a UV fixed point at t_c and a IR fixed point at $t = 0$. In $D = 2$ dimensions this theory is asymptotically free.

It is instructive to compute the leading corrections in the $1/N$ expansion.

To lowest order in the $1/N$ expansion, the partition function is

$$\begin{aligned} \mathcal{Z} &= \exp(-NS_{\text{eff}}[\alpha_c]) \\ &\times \int \mathcal{D}\tilde{\alpha} \exp\left(-\frac{1}{2}\int d^D x \int d^D y \; \tilde{\alpha}(x)\Pi(x-y|\alpha_c)\tilde{\alpha}(y)\right) \\ &\times \int \mathcal{D}\mathcal{A}_\mu \exp\left(-\frac{1}{2}\int d^D x \int d^D y \; \mathcal{A}_\mu(x)\Pi_{\mu\nu}(x-y|\alpha_c)\mathcal{A}_\nu(y)\right) \\ &\times\left[1+O\left(\frac{1}{N}\right)\right] \end{aligned} \tag{17.81}$$

where the kernel $\Pi(x-y|\alpha_c)$ is

$$\Pi(x-y|\alpha_c) = G(x-y|\alpha_c)G(y-x|\alpha_c) \tag{17.82}$$

which is clearly a bubble diagram. Here, $G(x-y|\alpha_c)$ is the propagator of the $\boldsymbol{z}$ fields, given in Eq.(17.74). The Fourier transform $\Pi(p)$ is

$$\Pi(p) = \int \frac{d^D q}{(2\pi)^D}\frac{1}{(q^2+\alpha_c)((q+p)^2+\alpha_c)} \tag{17.83}$$

On the other hand, the polarization tensor $\Pi_{\mu\nu}(x-y)$ is the correlation function of the currents of the $\boldsymbol{z}$ fields. By gauge invariance it must be transverse, i.e. it should obey the Ward identity

$$\partial_\mu^x \Pi_{\mu\nu}(x-y) = 0 \tag{17.84}$$

Explicitly, $\Pi_{\mu\nu}(x-y)$ is given by

$$\Pi_{u\nu}(x-y) = \partial_\mu^x G(x-y|\alpha_c)\partial_\nu^x G(y-x|\alpha_c) - 2\delta_{\mu\nu}G(x,x|\alpha_c) \tag{17.85}$$

In momentum space this kernel is given by

$$\Pi_{\mu\nu}(p) = \int \frac{d^D q}{(2\pi)^D}\frac{(2q_\mu+p_\mu)(2q_\nu+p_\nu)}{(q^2+\alpha_c)((q+p)^2+\alpha_c)} - 2\delta_{\mu\nu}\int \frac{d^D q}{(2\pi)^D}\frac{1}{(q^2+\alpha_c)} \tag{17.86}$$

At long distances, $p \to 0$, $\Pi_{\mu\nu}(p)$ behaves as

$$\Pi_{\mu\nu}(p) \simeq \frac{1}{48\pi\alpha_c}(p^2\delta_{\mu\nu} - p_\mu p_\nu) \tag{17.87}$$

where $\alpha_c = m_z^2$ is the mass squared of the z particles. Hence, in the massive phase, $\alpha_c \neq 0$, the low-energy effective action of the gauge field $\mathcal{A}_\mu$ has the Maxwell form

$$S_{\text{eff}}[\mathcal{A}_\mu] = C\frac{N}{4\alpha_c}\int d^D x \; \mathcal{F}_{\mu\nu}^2(x) \tag{17.88}$$

where $\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$ is the field strength of this "emergent" gauge field, and $C = \frac{1}{48\pi}$.

The $\boldsymbol{z}$ fields couple minimally to the gauge field $\mathcal{A}_\mu$. Thus, we see that in the symmetric phase (with $t > t_c$) the $\boldsymbol{z}$ particles will experience a long-range "Coulomb" interaction, and should form gauge-invariant bound states of the form $z^*_\alpha z_\beta$. This is particularly strong in $D = 2$ space-time dimensions. In two dimensions, as in the case of the $O(N)$ non-linear sigma model, the theory is asymptotically free and the coupling constant flows to large values. The theory is in a phase dominated by non-perturbative effects. In the large-N limit we find that $t_c = 0$ and the fields are massive.

However there is more than that. In $D = 2$ dimensions (i.e. 1+1 Minkowski spacetime) the Coulomb interaction is *linear*, i.e. $\propto |x-y|$, where x and y are the spatial coordinates of the z particles. This can be checked by computing the expectation value of the Wilson loop operator, corresponding to the worldlines of a heavy particle-antiparticle pair, see section 9.7. As we know the Coulomb interaction in 1+1 dimensions becomes $V(R) = \sigma R$. This is a *confining* potential. Hence, not only these bound states are very tight but there are no $\boldsymbol{z}$ particles in the spectrum. We will return to the problem of confinement in gauge theory in a later chapter.

To study the broken symmetry state, with $t < t_c$, we need to modify our approach somewhat. Let us call the component $z_1 = z_\parallel$ and the $N-1$ remaining (complex) components $\boldsymbol{z}_\perp$. The Lagrangian now is

$$\mathcal{L} = \frac{1}{g}|D_\mu[\mathcal{A}]z_\parallel|^2 + \frac{1}{g}|D_\mu[\mathcal{A}]\boldsymbol{z}_\perp|^2 + \frac{\alpha}{g}(|z_\parallel|^2 + |\boldsymbol{z}_\perp|^2 - 1) \tag{17.89}$$

In order to have a well defined large-N limit we now set $g = g_0(N-1)$ (as we did in the case of the non-linear sigma model). We can now integrate out the $N-1$ complex transverse fields, $\boldsymbol{z}_\perp$ and obtain the effective action

$$\begin{aligned} S_{\text{eff}}[\mathcal{A},\alpha,\rho] =&(N-1)\operatorname{tr}\ln\left(-D[\mathcal{A}]^2+\alpha\right) - \frac{N-1}{g_0}\int d^Dx\, \alpha(x) \\ &+\frac{N-1}{g_0}\int d^Dx\Big[|D_\mu[\mathcal{A}]z_\parallel|^2 + \alpha|z_\parallel|^2\Big] \end{aligned} \tag{17.90}$$

This effective action is invariant under $U(1)$ gauge transformations, and require that we fix the gauge. Here it is convenient to use the unitary gauge, $z_\parallel = \rho \in \mathbb{R}^+$. The path integral for the field ρ (whose action is given in second line of Eq.(17.90)) becomes

$$Z[\mathcal{A},\alpha] = \int \mathcal{D}\rho\, \exp\left(-\frac{(N-1)}{g_0}\int d^Dx\, \left[(\partial_\mu\rho)^2 + \rho^2\mathcal{A}_\mu^2 + \alpha\rho^2\right]\right) \tag{17.91}$$

The saddle point equation for α now is

$$\frac{1}{g_0} - \frac{\rho^2}{g_0} = G(x, x|\alpha) \tag{17.92}$$

and the saddle point equation for ρ is given by

$$-\partial^2 \rho + \rho \mathcal{A}_\mu^2 + \alpha \rho = 0 \tag{17.93}$$

We will see solutions with $\mathcal{A} = 0$ (as before). For $g_0 < g_0^c$, i.e. $t < t_c$, we will set $\alpha_c = 0$ and seek a solution with ρ constant,

$$\rho_c = \left(1 - \frac{t}{t_c}\right)^\beta \tag{17.94}$$

with $\beta = 1/2$, in the $N \to \infty$ limit.

Thus, in the broken symmetry in $D > 2$, $\alpha_c = 0$ and the fields $\boldsymbol{z}_\perp$ are massless, i.e. the $2(N-1)$ Goldstone bosons. In this phase the gauge field $\mathcal{A}_\mu$ is massive, and the mass is ρ_c^2. So, in the broken symmetry phase the gauge field is "Higgsed". In chapter 18 we will discuss the Higgs mechanism. In chapter 19 we will discuss the role of topology in field theory and return to this model to discuss its instantons (and solitons) and their role.

17.4 The Gross-Neveu model in the large N limit

The Gross-Neveu model is a theory of interacting massless Dirac fermions in 1 + 1 dimensions whose Lagrangian (in the Minkowski metric) is (Gross and Neveu, 1974)

$$\mathcal{L} = \bar{\psi}_a i \not{\partial} \psi_a + \frac{g}{2} \left(\bar{\psi}_a \psi_a\right)^2 \tag{17.95}$$

Here the Dirac fermions are bispinors and the index $a = 1, \dots, N$ labels the fermionic "flavors". The theory has a global $SU(N)$ flavor symmetry,

$$\psi \mapsto U\psi \tag{17.96}$$

where $U \in SU(N)$. This theory is also invariant under a *discrete* chiral transformations

$$\psi \mapsto \gamma_5 \psi \tag{17.97}$$

where γ_5 is a hermitian 2×2 (with $\gamma_5^2 = I$) Dirac matrix that anticommutes with the Dirac matrices γ_μ. Hence, we have the algebra,

$$\left\{\gamma_\mu, \gamma_\nu\right\} = 2 g_{\mu\nu} I, \quad \left\{\gamma_5, \gamma_\mu\right\} = 0 \tag{17.98}$$

with I being the 2×2 identity matrix and $g_{\mu\nu} = \text{diag}(1, -1)$ the metric of $D = 2$-dimensional Minkowski space-time.

The Gross-Neveu model has a $U(N)$ flavor symmetry and a $\mathbb{Z}_2$ chiral symmetry. The fermion mass bilinear is odd under the discrete chiral transformation, Eq.(17.97), i.e.

$$\bar{\psi}\psi \mapsto -\bar{\psi}\psi \tag{17.99}$$

We will see that in the Gross-Neveu model the discrete chiral symmetry is broken spontaneously and there is a dynamical generation of a mass for the fermions.

We can also define a chiral version of the Gross-Neveu model whose Lagrangian is

$$\mathcal{L} = \bar{\psi}_a i\partial\!\!\!/ \psi_a + \frac{g}{2}\left((\bar{\psi}_a\psi_a)^2 - (\bar{\psi}_a\gamma_5\psi_a)^2 \right) \tag{17.100}$$

which, in addition to the global $SU(N)$ flavor symmetry, now has a $U(1)$ chiral symmetry under the transformations

$$\psi \mapsto e^{i\theta\gamma_5}\ \psi \tag{17.101}$$

The fermion bilinears can be put together into a two-component real vector field $(\bar{\psi}\psi, i\bar{\psi}\gamma_5\psi)$ that transforms as rotation by a global angle 2θ under the global $U(1)$ chiral symmetry. We will see that this symmetry is (almost) spontaneously broken.

We will see that in $D = 2$ space-time dimensions the Gross-Neveu model (both chiral and non-chiral) are asymptotically free and exhibit dynamical mass generation. As presented here, the non-chiral Gross-Neveu can also be defined in higher dimensions with the proviso that it is no longer renormalizable, meaning that with a suitable regularization it has a phase transition. In $D = 4$ dimensions this model is closely related to the Fermi theory of weak interactions. The chiral version of the model is closely related to the massless Thirring model (in $D = 2$) and to the Nambu-Jona-Lasinio model (in $D = 4$) (Nambu and Jona-Lasinio, 1961).

17.4.1 The non-chiral Gross-Neveu model

Both versions of the Gross-Neveu model can be solved in the large-N limit. We begin with the non-chiral model, Eq.(17.95), by decoupling the four-fermion interaction term using a real scalar field σ. Upon scaling the coupling constant $g = g_0/N$, Lagrangian now is

$$\mathcal{L} = \bar{\psi}_a i\partial\!\!\!/ \psi_a - \sigma\bar{\psi}_a\psi_a - \frac{N}{2g_0}\sigma^2 \tag{17.102}$$

Now we see that the discrete chiral symmetry is equivalent to the $\mathbb{Z}_2$ ("Ising") symmetry $\sigma(x) \mapsto -\sigma(x)$. Moreover, if the $\langle\sigma\rangle \neq 0$, the Ising symmetry would to be broken spontaneously and $\langle\bar{\psi}\psi\rangle \neq 0$ as well. Furthermore, a non-vanishing value for $\langle\sigma\rangle$ means that the fermions are became dynamically massive. Thus, this is a theory of dynamical mass generation by the spontaneous breaking of the chiral symmetry.

We now proceed to study this theory in the large-N limit. To this end we first integrate out the fermionic fields and obtain the following effective action for the σ field, S_{eff}, with

$$S_{\text{eff}} = -i\text{tr}\ln\left(i\partial\!\!\!/ - \sigma\right) - \int d^D x\, \frac{\sigma^2}{2g_0} \tag{17.103}$$

where the first term now arises from the fermion determinant, i.e. from summing over fermion bubble diagrams. The partition function now is

$$Z = \int \mathcal{D}\sigma \exp(iNS_{\text{eff}}[\sigma]) \tag{17.104}$$

Thus, as in the examples of the non-linear sigma model and the $\mathbb{CP}^{N-1}$ models, the large-N limit of the theory is the semiclassical approximation of a field that couples to a composite operator which in the case of the Gross-Neveu model is the fermion mass bilinear, $\bar{\psi}\psi$. The main difference is that the Gross-Neveu model is a fermionic theory which is the reason for the negative sign in front of the first term in the effective action, aside from the fact that the determinant involves the Dirac and not the Klein-Gordon operator. Notice that we are working in Minkowski spacetime (hence the factor of i in the first term of the effective action).

The stationary (saddle-point) equation is

$$\frac{\partial S_{\text{eff}}}{\partial\sigma(x)} = -i\langle x|\frac{1}{i\partial\!\!\!/ - \sigma_c}|x\rangle - \frac{\sigma_c}{g_0} = 0 \tag{17.105}$$

where σ_c is the saddle-point (uniform) value of the field $\sigma(x)$, where we can identify

$$\mathcal{S}_{ab}(x-y;m) = -i\langle x|\frac{1}{i\partial\!\!\!/ - m}|y\rangle\delta_{ab} \tag{17.106}$$

with the Feynman propagator for a Dirac field with mass $m \equiv \sigma_c$.

Using the momentum space form for the Dirac propagator we can readily write the saddle-point equation as

$$\text{tr}\int \frac{d^D p}{(2\pi)^D}\frac{-i}{p\!\!\!/ - \sigma_c} = \frac{\sigma_c}{g_0} \tag{17.107}$$

where the usual Feynman contour prescription has been assumed, and the trace runs over the Dirac indices. This result, together with the definition of the propagator, implies that at $N = \infty$ we can identify the chiral condensate with

$$\langle \bar{\psi}\psi \rangle = \frac{\sigma_c}{g_0} \tag{17.108}$$

Hence we have a chiral condensate if $\sigma_c \neq 0$.

We now use the Dirac algebra and a Wick rotation of the integration path into the complex plane, $ip_0 \mapsto p_D$, to write the saddle-point equation as

$$2 \int \frac{d^D p}{(2\pi)^D} \frac{\sigma_c}{p^2 + \sigma_c^2} = \frac{\sigma_c}{g_0} \tag{17.109}$$

Therefore, in this theory either $\sigma_c = 0$, and the chiral symmetry is unbroken, or we have a spontaneously broken discrete chiral symmetry state with $\sigma_c \neq 0$. In this case the value of σ_c is the solution of the equation

$$2 \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + \sigma_c^2} = \frac{1}{g_0} \tag{17.110}$$

which is identical to the saddle point equation we found in the $\mathbb{CP}^{N-1}$ model, c.f. Eq.(17.73). Therefore it has the same solution.

Here too, we will first define a renormalized dimensionless coupling constant t, such that $g_0 = t\kappa^{-\epsilon} Z$, where $\epsilon = D - 2$. The resulting beta function for the dimensionless coupling constant t has the same form as in the non-linear sigma model and the $\mathbb{CP}^{N-1}$ model

$$\beta(t) = \epsilon t - \epsilon \frac{t^2}{t_c} \tag{17.111}$$

but with a value of t_c which is 1/2 of the value for the non-linear sigma model, Eq.(17.44).

We conclude that, for $t > t_c$ the fermion of the theory has the dynamically generated mass

$$m = \sigma_c = \kappa \left(1 - \frac{t_c}{t}\right)^{1/\epsilon} \tag{17.112}$$

and it remains massless for $t < t_c$. Notice that, at $N = \infty$, the dynamically generated mass m and the chiral condensate $\langle \bar{\psi}\psi \rangle$ are related by

$$\langle \bar{\psi}\psi \rangle = \frac{m}{g_0} \tag{17.113}$$

Here too, the case of $D = 2$ is special. Indeed, for $D = 2$ the theory is asymptotically free, and has a dynamically generated mass for all values of the coupling constant which, as $N \to \infty$, is

$$m = \sigma_c = \kappa \exp\left(-\frac{\pi}{t}\right) \tag{17.114}$$

An alternative way to understand what happens is to compute the effective potential for a constant value of the field σ. At $N = \infty$, the partition function is just

$$Z[\sigma_c] = \exp\left(-iNU(\sigma_c)\right) \tag{17.115}$$

where σ_c is the constant value of the field σ that minimizes the potential

$$U(\sigma_c) = i\mathrm{tr}\ln\left(i\partial\!\!\!/ + \sigma_c\right) + \int d^2x \frac{\sigma_c^2}{2g_0} \tag{17.116}$$

Using the properties of the 2D Dirac gamma matrices we have

$$\mathrm{tr}\ln(i\partial\!\!\!/ - \sigma_c) = \mathrm{tr}\ln(\partial^2 + \sigma_c^2) \tag{17.117}$$

Then, after a Wick rotation, the potential becomes

$$U(\sigma_c) = \mathrm{tr}\ln(-\nabla^2 + \sigma_c^2) + V\frac{\sigma_c^2}{2g_0} \tag{17.118}$$

where $V = L^2$ is the volume of $D = 2$ Euclidean spacetime. Hence, the

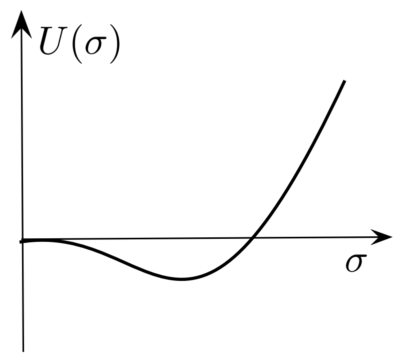


Figure 17.6 Effective potential for the Gross-Neveu model.

problem reduces to the computation of the determinant of the Euclidean Klein-Gordon operator. Using the expression for the determinant of the 2D Euclidean Klein-Gordon operator computed using the ζ-function regularization in Section 8.8, see Eq.(8.212), we can write the potential $U(\sigma)$ as

$$U(\sigma) = V\frac{\sigma^2}{4\pi}\left[\ln\left(\frac{\sigma^2}{\kappa^2}\right) - 1\right] + V\frac{\sigma^2}{2g_0} \tag{17.119}$$

This potential, shown in Fig.17.6, has a minimum at the value of σ_c obtained in Eq. (17.114).

The relation of Eq.(17.108) allows us to identify the fluctuations of the field σ with the fluctuations of the chiral condensate. Thus, the propagator of the σ field is the propagator of the composite operator $\bar{\psi}\psi$. Expanding the effective action to quadratic order in the fluctuations about its expectation value, $\tilde{\sigma}(x) = \sigma(x) - \sigma_c$, we find

$$S_{\text{eff}}[\tilde{\sigma}] = -\frac{N}{2}\int d^2x \int d^2y\, \tilde{\sigma}(x)K(x-y)\tilde{\sigma}(y) \tag{17.120}$$

where the kernel is

$$K(x-y) = \text{tr}\left(S_F(x,y)S_F(y,x)\right) + \frac{1}{g_0}\delta(x-y) \tag{17.121}$$

Here $S_F(x,y)$ is the Feynman propagator for a massive Dirac field in two dimensions, which, upon a Wick rotation, falls off exponentially on distances long compared to $\xi = \sigma_c^{-1}$. This result also implies that the correlator of the fluctuations of the composite operator $:\bar{\psi}\psi(x):=\bar{\psi}\psi(x) - \langle\bar{\psi}\psi(x)\rangle$, where the chiral condensate $\langle\bar{\psi}\psi\rangle$ is given by Eq.(17.113), also falls off exponentially with distance,

$$\langle:\bar{\psi}\psi(x)::\bar{\psi}\psi(y):\rangle \sim \exp(-|x-y|/\xi) \tag{17.122}$$

It is straightforward to see that, again in the $N \to \infty$ limit, the composite operator behaves as a scalar bound state with mass twice the fermion mass. Hence at $N = \infty$ the composite operator represents a bound state on threshold, i.e. the binding energy is $O(1/N)$.

17.4.2 The chiral Gross-Neveu model

We will close with a brief discussion of the *chiral* Gross-Neveu model whose Lagrangian is given in Eq.(17.100). This Lagrangian is invariant under the global $U(1)$ chiral transformations of Eq.(17.101). This will lead to important changes. Since we know that in the non-chiral case the global discrete chiral symmetry is spontaneously broken we may suspect that this may also be the case in the chiral Gross-neveu model as well. We will see that in 2D spontaneous breaking of a $U(1)$ symmetry is subtle and as a result this claim is almost (but not completely) correct.

To study this theory in its large-N limit (here N is the number of Dirac fermion flavors) we will proceed as before and use a Hubbard-Stratonovich decoupling of the quartic fermionic interactions. Since the Lagrangian has two quartic terms we will need two real scalar fields, which we will denote by $\sigma(x)$ and $\pi(x)$, respectively. The partition function for $D = 2$ now is

$$Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}\sigma\mathcal{D}\pi \exp\left(iS(\bar{\psi},\psi,\sigma,\pi)\right) \tag{17.123}$$

where the action now is

$$S = \int d^2x \bar{\psi} \left(i\partial\!\!\!/ + \sigma(x) + i\pi(x)\gamma_5 \right) \psi - \frac{N}{2g_0} \int d^2x \left(\sigma^2(x) + \pi^2(x) \right) \quad (17.124)$$

Under the continuous chiral symmetry of Eq.(17.101) the two component real field (σ, π) transforms as a rotation by a global angle 2θ.

Again, we now integrate out the fermions and obtain the effective action for the fields σ and π,

$$S_{\text{eff}}[\sigma, \pi] = -iN \text{tr} \ln \left(i\partial\!\!\!/ + \sigma(x) + i\pi(x)\gamma_5 \right) - \frac{N}{2g_0} \int d^2x \left(\sigma^2(x) + \pi^2(x) \right) \quad (17.125)$$

The $U(1)$ symmetry requires that the effective potential can only depend on $\sigma_c^2 + \pi_c^2$, where σ_c and π_c are the solutions of the saddle point equations. After tracing over the Dirac indices and a Wick rotation of the integration contours, the saddle point equations are

$$2\sigma_c \int \frac{d^2p}{(2\pi)^2} \frac{1}{p^2 + \sigma_c^2 + \pi_c^2} = \frac{\sigma_c}{g_0}, \quad 2\pi_c \int \frac{d^2p}{(2\pi)^2} \frac{1}{p^2 + \sigma_c^2 + \pi_c^2} = \frac{\pi_c}{g_0} \quad (17.126)$$

Therefore, if (σ_c, π_c) is a solution of these saddle-point equations, any other uniform configuration obtained by a rotation is also a solution. Hence, any solution will break the continuous chiral symmetry spontaneously. Another way to see that this must be the case is to compute the effective potential $U(\sigma, \pi)$. A simple calculation shows that the result is the same as for the non-chiral Gross-Neveu model of Eq.(17.119) with the replacement $\sigma \to (\sigma^2 + \pi^2)^{1/2}$. Thus, the effective potential has the standard "Mexican hat" shape of a system with a $U(1)$ global symmetry.

In $D = 2$ the chiral Gross-Neveu model is asymptotically free, as is the non-chiral model. In fact, the beta function is the same for both models. I addition, the chiral model also has dynamical mass generation in $D = 2$ dimensions (and after a phase transition for $D > 2$). The main difference between the chiral and the non-chiral models is the way the chiral symmetry is broken. Since the spontaneously broken chiral symmetry is continuous we expect that there should be an associated Goldstone boson.

A straightforward way to see this is to rewrite the σ and π fields in terms on an amplitude field ρ and a phase field θ,

$$\sigma(x) + i\pi(x)\gamma_5 = \rho(x) \exp(i\theta(x)\gamma_5) \quad (17.127)$$

Clearly the amplitude field $\rho(x)$ will be massive and, at $N = \infty$, is going to be pinned at the value $\rho_c = (\sigma_c^2 + \pi_c^2)^{1/2}$, which is the same as the mass on the non-chiral model. On the other hand, the phase field $\theta(x)$ must be the Goldstone boson. Hence, the symmetry dictates that the effective low-energy action must depend only on the derivatives of θ. To leading order in an expansion in powers of the inverse mass, the low-energy action is

$$S_{\text{eff}}[\theta] = \int d^2x \frac{N}{4\pi} (\partial_\mu \theta(x))^2 \tag{17.128}$$

which shows that the phase field is indeed massless. On the other hand, the propagator of a massless scalar field in $D = 2$ is

$$G(x-y) = \frac{1}{2N} \ln(x-y)^2 \tag{17.129}$$

which does not decay at long distances. In this sense, the field θ does not describe a physically meaningful excitation, and "does not exist" (Coleman, 1973).

However, it is easy to see that, in the large N limit, composite operators such as the fermion bilinears

$$\bar{\psi}(1 \pm \gamma_5)\psi(x) = \rho(x) \exp(\mp i\theta(x)) \tag{17.130}$$

exhibit a power-law decay as a function of distance

$$\begin{aligned} \langle \bar{\psi}(1+\gamma_5)\psi(x)\bar{\psi}(1-\gamma_5)\psi(y)\rangle =& \langle \rho(x)\exp(i\theta(x))\rho(y)\exp(-i\theta(y))\rangle \\ \propto& \frac{\rho_c^2}{|x-y|^{1/N}} \end{aligned} \tag{17.131}$$

albeit with a non-trivial exponent $\propto 1/N$. Thus, the correlator of the fermion mass terms do not approach a constant at infinity, and in this sense there is no chiral condensate. This means that in $D = 2$, although there is a dynamical mass generation, the chiral symmetry is almost (but not quite) spontaneously broken. In chapter 19 we will see that this behavior corresponds to a line of fixed points, and not to a broken symmetry state. We will also return to this point in chapter 21 where we discuss conformal field theories.This behavior is a manifestation of the Mermin-Wagner Theorem (Mermin and Wagner, 1966; Hohenberg, 1967) (known as Coleman's Theorem in high energy physics (Coleman, 1973)) that states that continuous global symmetries in $D = 2$ classical Statistical Mechanics and in 1+1 dimensional Quantum Field Theory cannot be spontaneously broken. On the other hand, for $D > 2$ and for g_0 larger than a critical value, the dynamical mass generation does correspond to a state with a spontaneously broken symmetry.

17.5 QED in the limit of large number of flavors

We will consider now the large N limit of quantum electrodynamics. The Euclidean action is

$$S = \int d^D x \left[\frac{1}{4e^2} F_{\mu\nu}^2 - \bar{\psi} i (\not{\partial} + i \not{A}) \psi \right] \tag{17.132}$$

This theory has a local $U(1)$ gauge invariance and a global $U(N_f) \times U(N_f)$ flavor symmetry. Here $D = 4 - \epsilon$. The coupling constant, i.e. the fine structure constant, is

$$\alpha = \frac{e^2}{4\pi} \kappa^{-\epsilon} \tag{17.133}$$

where κ is the renormalization scale. The one-loop beta-function for theory is

$$\beta(\alpha) = -\epsilon\alpha + \frac{2N_f}{3\pi} \alpha^2 + O(\alpha^3) \tag{17.134}$$

(actually, the beta function is known to 4 loop order). In $D = 4$ dimensions in the IR the theory flows to $\alpha \to 0$, just as in the case of ϕ^4 theory. So it is trivial in the iR. Now, for $D < 4$, the beta function has a finite fixed point at

$$e_*^2 = 24\pi^2 \frac{\epsilon}{4N_f} \kappa^{\epsilon} \tag{17.135}$$

Thus, the coupling constant flows to a finite value in the IR where the theory becomes scale-invariant and non-trivial.

To proceed with the large N_f limit we integrate out the fermions and write the partition function (in the Euclidean signature)

$$Z = \int \mathcal{D} A_\mu i \exp\left(-S_{\text{eff}}[A_\mu]\right) \tag{17.136}$$

where

$$S_{\text{eff}}[A_\mu] = \int d^D x \left[\frac{1}{4e^2} F_{\mu\nu}^2 - N_f \operatorname{tr} \ln\left(\not{\partial} + i\not{A}\right) \right] \tag{17.137}$$

Upon rescaling the coupling constant

$$e^2 = \frac{e_0^2}{N_f} \tag{17.138}$$

we obtain, as before, an action of the form

$$S_{\text{eff}}[A_\mu] = N_f \left[\int d^D x \frac{1}{4e_0^2} F_{\mu\nu}^2 - \operatorname{tr} \ln\left(\not{\partial} + i\not{A}\right) \right] \tag{17.139}$$

In the limit $N_f \to \infty$ the partition function is dominated by the semiclassical configurations. The saddle-point equations

$$\frac{\delta S_{\text{eff}}[A]}{\delta A_\mu(x)} = 0 \tag{17.140}$$

are trivially satisfied by $A_\mu = 0$. The leading $1/N_f$ corrections are obtained by expanding to quadratic order in A_μ,

$$Z[A] = \left[\text{Det}\left(i\not{\partial}\right)\right]^{N_f} \exp\left(-\frac{N_f}{2}\int d^D x \int d^D y \; A_\mu(x) K_{\mu\nu}(x-y) A_\nu(y)\right) \tag{17.141}$$

where the kernel $K_{\mu\nu}(x-y)$ is given by (in momentum space)

$$K_{\mu\nu}(p) = \left[p^2\delta_{\mu\nu} - p_\mu p_\nu\right] K(p) \tag{17.142}$$

where

$$K(p) = \frac{1}{e_0^2} + \frac{D-2}{2(D-1)}\left[b(D)p^{D-4} - a(D)\Lambda^{D-4}\right] + O(\Lambda^{-2}) \tag{17.143}$$

where Λ is the UV momentum cutoff,

$$b(D) = -\frac{\pi}{\sin\left(\frac{\pi D}{2}\right)} \frac{\Gamma^2(D/2)}{\Gamma(D-1)} S_D \tag{17.144}$$

is universal, and $a(D)$ is a correction to scaling that depends on the choice of regularization (and, hence, is not universal).

The fixed point is determined by cancelling out the correction to scaling against the bare Maxwell term. The fixed point for $D < 4$ thus determined is located at the value of the charge

$$e_*^2 = \frac{2(D-1)}{(D-2)a(D)} \frac{\Lambda^\epsilon}{N_f} \tag{17.145}$$

The scale-invariant effective action at the fixed point is the non-local expression

$$S_{\text{eff}}[A_\mu] = \frac{N_f}{4}\frac{(D-2)}{(D-1)} b(D) \int d^D x \int d^D y F_{\mu\nu}(x) G_D(x-y) F_{\mu\nu}(y) \tag{17.146}$$

where $G(x-y)$ is given by

$$G_D(x-y) = \langle x| \frac{1}{(-\partial^2)^{\frac{4-D}{2}}} |y\rangle \tag{17.147}$$

and its Fourier transform is

$$G(p) = p^{D-4} \tag{17.148}$$

In $D = 2$, the model with $N_f = 1$ is known as the Schwinger model (Schwinger, 1962). One can check that the large N_f analysis predicts that in $D = 2$ the effective action is the same as that of a massive scalar field with mass squared $N_f e^2/\pi$. This result agrees with a direct analysis of the Schwinger model in $D = 2$ spacetime dimensions using the chiral anomaly or, what is the same, bosonization, see section 20.9.1. There is a subtlety here, that the large N_f approach superficially misses, and is that in addition to the massive scalar there also $N_f - 1$ massless scalars. This result, easily found in bosonization (see section 20.3), matters in the realization of chiral symmetry, i.e. in the behavior of fermion bilinears.

17.6 Matrix sigma models in the large rank limit

The discussion of the previous sections suggests that theories become simpler and solvable in a suitable large-N limit. We will now see that indeed in this limit theories do become simpler but they are not as simple enough to be solvable as in the cases we have seen. As we will see, the reason for the complexity can be traced back to the fact that in the theories that we have considered the number of Lagrange multiplier fields (Hubbard-Stratonovich fields) is independent of N, which allows for a simple $N \to \infty$ limit. From a perturbative point of view we were able to achieve simplicity since, although the number of diagrams grows with the order of perturbation theory, they do not grow as fast as N. Thus, in the cases that we have examined, the corrections to the large N limit are down by powers of $1/N$.

We will now consider theories in which the scalar field transforms as a rank N *tensor* (rather than a *vector*). For example let $\phi_{ij}(x)$ be an $N \times N$ real matrix field (here $i, j = 1, \ldots, N$) which transforms under the global symmetry group $O(N) \times O(N)$. A non-linear sigma model can be defined by imposing the constraint that the field ϕ is an $O(N)$ rotation matrix and, as such, it must obey the local constraint that the inverse must be its transpose, i.e. $(\phi(x)^{-1})_{ij} = \phi(x)_{ji}$. The partition function must again have a local constraint which now is

$$Z = \int \mathcal{D}\phi \prod_x \delta(\phi_{ij}(x)\phi_{kj}(x) - \delta_{ij}) \exp(-S[\phi]) \tag{17.149}$$

We can now use a representation of the delta function

$$\prod_x \delta(\phi_{ij}(x)\phi_{kj}(x) - \delta_{ij}) = \int \mathcal{D}\lambda_{ij}(x) \exp(i \int d^D x \lambda_{ik}(x)(\phi_{ij}(x)\phi_{kj}(x) - \delta_{ik})) \tag{17.150}$$

Hence, the matrix-valued constraint requires that the Lagrange multiplier field should also be a matrix of the same rank as the field itself. This means that the rank of the Lagrange multiplier field diverges as $N \to \infty$. In contrast, in the theories with vector symmetries (that we have discussed in this chapter) the rank of the Lagrange multiplier field is fixed and independent of N. The same considerations apply to all matrix-valued scalar fields, e.g. the principal chiral models on a Lie group G, or for general Grassmanian manifolds, and to tensor (non-abelian) generalizations of the Gross-Neveu model. We will encounter a similar structure in the case of non-abelian Yang-Mills gauge theory.

Another, and simpler, example is a matrix scalar field theory, in which the field $\phi_{ij}(x)$, with $i, j = 1, \ldots, N$, is a real symmetric matrix, $\phi_{ij} = \phi_{ji}$. In this case, the theory will have a global $O(N)$ symmetry. In another class of theories of this type, the field the field is an $N \times N$ complex hermitian matrix, $\phi_{ij} = \phi^*_{ji}$, and the global symmetry is $SU(N)$. Finally, a third class consists of a theory on $N \times N$ complex matrices, and the global symmetry is $SU(N) \times SU(N)$. Let us consider a theory for a $N \times N$ matrix field ϕ (here the trace acts after matrix multiplication). The general form of the (Euclidean) Lagrangian is

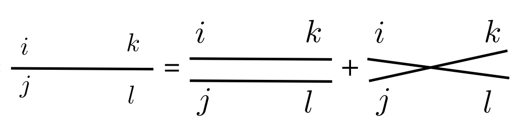
$$\mathcal{L} = \frac{1}{2}\mathrm{tr}(\partial_\mu \phi \partial_\mu \phi^\dagger) + \frac{m^2}{2}\mathrm{tr}(\phi\phi^\dagger) + \alpha\frac{g_4}{N}\mathrm{tr}(\phi\phi^\dagger\phi\phi^\dagger) + \ldots \tag{17.151}$$

where $\alpha = 1$ if the field is a real symmetric matrix, $\alpha = 2$ if it is hermitian, and $\alpha = 4$ if it is complex. If the matrix field is real symmetric a cubic term is also allowed, but not in the other cases. Here g_4 is the coupling constant of the quartic term. Similarly, the coupling constant for an allowed cubic term will be denoted by g_3, etc. In order to obtain a finite large-N limit we will scale the coupling constant of the trilinear term by $1/\sqrt{N}$, the quartic coupling by $1/N$, etc.

The propagator of the matrix field has the form

$$G_{ij|kl}(x-y) = \langle \phi_{ij}(x)\phi_{kl}(y)\rangle \tag{17.152}$$

We will use a "double-line" representation to track the propagation of the indices, introduced by G. 't Hooft in the context of Yang-Mills gauge theory ('t Hooft, 1974). In this picture, the the two lines of the free propagator are unoriented if the matrix field is real symmetric, and can be represented as

$G_{ij|kl}(x-y) =$  (17.153)

If the matrix field is hermitian the two lines are oppositely oriented, and the crossed term of Eq. (17.153) is absent in this case. In the case of a complex matrix both lines have the same orientation. The trilinear and quartic vertices are represented in Fig.17.7 using the double line representation.

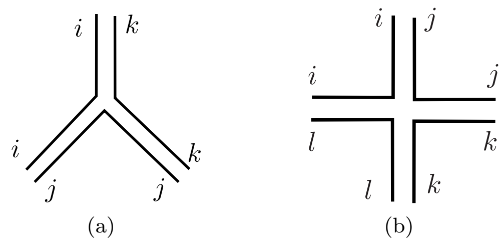


Figure 17.7 a)The trilinear vertex coupling in $O(N) \times O(N)$ matrix field theory has a weight of $1/\sqrt{N}$, and b) the quartic vertex has a weight $1/N$.

We will now see that the expansion in Feynman diagrams has a topological character. We will follow the work by 't Hooft ('t Hooft, 1974) (see also the work by Brézin and coworkers Brézin et al. (1978)). A general Feynman diagram has P propagators, V vertices (of different types) and I closed internal loops. We will denote by V_3 is the number of trilinear vertices, V_4 the number of quartic vertices, etc., with g_3, g_4, etc., the associated coupling constants. Then, for a vacuum diagram, we must have

$$2P = 3V_3 + 4V_4 + \dots \tag{17.154}$$

Each internal loop with a given index is geometrically the face of a polyhedron. Then, the Euler relation says that

$$V - P + I = \chi = 2 - 2H \tag{17.155}$$

where χ is the Euler character of the surface, H is the number of holes (the genus) on the surface on which the polyhedron is drawn (zero for a plane or sphere, one for a torus, etc.) With these definitions, since each closed loop contributes a factor of N, the contribution of the diagrams is proportional to

$$g_3^{V_3} g_4^{V_4} \dots N^I = (g_3\sqrt{N})^{V_3}(g_4 N)^{V_4} \dots N^{2-H} \tag{17.156}$$

Therefore, provided each coupling is scaled by the appropriate power of N (e.g. $g_r \propto N^{1-r/2}$), the vacuum energy (in units of N^2) has a finite large-N limit given by the diagrams with $H = 0$ (no handles, e.g. the sphere). The leading correction, which is $O(1/N^2)$, is given by diagrams drawn on the torus, and so forth. This means that, in the $N \to \infty$ limit, the planar diagrams (those with $H = 0$) yield the exact answer. This also means that the

$1/N$ expansion of these theories, is a topological expansion, i.e. an expansion in powers of $1/N^{2H}$, where H is the genus of the surface on which the diagrams are drawn.

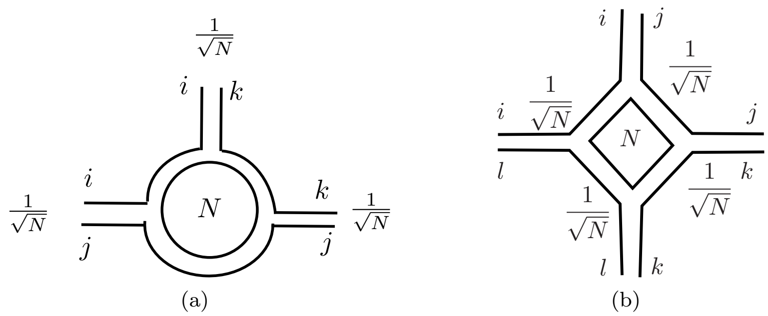


Figure 17.8 a) One-loop planar diagram contribution to the trilinear vertex coupling in $O(N)\times O(N)$ matrix field theory has a weight of $(1/\sqrt{N})^3 N = 1/\sqrt{N}$, and b) one-loop planar diagram contribution to the quartic vertex has a weight $(1/\sqrt{N})^4 N = 1/N$.

In Fig.17.8 we show one loop contributions to the trilinear and quartic vertices. Notice that in these one-loop planar diagrams (i.e. the propagator lines not cross) the internal loop contributes a factor of N from the summation over the internal index, while each of the three vertex insertions contribute with a factor of $g_3/\sqrt{N}$, where g_3 is the coupling constant for the cubic term of the action. Thus the overall contribution is $(1/\sqrt{N})^3 N = 1/\sqrt{N}$ for the trilinear vertex, and $(1/\sqrt{N})^4 N = 1/N$ for the quartic vertex. Thus, these diagrams are of the same order in N as the bare vertex itself but of order g_3^3 and g_3^4 in the coupling constant, respectively. It is now easy to see that subdividing the internal loop in the diagram by stretching a pair of lines ending at a pair of trilinear vertices leads to a diagram with two loops but *of the same order in* N (i.e. $1/\sqrt{N}$) but of higher order in g_3. The same is true for the quartic vertex.

We can now repeat this process an indefinite number of times and each insertion gets an extra factor on N for the new loop and a factor of $g_3^2(1/\sqrt{N})^2$ for the two trilinear vertices. Hence, in the large-N limit we must sum over *all planar diagrams* of this type of the same order in $1/N$ but of increasing order in the coupling constant g_3. On the other hand, if one of the internal propagator double lines were to be crossed (as in the second term of Eq.(17.153)) then the factor of N will disappear while the overall factor of $1/N^{3/2}$ will remain. Hence, such a non-planar diagram is down by one factor of $1/N$ relative to the planar diagrams. The moral is that the leading order in $1/N$ has diagrams of *all orders* in the coupling constants $g_3, g_4, \ldots$. In this sense, this theory is non-perturbative.

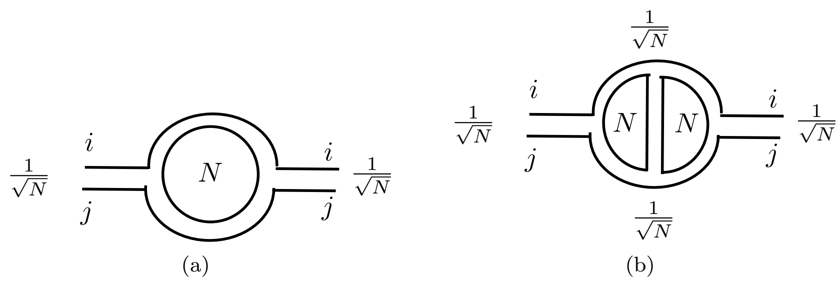


Figure 17.9 Perturbative contributions to the propagator of the $O(N) \times O(N)$ matrix field theory: a) the one loop planar diagram, and b) a two loop planar diagram.

Let us now examine the corrections to the propagator. In Fig.17.9a we show a one-loop contribution. The internal loop has a weight of N and the two vertices contribute with $(1/\sqrt{N})^2$. Thus this diagram has a contribution of $1 = N^0$. On the other hand the diagram of Fig.17.9b has two internal loops and four trilinear vertices. Its weight is $N^2(1/\sqrt{N})^4 = 1$. Hence, it is of the same order as the "leading" diagram. Clearly, here too we can continue with this process ad infinitum by inserting inside each close loop a double propagator line stretched between two trilinear vertices. The resulting diagram is of the same order (1 in this case) in the $1/N$ expansion.

A similar analysis can be made for the four-point function. Some of the contributing planar diagrams are shown in Fig.17.10a-c. The tree-level diagram shown in Fig.17.10a is of order $(1/\sqrt{N})^2 = 1/N$. Therefore at $N = \infty$ the particles of this theory are not interacting and all scattering appears at order $1/N$. The one-loop planar diagram Fig.17.10b also is of order $(1/\sqrt{N})^2 = 1/N$, as is the two-loop planar diagram Fig.17.10c.

As these diagrams show, the large N limit is, in some sense, a theory in which both ladder diagrams and bubble diagrams are summed over consistently. Even though some significant simplification has been achieved by taking the large N limit, the theory is still highly non-trivial, and is still poorly understood.

This analysis leads to a picture for a general planar diagram as a set of nodes (the coordinates of the trilinear vertices) linked to each other (and to the external points) in all possible planar ways by propagator lines. The resulting class of Feynman diagrams have the form of a"fishnet". Such a diagram can also be interpreted as a picture of a tessellated surface in which

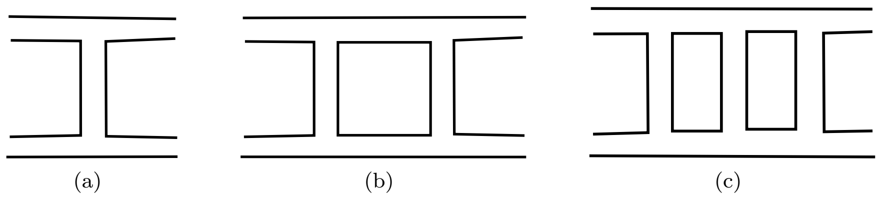


Figure 17.10 Perturbative contributions to the quartic vertex of the $O(N) \times O(N)$ matrix field theory: a) tree–level, b) one loop planar diagram, and c) a two loop planar diagram.

the propagators are the edges of polygons meeting at vertices. In the limit in which the number of insertions goes to infinity the surface approaches a smooth limit. This a general "one-loop" planar diagram is a *sum over all possible surfaces* anchored on the external loop. We will see in our discussion of gauge theory that this picture is equivalent to a type of *string theory*.

17.7 Yang-Mills gauge theory with a large number of colors

We now turn to the important case of the case of gauge theories. In the case of Yang-Mills, this theory is non-abelian and the gauge fields take values in the algebra of a compact Lie group such as $U(N_c)$, known as the color group. In the large N_c limit, $SU(N_c)$ and $U(N_c)$ are essentially equivalent. Using $U(N_c)$ as the color group simplifies the analysis. Thus, it is natural to ask how does this theory behave in the limit of a large number of colors, $N_c \to \infty$. However, when coupled to fermions (quarks) one can also consider the regime in which their flavor symmetry group (a global symmetry of the theory) is $SU(N_f)$, and one may also consider the limit $N_f \to \infty$. These two limits lead to theories with very different character.

17.7.1 Yang-Mills planar diagrams

Let us consider first pure Yang-Mills theory with a color group $U(N_c)$ in the limit $N_c \to \infty$. The Yang-Mills gauge field is a matrix-valued vector field that takes values on the algebra of $U(N_c)$ and, as such, can be written as $A_\mu^{ij}(x) = A_\mu^a(x) t_a^{ij}$, where $i, j = 1, \dots, N_c$, and t_a^{ij} are the N_c^2 generators of $U(N_c)$ in the fundamental representation. In other words, since the Yang-Mills gauge field takes values in the algebra, and it is in the adjoint representation of

$U(N_c)$. The Yang-Mills action is (dropping gauge-fixing terms)

$$S_{YM} = \frac{1}{4} \int d^4x \, (F_{\mu\nu})^i_j \, (F^{\mu\nu})^j_i \tag{17.157}$$

where

$$(F_{\mu\nu})^i_j = i\left([D_\mu, D_\nu]\right)^i_j \tag{17.158}$$

is the Yang-Mills field strength, and

$$D^{ij}_\mu = \delta_{ij}\partial_\mu + i\frac{g}{\sqrt{N_c}} A^{ij}_\mu \tag{17.159}$$

is the covariant derivative in the fundamental representation of $U(N_c)$; here $i, j = 1, \ldots, N_c$. Here we have rescaled the Yang-Mills coupling as $g \to g/\sqrt{N_c}$, and g is now called the 't Hooft coupling constant ('t Hooft, 1974).

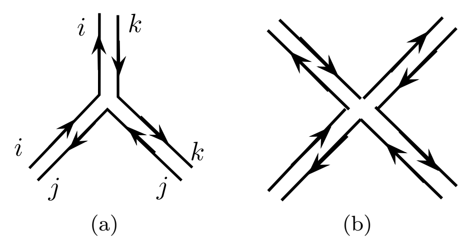


Figure 17.11 Yang-Mills vertices in the double-line notation: a) the cubic vertex, and b) the quartic vertex .

Since the gauge fields are $U(N_c)$ matrices it is natural to use a "double line" representation for its propagator, and to write it in the following form

$$\langle A^{ij}_\mu(x) A^{kl}_\nu \rangle = \delta^l_i \delta^j_k \mathcal{D}_{\mu\nu}(x-y) \equiv \begin{matrix} i \quad\quad l \\ \rightleftarrows \\ j \quad\quad k \end{matrix} \tag{17.160}$$

This way to represent the propagator simply follows the way the color indices are contracted to each other in the Feynman diagrams.

We can also rewrite the cubic and quartic vertices in the double line notation, as shown in Fig.17.11 a and b. Each cubic vertex has a weight of $1/\sqrt{N_c}$ and the quartic vertex also has a weight of $1/N_c$. On the other hand, each loop contributes with a factor of N_c. It is easy to see that, just as in the example of the matrix-valued scalar field, in the large N_c limit the perturbative expansion of the propagator of the gauge field consists in the sum of all possible planar diagrams. To see this we will regard each closed loop as a polygon. Then, the perturbation theory rules will tell us how to fit the polygons together. Let us count the N_c-dependence of a diagram. Each diagram will have V vertices, E edges (the propagators) and F faces (the polygons) and will have an overall weight of

$$N_c^{V-E+F} = N_c^{\chi} \tag{17.161}$$

where

$$\chi = V - E + F \tag{17.162}$$

is a topological invariant known as the *Euler character* of the two-dimensional surface. The diagram, as before, is a tessellation of the surface. Therefore, the weight in N_c of a diagram is given by the Euler character of the surface! However, for a connected orientable surface, the Euler character χ is

$$\chi = 2 - 2H - B \tag{17.163}$$

where H is the number of handles of the surface and B is the number of boundaries (or holes). For an oriented surface without boundaries, $B = 0$ and $H = g$ where g (not to be confused with the coupling constant!) is known as the *genus* of the surface. For the the vacuum diagrams, which do not have edges, we have

$$\chi = 2 - 2g \tag{17.164}$$

Thus, the $1/N_c$ expansion for the vacuum diagrams is a sum over closed surfaces with increasing genus. The leading term is the sphere which has no handles, and hence $g = 0$. The weight for the sphere is N_c^2. The next contribution is the torus which has $g = 1$ and hence $\chi = 0$. Such diagrams scale as $N_c^0 = 1$. Thus, the $1/N_c$ expansion is a sum over surfaces of different topologies! On the other hand, diagrams with quark loops, such as the example shown in Fig.17.12, have one boundary (the quark loop) and hence for them $B = 1$. Therefore, compared to the vacuum diagrams, for diagrams with one quark loop the largest value of $\chi = 1$ and their weight is, at most, N_c.

17.7.2 QCD strings and confinement

In other terms, in the $N_c \to \infty$ limit only planar diagrams with simple topology contribute to the partition function as well as to the correlators, and, therefore, to all physical amplitudes. Although, as we will see, these observations imply that the theory is simpler in this limit, it is by no means as trivial to solve it as in the "vector" large-N limits discussed in this chapter. With some provisos, to this date this problem remains largely unsolved.

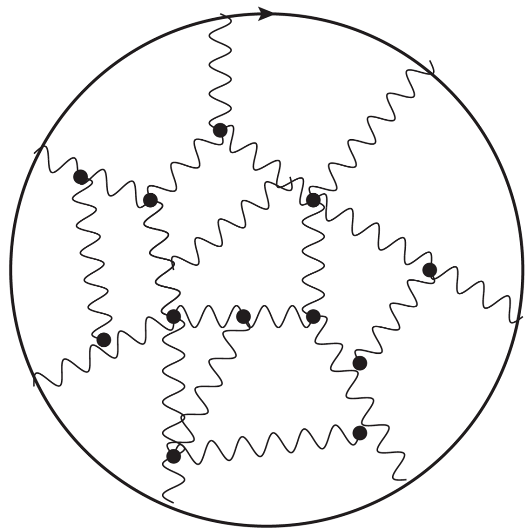


Figure 17.12 A quark loop "fishnet" diagram.

The only controlled solutions are in $1+1$ dimensions where the dynamics of the gauge field is much more trivial.

This is the context in which 't Hooft proposed a solution ('t Hooft, 1974). He considered Yang-Mills theory with fermions $1+1$ dimensions in the large-N_c limit, and showed that in that case it is a confining theory: i.e. the energy to separate a quark-anti-quark pair over a distance R is $V(R) = \sigma R$, where σ is the string tension. Hence, quarks do not exist as asymptotic states and are confined, and the spectrum of states consists of color-singlet bound states, i.e. mesons, hadrons, glue-balls, etc. In fact, in $1+1$ dimensions it is possible to solve the theory even at finite N using non-abelian bosonization methods (whose extension to higher dimensions presently are not known). A similar result is found in the Schwinger model, i.e. quantum electrodynamics also in 1+1 dimensions. Another way to address the problem of a strongly coupled gauge theory is Lattice Gauge Theory which, at the expense of having an explicit Lorentz invariance, can be done in any dimension. We will discuss this approach in another chapter.

The fact that the $1/N_c$ expansion can be related to a sum over surfaces of different topology suggests an alternative physical picture in terms of propagating *strings*. In this picture the total contribution of a single quark loop diagram is a sum over the contributions of all possible surfaces whose boundary is the loop itself. For example if we consider the regime in which the quarks are very heavy. In this case, the diagrams compute the expectation value of a Wilson loop $W[\gamma]$

$$W[\gamma] = \left\langle \mathrm{tr}_F \exp\left(i \oint_\gamma dx_\mu A^\mu \right) \right\rangle \tag{17.165}$$

where γ is the loop and F tells us that the quarks are in the fundamental representation of $U(N_c)$. A Wilson loop represents a process in which a pair of a heavy quark and a heavy anti-quark which are created in the remote past and are annihilated in the future. Thus, we can reinterpret the sum over surfaces as the path-integral of a one-dimensional object, i.e. a string, stretching from the quark to the anti-quark which, as time goes by, sweeps over the surface. This picture then suggests that Yang-Mills theory can be understood as a type of string theory known as QCD strings.

The path integral for a particle is a sum over its histories with a weight given in terms of the action. The Feynman path-integral for the amplitude of a particle to go from $X^\mu(0)$ at time 0 to $X^\mu(\tau)$ at time τ is

$$\langle X^\mu(\tau) | X^\mu(0) \rangle = \int \mathcal{D}X^\mu(t) \exp\Big[- S(X^\mu, \dot{X}^\mu) \Big] \tag{17.166}$$

In the case of a free relativistic particle of mass m, the action is (in the Euclidean signature)

$$S = \int_0^\tau dt\, mc \sqrt{(\dot{X}^\mu)^2} \tag{17.167}$$

and is proper length of the history of the particle.

A string is a curve in spacetime and is given as a map from a two-dimensional worldsheet labelled by (σ, τ) to Minkowski spacetime of the form $X^\mu(\sigma, \tau)$. The (Euclidean) path-integral for a string has the same form as for the particle. For a relativistic string the action is

$$S = \int d\sigma \int d\tau\, T \sqrt{\det[\partial_a X^\mu \partial_b X_\mu]} \tag{17.168}$$

which is the proper area of the surface swept by the string. Here T here denotes the string tension, and the worldsheet indices are $a, b = \sigma, \tau$.

Let us assume that the string ansatz is correct and examine the string path integral. In the case of a particle, the path integral is dominated by the history with minimal proper length. This is the classical trajectory. The weight of the path integral is then determined by the action of the classical trajectory which is proportional to the minimal displacement is spacetime. Likewise, in the case of the string the path integral is dominated by a history corresponding to a *minimal surface*. It is then obvious that, if these assumptions are correct, that the effective potential for a quark-antiquark pair must be linear in their separation and that the coefficient T of the string action, Eq.(17.168) is indeed the string tension. This picture is expected to hold if Yang-Mills theory in the large-N_c limit is a confining theory.

If this picture is correct, in physical terms it means that the chromoelectric field configuration created by a static quark-antiquark pair does not have the dipolar configuration of classical electrodynamics but, rather, that the field lines are compressed by the actual vacuum state of this strongly interacting theory to a long "sausage", as schematically shown in Fig.17.13. In other terms, this picture requires that the Yang-Mills vacuum should *expel* the chromoelectric fields much in the same way as a superconductor expels magnetic fields. In the case of a superconductor flux expulsion (the Meissner effect) results from the fact that the ground state is a condensate of Cooper pairs, an *electric charge condensate*. Hence, in the Yang-Mills case, for the chromoelectric field to be expelled, the vacuum must be a *magnetic condensate*, i.e. a condensate of magnetic monopoles (or "dual superconductor").

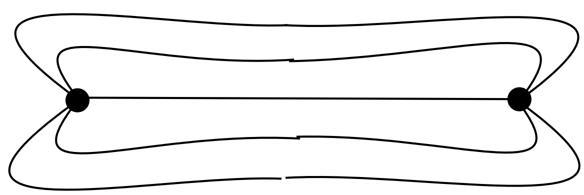


Figure 17.13 Qualitative picture of a QCD string: the quark and the anti-quark create a chromo-electric field that is compressed by the Yang-Mills vacuum into a long sausage shape of length R.

17.7.3 The planar limit and the Maldacena conjecture

Although the surviving diagrams of the large-N_c limit suggest the string picture, it is far from obvious how to derive an effective string action from the sum of planar diagrams. A possible solution of this problem has emerged not from directly summing the planar diagrams but, unexpectedly, from the Maldacena Conjecture, which is based on String Theory. Although String Theory is beyond the scope of this book, I will highlight the main arguments that led to a connection to the large-N_c limit of a gauge theory (Maldacena, 1998).

The argument goes as follows. Maldacena considered a ten-dimensional String Theory on a ten-dimensional spacetime compactified to $S_5 \times AdS_5$

where S_5 is a five-dimensional sphere, S_5, and AdS_5 is a five-dimensional anti-de Sitter (AdS) spacetime, which is a space of constant negative curvature. In $d+2$ dimensions, the metric of AdS_{d+2} spacetime is

$$ds^2 = \frac{R^2}{r^2}\left(-dt^2 + d\boldsymbol{x}^2 + dr^2\right) \tag{17.169}$$

where R is the curvature of AdS, and $0 \leq r < \infty$ is the "radial" coordinate. An AdS spacetime has a boundary at infinity (along the fifth dimension of AdS_5) which behaves as a flat four-dimensional Minkowski spacetime. Maldacena showed that the classical limit of String Theory on this spacetime, which is a supergravity theory, is "dual" to the strong coupling limit of a version of Yang-Mills theory at the four dimensional boundary. However, it is not quite a "plain vanilla" Yang-mills theory but the large-N_c limit of a super-Yang-Mills theory, a supersymmetric version of this theory.

The way this mapping works is as follows. One considers a theory on AdS_5 as a spacetime. In Maldacena's original argument it was a theory of supergravity but this was later extended to other (simpler) theories. One solves the classical equations of motion of this theory. This involves, for example, solving classical Yang-mills theory coupled to classical Einstein gravity, imposing the condition that the spacetime remains asymptotically AdS_5. The Maldacena mapping states that the boundary values of the solutions at infinity yields the expectation values of physical observables in the dual quantum field theory. Thus, the boundary values of the gravitational field yields the expectation value of the energy-momentum tensor operator of the quantum field theory. Similarly, the boundary values of the gauge field leads to the expectation value of the associated conserved currents in the quantum field theory. In this picture the behavior deep in AdS space is viewed as the IR behavior of the field theory. Conversely, the boundary values yield the UV behavior of the (boundary) theory. In this sense, the fifth dimension of AdS_5 plays the role of the renormalization group flow, and there is a one-to-one correspondence between the gravity theory in the bulk AdS and the strongly coupled gauge field theory defined on the boundary. This correspondence is the *Holography Principle* of 't Hooft and Susskind('t Hooft, 1993; Susskind, 1995).

More specifically, consider a theory that has some quantum fields (observables), that we will denote by $\mathcal{O}$, defined at the boundary of AdS_5. Let us consider, for simplicity, a scalar field ϕ defined on AdS_5 whose equations of motion are the covariant Laplace equation, $D_i D^i \phi = 0$, where D_i is the covariant derivative on AdS_5. Let ϕ_0 be the value of the solution of this Laplace equation at the boundary of AdS_5. We will regard ϕ_0 as the source of the field $\mathcal{O}$. Hence, there will be a coupling in the boundary QFT of the form $\int_{\mathcal{M}_4} \phi_0 \mathcal{O}$, where $\mathcal{M}_4$ is four-dimensional Minkowski spacetime. The quantity

$$Z[\phi_0] = \left\langle \exp\left(\int_{\mathcal{M}_4} \phi_0 \mathcal{O} \right) \right\rangle \tag{17.170}$$

is the generating functional of all the correlators of the field $\mathcal{O}$.

Let us now consider the partition function $Z_S[\phi_0]$ in the five-dimensional (supergravity) theory computed with the boundary condition ϕ_0 at AdS_5 infinity. In the limit in which the classical partition function is

$$Z_S[\phi_0] = \exp\left(-I_S[\phi]\right) \tag{17.171}$$

where $I_S[\phi]$ is the classical action for the field ϕ with boundary value ϕ_0. In this language, the conjecture becomes the identity (Witten, 1998; Gubser et al., 1998)

$$\left\langle \exp\left(\int_{\mathcal{M}_4} \phi_0 \mathcal{O} \right) \right\rangle = \exp\left(-I_S[\phi]\right) \tag{17.172}$$

where the left hand side is computed in the strongly coupled QFT, and the right hand side is computed in the classical theory on AdS_5. This dictionary has been extended (or further conjectured) to hold even in cases in which there is no String Theory from which it may descend.

If this conjecture is correct (and it has survived many checks since it was formulated in 1998) computations on a classical theory of gravity on AdS_5 can be mapped to the strong coupling limit of a gauge theory in the large-N limit, known as the large-N limit of super Yang-Mills theory with $\mathcal{N} = 4$ supersymmetries. For technical, but crucial, reasons the conjecture requires that the theory at the boundary be conformally invariant. In another chapter we will discuss conformal invariance in quantum field theory. There we will see that conformal invariance requires that the theory should not have any scales. Now, if what we have discussed here is correct, Yang-Mills theory (and QCD) is expected to be a theory with a dynamically scale, the confinement scale, so it is not scale invariant. Furthermore, it is not supersymmetric. Thus, for the Maldacena Conjecture to explain confinement it must also hold in a theory that breaks the supersymmetry and somehow that is not conformally invariant. Such deformations of this theory have been constructed, but are technically beyond the scope of this book.