Dynamics of $N = 1$ Supersymmetric Field Theories in Four Dimensions

Nathan Seiberg

Introduction

The exact solutions of supersymmetric field theories have four distinct applications:

1. They teach us about the dynamics of strongly coupled field theories. Here the main lesson is the role of electric-magnetic duality in the dynamics.
2. The ideas from field theory have extensions to string theory whose dynamics is richer. This has already led to a revolution in our understanding of string theory.
3. Supersymmetry might be discovered experimentally in the next generation of experiments. In this case, the dynamics of supersymmetric field theories is of direct phenomenological interest, especially for the question of how supersymmetry is spontaneously broken leading to a non-supersymmetric spectrum.
4. Witten's topological field theory is based on supersymmetry. Understanding these theories has applications to mathematics.

In these lectures we will focus on the first application. Other applications were described by other lecturers.

Exact solutions play a crucial role in physics. It is often the case that a simple model exhibits the same phenomena which are also present in more complicated examples. The exact solution of the simple model then teaches us about more generic situations. The main point of these lectures is to show how four dimensional supersymmetric quantum field theories can play a similar role as laboratories and testing grounds for ideas in more generic quantum field theories. This follows from the fact that these theories are more tractable than ordinary, non-supersymmetric theories, and many of their observables can be computed exactly. Yet, it turns out that these theories exhibit explicit examples of various phenomena in quantum field theory. Some of them had been suggested before without an explicit realization and others are completely new.

The generic field theory in four dimensions has no supersymmetry at all. With current techniques its dynamics is difficult to control. With more supersymmetry

---

Notes by Siye Wu

the theories are less generic. The most constrained theories have $N = 4$ supersymmetry. A more general class of theories has $N = 2$. An even more general class has the minimal amount of supersymmetry: $N = 1$ supersymmetry. These lectures are devoted to the analysis of this largest class. (The theories with $N = 1$ supersymmetry are also the relevant ones for the third application above—particle physics phenomenology.) Since this class includes more generic theories than the more special $N = 2$ or $N = 4$ theories, the dynamics which is being exhibited is richer; it includes phenomena which are impossible with more supersymmetry. Therefore these theories are more useful for our first application. They are the most generic, yet still amenable to exact analysis.

In Lecture 1 we discuss the simplest theories (Wess-Zumino models) and prove their non-renormalization theorem. We will then turn to a discussion of gauge theories. The key feature is the existence of flat directions of the classical potential leading to a moduli space of classical ground states. In Lecture 2 we analyze some of the quantum effects in these gauge theories. We show how the moduli space of classical vacua can disappear (by generating a potential) in the quantum theory. Alternatively, the moduli space is deformed. We will also see an example where the space is not deformed, but the massless particles at the singularities are different in the quantum theory than in the short distance classical theory. Finally, in Lecture 3 we complete the analysis of these theories by uncovering a new phenomenon—electric-magnetic duality in $N = 1$ theories. It generalizes previously noticed dualities in $N = 4$ and $N = 2$ theories and explains all the phenomena which we discuss in the earlier lectures.

LECTURE 1
Basic Aspects of $N = 1$ QCD

We will study various supersymmetric theories in 4 dimensions, i.e., the underlying bosonic spacetime is the Minkowski space $\mathbb{R}^{1,3}$. Supersymmetry put constraints on the geometry setting of the field theory. In fact, $N = 0, 1, 2, 4$ supersymmetry corresponds to having a target space of Riemannian, Kähler, hyperKähler (or special Kähler) and flat geometry, respectively.

§1.1. Wess-Zumino model

Wess-Zumino model is a simple example to demonstrate the principle of holomorphy and the non-renormalization theorem. It is an $N = 1$ supersymmetric theory on $\mathbb{R}^{1,3|4}$ with a multiplet of chiral superfields Φ valued in a complex vector space R. In components, we have

$$\Phi = \phi + \theta\,\psi + \theta^2 F, \tag{1.1}$$

where ϕ, ψ and F are complex boson, Weyl fermion and the auxiliary field, respectively, all valued in R. The most general Lagrangian (density) with at most two derivatives is

$$\begin{aligned}\mathcal{L} &= \int d^4\theta\, K(\Phi, \bar{\Phi}) + \left(\int d^2\theta\, W(\Phi) + \text{c.c.}\right)\\ &= g(\partial\phi, \bar{\partial}\bar{\phi}) + g(\not{\partial}\psi, \bar{\psi}) + g(F, \bar{F}) + (\text{Hessian}(W)_\phi(\psi, \psi) + \partial W(F) + \text{c.c.}).\end{aligned} \tag{1.2}$$

Here K is a Kähler potential on R, g, the corresponding metric, and W, a holomorphic function on R called the superpotential. After integrating out the auxiliary field F by its (algebraic) equation of motion, we get a bosonic potential $V = g(\partial W, \overline{\partial W})$. Supersymmetry is unbroken if and only if there exists a solution to the equation $\partial W = 0$.

We consider the following examples. In all the cases considered, let K be a Kähler potential on R such that the metric is flat.

Example 1. Φ is a single chiral superfield valued in $\mathbb{C}$ and $W = \frac{1}{2}m\Phi^2$ ($m \in \mathbb{C}$). This is a free theory in which the boson ϕ and the fermion ψ are of equal mass $|m|$. The free action contains terms $m\psi\psi + |m|^2|\phi|^2$. The phase of $m \in \mathbb{C}$, and more generally that of W, can be rotated away by a $U(1)$ rotation of the fermionic coordinate θ (called the *R-symmetry*). We denote this $U(1)$ group by $U(1)_R$. By (1.1), the $U(1)_R$ rotation produces a redefinition of ψ in component language.

Example 2. Φ is valued in $\mathbb{C}$ and $W = \frac{1}{2}m\Phi^2 + \frac{1}{3}\lambda\Phi^3$ ($m, \lambda \in \mathbb{C}$, $\lambda \neq 0$). This is the original Wess-Zumino model with the Yukawa coupling $\lambda\phi\psi\psi$ and the quartic interaction $|\lambda|^2|\phi|^4$. The equation $\partial W = m\Phi + \lambda\Phi^2 = 0$ has two solutions: $\Phi = 0$ and $\Phi = -\frac{m}{\lambda}$, corresponding to two classical ground states. The bosonic potential of ϕ is shown in Figure 1.1. Under the transformation $\Phi \mapsto -\frac{m}{\lambda} - \Phi$ which interchanges the two ground states, $W \mapsto -W$ up to an additive constant. The sign of W can be rotated away by a $U(1)_R$ symmetry which maps ψ to $i\psi$. Therefore we have a $\mathbb{Z}_4$ discrete symmetry spontaneously broken to $\mathbb{Z}_2$: the generator of $\mathbb{Z}_4$ relates the two ground states 0 and $-\frac{m}{\lambda}$; its square acts as a 2π rotation in $Spin(1,3)$ and is therefore unbroken.

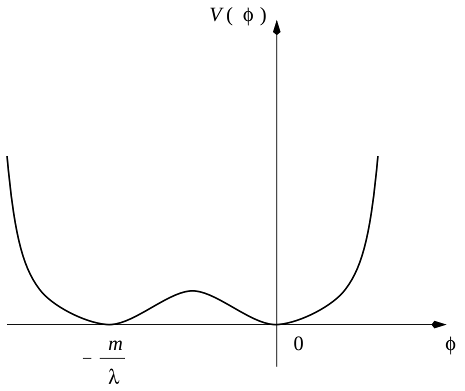

Figure 1.1. The bosonic potential of Example 2, with two minima

Example 3. Φ consists of two complex valued fields L and H (which stand for light and heavy, respectively). The superpotential is $W = \frac{\lambda}{2}LH^2$. The minimum of the bosonic potential is achieved at $H = 0$ and arbitrary L. The field L is massless (light) whereas H has mass $\frac{\lambda}{2}\langle L\rangle$; H is heavy except when $\langle L\rangle = 0$. The moduli space of vacua is shown in Figure 1.2 These (classical) vacua are inequivalent - there are no symmetries that relate them. The metric on the L-space is given by $K_{\text{eff}} = \bar{L}L$, which is the restriction of the original Kähler potential K to the L-space.

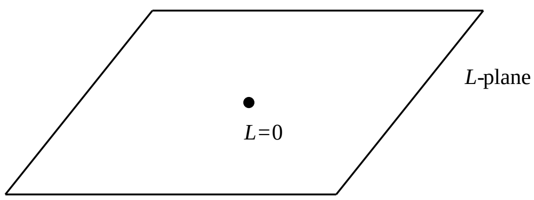

Figure 1.2. The moduli space of vacua of Example 3 with H being massless at $L = 0$

Example 4. Φ consists of three chiral superfields X, Y and Z, and the superpotential is $W = XYZ$. The minimum is achieved at $X = Y = 0$, $Y = Z = 0$ and

$Z = X = 0$. The moduli space of vacua has three branches, as shown in the Figure 1.3. Even in the classical theory, there is a singularity $X = Y = Z = 0$, where new massless fields appear.

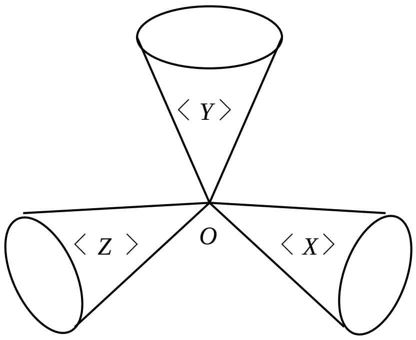

Figure 1.3. The singularity of the moduli space in Example 4

At the quantum level, all theories considered above are renormalizable (but not asymptotically free) and make sense perturbatively in 4 dimensions. The non-renormalization theorem says that the superpotential is not changed by the quantum effects. Although the Wess-Zumino model in 4 dimensions is not an interesting theory non-perturbatively and the standard way to prove the theorem in perturbation theory is to use Feynman diagrams, we are looking for a conceptual and non-perturbative proof which will shed light on other theories.

We need a digression on two notions of effective actions. First, the *1PI* (one-particle irreducible) *effective action* is obtained by adding sources to the (classical) Lagrangian, performing the entire path integral, and doing a Legendre transformation. This procedure fails if the integral diverges in the infrared. On the other hand, in defining the *Wilsonian effective action*, we separate a field, say Φ, into the high energy modes Φ_H and the low energy modes Φ_L in the path integral, i.e., $\int D\Phi = \int D\Phi_H\, D\Phi_L$, do the integration $\int D\Phi_H$ only, and obtain a Lagrangian for Φ_L. Here integrating over Φ_H is putting a cut-off. The infrared difficulty is absent in the Wilsonian effective action. When there are no interacting massless fields, the two notions are identical.

The effective action is subject to constraints of symmetries present in the classical theory. In addition, supersymmetry requires that the superpotential is holomorphic in the effective fields. This is the *principle of holomorphy.* Finally, the effective action is constrained at various limits where the theory is in the perturbative or classical regimes.

A key technique in our proof of non-renormalization theorem is to think of all coupling constants, such as m, λ in Example 2, as expectation values of some background superfields. (We assume that the extended theory make sense and is still supersymmetric at the quantum level.) As a result of holomorphy, the superpotential is holomorphic not only in the effective fields, but also in the coupling constants.

We consider Example 2, in which the superpotential is $W = \frac{1}{2}m\Phi^2 + \frac{1}{3}\lambda\Phi^3$ classically. To find the (Wilsonian) effective superpotential, we regard m, λ as background superfields and assign the following $U(1) \times U(1)_R$ charges to the fields:

$$\begin{array}{c|cc} & U(1) \quad \times & U(1)_R \\ \hline \Phi & 1 & 1 \\ m & -2 & 0 \\ \lambda & -3 & -1 \end{array} \tag{1.3}$$

$U(1)_R$ is the rotation on the fermionic coordinates in superspace. W is invariant under the first $U(1)$ factor and has $U(1)_R$ charge 2, as required by supersymmetry. By holomorphy, W_{eff} is holomorphic not only in Φ, but also in m, λ. (If supersymmetry was unbroken for $W(\Phi)$ but broken for $W(\Phi, m, \lambda)$, this argument would fail.) So

$$\begin{aligned} W_{\text{eff}}(\Phi, m, \lambda) &= f(m\Phi^2, \lambda\Phi^3) \quad \text{(from the first } U(1) \text{ symmetry)} \\ &= m\Phi^2 h\left(\frac{\lambda\Phi}{m}\right) \quad \text{(from } R\text{-charge of } W \text{ is 2)} \end{aligned} \tag{1.4}$$

for some yet unknown holomorphic functions f and h. We can now forget that m, λ are background fields. Next, we consider the limits. If $|\lambda| << 1$, perturbation theory should be valid. So

$$m\Phi^2 h\left(\frac{\lambda\Phi}{m}\right) = \frac{1}{2}m\Phi^2 + \frac{1}{3}\lambda\Phi^3 + \sum_{n=2}^{\infty} a_n \frac{\lambda^n \Phi^{n+2}}{m^{n-1}}. \tag{1.5}$$

The n-th term must come from the diagram in Figure 1.4 with n vertices and $n+2$ external legs, with each internal line contributing a factor of m^{-1}. No loop could appear because each loop would increase the power of λ by 1 but not that of Φ. However for $n \geq 2$, the diagram is not 1PI and therefore can not contribute. So $W_{\text{eff}} = W$, i.e., there is no quantum correction to the superpotential. This is the *non-renormalization theorem.* Notice however there is no claim about the quantum correction to the metric K.

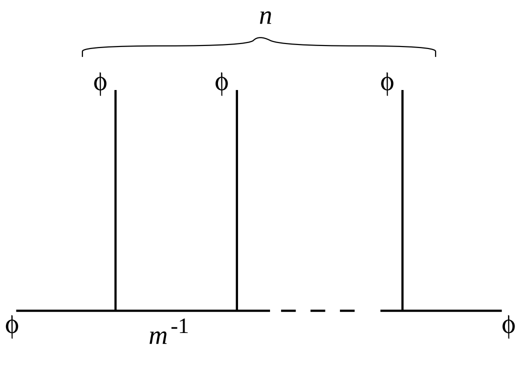

Figure 1.4. The diagram for the n-th term in formula (1.5)

Theories with more fields behave similarly. Consider a variant of Example 3, a theory with two chiral superfields L, H and a superpotential $W = \frac{1}{2}mH^2 + \lambda L^2 H$. Classically, the minimum is reached at $\frac{\partial W}{\partial H} = mH + \lambda L^2 = 0$, $\frac{\partial W}{\partial L} = 2\lambda LH = 0$, i.e.,

$L = H = 0$. So L is a light field. Fixing the value of L, the minimum of W is reached at $H = -\frac{\lambda L^2}{m}$, and we get $W_{\text{eff}} = -\frac{\lambda^2}{2m}L^4$. This procedure is called *integrating out* the heavy field H. The effective superpotential W_{eff} corresponds to the tree diagram in Figure 1.5. Non-renormalization theorem in this case says that W is renormalized by the tree diagram only and not by loop diagrams. In fact, we again regard m, λ as background fields and assign the following $U(1)_H \times U(1)_L \times U(1)_R$ charges to the fields

(1.6)

	$U(1)_H$	$\times$ $U(1)_L$	$\times$ $U(1)_R$
H	1	0	1
L	0	1	1
m	-2	0	0
λ	-1	-2	-1

Invariance under $U(1)_H \times U(1)_L$ and holomorphy imply that

$$W_{\text{eff}}(L, m, \lambda) = h\left(\frac{\lambda^2 L^4}{m}\right) \tag{1.7}$$

for some holomorphic function h. Using the fact that W_{eff} has R-charge 2 or mass dimension 3, we conclude that $W_{\text{eff}} \sim \frac{\lambda^2}{m}L^4$.

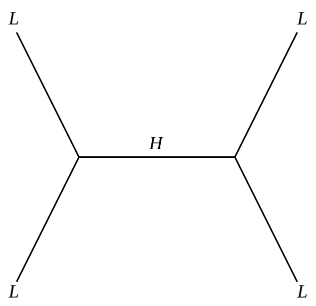

Figure 1.5. The tree diagram that produces the interaction L^4

Several remarks are in order.

1. Supersymmetry is crucial in the above arguments; without holomorphy, we can not rule out terms such as $e^{-1/\bar{\lambda}\lambda}$.
2. Strictly speaking, we need to show the existence of a regularization scheme that respects all the symmetries. Here we simply assume its existence without discussing the details.
3. These arguments are valid both in perturbation theory and non-perturbatively.
4. If supersymmetry is unbroken classically, it is unbroken quantum mechanically because $W_{\text{eff}} = W$.

5. In usual quantum field theories, we have to include in the Lagrangian all terms compatible to renormalizability and the symmetries. For supersymmetric theories, this is not necessary because renormalization does not generate new terms in the superpotential.
6. Supersymmetry is the only known way to have widely separated scales in a theory which are not ruined by quantum corrections. Therefore the gauge hierarchy problem could be explained if supersymmetry holds in Nature.
7. The kinetic term is usually very complicated. However, in $N = 2$ theories (for example, $N = 2$ super Yang-Mills), it is possible to determine the kinetic term by an extra supersymmetry.
8. We will see examples of supersymmetric gauge theories in which the non-renormalization theorem can be avoided: new terms are generated non-perturbatively. Even so, the constraints of symmetries and holomorphy are strong enough to determine the superpotential completely.

§1.2. Pure supersymmetric gauge theory

Recall that an $N = 1$ gauge field is a connection (with some constraints on the curvature) on a principal G-bundle P over the superspace $\mathbb{R}^{1,3|4}$. In components, a constrained connection on the superspace consists of a gluon field A (a connection on the ordinary spacetime $\mathbb{R}^{1,3}$) and a gluino field λ (a section of the spinor bundle over $\mathbb{R}^{1,3}$ twisted by the adjoint bundle). Let W_α be the superfield strength and F, the usual field strength (curvature of A). The Lagrangian density is

$$\mathcal{L} = \frac{\tau}{4\pi i}\int d^2\theta \,\mathrm{tr}\, W_\alpha W^\alpha + \text{c.c.} = \frac{1}{g^2}(\mathrm{tr}\, F\wedge *F + \lambda\, i\,\not{D}\lambda) - \frac{i\theta}{8\pi^2}\mathrm{tr}\, F\wedge F, \tag{1.8}$$

where g is the coupling constant, θ is the θ-angle, and $\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}$.

Classically, there is a $U(1)_R$ symmetry. Its action on the gluino field λ is the chiral rotation $\lambda \mapsto e^{i\alpha\gamma_5}\lambda$ ($e^{i\alpha} \in U(1)_R$). Here γ_5 is ± 1 on the positive (negative) chiral spinors, respectively, and can be regarded as an operator acting on the space of sections of the twisted spinor bundle. The associated current j_5 has an anomaly

$$\partial_\mu j_5^\mu = -\frac{c_2(G)}{32\pi^2} *\mathrm{tr}\, F\wedge F, \tag{1.9}$$

which perturbatively comes from the Feynman diagram in Figure 1.6. Here $c_2(G)$ is the second Casimir of the gauge group G. Numerically, it is equal to $2h$, h being the dual Coxeter number of G. Consequently, $U(1)_R$ is not a symmetry at the quantum level. This is called the *axial anomaly*. In the path-integral language, a classical symmetry survives at the quantum level if it preserves the path-integral measure in addition to the classical action. Under $e^{i\alpha} \in U(1)_R$, the measure $D\lambda$ transforms as

$$D\lambda \mapsto D\lambda\,(\det e^{i\alpha\gamma_5})^{-1} = D\lambda\, e^{-i\alpha\,\mathrm{tr}\gamma_5} = D\lambda\, e^{-i\alpha\,\mathrm{ind}\,\not{D}_{\mathrm{adj}}}. \tag{1.10}$$

Since the index of the Dirac operator twisted by the adjoint bundle is always an integral multiple of $c_2(G) = 2h$, the measure $D\lambda$ is invariant only under the $\mathbb{Z}_{2h}$ subgroup of $U(1)_R$. Thus the symmetry that remains at the quantum level is $\mathbb{Z}_{2h}$. For $G = SU(N_c)$ (N_c is called the *number of colors*), $h = N_c$ and $c_2 = 2N_c$. Hence the symmetry is $\mathbb{Z}_{2N_c}$.

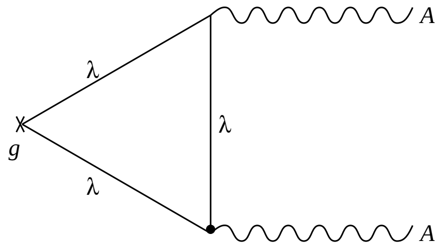

Figure 1.6. A Feynman diagram that causes the axial anomaly

The standard lores are that the theory has a mass gap, that it has confinement, i.e., $\langle W(C)\rangle \sim e^{-\text{Area}}$ for a Wilson loop in the fundamental representation, and that $\langle\lambda\lambda\rangle \neq 0$. The non-zero value of $\langle\lambda\lambda\rangle$ spontaneously breaks the $\mathbb{Z}_{2h}$ symmetry to $\mathbb{Z}_2$. The non-trivial element of $\mathbb{Z}_2$ can be identified as the 2π rotation in $Spin(1,3)$ because its action is equal to $(-1)^F$. Therefore this $\mathbb{Z}_2$ subgroup is unbroken. Consequently the theory has h vacua (N_c if $G = SU(N_c)$). In the k-th vacuum,

$$\langle\lambda\lambda\rangle_k \sim e^{\frac{2\pi i}{h}k}\Lambda^3, \tag{1.11}$$

where Λ is the mass scale generated by renormalization, which we now explain.

Due to the regularization procedure, a quantum field theory depends not only on the classical Lagrangian, but also on a cutoff μ. If we change μ but keep the theory fixed, the coupling constant g then depends on μ. The change of g is governed by

$$\mu\frac{dg}{d\mu} = \beta(g), \tag{1.12}$$

where the $\beta(g)$ is the *β-function*. Up to 1-loop,

$$\beta(g) = -\frac{b_0}{(4\pi)^2}g^3 \tag{1.13}$$

for some number b_0. If $b_0 \neq 0$, a renormalization group invariant mass scale Λ given by the equation

$$\left(\frac{\Lambda}{\mu}\right)^{b_0} = e^{-\frac{8\pi^2}{g(\mu)^2}} \tag{1.14}$$

is generated. This is called *dimensional transmutation*. Notice that Λ is intrinsic to the theory; it is not a cutoff. Because Λ has mass dimension 1, the quantum theory is not scale invariant even if the classical theory is.

In gauge theories, it is convenient to define Λ as a complex number satisfying

$$\left(\frac{\Lambda}{\mu}\right)^{b_0} = e^{-\frac{8\pi^2}{g(\mu)^2}+i\theta} = e^{2\pi i\tau(\mu)}. \tag{1.15}$$

The exponent in (1.15) is simply the value of the classical action for one-instanton. Moreover, any quantity holomorphic in τ, which appears in the Lagrangian density (1.8), is also holomorphic in Λ. Of course Λ defined by (1.15) is not single valued. If a physical quantity expressed in term of Λ is not single valued, it is often because there are many vacua in the theory. For $N = 1$ pure gauge theory, $b_0 = \frac{3}{2}c_2(G) =$

$3h$. Λ^3 is h-valued and there are h vacua in the theory. Fixing a choice of Λ^3, $\langle\lambda\lambda\rangle$ takes values according to (1.11) in the h vacua.

We need a digression on confinement. Recall that the Wilson loop $W_R(C) = \mathrm{tr}_R\, \mathrm{P}\exp\oint_C A$ is the trace in the representation R of G of the holonomy around a closed curve C. For the adjoint representation, we always have $\langle W_{\mathrm{adj}}(C)\rangle \sim e^{-\mathrm{Perimeter}(C)}$. This is because it is possible to screen the charges by the gluons. So the potential is independent of the distance. Likewise for a theory with fundamental quarks that transform in a representation where the center of G acts faithfully, we have $\langle W_R(C)\rangle \sim e^{-\mathrm{Perimeter}(C)}$ for any representation R because the quarks can screen any charges. In both cases, the Wilson loop obeys the perimeter law and the charges are not confined. However in a theory without fundamental quarks, the Wilson loop in fundamental representation may exhibit confinement, in which case we have the area law $\langle W_R(C)\rangle \sim e^{-\mathrm{Area(D)}}$, where D is the disk bounded by C.

§1.3. Supersymmetric QCD

Supersymmetric QCD is a theory of $N = 1$ matter coupled to $N = 1$ gauge theory. It is labeled by a compact Lie group G, a complex representation R of G (where the quarks take value), and a classical superpotential W, which is a G-invariant function on the representation space R. Recall that we have a principal G-bundle $P \to \mathbb{R}^{1,3|4}$ in $N = 1$ gauge theory. Let R be an Hermitian vector space with a unitary representation of G. An $N = 1$ matter (or quark) field Q is a chiral section of the associated bundle $P \times_G R \to \mathbb{R}^{1,3|4}$. The superpotential $W(Q)$ may contain a mass term for Q. It is at most cubic in Q if the theory is to be renormalizable.

Our goal is to find the physics of the theory and its dependence on the bare mass. Does the theory have confinement? Which phase is the theory in? If the theory flow to a non-trivial infrared fixed point (a conformal field theory in four dimensions), what are the anomalous dimensions of the operators?

§1.3.1. Classical theory

We start with the case $W = 0$. Since Q couples to the $N = 1$ gauge theory, the potential for the bosonic part of Q (also denoted by Q) is $V(Q) = |D(Q)|^2$, where D is the moment map of the Hamiltonian G-action on R. The minimum is reached at $Q \in D^{-1}(0)$ (which is non-empty if G is semi-simple). Modulo the gauge transformations (the G-action), the moduli space of classical vacua $\mathcal{M}$ is the symplectic quotient $D^{-1}(0)/G$, which is an open dense set in the complex quotient $R/G^{\mathbb{C}}$.

In the lectures, we will consider theories with the gauge group $G = SU(N_c)$. Let C be the fundamental representation (or the defining representation) of G. C is an Hermitian vector space of dimension N_c with a complex volume form v_c preserved by $G = SU(C)$ (and $G^{\mathbb{C}} = SL(C)$ as well). The matter field consists of N_f copies of quarks in the fundamental representation C and N_f copies of anti-quarks in the conjugate representation $\bar{C} \cong C^*$. N_f is called the *number of flavors*. There are altogether $2N_fN_c$ chiral superfields. It is convenient to introduce two Hermitian vector spaces F, F' of complex dimension N_f and let $R = \mathrm{Hom}(F, C) \oplus \mathrm{Hom}(C, F')$. The values of the quark, anti-quark fields are $Q \in \mathrm{Hom}(F, C)$, $Q' \in \mathrm{Hom}(C, F')$, respectively, or $(Q, Q') \in R$.

The action of the group $G = SU(C)$ on R is Hamiltonian and the moment map $D\colon R \to \mathfrak{su}(C)^*$ is

$$\langle D(Q, Q'), X\rangle = \mathrm{tr}_F Q^\dagger X Q - \mathrm{tr}_{F'} Q' X Q'^\dagger \tag{1.16}$$

for any $X \in \mathfrak{su}(C)$. Under the isomorphism $\mathfrak{su}(C)^* \cong \mathfrak{su}(C)$, D takes values in $\mathfrak{su}(C)$ and is given by

$$D(Q, Q') = \text{traceless part of } i(QQ^\dagger - Q'^\dagger Q'). \tag{1.17}$$

So $(Q, Q') \in D^{-1}(0)$ if and only if $QQ^\dagger - Q'^\dagger Q' = rI_C$ for some $r \in \mathbb{R}$. For $N_f < N_c$, we have $r = 0$. There exist unitary bases of C, F, F' such that

$$Q \sim \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_{N_f} \\ & & \end{pmatrix} \sim Q'^\dagger \tag{1.18}$$

for some $a_k \geq 0$ $(1 \leq k \leq N_c)$. Generically, these $a_k > 0$ and the gauge group $SU(N_c)$ is broken to $SU(N_c - N_f)$, the isotropy group of (Q, Q'). For $N_f \geq N_c$, r can be non-zero. There exist unitary bases such that

$$Q \sim \begin{pmatrix} a_1 & & & \\ & \ddots & & \\ & & a_{N_c} & \\ & & & \end{pmatrix}, \quad Q'^\dagger \sim \begin{pmatrix} a'_1 & & & \\ & \ddots & & \\ & & a'_{N_c} & \\ & & & \end{pmatrix} \tag{1.19}$$

for some $a_k \geq 0$, $a'_k \geq 0$ $(1 \leq k \leq N_c)$ satisfying $a_k^2 - {a'_k}^2 = r$ (independent of k). Generically, the gauge group is completely broken.

We now construct gauge invariants from Q and Q'. First, we have the *meson* field $M = Q'Q$, which is clearly invariant under $SL(C)$. If $N_f \geq N_c$, then we have additional invariants $B = Q^* v_c \in \wedge^{N_c} F^*$ and $B' = Q'_* v_c^* \in \wedge^{N_c} F'$, called the *baryon* fields. They are subject to the constraints

$$B, B' \text{ are decomposable} \tag{1.20}$$

and

$$\wedge^{N_c} M = B \otimes B'. \tag{1.21}$$

If $N_f > N_c$, then further constraints

$$M \wedge B = 0, \qquad B' \wedge M = 0 \tag{1.22}$$

shall be satisfied. In particular, if $N_f = N_c$, then (1.20) is vacuous and (1.21) reduces to

$$\det M = B \otimes B'. \tag{1.23}$$

If $N_f = N_c + 1$, then B, B' can be identified as elements of $F \otimes \wedge^{N_f} F^*$, $\wedge^{N_f} F' \otimes F'^*$, respectively. In this case, (1.20) is again vacuous and (1.22) reduces to

$$MB = 0, \qquad B'M = 0. \tag{1.24}$$

The gauge invariant quantities M, B and B' subject to the constraints (1.20), (1.21) and (1.22) provide a good coordinate system on an open dense subset of $R/SL(C)$ and hence of $\mathcal{M}$. This is because, first, there exist Q, Q' such that $M =$

$Q'Q$, $B = Q^* v_c$ and $B' = Q'_* v_c^*$ if M, B, B' satisfy these constraints,[1] and secondly, generically, the gauge invariants M and B, B' (if $N_f \geq N_c$) determine Q, Q' up to a (complexified) gauge transformation.[2] As a consistency check, when $N_f < N_c$, the complex dimension of a generic $G^{\mathbb{C}}$-orbit in R is $\dim_{\mathbb{C}}(SL(N_c)/SL(N_c - N_f)) = 2N_f N_c - N_f^2$. Therefore the quotient $\mathcal{M}$ has complex dimension $\dim_{\mathbb{C}} R - (2N_f N_c - N_f^2) = N_f^2$ and is labeled by N_f^2 gauge invariants M.

§1.3.2. Quantum theory

With an equal number of quarks and anti-quarks in the representation $R = \mathrm{Hom}(F, C) \oplus \mathrm{Hom}(C, F')$, the quantum theory is free from gauge and gravitational anomalies. (We shall come to this point later.) But the quantum theory is different from the classical theory in several aspects.

First, classically the global symmetry is $U(F) \times U(F') \times U(1)_X$, where $U(1)_X$ is the $U(1)$ rotation on the fermionic coordinate θ in the superspace $\mathbb{R}^{1,3|4}$. (The notation $U(1)_R$ is reserved for an anomaly-free group.) The group $U(1)_X$ and the $U(1)$ subgroups of $U(F)$ and $U(F')$ are anomalous at the quantum level. Let $U(1)_A$, $U(1)_B$ be the anti-diagonally and diagonally embedded $U(1)$ subgroups in $U(F) \times U(F')$, respectively. The weights of these $U(1)$ groups on the fermionic fields and on the path-integral measure (in k-instanton background), calculated similarly as in (1.10), are given by

(1.25)

	$U(1)_A$	$\times$	$U(1)_B$	$\times$	$U(1)_X$
ψ_Q	-1		-1		-1
$\psi_{Q'}$	-1		1		-1
λ	0		0		1
$D\lambda D\psi_Q D\psi_{Q'}$	$2N_f k$		0		$-2(N_c - N_f)k$

So the group $U(1)_B$ is anomaly free. Let $U(1)_R$ be a linear combination of $U(1)_A$ and $U(1)_X$ under which the fermionic measure in (1.25) is invariant. The generator is chosen so that it projects to that of $U(1)_X$. Thus a weight of $U(1)_R$ is fractional in general and is related to those of $U(1)_A$ and $U(1)_X$ by

$$R = X - \frac{N_f - N_c}{N_f} A. \tag{1.26}$$

[1] This is obvious if $N_f < N_c$, when B, B' do not exist and if $N_f = N_c$, when the constraint is (1.23). If $N_f > N_c$, then B (if non-zero) determines a proper subspace $F_0 \subset F$ of dimension N_c. On can choose an isomorphism $Q_0 \colon F_0 \to C$ such that $\wedge^{N_c} Q_0 = B$, let $Q' = M Q_0^{-1}$ and extend Q_0 to Q on F by setting it to be 0 on $F_0^{\perp}$.

[2] When $N_f < N_c$, then generically $\ker Q = 0$, $\mathrm{im} Q$ is a proper subspace in C, $\mathrm{im} Q' = F'$ and $C = \mathrm{im} Q \oplus \ker Q'$. If $Q'_1 Q_1 = Q'Q$ for generic variables, then there is an invertible map $U \colon \mathrm{im} Q \to \mathrm{im} Q_1$ such that $Q_1 = UQ$. Therefore $Q'_1 U = Q'$ on $\mathrm{im} Q$. Extend U by adding an invertible map $\ker Q' \to \ker Q'_1$ such that $U \in SL(C)$. Then $Q'_1 U = Q'$ on C. When $N_f \geq N_c$, then generically $\mathrm{im} Q = C$ and $\ker Q' = 0$. If $Q'_1 Q_1 = Q'Q$, $Q_1^* v_c = Q^* v_c$, $Q'_{1*} v_c^* = Q'_* v_c^*$ for generic variables, then there exists $U \in GL(C)$ defined as follows: for any $x \in C$, choose $y \in F$ such that $Qy = x$, and set $Ux = Q_1 y$. U is well-defined because $\ker Q = \ker Q_1$ (from $Q_1^* v_c = Q^* v_c$). Since $Q_1^* v_c = \det U \, Q^* v_c$, we conclude that $U \in SL(C)$. It is easy to see $Q'_1 = Q' U^{-1}$.

The global symmetry of the quantum theory is $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$. The matter fields Q, Q' (together with their composites M, B, B') transform according to

(1.27)

	$SU(F)$	$\times$	$SU(F')$	$\times$	$U(1)_B$	$\times$	$U(1)_R$
Q	F^*		$\mathbf{1}$		-1		$\frac{N_f-N_c}{N_f}$
Q'	$\mathbf{1}$		F'		1		$\frac{N_f-N_c}{N_f}$
M	F^*		F'		0		$2\frac{N_f-N_c}{N_f}$
B	$\wedge^{N_c} F^*$		$\mathbf{1}$		$-N_c$		$N_c - \frac{N_c^2}{N_f}$
B'	$\mathbf{1}$		$\wedge^{N_c} F'$		N_c		$N_c - \frac{N_c^2}{N_f}$

Here $\mathbf{1}$ is the trivial representation. The representations of the $U(1)$ groups are labeled by their weights.

Secondly, whereas the classical Lagrangian depends only on a dimensionless coupling constant g, we need to introduce a mass scale Λ to define the quantum theory; this is due to dimensional transmutation discussed earlier. The running of the coupling constant is governed by the β-function as in (1.12) and (1.13). For $N = 1$ supersymmetric QCD with gauge group G and matter representation R,

$$b_0 = \tfrac{3}{2}c_2(G) - \tfrac{1}{2}c_2(R), \tag{1.28}$$

where $c_2(R)$ is the value of the Casimir operator in R. In the theories we are considering, $G = SU(C)$ and $R = \mathrm{Hom}(F, C) \oplus \mathrm{Hom}(C, F')$. So $b_0 = 3N_c - N_f$. The theory is asymptotically free if $N_f < 3N_c$. Λ^{b_0} is a scalar under $U(1)_B$. However, since the $U(1)$ subgroups of $U(F)$ and $U(F')$ shift the θ-angle by units of $-N_f$ and N_f, respectively, it is more appropriate to regard Λ^{b_0} in (1.15) as an element in a suitable complex line, i.e.,

$$\Lambda^{b_0} \in \wedge^{N_f} F^* \otimes \wedge^{N_f} F'. \tag{1.29}$$

Thirdly, we have a moduli space of classical vacua $\mathcal{M}$ and want to know whether the vacuum degeneracy will be lifted after quantization. There are several possibilities. The space of vacua could disappear or be modified. If the moduli space remains the same, the metric on it could be changed. We will see examples of each case.

Consider the region in $\mathcal{M}$ where all $a_k >> \Lambda$. The gauge group $SU(N_c)$ is broken to $SU(N_c - N_f)$ if $N_f < N_c - 1$ and is completely broken if otherwise. The light fields are the massless fields M and B, B' (if $N_f \geq N_c$) and the $N = 1$ pure $SU(N_c - N_f)$ gauge fields (if $N_f < N_c - 1$), which are uncoupled to the former. (Gauge bosons in the directions complement to $SU(N_c - N_f)$ have acquired masses of the order of a_k through spontaneous symmetry breaking.) Since the $N = 1$ $SU(N_c - N_f)$ gauge theory has $N_c - N_f$ vacua, we obtain an $N_c - N_f$ cover of the space $\mathcal{M}$. The $N = 1$ pure gauge theory has a mass gap. Going to a lower energy, we have only M, and B, B' (if $N_f \geq N_c$), which are combinations of Q, Q'. We wish to find out whether an effective superpotential W_{eff} could be generated for these light fields. The only holomorphic combination of Q, Q' that is invariant under both the gauge and the global $SU(F) \times SU(F') \times U(1)_B$ symmetry is $\det Q'Q$. Since W_{eff} has $U(1)_R$ charge 2, we get $W_{\text{eff}} \sim (\det Q'Q)^{-1/(N_c - N_f)}$. Since W_{eff} has mass dimension 3 and the only parameter in our theory which has a mass dimension is Λ, we get

$$W_{\text{eff}}(Q, Q') = b_{N_c,N_f} \left(\frac{\Lambda^{3N_c - N_f}}{\det Q'Q} \right)^{\frac{1}{N_c - N_f}} \tag{1.30}$$

for some coefficient b_{N_c,N_f}. (Both $\Lambda^{3N_c-N_f} = \Lambda^{b_0}$ and $\det Q'Q$ are in $\wedge^{N_f} F^* \otimes \wedge^{N_f} F'$. Therefore their ratio is a number.) By holomorphy, (1.30) should also be valid in $\mathcal{M}$ away from infinity. We emphasize here that (1.30) is derived using constraints of global symmetry, holomorphy and the limit far out in the moduli space $\mathcal{M}$. Note that the form of an effective potential need not be constrained by the renormalizability condition.

For $N_f > N_c$, $\det Q'Q = 0$. For $N_f = N_c$, the exponent in (1.30) is singular. So $W_{\text{eff}} = 0$ in both cases. Thus the vacuum degeneracy is not lifted, and we have a moduli space of vacua in the quantum theory. However, the quantum moduli space may be modified away from infinity. For $N_f < N_c$, we have a superpotential (1.30) compatible with all symmetries; this is unlike the Wess-Zumino model. If $b_{N_c,N_f} \neq 0$, then there is no non-renormalization theorem which is valid non-perturbatively for supersymmetric theories with non-Abelian gauge fields.

In $N = 2$ gauge theory, the $N = 1$ supersymmetry already forbids a superpotential. In fact, the pure $N = 2$ gauge theory is an $N = 1$ supersymmetric QCD with matter field Q in the adjoint representation. Since the fermions λ and ψ_Q are both in adjoint representation, and have opposite charges under the $U(1)_R$ rotation in $N = 1$ superspace, this $U(1)_R$ is anomaly-free. So Q has R-charge 0 and it is not possible to have a superpotential holomorphic in Q that has an R-charge 2.

LECTURE 2
Quantum Behavior of Super QCD: N_f Small

We review some aspects of $N = 1$ supersymmetric QCD with N_c number of colors and N_f number of flavors covered in the last lecture. The gauge group is $G = SU(C)$ where C is an Hermitian vector space of dimension N_c equipped with a complex volume form v_c. The matter (quark) fields Q and Q' take values in $\mathrm{Hom}(F, C)$ and $\mathrm{Hom}(C, F')$, respectively. Here F, F' are Hermitian vector spaces of dimension N_f. Under $SU(C)$, the quarks Q, Q' transform as C and $C^* \cong \bar{C}$, respectively. The (anomaly-free) global symmetry is $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$, where $U(1)_B$ is the diagonally embedded $U(1)$ subgroup in $U(F) \times U(F')$, and $U(1)_R$ is a linear combination of the anti-diagonal $U(1)$ subgroup in $U(F) \times U(F')$ and the $U(1)$ rotation in the superspace. Under this global symmetry group, Q, Q' transform as (1.27).

The moduli space $\mathcal{M}_c$ of classical vacua can be parametrized by gauge invariant quantities $M = Q'Q$ and (if $N_f \geq N_c$) $B = Q^* v_c$, $B' = Q'_* v_c^*$ constructed from Q, Q'. In fact, an open dense subset of $\mathcal{M}_c$ is equivalent to $\{M\}$ if $N_f < N_c$, $\{M, B, B' \,|\, \det M = B \otimes B'\}$ if $N_f = N_c$, and $\{M, B, B' \,|\, \wedge^{N_c} M = B \otimes B', MB = B'M = 0\}$ if $N_f = N_c + 1$. This classical moduli space may be modified in the quantum theory. We showed using constraints of global symmetry, holomorphy that for $N_f \geq N_c$, no superpotential can be dynamically generated, whereas for $N_f < N_c$, the only possible form is (1.30). The central question question is whether the coefficient b_{N_c,N_f} is non-zero when $N_f < N_c$. We shall give an outline of the computation of b_{N_c,N_f}.

§2.1. $N_f = N_c - 1$

Consider the case $N_f = N_c - 1$. The gauge group $SU(C)$ is completely broken at a generic point of $\mathcal{M}_c$. Having no massless non-Abelian gluons, the theory is not strongly interacting at low energies. Moreover, there is no branch cut in the superpotential (1.30). This suggests that the theory does not have multiple vacua. In this case, the superpotential is generated by instanton for the following three reasons.

Suppose in a field theory with Euclidean action $S(\phi)$ we do perturbation around an instanton ϕ_0, which a solution of the equation of motion $\frac{\delta S}{\delta \phi}\Big|_{\phi_0} = 0$. Then

$$S(\phi) = S(\phi_0) + \frac{1}{2}\frac{\delta^2 S}{\delta \phi^2}\Big|_{\phi_0} + \cdots . \tag{2.1}$$

The result of any perturbative path integral computation in the instanton background ϕ_0 is of the form

$$e^{-S(\phi_0)}\frac{\det_F}{\det_B}\cdots , \tag{2.2}$$

where $\det_B$, $\det_F$ are the determinants of the operators in $\frac{\delta^2 S}{\delta \phi^2}\Big|_{\phi_0}$ acting on the bosonic, fermionic fields, respectively. In gauge theories, the instanton contribution is non-zero and is proportional to

$$e^{-S(\phi_0)} = e^{-\frac{8\pi^2}{g(\mu)^2}+i\theta} = e^{2\pi i \tau(\mu)}. \tag{2.3}$$

So (2.2) is proportional to $\Lambda^{b_0} = \Lambda^{3N_c - N_f}$ by (1.15); the appropriate power of μ comes from regularizing the determinants and other factors. The superpotential (1.30) has the same power of Λ when $N_f = N_c - 1$. Thus for $N_f = N_c - 1$, the superpotential is instanton generated, and the coefficient b_{N_c,N_c-1} is non-zero. We can choose a renormalization subtraction scheme such that $b_{N_c,N_c-1} = 1$. In other words, the non-zero coefficient is absorbed in the mass scale Λ. Here we see how in a supersymmetric theory, the instanton action, β-function and anomaly are interrelated.

Secondly, in instanton perturbation theory outlined above, because of the zero modes of the Dirac operator, an instanton absorbs $2N_c$ of the λ and N_f each of the ψ_Q and $\psi_{Q'}$. This can be represented graphically by an instanton vertex (Figure 2.1(a)). On the other hand, the Lagrangian contains interactions $\lambda\phi_Q Q$ and $\lambda\psi_{Q'}Q'$ which are represented by the usual vertices (Figure 2.1(b)). In the presence of expectation values $\langle Q\rangle$, $\langle Q'\rangle$, the external Q, Q' lines can be absorbed. When $N_f = N_c - 1$, it is possible to form diagrams with one external line of ψ_Q and $\psi_{Q'}$. (An example is shown in Figure 2.1(c).) The existence of such diagrams is the hallmark of having a superpotential.

Finally, if $N_f = N_c - 1$, instanton computations are reliable and yields finite results. In theories with strong interaction, small instantons correspond to a small coupling constant g (by asymptotic freedom), whereas a large instanton size means strong coupling. When the instanton size becomes large, instantons overlap and have long range interactions. The integration over instanton size diverges. In our theory, at the energy much higher than $\langle Q\rangle$, $\langle Q'\rangle$, the gauge group $SU(C)$ is unbroken, so g is small by asymptotic freedom. It turns out that our theory does not reach strong coupling in the infrared as well. This is because when the energy is smaller than $\langle Q\rangle$, $\langle Q'\rangle$, the gauge group is completely broken. So the instantons can not grow and our theory is free from the usual troubles in the instanton calculations.

Higher order corrections beyond one instanton contribution would have a dependence on Λ not allowed by symmetry and therefore are not present. So when $N_f = N_c - 1$, the one instanton contribution gives the exact answer of the effective superpotential. In this case, we see that there is no non-renormalization theorem because a non-zero superpotential is generated non-perturbatively.

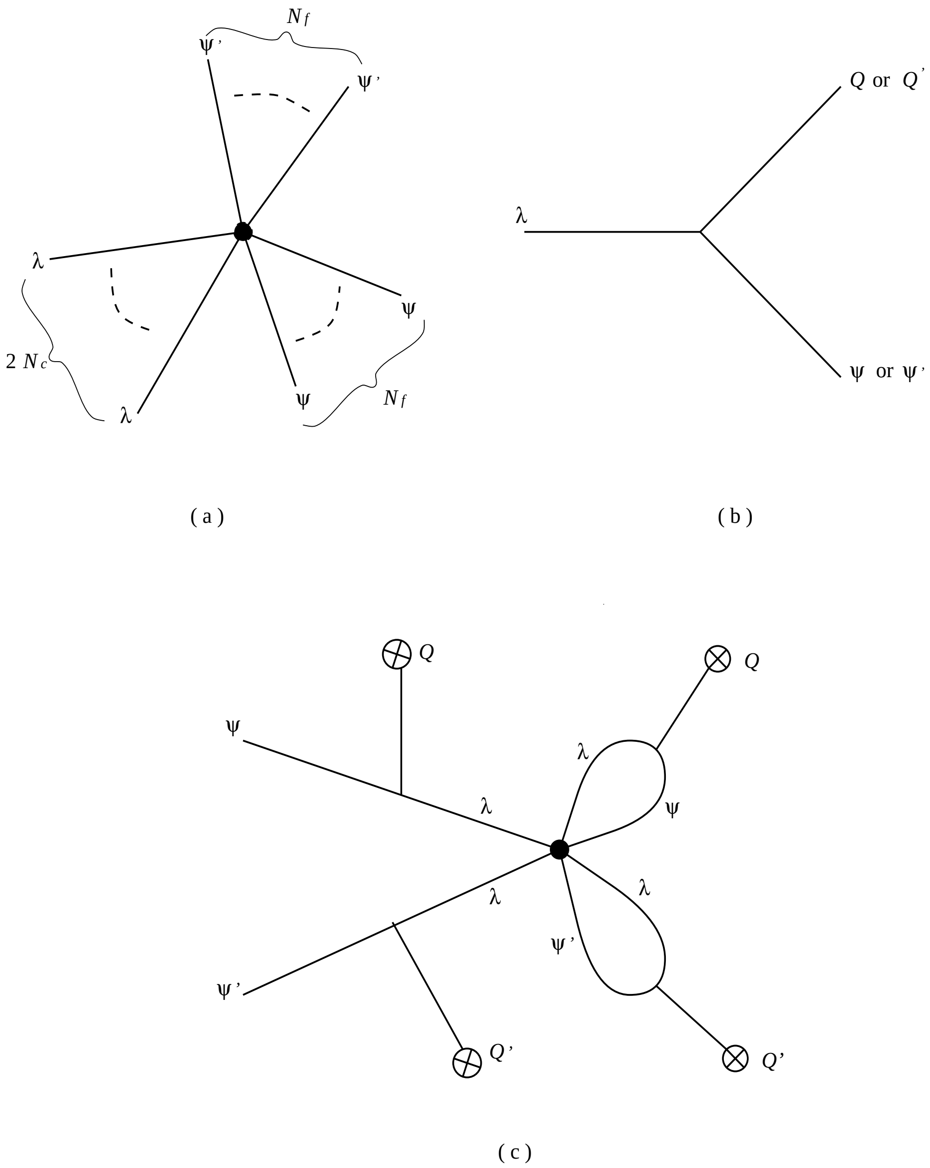

Figure 2.1. (a) The instanton vertex; (b) Yukawa interaction vertex; (c) A diagram with one external line of ψ_Q and $\psi_{Q'}$

§2.2. $N_f < N_c - 1$

We start with the theory with N_c-1 number of flavors, i.e., $\dim F = \dim F' = N_c-1$ and $Q \in \mathrm{Hom}(F, C)$, $Q' \in \mathrm{Hom}(C, F')$. We then add to it a mass term. At short distances (ultraviolet), the mass term is the tree level superpotential

$$W_{\text{tree}} = \operatorname{tr} mQ'Q, \tag{2.4}$$

where $m \in \mathrm{Hom}(F', F)$ is of rank $N_c - N_f - 1$, N_f being the number of light flavors. At long distances (infrared), we shall have an effective superpotential $W_{\text{eff}}(M, \Lambda, m)$ which reduces to (1.30) for $N_c - 1$ flavors when $m = 0$. (Recall that we have chosen a subtraction scheme such that $b_{N_c, N_c-1} = 1$.) Since W_{eff} is holomorphic in m, is invariant under $SU(F) \times SU(F')$, and has R-charge 2, mass dimension 3, it must be of the form

$$W_{\text{eff}} = \frac{\Lambda^{2N_c+1}}{\det M} + \operatorname{tr} mM. \tag{2.5}$$

When the energy drops below m, the effective theory has N_f number of flavors. Take $F = \hat{F} \oplus \hat{F}^{\perp}$ and $F' = \hat{F}' \oplus \hat{F}'^{\perp}$ with $\dim \hat{F} = \dim \hat{F}' = N_f$ such that

$$m = \begin{pmatrix} 0 & 0 \\ 0 & m^{\perp} \end{pmatrix}, \tag{2.6}$$

where $m^{\perp} \in \mathrm{Hom}(\hat{F}^{\perp}, \hat{F}'^{\perp})$ is non-degenerate.

The mass scales $\hat{\Lambda}$ of the infrared theory and Λ of the ultraviolet theory are related by the renormalization group flow. At high energy (large μ), $b_0 = 2N_c + 1$. So the coupling constant $g(\mu)$ runs according to

$$e^{-\frac{8\pi^2}{g(\mu)^2} + i\theta} = \left(\frac{\Lambda}{\mu}\right)^{2N_c+1}. \tag{2.7}$$

At low energy (small μ), $b_0 = 3N_c - N_f$. So $g(\mu)$ flows as

$$e^{-\frac{8\pi^2}{g(\mu)^2} + i\theta} = \left(\frac{\hat{\Lambda}}{\mu}\right)^{3N_c-N_f}. \tag{2.8}$$

Matching $g(\mu)$ at a typical mass scale $\mu = (\det m^{\perp})^{1/(N_c - N_f)}$ (see Figure 2.2), we get in a specific subtraction scheme (called $\overline{DR}$ scheme)

$$\hat{\Lambda}^{3N_c - N_f} = \Lambda^{2N_c+1} \det m^{\perp}. \tag{2.9}$$

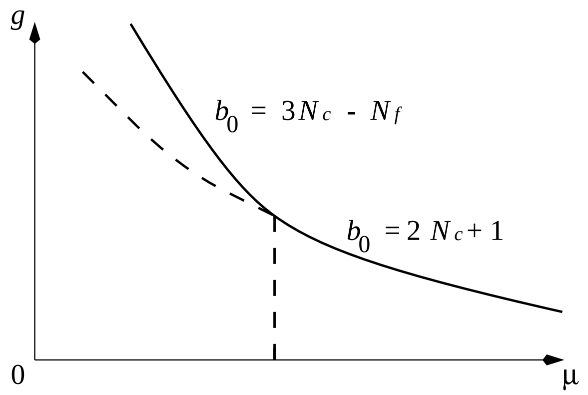

Figure 2.2. Matching of renormalization group flows at high and low energies

Since (2.5) is an effective potential, we can integrate out the heavy quarks classically. The light components of

$$M = \begin{pmatrix} \hat{M} & M_{12} \\ M_{21} & M^{\perp} \end{pmatrix} \tag{2.10}$$

are in $\hat{M} \in \mathrm{Hom}(\hat{F}, \hat{F}')$. We integrate out the heavy components M_{12}, M_{21} and $M^{\perp}$ of M. Fixing $\hat{M}$, the minimum of W_{eff} in (2.5) is reached at $M_{12} = M_{21} = 0$ and

$$M^{\perp} = \left(\frac{\Lambda^{2N_c+1} \det m^{\perp}}{\det \hat{M}} \right)^{\frac{1}{N_c - N_f}} (m^{\perp})^{-1}. \tag{2.11}$$

Substituting these values into (2.5) and using (2.9), we get

$$W_{\text{eff}}(\hat{M}) = (N_c - N_f) \left(\frac{\hat{\Lambda}^{3N_c - N_f}}{\det \hat{M}} \right)^{\frac{1}{N_c - N_f}}. \tag{2.12}$$

This is of the same form as (1.30). Moreover we conclude that in our renormalization scheme, $b_{N_c,N_f} = N_c - N_f \neq 0$ when $N_f < N_c$.

We also want to find out the physical mechanism that generates the superpotential (1.30) when $N_f < N_c - 1$ without embedding the theory into a larger one. At a generic point in $\mathcal{M}$, expectation values $\langle Q \rangle$, $\langle Q' \rangle$ break the $SU(N_c)$ gauge symmetry to $SU(N_c - N_f)$. The gauge bosons not in the unbroken subgroup acquire a mass classically. At an energy below this mass scale, we have an effective theory whose Lagrangian density is

$$\mathcal{L}_{\text{eff}} = \int d^4\theta\, K(M, M^{\dagger}) + \left(\int d^2\theta\, \frac{\tau(M)}{4\pi i} \mathrm{tr}\, W_{\alpha} W^{\alpha} + \text{c.c.} \right), \tag{2.13}$$

where K is a Kähler potential on $\mathcal{M}$ which determines the kinetic term of M, W_α is the superfield strength of the $SU(N_c - N_f)$ gauge theory, and $\tau(M) = \frac{\theta}{2\pi} + \frac{4\pi i}{g(M)^2}$ is determined by the coupling constant $g(M)$ at scale M. Using the 1-loop β-function, $\tau(M) \sim \log\det M$. The pure $SU(N_c - N_f)$ gauge theory has a mass gap and depends on a scale $\Lambda_{SU(N_c - N_f)}$. Below this mass scale, we have M only. Again, $\Lambda_{SU(N_c-N_f)}$ is related to Λ by matching two renormalization group flows at $\mu = (\det M)^{1/2N_f}$. In fact,

$$\Lambda^{3N_c - N_f} = \Lambda^{3(N_c - N_f)}_{SU(N_c - N_f)} \det M, \tag{2.14}$$

where the same subtraction scheme is used. We claim that the effective superpotential of M is generated by the pure gauge theory. Indeed, the leading order term of $W_\alpha W^\alpha$ is $\lambda\lambda$. According to (1.11),

$$\langle \lambda\lambda \rangle \sim \Lambda^3_{SU(N_c - N_f)}. \tag{2.15}$$

Using (2.14), we get

$$\langle \lambda\lambda \rangle \sim \left(\frac{\Lambda^{3N_c - N_f}}{\det M} \right)^{\frac{1}{N_c - N_f}}. \tag{2.16}$$

Therefore the bosonic potentials from (1.30) and (2.13) agree:

$$\frac{\partial W_{\text{eff}}(M)}{\partial M} \sim \frac{\partial \tau(M)}{\partial M} \langle \lambda\lambda \rangle. \tag{2.17}$$

That $b_{N_c,N_f} \neq 0$ also confirms that $\langle \lambda\lambda \rangle \neq 0$ in the pure gauge theory.

The bosonic potential of Q, Q' is qualitatively illustrated in Figure 2.3(a). We see that this potential pushes the vacuum to infinity. Therefore theories with $N_f < N_c$ do not have a vacuum. We can add a bare mass to the theory and recover the original theory by taking the limit as the mass goes to zero. If $N_f < N_c$, the effective superpotential is

$$W_{\text{eff}} = (N_c - N_f)\left(\frac{\Lambda^{3N_c - N_f}}{\det M}\right)^{\frac{1}{N_c - N_f}} + \operatorname{tr} mM. \tag{2.18}$$

The bosonic potential now looks like Figure 2.3(b). The minimum is reached at

$$\langle M \rangle = (\Lambda^{3N_c - N_f} \det m)^{\frac{1}{N_c}}\, m^{-1}. \tag{2.19}$$

The N_c solutions (in the choices of the N_c-th root) correspond to the N_c vacua of the pure gauge theory. Indeed as $m \to 0$, all the eigenvalues of M go to infinity.

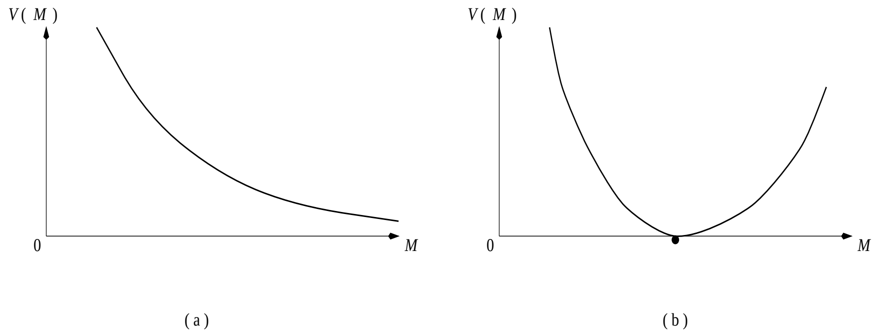

Figure 2.3. The bosonic effective potential in supersymmetric QCD: (a) $m = 0$; (b) $m \neq 0$

The result (2.19) is consistent with all the symmetries. We claim that it is also valid in the range $N_f \geq N_c$, where no superpotential is generated. To show this, we imagine adding a huge mass to some of the quarks, integrating them out, and reducing to a problem of $N_f < N_c$. Since $\langle M \rangle$ is holomorphic in m, (2.19) for $N_f \geq N_c$ follows from the case $N_f < N_c$. For $N_f \geq N_c$, the limit of $\langle M \rangle$ as $m \to 0$ depends on the way the limit is taken. This is not a surprise because there is a moduli space of vacua when $m = 0$. With different ways of taking the limit, the massive theory lands on different points on the moduli space.

§2.3. $N_f = N_c$

As before, we first add a mass term $\operatorname{tr} mM$, then take the limit $m \to 0$. If $m \in \operatorname{Hom}(F', F)$ is invertible then $\langle M \rangle$ is given by (2.19). The invariance under the $U(1)_B$ global symmetry implies that $\langle B \rangle = \langle B' \rangle = 0$. These expectation values lie on the surface

$$\det M - B \otimes B' = \Lambda^{2N_c}. \tag{2.20}$$

Because (2.20) is independent of m, the quantum ground states are not on the classical moduli space $\mathcal{M}_{\text{c}} = \{M, B, B' \,|\, \det M - B \otimes B' = 0\}$ even in the limit $m \to 0$. We call $\mathcal{M}_{\text{q}} = \{M, B, B' \,|\, \det M - B \otimes B' = \Lambda^{2N_c}\}$ the *quantum moduli space.*

The constraint (2.20) is consistent with all the symmetries. We postulate that the low energy effective theory consists of fields M, B, B' constrained on $\mathcal{M}_{\mathrm{q}}$. Unlike the classical moduli space $\mathcal{M}_{\mathrm{c}}$, the quantum moduli space $\mathcal{M}_{\mathrm{q}}$ has no singularities near the origin because of the correction Λ^{2N_c}, as shown in Figure 2.4. This suggests that there are no extra massless particles (other than the modes in the flat directions) – a point we shall check later by 't Hooft's anomaly matching condition. If so, the original gluons are not light, even near the origin, the region of strong coupling. Thus the gluons are confined. Far out along the flat directions (when M is big), the quantum correction is small. There the classical moduli space $\mathcal{M}_{\mathrm{c}}$ is a good approximation to $\mathcal{M}_{\mathrm{q}}$, as expected. Again there are no massless gluons. But in that region, it is more natural to interpret it as Higgs phenomena, in which the gauge group is completely broken by $\langle M\rangle$ and all the gauge bosons acquire a mass. Therefore there is a continuous interpolation between the confinement and the Higgs descriptions.

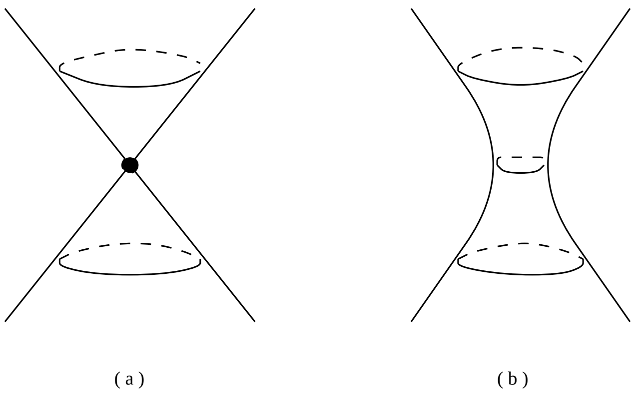

Figure 2.4. (a) The classical moduli space has a singularity at the origin; (b) The quantum moduli space is smooth.

We also remark on chiral symmetry breaking. The global symmetry $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$ is unbroken at the origin of the classical moduli space $\mathcal{M}_{\mathrm{c}}$. But at any point on the quantum moduli space $\mathcal{M}_{\mathrm{q}}$, the symmetry is always broken to a proper subgroup H, the isotropy subgroup, which varies according to the point. The massless particles live in the tangent space of $\mathcal{M}_{\mathrm{q}}$ and form a representation of H. For example, at $M = 0$, $B = B' = \sqrt{-\Lambda^{2N_c}}$, we have $H = SU(F) \times SU(F') \times U(1)_R$. So we have confinement with chiral symmetry breaking.

Since the values of M B, B' in the low energy effective theory are constrained on $\mathcal{M}_{\mathrm{q}}$, the theory can be described by

$$W_{\text{eff}} = \frac{1}{\Lambda^{2N_c-1}} A\,(B \otimes B' - \det M + \Lambda^{2N_c}), \tag{2.21}$$

where A is a superfield (of mass dimension 2) that imposes the constraint (2.20) Now M, B, B' can take arbitrary values. We check that our theory reduces to supersymmetric QCD with $N_f < N_c$ after integrating out the quarks to which a bare mass is added. The effective potential with a mass term is

$$W_{\text{eff}} = \frac{1}{\Lambda^{2N_c-1}} A\,(B \otimes B' - \det M + \Lambda^{2N_c}) + \operatorname{tr} mM, \tag{2.22}$$

where $m \in \operatorname{Hom}(F', F)$ is of rank $N_c - N_f$. (N_f is the number of light flavors.) Write $F = \hat{F} \oplus \hat{F}^{\perp}$ and $F' = \hat{F}' \oplus \hat{F}'^{\perp}$ with $\dim \hat{F} = \dim \hat{F}' = N_f$ such that m has the form (2.6). Again the light components in M of the form (2.10) are in $\hat{M} \in \operatorname{Hom}(\hat{F}, \hat{F}')$. Integrating out the rest of M and B, B' in (2.22), we get

$$\hat{W}_{\text{eff}} = (N_c - N_f) \left(\frac{\hat{\Lambda}^{3N_c - N_f}}{\det \hat{M}} \right)^{\frac{1}{N_c - N_f}}. \tag{2.23}$$

This is precisely the superpotential (1.30) with the correct coefficient. Here $\hat{\Lambda}^{3N_c-N_f}$ $= \Lambda^{2N_c} \det m^{\perp}$, which is again consistent with the renormalization group flow.

§2.4. $N_f = N_c + 1$

Again, we perturb the theory by adding a mass term $\operatorname{tr} mM$. Then we have $\langle B \rangle =$ $\langle B' \rangle = 0$ and using (2.19), $\wedge^{N_c} \langle M \rangle = \Lambda^{2N_c-2} \det m\,(\wedge^{N_c} m^{-1}) = \Lambda^{2N_c-2} m$. In the limit $m \to 0$, the expectation values satisfy the constraints (1.21) and (1.24) of the classical configurations. However, for $m \neq 0$, they are not on the classical moduli space $\mathcal{M}_c$. In fact, all the classical values of M with $B = B' = 0$ can be obtained by taking an appropriate limit $m \to 0$. Unlike the case $N_f = N_c$, there is not a constraint which $\langle B \rangle$, $\langle B' \rangle$ and $\langle M \rangle$ satisfy for all values of m. We propose that the long distance effective theory of the original theory without a bare mass is a Wess-Zumino model with the unconstrained fields M, B, B' and a superpotential

$$W_{\text{eff}}(M, B, B') = \frac{1}{\Lambda^{2N_c-1}} (B' M B - \det M). \tag{2.24}$$

The stationary points of W_{eff} are at $\frac{\partial W}{\partial M} = B \otimes B' - \wedge^{N_c} M = 0$, $\frac{\partial W}{\partial B} = B' M = 0$, $\frac{\partial W}{\partial B'} = B' M = 0$; these are precisely the constraints that every classical configuration satisfies. However, the moduli space is interpreted differently: it is embedded into a larger space. As a result, more particles become light in the effective theory. For example, at $M = 0$, B, B' are the new massless fields.

Similar to the theory with $N_f = N_c$, we see confinement at $M = 0$ since there are no massless gluons. There is again a continuum between the confinement description near the origin and the Higgs description in the semiclassical region. However, at $M = B = B' = 0$, the full global symmetry is unbroken. So we can have confinement without chiral symmetry breaking.

We check that by adding a mass term to the superpotential (2.24), i.e,

$$W_{\text{eff}}(M, B, B') = \frac{1}{\Lambda^{2N_c-1}} (B' M B - \det M) + \operatorname{tr} mM, \tag{2.25}$$

the theory with $N_f = N_c$ can be recovered. Here m is of rank 1. Again we write m in the form (2.6). The light components in

$$M = \begin{pmatrix} \hat{M} & * \\ * & A \end{pmatrix}, \quad B = \begin{pmatrix} * \\ \hat{B} \end{pmatrix}, \quad B' = \begin{pmatrix} * \,, & \hat{B}' \end{pmatrix} \tag{2.26}$$

are $\hat{M}$, A, $\hat{B}$, $\hat{B}'$. Integrating out the rest, we get

$$W_{\text{eff}}(\hat{M}, A, \hat{B}, \hat{B}') = \frac{1}{\Lambda^{2N_c-1}} A\,(\hat{B} \otimes \hat{B}' - \det \hat{M} + \hat{\Lambda}^{2N_c}), \tag{2.27}$$

where $\hat{\Lambda}^{2N_c} = m^1\Lambda^{2N_c-1}$. This agrees with the superpotential for $N_f = N_c$ after a rescaling of the field A.

We shall further check the postulate on the field content of the low energy theory by 't Hooft's anomaly matching condition.

§2.5. 't Hooft anomaly matching condition

A low energy theory may have massless fermions very different from the fundamental theory. The anomaly matching condition of 't Hooft is a highly non-trivial test on the compatibility of the two theories. We first introduce the condition, then apply it to check some of the claims we made about supersymmetric QCD with $N_f = N_c$ and $N_f = N_c + 1$.

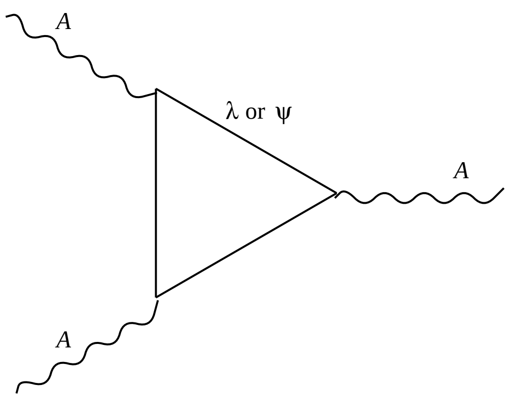

Figure 2.5. The triangular Feynman that causes the gauge anomaly

Any gauge theory with Weyl fermions may suffer from anomalies. If so, the quantum theory is ill-defined. Suppose the gauge group is G and the fermions are in the representation R of G. Perturbatively, the gauge anomaly comes from the triangular Feynman diagram in Figure 2.5. This is proportional to a totally symmetric cubic form d_R on the Lie algebra $\mathfrak{g}$ of G given by

$$d_R(x, y, z) = \mathrm{tr}_R x\{y, z\} \qquad (x, y, z \in \mathfrak{g}). \tag{2.28}$$

$d_R = 0$ if R is a real or quaternionic (pseudo-real) representation. For simple Lie groups, $d_R \neq 0$ only if $\mathfrak{g} = \mathfrak{su}(N_c)$ ($N_c \geq 3$). In the path-integral language, the fermionic path integral is the determinant of the Dirac operator $\not{D}_R^+$ (twisted by R) on the spacetime 4-manifold X. This determinant is a section of the determinant line bundle Det $\not{D}_R^+$ over $\mathcal{A}/\mathcal{G}$, the space $\mathcal{A}$ of connections modulo the group $\mathcal{G}$ of gauge transformations. By local family index theorem, the curvature $\mathcal{F}_R$ of Det $\not{D}_R^+$ evaluated on two tangent vectors B, B' at $A \in \mathcal{A}$ is given by

$$\mathcal{F}_R(B, B') = \frac{-i}{24\pi^3}\int_X d_R(F, F, \Delta_A^{-1} *[B, *B']) + d_R(F, B, B'), \tag{2.29}$$

where F is the curvature of $A \in \mathcal{A}$ and Δ_A, the Laplacian acting on $T_A\mathcal{A}$. The quantum theory makes sense only when the bundle Det $\not{D}_R^+$ is canonically trivial. This requires $\mathcal{F}_R = 0$, hence $d_R = 0$, in agreement with the perturbative argument. Strictly speaking, (2.29) is true only when $\hat{A}(X) = 1$. In general, $\mathcal{F}_R$ contains terms proportional to the linear form tr_R on $\mathfrak{g}$ (which is non-zero only if G has a $U(1)$ factor). This is the mixed gauge and gravitational anomaly. So the quantum theory is free from gauge and gravitational anomalies if and only if the representation R is arranged so that $d_R = \mathrm{tr}_R = 0$. For example, if $G = SU(C)$, it is easy to check that $d_C = -d_{\bar{C}}$ and $d_{\mathrm{adj}} = 0$. So $N = 1$ supersymmetric QCD with equal number of quarks and anti-quarks is free from anomalies.

Recall that the global symmetry of our theory is $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$. At a particular point in the moduli space, this symmetry may be broken to a subgroup H. The massless fermions will form a representation R of H. H is not gauged; if we gauge it (adding gauge fields interacting with the fermions), there will be anomalies proportional to d_R and tr_R. But imagine that in the gauged theory we add some massless fermions in another representation R_0 that couple only to the gauge fields of H. The representation R_0 is chosen so that the combined theory is anomaly free, i.e., $d_R + d_{R_0} = 0$, $\mathrm{tr}_R + \mathrm{tr}_{R_0} = 0$. Suppose our original theory with representation R has a low energy description, in which the fermions are in representation R'. Then the low energy description of the combined theory is anomaly free because it comes from an anomaly free ultraviolet theory. Since the additional massless fermions stay the same, we have $d_{R'} + d_{R_0} = 0$, $\mathrm{tr}_{R'} + \mathrm{tr}_{R_0} = 0$. Therefore

$$d_R = d_{R'}, \quad \mathrm{tr}_R = \mathrm{tr}_{R'}. \tag{2.30}$$

This is the 't Hooft anomaly matching condition. It is a stringent check on whether two theories with fermions in representations R and R' describe the same low energy physics.

Now we check the claimed solutions of the low energy theories with $N_f = N_c$ and $N_f = N_c + 1$ against 't Hooft's criterion. Using (1.27), the fermions in the ultraviolet theory transform under the global symmetry according to

(2.31)

	$SU(F)$	$\times$	$SU(F')$	$\times$	$U(1)_B$	$\times$	$U(1)_R$
ψ_Q	F^*		$\mathbf{1}$		-1		$-\frac{N_c}{N_f}$
$\psi_{Q'}$	$\mathbf{1}$		F'		1		$-\frac{N_c}{N_f}$
λ	$\mathbf{1}$		$\mathbf{1}$		0		1

In the infrared theory, the fermionic components ψ_M, ψ_B, $\psi_{B'}$ of the composite fields M, B, B' transform as

(2.32)

	$SU(F)$	$\times$	$SU(F')$	$\times$	$U(1)_B$	$\times$	$U(1)_R$
ψ_M	F^*		F'		0		$1 - 2\frac{N_c}{N_f}$
ψ_B	$\wedge^{N_c} F^*$		$\mathbf{1}$		$-N_c$		$N_c - 1 - \frac{N_c^2}{N_f}$
$\psi_{B'}$	$\mathbf{1}$		$\wedge^{N_c} F'$		N_c		$N_c - 1 - \frac{N_c^2}{N_f}$

We can calculate the cubic and the trace forms on the Lie algebra $\mathfrak{su}(F) \oplus \mathfrak{su}(F') \oplus \mathfrak{u}(1)_B \oplus \mathfrak{u}(1)_R$ in both the ultraviolet and the infrared theories. The non-zero combinations are given by

(2.33)

	ultraviolet	infrared
$\mathfrak{su}(F)^3$	$N_c\,(-d_F)$	$N_f\,(-d_F) + (-d_{\wedge^{N_c}(F)})$
$\mathfrak{su}(F')^3$	$N_c\,d_{F'}$	$N_f\,d_{F'} + d_{\wedge^{N_c}F'}$
$\mathfrak{su}(F)^2\mathfrak{u}(1)_B$	$N_c\,c_2(F)\cdot(-1)$	$c_2(\wedge^{N_c}F)\cdot(-N_c)$
$\mathfrak{su}(F')^2\mathfrak{u}(1)_B$	$N_c\,c_2(F')\cdot 1$	$c_2(\wedge^{N_c}F')\cdot N_c$
$\mathfrak{su}(F)^2\mathfrak{u}(1)_R$	$N_c\,c_2(F)\cdot(-\frac{N_c}{N_f})$	$N_f\,c_2(F)\cdot(1-\frac{2N_c}{N_f}) + c_2(\wedge^{N_c}F)\cdot(N_c-1-\frac{N_c^2}{N_f})$
$\mathfrak{su}(F')^2\mathfrak{u}(1)_R$	$N_c\,c_2(F')\cdot(-\frac{N_c}{N_f})$	$N_f\,c_2(F')\cdot(1-\frac{2N_c}{N_f}) + c_2(\wedge^{N_c}F')\cdot(N_c-1-\frac{N_c^2}{N_f})$
$\mathfrak{u}(1)_B^2\mathfrak{u}(1)_R$	$N_cN_f[(-1)^2+1^2]\cdot(-\frac{N_c}{N_f})$	$2\binom{N_f}{N_c}[(-N_c)^2+N_c^2]\cdot(N_c-1-\frac{N_c^2}{N_f})$
$\mathfrak{u}(1)_R^3$	$2N_cN_f(-\frac{N_c}{N_f})^3 + (N_c^2-1)1^3$	$N_f^2(1-\frac{2N_c}{N_f})^3 + 2\binom{N_f}{N_c}(N_c-1-\frac{N_c^2}{N_f})^3$
$\mathfrak{u}(1)_R$	$2N_cN_f(-\frac{N_c}{N_f}) + (N_c^2-1)1$	$N_f^2(1-\frac{2N_c}{N_f}) + 2\binom{N_f}{N_c}(N_c-1-\frac{N_c^2}{N_f})$

When $N_f = N_c$, the quantum moduli space $\mathcal{M}_{\mathrm{q}}$ is defined by the constraint (2.20). At the point $M = 0$, $B = B' = \sqrt{-\Lambda^{2N_c}}$, the global symmetry group is broken to $H = SU(F)\times SU(F')\times U(1)_R$. Using the condition $N_f = N_c$, it is straightforward to check that for $\mathfrak{su}(F)^3$, $\mathfrak{su}(F')^3$, $\mathfrak{su}(F)^2\mathfrak{u}(1)_R$ and $\mathfrak{su}(F')^2\mathfrak{u}(1)_R$, the entries in the left and right columns match. However for $\mathfrak{u}(1)_R^3$ and $\mathfrak{u}(1)_R$, the entry in left column is $-(N_c^2+1)$ whereas as that in the right is $-(N_c^2+2)$. The discrepancy is because M, B, B' are constrained on $\mathcal{M}_{\mathrm{q}}$. Alternatively, we can introduce a Lagrange multiplier A as in (2.22). A is invariant under $SU(F)\times SU(F')\times U(1)_B$. Since the R-charges of M, B, B' are 0 and that of W_{eff} should be 2, the R-charge of A is 2. So ψ_A has R-charge 1 and contributes 1^3 and 1 to $\mathfrak{u}(1)_R^3$ and $\mathfrak{u}(1)_R$, respectively, in the right column. Thus we have the matching $-(N_c^2+1) = -(N_c^2+2)+1$. This computation confirms that for $N_f = N_c$, the fields M, B, B' are constrained on $\mathcal{M}_{\mathrm{q}}$ and that there are no additional light fields.

For $N_f = N_c+1$, consider the point $M = B = B' = 0$ on the moduli space, where the full global symmetry group is unbroken. Using $N_f = N_c+1$, it is straightforward to check that in (2.33), all the corresponding entries in the two columns match. This shows that in the low energy theory, M, B, B' are indeed unconstrained and that they are the only light fields.

LECTURE 3
Quantum Behavior of Super QCD: N_f Large

In the previous lectures, we studied $N = 1$ supersymmetric QCD with $N_f \leq N_c + 1$. More concretely, we have an $N = 1$ gauge theory with gauge group $G = SU(C)$ coupled to $N = 1$ matter which consists of the quarks $Q \in \mathrm{Hom}(F, C)$ and the anti-quarks $Q' \in \mathrm{Hom}(C, F')$. $N_c = \dim C$ is called the number of colors and $N_f = \dim F = \dim F'$, the number of flavors. The anomaly-free global symmetry of the theory is $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$, under which Q, Q' transform as (1.27). The moduli space $\mathcal{M}_c$ of classical ground states are parametrized by M, B, B', which are also the fields in the low energy effective theory. Their transformations under the global symmetry are also given by (1.27).

If $N_f < N_c$, a non-zero superpotential (1.30) is dynamically generated. Consequently, the quantum theory has no vacuum. If $N_f = N_c$, the classical moduli space $\mathcal{M}_c$ is modified in the quantum theory. The light fields M, B, B' satisfy the constraint (2.20). If $N_f = N_c + 1$, the quantum moduli space is the same as $\mathcal{M}_c$. But all the components of M, B, B' become physical in the low energy theory. We checked these statements on the low energy field theory using 't Hooft's anomaly matching condition.

In this lecture, we will study the theory for all ranges of N_f. If $N_f \geq 3N_c$, then the β-function $\beta(g) > 0$, so the theory is infrared free. (If $N_f = 3N_c$, the 1-loop β-function vanishes but the 2-loop contribution is positive. See below.) At long distances, we see elementary degrees of freedom: the massless quarks and gluons. We will show that when $\frac{3N_c}{2} < N_f < 3N_c$, the theory has a non-trivial infrared fixed point, which is a superconformal field theory in four dimensions. When $N_c + 2 < N_f < \frac{3N_c}{2}$, the theory is strongly coupled at the infrared. But there is a dual description, the magnetic theory, which is infrared free.

§3.1. $\frac{3N_c}{2} < N_f < 3N_c$: non-trivial infrared fixed points

Up to 2-loop, the β-function is

$$\beta(g) = -\frac{g^3}{16\pi^2}(3N_c - N_f) + \frac{g^5}{128\pi^4}\left(2N_cN_f - 3N_c^2 - \frac{N_f}{N_c}\right) + o(g^7). \tag{3.1}$$

Note that the 2-loop term is positive when $N_f = 3N_c$. Consider the limit $N_c, N_f \to \infty$ while $\frac{N_f}{N_c} = 3 - \epsilon$ is fixed for some $\epsilon > 0$. The β-function is shown in Figure 3.1. If ϵ is sufficiently small, the higher order terms in $\beta(g)$ are negligible. So there is a solution g_* of $\beta(g) = 0$ with

$$g_*^2 N_c = \frac{8\pi^2}{3}\epsilon + o(\epsilon^2). \tag{3.2}$$

The coupling constant runs and stops at this point in the infrared.

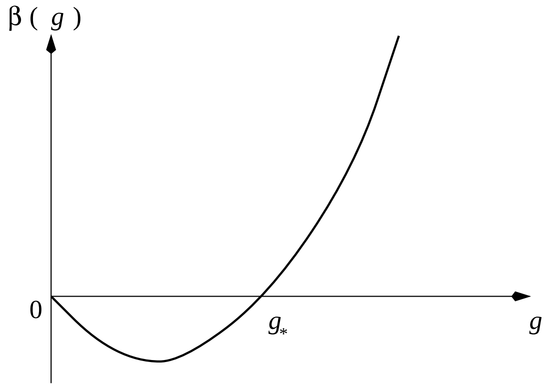

Figure 3.1. The behavior of β-function to two loops

At g_*, the we have a scale invariant and supersymmetric field theory, i.e., a superconformal field theory in four dimensions. It is moreover an non-trivial interacting theory; this will be confirmed by computing the mass dimension of certain fields later. At such a fixed point, the notion of particle is ill-defined. Because of scale invariance, there is no mass gap. The energy spectrum contains a vacuum and a continuum above it. So massless particles can decay arbitrarily since they are not free at long distances. Consequently, there are no well-defined one-particle states in the Hilbert space and there is no S-matrix.

Unlike in two dimensions, the conformal algebra in four dimensions is finite dimensional. Yet it is still possible to extract information from it. We consider radial quantization, i.e., regard the Euclidean spacetime $\mathbb{R}^4 - \{0\}$ as an expanding continuum of S^3. There is a one-to-one correspondence between operators and states in radial quantization. Let $|0\rangle$ be the vacuum. Then an operator $\mathcal{O}$ defines a state $|\mathcal{O}\rangle = \mathcal{O}|0\rangle$ in the Hilbert space. Since the theory is unitary, the operators form a unitary representation of the superconformal algebra.

Recall that the Lie algebra of the conformal group in spacetime of dimension $d \geq 3$ is generated by translations $P_\mu = \partial_\mu$, Lorentz generators $M_{\mu\nu} = x_\mu\partial_\nu - x_\nu\partial_\mu$, dilation $D = -x^\mu\partial_\mu$ and special inversions $K_\mu = 2x_\mu x^\nu\partial_\nu - x^2\partial_\mu$. P_μ, $M_{\mu\nu}$ form the standard Poincaré Lie algebra. The commutation relations of other generators are

$$[D, P_\mu] = P_\mu, \quad [D, K_\mu] = -K_\mu, \quad [K_\mu, P_\nu] = 2(\eta_{\mu\nu}D - M_{\mu\nu}). \tag{3.3}$$

We look for representations in which P_μ are the creation operators and K_μ are the annihilation operators, and $P_\mu^\dagger = K^\mu$. There is a vacuum or highest weight vector $|\mathcal{O}\rangle$ such that it is an eigenvector of D (whose eigenvalue is also denoted by D) and $K_\mu|\mathcal{O}\rangle = 0$. Assume that $\mathcal{O}$ is a scalar operator, i.e., $M_{\mu\nu}|\mathcal{O}\rangle = 0$. Using (3.3), the square of the norm of $P_\mu|\mathcal{O}\rangle$ is

$$\langle\mathcal{O}|K^\mu P_\mu|\mathcal{O}\rangle = 2dD. \tag{3.4}$$

Therefore unitarity implies that $D \geq 0$ and $D = 0$ if and only if $P_\mu|\mathcal{O}\rangle = 0$, i.e., $\mathcal{O}$ is a constant operator. Furthermore, using (3.3), the square of the norm of $P^\mu P_\mu|\mathcal{O}\rangle$ is

$$\langle\mathcal{O}|K^\nu K_\nu P^\mu P_\mu|\mathcal{O}\rangle = 8dD(D - \tfrac{d-2}{2}). \tag{3.5}$$

For $d = 4$, unitarity implies that $D \geq 1$ and $D = 1$ if and only if $P^\mu P_\mu|\mathcal{O}\rangle = 0$, i.e., $\mathcal{O}$ corresponds to a free field.

When $d = 4$, the superconformal algebra has extra generators Q_α, $\bar{Q}_{\dot\alpha}$ of conformal dimension $\frac{1}{2}$ and S^α, $\bar{S}^{\dot\alpha}$ of conformal dimension $-\frac{1}{2}$. The latter are the supersymmetric partners of $K_m u$, i.e., $[Q_\alpha, K^\mu] = \sigma^\mu_{\alpha\dot\alpha}\bar{S}^{\dot\alpha}$ and $[\bar{Q}_{\dot\alpha}, K^\mu] = \bar\sigma^\mu_{\dot\alpha\alpha}S^\alpha$. Moreover, we have[1]

$$\{Q_\alpha, S^\beta\} = (\sigma^{\mu\nu})_\alpha{}^\beta M_{\mu\nu} + 2\delta_\alpha{}^\beta(D - \tfrac{3}{2}R), \tag{3.6}$$

where R is the generator of the $U(1)_R$ transformation (whose eigenvalue is also denoted by R). Again, we look for highest weight representations such that $P_\mu^\dagger = K^\mu$ and $Q_\alpha^\dagger = S^\alpha$. An operator $\mathcal{O}$ is called chiral if $\bar{Q}_{\dot\alpha}|\mathcal{O}\rangle = 0$. Again we assume that $\mathcal{O}$ is a scalar operator. If $|\mathcal{O}\rangle$ is also a highest weight vector, i.e., $K_\mu|\mathcal{O}\rangle = 0$ and $S^\alpha|\mathcal{O}\rangle = 0$, then the square of the norm of $Q_\alpha|\mathcal{O}\rangle$ is

$$\langle\mathcal{O}|S^\alpha Q_\alpha|\mathcal{O}\rangle = 4(D - \tfrac{3}{2}R). \tag{3.7}$$

So $D \geq \frac{3}{2}R$ and the equality holds for chiral operators only.

A consequence of the above study is that the chiral operators form a ring. Consider the product of two chiral operators $\Phi_1(x_1)$ and $\Phi_2(x_2)$. When x_1 is close to x_2, we have an expansion

$$\Phi_1(x_1)\Phi_2(x_2) \sim \sum_i (x_1 - x_2)^{\alpha_i}\Phi_i(x_2). \tag{3.8}$$

All the operators Φ_i in the expansion have $R(\Phi_i) = R(\Phi_1) + R(\Phi_2)$. So

$$D(\Phi_i) \geq \tfrac{3}{2}(R(\Phi_1) + R(\Phi_2)) = D(\Phi_1) + D(\Phi_2). \tag{3.9}$$

Therefore all the exponents $\alpha_i \geq 0$. In the limit $x_1 \to x_2$, the surviving terms terms have $\alpha_i = 0$ and saturate the inequality (3.9). These are the chiral operators. So chiral operators form a ring.

We return to the study of the non-trivial fixed point g_* with $\beta(g_*) = 0$. The dimension of the chiral fields can be computed by their R-charges. For example, using R-charges in (1.27), we get, in the infrared limit,

$$D(M) = \frac{3}{2}R(M) = 3\frac{N_f - N_c}{N_c} \tag{3.10}$$

and

$$D(B) = D(B') = \frac{3}{2}\left(N_c - \frac{N_c^2}{N_f}\right). \tag{3.11}$$

[1] See for example Phys. Rep. 128 (1985) 39.

The meson M has dimension 2 in the ultraviolet. The difference with (3.10) is called the *anomalous dimension*. Perturbatively, the anomalous dimension of mass is

$$\gamma(g) = -\frac{g^2}{8\pi^2}\frac{N_c^2 - 1}{N_c} + o(g^4). \tag{3.12}$$

At the infrared fixed point g_* given by (3.2), $\gamma(g_*) = -\frac{\epsilon}{3} + o(\epsilon^2)$. In the perturbative expansion, (3.10) is $D(M) = 2 - \frac{\epsilon}{3} + o(\epsilon^2)$. This matches $D(M) = 2 + \gamma(g_*)$.

The requirement $D(M) \geq 1$ implies that $N_f \geq \frac{3}{2}N_c$. For $N_f \geq 3N_c$, the theory is free at the infrared. We claim that for $\frac{3}{2}N_c < N_f < 3N_c$, the theory is an interacting superconformal field theory in the infrared limit. The fixed point is non-trivial because $D(M) > 1$ in this range. If $N_f = \frac{3}{2}N_c$ (for N_c even), then $D(M) = 1$. So M (and perhaps the whole low energy theory) is free. If $N_f < \frac{3}{2}N_c$, (3.10) would imply $D(M) < 1$. Therefore the picture breaks down and new physics should emerge.

§3.2. $N_c + 2 \leq N_f \leq \frac{3N_c}{2}$: infrared free magnetic theory

Associated with the original supersymmetric QCD which we now call the *electric theory*, we introduce a dual theory when $N_f \geq N_c + 2$, called the *magnetic theory*. The dual theory has a gauge group $\tilde{G} = SU(\tilde{C})$, where $\tilde{C}$ is an Hermitian vector space of dimension $\tilde{N}_c = N_f - N_c$ and is equipped with a complex volume form $\tilde{v}_c \in \wedge^{\tilde{N}_c}\tilde{C}^*$ invariant under $\tilde{G}$ (and $\tilde{G}^{\mathbb{C}} = SL(\tilde{C})$). The quarks and anti-quarks are $\tilde{Q} \in \mathrm{Hom}(\tilde{F}, \tilde{C})$ and $\tilde{Q}' \in \mathrm{Hom}(\tilde{C}, \tilde{F}')$, respectively, where

$$\tilde{F} = F^* \otimes (\wedge^{N_f} F)^{1/\tilde{N}_c}, \quad \tilde{F}' = F'^* \otimes (\wedge^{N_f} F')^{1/\tilde{N}_c} \tag{3.13}$$

are the spaces of dual flavors (defined up to a root of unity). In addition, the dual theory has a meson $M \in \mathrm{Hom}(F, F')$, which comes from the electric theory but is regarded as an elementary field here. In the electric theory, M has mass dimension 2 in the ultraviolet and acquires an anomalous dimension in the infrared. Therefore in the magnetic theory, it is more natural to regard the field $\mu^{-1}M \in \mathrm{Hom}(\tilde{F}', \tilde{F})$ of mass dimension 1 as the elementary meson, where

$$\mu \in (\wedge^{N_f} F^* \otimes \wedge^{N_f} F')^{1/\tilde{N}_c} \tag{3.14}$$

is a parameter of mass dimension 1. M is uncoupled to the $SU(\tilde{C})$ gauge fields but interact with the dual quarks through a superpotential

$$W = \frac{1}{\mu}\mathrm{tr}_F\, \tilde{Q}'\tilde{Q}M. \tag{3.15}$$

Here $\tilde{Q}'\tilde{Q} \in \mathrm{Hom}(\tilde{F}, \tilde{F}') \cong \mathrm{Hom}(F', F) \otimes (\wedge^{N_f} F^* \otimes \wedge^{N_f} F')^{1/\tilde{N}_c}$ is the meson of the magnetic theory. W has mass dimension 3.

Under the global symmetry $SU(F) \times SU(F') \times U(1)_B \times U(1)_R$ of the electric theory, the fields $\tilde{Q}$, $\tilde{Q}'$, M transform as, according to (3.13) and (1.26),

(3.16)

	$SU(F)$	$\times$	$SU(F')$	$\times$	$U(1)_B$	$\times$	$U(1)_R$
$\tilde{Q}$	F		$\mathbf{1}$		$-\frac{N_c}{\tilde{N}_c}$		$\frac{N_c}{N_f}$
$\tilde{Q}'$	$\mathbf{1}$		F'^*		$\frac{N_c}{\tilde{N}_c}$		$\frac{N_c}{N_f}$
M	F^*		F'		0		$2\frac{N_f-N_c}{N_f}$

So the superpotential (3.15) has R-charge $2\frac{N_f-N_c}{N_f}+2\frac{N_f-\tilde{N}_c}{N_f}=2$, as required. Their fermionic components $\psi_{\tilde{Q}}$, $\psi_{\tilde{Q}'}$, ψ_M and the gluino field $\tilde{\lambda}$ (of the dual theory) transform as

(3.17)

	$SU(F)$	$\times$	$SU(F')$	$\times$	$U(1)_B$	$\times$	$U(1)_R$
$\psi_{\tilde{Q}}$	F		$\mathbf{1}$		$-\frac{N_c}{\tilde{N}_c}$		$\frac{N_c}{N_f}-1$
$\psi_{\tilde{Q}'}$	$\mathbf{1}$		F'		$\frac{N_c}{\tilde{N}_c}$		$\frac{N_c}{N_f}-1$
ψ_M	F^*		$\tilde{F}$		0		$1-2\frac{N_c}{N_f}$
$\tilde{\lambda}$	$\mathbf{1}$		$\mathbf{1}$		0		1

It is easy to check that the measure $D\psi_{\tilde{Q}}\, D\psi_{\tilde{Q}'}\, D\psi_{\tilde{\lambda}}$ of the fermions that are coupled to the $SU(\tilde{C})$ gauge fields is invariant under the global symmetry. Therefore the anomaly-free global symmetry of the dual theory is $SU(F)\times SU(F')\times U(1)_B\times U(1)_R$ as well.

We argue that the electric and magnetic theories describe the same physics at least at long distances. First, we already see that both have the same anomaly-free global symmetry $SU(F)\times SU(F')\times U(1)_B\times U(1)_R$. That they do not have the same gauge group is not a problem. Strictly speaking, gauge symmetry is not a symmetry, but a redundancy in the description of the theory. So gauge symmetry is not fundamental. In fact, the gauge fields in the magnetic theory are composites of the electric degrees of freedom, and vice versa.

Secondly, the electric and magnetic theories have the same gauge invariant operators, though their field contents seem rather different. We construct the mapping of gauge invariant chiral operators. The composite meson $Q'Q$ of the electric theory corresponds to the elementary meson M of the magnetic theory, whereas the composite meson $\tilde{Q}'\tilde{Q}$ of the magnetic theory vanishes by the equation of motion of M. The baryons of the magnetic theory are $\tilde{B}=\tilde{Q}^*\tilde{v}_c\in\wedge^{\tilde{N}_c}\tilde{F}^*\cong\wedge^{N_c}F^*$ and $\tilde{B}'=\tilde{Q}'_*\tilde{v}_c^*\in\wedge^{\tilde{N}_c}\tilde{F}'\cong\wedge^{N_c}F'$. Thus they have the same quantum numbers as B, B' of the electric theory, which we shall identify. In fact, taking into account (1.29), (3.14) as well as their mass dimensions, the only possible relations are

$$B=i^{\tilde{N}_c+1}\Lambda^{b_0/2}\mu^{-\tilde{N}_c/2}\tilde{B},\quad B'=i^{\tilde{N}_c+1}\Lambda^{b_0/2}\mu^{-\tilde{N}_c/2}\tilde{B}',\tag{3.18}$$

where the numerical factors are chosen to fit future calculations. However, $\tilde{Q}$, $\tilde{Q}'$ are not polynomials of Q, Q' and vice versa because the $U(1)_B$ charges of $\tilde{Q}$, $\tilde{Q}'$ are fractional. Rather, the dual quarks $\tilde{Q}$, $\tilde{Q}'$ should be interpreted as solitons of Q, Q' in the electric theory, and vice versa.

Thirdly, we check the 't Hooft anomaly matching condition. Under the global symmetry, the fermions in the electric theory transform as (2.31). Their contribution to anomaly is given by the left column of (2.33). In the magnetic theory, the fermions transform as (3.17). Their contribution to anomaly is given by

(3.19)

	magnetic
$\mathfrak{su}(F)^3$	$\tilde{N}_c d_F + N_f(-d_F)$
$\mathfrak{su}(F')^3$	$\tilde{N}_c(-d_{F'}) + N_f d_{F'}$
$\mathfrak{su}(F)^2\mathfrak{u}(1)_B$	$\tilde{N}_c\, c_2(F)\cdot(-\frac{N_c}{\tilde{N}_c})$
$\mathfrak{su}(F')^2\mathfrak{u}(1)_B$	$\tilde{N}_c\, c_2(F')\cdot\frac{N_c}{\tilde{N}_c}$
$\mathfrak{su}(F)^2\mathfrak{u}(1)_R$	$\tilde{N}_c\, c_2(F)\cdot(\frac{N_c}{N_f}-1) + N_f\, c_2(F)\cdot(1-2\frac{N_c}{N_f})$
$\mathfrak{su}(F')^2\mathfrak{u}(1)_R$	$\tilde{N}_c\, c_2(F')\cdot(\frac{N_c}{N_f}-1) + N_f\, c_2(F')\cdot(1-2\frac{N_c}{N_f})$
$\mathfrak{u}(1)_B^2\mathfrak{u}(1)_R$	$\tilde{N}_c N_f[(-\frac{N_c}{\tilde{N}_c})^2 + (\frac{N_c}{\tilde{N}_c})^2]$
$\mathfrak{u}(1)_R^3$	$2\tilde{N}_c N_f(\frac{N_c}{N_f}-1)^3 + N_f^2(1-2\frac{N_c}{N_f})^3 + (\tilde{N}_c^2-1)1^3$
$\mathfrak{u}(1)_R$	$2\tilde{N}_c N_f(\frac{N_c}{N_f}-1) + N_f^2(1-2\frac{N_c}{N_f}) + (\tilde{N}_c^2-1)1$

It is straightforward to check that these two columns match.

In the range $N_c+2 \le N_f \le \frac{3N_c}{2}$ (not yet studied except for $N_f = \frac{3}{2}N_c$), we have $\tilde{N}_c \ge 2$ and $N_f \ge 3\tilde{N}_c$, therefore the magnetic theory is infrared free. So the magnetic theory is a better description in this range, where the electric theory is hard to study because of the strong coupling in the infrared. In particular, at $N_f = \frac{3N_c}{2}$ (provided that N_c is even), we calculated $D(M)=1$ in the electric theory. This confirms that M is a free field in the infrared, a consequence of of the magnetic description. Generally, as N_f decreases, the coupling in the electric theory gets stronger until the theory confines when $N_f \le N_c+1$, whereas the coupling of the magnetic theory gets weaker and the theory becomes infrared free when $N_f \le \frac{3N_c}{2}$. The infrared behavior of the theory in various ranges is summarized in the following table.

	electric theory	magnetic theory
$N_c \ge 2,\ N_f \ge 3N_c$ $(\tilde{N}_c+2 \le N_f \le \frac{3\tilde{N}_c}{2})$	infrared free	strongly coupled
$\frac{3N_c}{2} < N_f < 3N_c$ $(\frac{3\tilde{N}_c}{2} < N_f < 3\tilde{N}_c)$	non-trivial infrared fixed point	non-trivial infrared fixed point
$N_c+2 \le N_f \le \frac{3N_c}{2}$ $(\tilde{N}_c \ge 2,\ N_f \ge 3\tilde{N}_c)$	strongly coupled	infrared free
$N_f \le N_c+1$	analysed in lectures 1, 2	——

§3.3. Further tests of duality

Dynamically, the electric theory has a scale Λ as in (1.29), where $b_0 = 3N_c - N_f$ determines the β-function. The scale $\tilde{\Lambda}$ of the magnetic theory must be related to that of the electric theory. Because Λ, $\tilde{\Lambda}$, μ all have mass dimension 1 and because of (1.29) (and its magnetic counterpart) and (3.14), the only possible relation among them is, up to a numerical factor,

$$\Lambda^{3N_c - N_f}\tilde{\Lambda}^{3\tilde{N}_c - N_f} = (-1)^{\tilde{N}_c}\mu^{N_f}. \tag{3.20}$$

The phase $(-1)^{\tilde{N}_c}$ will be needed in a number of places below. This, together with (1.15), shows that if the electric theory is strongly coupled, then the magnetic theory is weakly coupled, and vice versa.

§3.3.1. Duality is an involution

We claim that the dual of the magnetic theory is the electric theory itself. The number of colors in the double dual theory is $N_f - \tilde{N}_c = N_c$. Therefore, its color space can be identified with C. Next, the flavor spaces are $\tilde{F}^* \otimes (\wedge^{N_f}\tilde{F})^{1/\tilde{N}_c} \cong F$ and $\tilde{F}'^* \otimes (\wedge^{N_f}\tilde{F}')^{1/\tilde{N}_c} \cong F'$, the same as the electric theory. The mass scale of the double dual theory is again Λ; this follows from (3.20), provided that

$$\tilde{\mu} = -\mu \in (\wedge^{N_F}\tilde{F}^* \otimes \wedge^{N_f}\tilde{F}')^{1/N_c} \cong (\wedge^{N_f}F^* \otimes \wedge^{N_f}F')^{1/\tilde{N}_c}. \tag{3.21}$$

The double dual theory has a new term $\frac{1}{\tilde{\mu}}\mathrm{tr}_{\tilde{F}}Q'Q\tilde{M}$ in the superpotential where $\tilde{M} \in \mathrm{Hom}(\tilde{F}, \tilde{F}')$ is a new meson field. This, combined with (3.15), gives the total superpotential

$$W = \frac{1}{\mu}\mathrm{tr}_F\, \tilde{M}(M - Q'Q). \tag{3.22}$$

The equation of motion of $\tilde{M}$ implies that $M = Q'Q$. Therefore we recover the electric theory.

§3.3.2. Deformation along the flat directions

So far, we have only studied the theory with $N_f \geq N_c + 2$ at the origin of the moduli space $\mathcal{M}$. Recall that in the electric theory, $\mathcal{M}$ is the space of M, B, B' subject to the classical constraints (1.20), (1.21), (1.22).

We show that the procedure of deforming M in $\mathcal{M}$ commutes with taking the dual. Suppose $\operatorname{rank} M = r$. We write

$$F = \hat{F} \oplus \hat{F}^\perp, \quad F' = \hat{F}' \oplus \hat{F}'^\perp \quad (\dim \hat{F}^\perp = \dim \hat{F}'^\perp = r) \tag{3.23}$$

such that M has the form

$$M = \begin{pmatrix} 0 & 0 \\ 0 & M^\perp \end{pmatrix}. \tag{3.24}$$

At low energies, the gauge group is broken to $SU(N_c - r)$, and there are $N_f - r$ light flavors. The mass scale $\hat{\Lambda}$ of the low energy effective theory is given by

$$\hat{\Lambda}^{3(N_c - r) - (N_f - r)} = \frac{\Lambda^{3N_c - N_f}}{\det M^\perp}. \tag{3.25}$$

In the dual description, the fundamental theory has $\tilde{N}_c$ colors and N_f flavors, and has a scale $\tilde{\Lambda}$ related to Λ by (3.20). Now $\mu^{-1}M \in \mathrm{Hom}(\tilde{F}', \tilde{F})$ gives mass to some components of the dual quarks $\tilde{Q}$, $\tilde{Q}'$. The low energy magnetic theory has $\tilde{N}_c$ colors and $N_f - r$ flavors. Its scale $\hat{\tilde{\Lambda}}$ is given by

$$\hat{\tilde{\Lambda}}^{3\tilde{N}_c-(N_f-r)} = \frac{\tilde{\Lambda}^{3\tilde{N}_c-N_f} \det M^{\perp}}{\mu^r}. \tag{3.26}$$

It follows from (3.20), (3.25), (3.26) that

$$\hat{\Lambda}^{3(N_c-r)-(N_f-r)}\hat{\tilde{\Lambda}}^{3\tilde{N}_c-(N_f-r)} = (-1)^{\tilde{N}_c}\mu^{N_f-r}. \tag{3.27}$$

Thus the low energy magnetic theory is dual to the low energy electric theory.

We further check that the magnetic theory has the same space of ground states $\mathcal{M}$. In the electric theory, the restriction $r = \operatorname{rank} M \leq N_c$ is a classical constraint. In the magnetic theory, the same constraints are enforced by non-perturbative quantum effects. In fact, if $r > N_c$, then the number of flavors in the low energy magnetic theory is $N_f - r < \tilde{N}_c$. In this case, a superpotential of the form (1.30) is generated, and consequently there is no vacuum. So generically, $r = N_c$ from both the electric and the magnetic points of view. In this case, the number of flavors of the low energy magnetic theory is $N_f - r = \tilde{N}_c$, equal to the number of colors. According to (2.20), the light components $\hat{\tilde{M}}$, $\hat{\tilde{B}}$, $\hat{\tilde{B}}'$ of the dual meson and baryons satisfy

$$\det \hat{\tilde{M}} - \hat{\tilde{B}} \otimes \hat{\tilde{B}}' = \hat{\tilde{\Lambda}}^{2\tilde{N}_c}. \tag{3.28}$$

Moreover, $\hat{\tilde{M}} = 0$ here because of the equation of motion of M. Using (3.26) and (3.20) when $r = N_c$, we obtain

$$\hat{\tilde{\Lambda}}^{2\tilde{N}_c} = \frac{\tilde{\Lambda}^{3N_c-N_f} \det M^{\perp}}{\mu^{N_c}} = \frac{(-1)^{\tilde{N}_c} \det M^{\perp}}{\Lambda^{3N_c-N_f}\mu^{-\tilde{N}_c}}. \tag{3.29}$$

By the mapping relation (3.18), we get

$$\hat{B} \otimes \hat{B}' = \det M^{\perp}. \tag{3.30}$$

This is precisely the constraint (1.23) in the electric theory. Other constraints are trivially satisfied.

§3.3.3. Mass deformation

We now check that mass deformation commutes with taking dual. Suppose a mass term $\operatorname{tr} mM$ with $\operatorname{rank} m = r$ is added to the electric theory. Again, we write (3.23) such that m has the form (2.6). The low energy theory has light quarks $\hat{Q} \in \operatorname{Hom}(\hat{F}, C)$, $\hat{Q}' \in \operatorname{Hom}(C, \hat{F}')$ and has a scale $\hat{\Lambda}$ given by

$$\hat{\Lambda}^{3N_c-N_f+r} = \Lambda^{3N_c-N_f} \det m^{\perp}. \tag{3.31}$$

M has the form (2.10), where $\hat{M}$ contains the light degrees of freedom. Let $\hat{\tilde{F}} = \hat{F}^* \otimes (\wedge^{N_f} F)^{1/\tilde{N}_c}$ and $\hat{\tilde{F}}' = \hat{F}'^* \otimes (\wedge^{N_f} F')^{1/\tilde{N}_c}$. Then the equations of motion of the heavy components of M imply that

$$\tilde{Q}'\tilde{Q} = \begin{pmatrix} \hat{\tilde{M}} & 0 \\ 0 & -\mu m^{\perp} \end{pmatrix} \tag{3.32}$$

for some light dual meson $\hat{\hat{M}} \in \mathrm{Hom}(\hat{\hat{F}}, \hat{\hat{F}}')$. This breaks the $SU(\tilde{N}_c)$ gauge symmetry to $SU(\tilde{N}_c - r)$. We assume that $\tilde{N}_c - r \geq 2$, i.e., $r \leq N_f - (N_c + 2)$. Integrating out the heavy components in M, $\tilde{Q}$, $\tilde{Q}'$, we obtain a low energy effective superpotential

$$\hat{W} = \frac{1}{\mu} \mathrm{tr}\, \hat{\hat{M}} \hat{M}. \tag{3.33}$$

This is exactly a magnetic theory of $\tilde{N}_c - r$ flavors and $N_f - r$ quarks. Its mass scale $\hat{\hat{\Lambda}}$ is given by

$$\hat{\hat{\Lambda}}^{3(\tilde{N}_c - r) - (N_f - r)} = \frac{\tilde{\Lambda}^{3\tilde{N}_c - N_f}}{(-\mu)^r \det m^{\perp}}. \tag{3.34}$$

From (3.31), (3.34) and (3.20), we obtain

$$\hat{\Lambda}^{3N_c - (N_f - r)} \hat{\hat{\Lambda}}^{3(\tilde{N}_c - r) - (N_f - r)} = (-1)^{\tilde{N}_c - r} \mu^{N_f - r}. \tag{3.35}$$

So the low energy theory of the magnetic theory after mass deformation is dual to that of the electric theory.

If $r = N_f - N_c - 1$, then in the low energy electric theory, the number of flavors is $N_f - r = N_c + 1$ and in the low energy magnetic theory, the gauge symmetry is completely broken since $\tilde{N}_c - r = 1$. The light baryons in the magnetic theory $\hat{\hat{B}} = \hat{\hat{Q}}$, $\hat{\hat{B}}' = \hat{\hat{Q}}'$ are related to $\hat{B}$ and $\hat{B}'$ by (3.18). The superpotential (3.15) can be written as

$$\frac{1}{\mu} \hat{\hat{Q}}' \hat{M} \hat{\hat{Q}} = \frac{1}{\hat{\Lambda}^{2N_c - 1}} \hat{B}' \hat{M} \hat{B}. \tag{3.36}$$

In addition, because the gauge symmetry is completely broken, instanton contribution should be included just as in the case $N_f = N_c - 1$. Now there are more than $\tilde{N}_c - 1$ zero modes of $\psi_{\tilde{Q}}$, $\psi_{\tilde{Q}'}$; these $N_c + 1$ extra zero modes are absorbed by the interaction (3.15), generating a term proportional to the $(N_c + 1)$-th power of $(\mu^{-1} M)_{\mathrm{light}}$ in the effective superpotential. So the instanton contribution is

$$\begin{aligned} \frac{\tilde{\Lambda}^{3\tilde{N}_c - N_f} \det(\mu^{-1} M)_{\mathrm{light}}}{\det(\tilde{Q}' \tilde{Q})_{\mathrm{heavy}}} &= \frac{\tilde{\Lambda}^{3\tilde{N}_c - N_f} \det(\mu^{-1} \hat{M})}{\det(-\mu m^{\perp})} = \frac{\tilde{\Lambda}^{3\tilde{N}_c - N_f} \det \hat{M}}{(-1)^{\tilde{N}_c + 1} \mu^{N_f} \det m^{\perp}} \\ &= -\frac{\det \hat{M}}{\Lambda^{3N_c - N_f} \det m^{\perp}} = -\frac{\det \hat{M}}{\hat{\Lambda}^{2N_c - 1}}, \end{aligned} \tag{3.37}$$

where (3.32), (3.20), (3.31) have been used. Combining (3.36) and (3.37), we get the total effective superpotential

$$W_{\mathrm{eff}} = \frac{1}{\hat{\Lambda}^{2N_c - 1}} (\hat{B}' \hat{M} \hat{B} - \det \hat{M}), \tag{3.38}$$

which is exactly (2.24).

There are many other checks of duality. We remark here that a theory with less flavors can be regarded as the deformation of a larger theory by a mass term. This duality is related to various phenomena in $N = 2$ theories and to Langlands duality in $N = 4$ theories.