An Introduction to Global Supersymmetry

Philip C. Argyres
Cornell University

Contents

Preface

These lecture notes provide an introduction to supersymmetry with a focus on the non-perturbative dynamics of supersymmetric field theories. It is meant for students who have had a one-year introductory course in quantum field theory, and assumes a basic knowledge of gauge theories, Feynman diagrams and renormalization on the physics side, and an aquaintance with analysis on the complex plane (holomorphy, analytic continuation) as well as rudimentary group theory ($SU(2)$, Lorentz group) on the math side. More adanced topics—Wilsonian effective actions, Lie groups and algebras, anomalies, instantons, conformal invariance, monopoles—are introduced as part of the course when needed. The emphasis will not be on comrehensive discussions of these techniques, but on their "practical" application.

The aims of this course are two-fold. The first is to introduce the technology of global supersymmetry in quantum field theory. The first third of the course introduces the $N = 1$ $d = 4$ superfields describing classical chiral and vector multiplets, the geometry of their spaces of vacua and the nonrenormalization rules they obey. A excellent text which covers these topics in much greater detail than this course does is S. Weinberg's *The Quantum Theory of Fields III: Supersymmetry* [1]. These notes try to follow the notation and conventions of Weinberg's book; they also try to provide alternative (often more qualitative) explanations for overlapping topics, instead of reproducing the exposition in Weinberg's book. Also, the student is directed to Weinberg's book for important topics not covered in these lectures (supersymmetric models of physics beyond the standard model and supersymmetry breaking), as well as to the original references.

The second second aim of these notes is to use our understanding of non-perturbative aspects of particular supersymmetric models as a window on strongly coupled quantum field theory. From this point of view, supersymmetric field theories are just especially symmetric versions of ordinary field theories, and in many cases this extra symmetry allows the exact determination of some non-perturbative properties of these theories. This gives us another context (besides lattice gauge theory and semi-classical expansions) in which to think concretely about non-perturbative quantum field theory in more than two dimensions. To this end, the last two-thirds of the course analyzes examples of strongly coupled supersymmetric gauge theories. We first discuss non-perturbative $SU(n)$ N=1 supersymmetric versions of QCD, covering cases with completely Higgsed, Coulomb, confining, and interacting conformal vacua. Next we describe d=4 theories with N=2 and 4 extended supersymmetry, central charges and Seiberg-Witten theory. Finally we end with a brief look at supersymmetry in other dimensions, describing spinors and supersymmetry algebras in various dimensions, 5-dimensional N=1 and 2 theories, and 6-dimensional N=$(2,0)$ and $(1,1)$ theories.

These notes owe a large intellectual debt to Nathan Seiberg: not only is his work the main focus of much of the course, but also parts of this course are modelled on two series of lectures he gave at the Institute for Advanced Study in Princeton in the fall of 1994 and at Rutgers University in the fall of 1995. These notes grew, more immediately, out of a graduate course on supersymmetry I taught at Cornell University in the fall semesters of 1996 and 2000. It is a pleasure to thank the students in these courses, and especially Zorawar Bassi, Alex Buchel, Ron Maimon, K. Narayan, Sophie Pelland, and Gary Shiu, for their many comments and questions. It is also a pleasure to thank my colleagues at Cornell—Eanna Flanagan, Kurt Gottfried, Tom Kinoshita, André Leclair, Peter Lepage, Henry Tye, Tung-Mow Yan, and Piljin Yi—for many helpful discussions. I'd also like to thank Mark Alford, Daniel Freedman, Chris Kolda, John March-Russell, Ronen Plesser, Al Shapere, Peter West, and especially Keith Dienes for comments on an earlier version of these notes. Thanks also to the physics department at the University of Cincinnati for their kind hospitality. Finally, this work was supported in part by NSF grant XXXX.

Philip Argyres

Ithaca, New York
January, 2001

Chapter 1

N=1 d=4 Supersymmetry

1.1 Why Supersymmetry?

Though originally introduced in early 1970's we still don't know how or if supersymmetry plays a role in nature. Why, then, have a considerable number of people been working on this theory for the last 25 years? The answer lies in the Coleman-Mandula theorem [2], which singles-out supersymmetry as the "unique" extension of Poincaré invariance in quantum field theory in more than two space-time dimensions (under some important but reasonable assumptions). Below I will give a qualitative description of the Coleman-Mandula theorem following a discussion in [3].

The Coleman-Mandula theorem states that *in a theory with non-trivial scattering in more than 1+1 dimensions, the only possible conserved quantities that transform as tensors under the Lorentz group (*i.e. *without spinor indices) are the usual energy-momentum vector* P_μ*, the generators of Lorentz transformations* $J_{\mu\nu}$*, as well as possible scalar "internal" symmetry charges* Z_i *which commute with* P_μ *and* $J_{\mu\nu}$*.* (There is an extension of this result for massless particles which allows the generators of conformal transformations.)

The basic idea behind this result is that conservation of P_μ and $J_{\mu\nu}$ leaves only the scattering angle unknown in (say) a 2-body collision. Additional "exotic" conservation laws would determine the scattering angle, leaving only a discrete set of possible angles. Since the scattering amplitude is an analytic function of angle (*assumption # 1*) it then vanishes for all angles.

We illustrate this with a simple example. Consider a theory of 2 free real bose fields ϕ_1 and ϕ_2:

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2. \tag{1.1}$$

Such a free field theory has infinitely many conserved currents. For example, it follows immediately from the equations of motion $\partial_\mu \partial^\mu \phi_1 = \partial_\mu \partial^\mu \phi_2 = 0$ that the series of currents

$$\begin{aligned} J_\mu &= (\partial_\mu \phi_2)\phi_1 - \phi_2 \partial_\mu \phi_1, \\ J_{\mu\rho} &= (\partial_\mu \phi_2)\partial_\rho \phi_1 - \phi_2 \partial_\mu \partial_\rho \phi_1, \\ J_{\mu\rho\sigma} &= (\partial_\mu \phi_2)\partial_\rho \partial_\sigma \phi_1 - \phi_2 \partial_\mu \partial_\rho \partial_\sigma \phi_1, \qquad etc. \end{aligned} \tag{1.2}$$

are conserved,

$$\partial^\mu J_\mu = \partial^\mu J_{\mu\rho} = \partial^\mu J_{\mu\rho\sigma} = 0, \tag{1.3}$$

leading to the conserved charges

$$Q_{\rho\sigma} = \int d^{d-1}\mathbf{x}\, J_{0\rho\sigma}, \qquad etc. \tag{1.4}$$

This does not violate the Coleman-Mandula theorem since, being a free theory, there is no scattering.

Now it is well known that there are interactions which when added to this theory still keep J_μ conserved; for example, any potential of the form $V = f(\phi_1^2 + \phi_2^2)$ does the job. However, the Coleman-Mandula theorem asserts that there are no Lorentz-invariant interactions which can be added so that the others are conserved (nor can they be redefined by adding extra terms so that they will still be conserved).

For suppose we have, say, a conserved traceless symmetric tensor charge $Q_{\rho\sigma}$. By Lorentz invariance, its matrix element in a 1-particle state of momentum p^μ and spin zero is

$$\langle p|Q_{\rho\sigma}|p\rangle \propto p_\rho p_\sigma - \frac{1}{d}\eta_{\rho\sigma} p^2. \tag{1.5}$$

($\eta_{\mu\nu}$ is the Minkowski metric, d is the space-time dimension.) Apply this to an elastic 2-body collision of identical particles with incoming momenta p_1, p_2, and outgoing momenta q_1, q_2, and assume that the matrix element of Q in the 2-particle state $|p_1 p_2\rangle$ is the sum of the matrix elements in the states $|p_1\rangle$ and $|p_2\rangle$:

$$\langle p_1 p_2|Q_{\rho\sigma}|p_1 p_2\rangle = \langle p_1|Q_{\rho\sigma}|p_1\rangle + \langle p_2|Q_{\rho\sigma}|p_2\rangle. \tag{1.6}$$

This should seem entirely reasonable for widely separated particle states, and is true if Q is "not too non-local" (*assumption # 2*)—say, the integral of a local current .

It is then follows that conservation of $Q^{\rho\sigma}$ together with energy momentum conservation implies

$$\begin{aligned} p_1^\rho p_1^\sigma + p_2^\rho p_2^\sigma &= q_1^\rho q_1^\sigma + q_2^\rho q_2^\sigma, \\ p_1^\rho + p_2^\rho &= q_1^\rho + q_2^\rho, \end{aligned} \tag{1.7}$$

and it is easy to check that the only solutions of these two equations have $p_1^\mu = q_1^\mu$ or $p_1^\mu = q_2^\mu$, *i.e.* zero scattering angle. For the extension of this argument to non-identical particles with spin, and a more detailed statement of the assumptions, see [2] or [1, section 25.B].

The Coleman-Mandula theorem does not apply to spinor charges, though. Consider a d=4 free theory of two real massless scalars ϕ_1, ϕ_2, and a massless four component Majorana fermion[1]

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - \frac{1}{2}\overline{\psi}\gamma^\mu\partial_\mu\psi \tag{1.8}$$

Again, an infinite number of currents with spinor indices, *e.g.*

$$\begin{aligned} S_{\mu\alpha} &= \partial_\rho(\phi_1 - i\phi_2)(\gamma^\rho\gamma_\mu\psi)_\alpha, \\ S_{\mu\nu\alpha} &= \partial_\rho(\phi_1 - i\phi_2)(\gamma^\rho\gamma_\mu\partial_\nu\psi)_\alpha, \qquad \textit{etc.}, \end{aligned} \tag{1.9}$$

are conserved. Now, as we will see in more detail in later lectures, after adding the interaction

$$V = g\overline{\psi}(\phi_1 + i\gamma_5\phi_2)\psi + \frac{1}{2}g^2(\phi_1^2 + \phi_2^2)^2 \tag{1.10}$$

to this free theory, $S_{\mu\alpha}$ (with correction proportional to g) remains conserved; however, $S_{\mu\nu\alpha}$ is never conserved in the presence of interactions.

We can see this by applying the Coleman-Mandula theorem to the anticommutators of the fermionic conserved charges

$$Q_\alpha = \int d^3\mathbf{x}\ S_{0\alpha}, \qquad Q_{\nu\alpha} = \int d^3\mathbf{x}\ S_{0\nu\alpha}. \tag{1.11}$$

Indeed, consider the anticommutator $\{Q_{\nu\alpha}, Q_{\nu\alpha}^\dagger\}$, which cannot vanish unless $Q_{\nu\alpha}$ is identically zero, since the anticommutator of any operator with its hermitian adjoint is positive definite. Since $Q_{\nu\alpha}$ has components of spin up to 3/2, the anticommutator has components of spin up to 3 by addition of angular momentum. Since the anticommutator is conserved if $Q_{\nu\alpha}$ is, and since the Coleman-Mandula theorem does not permit conservation of an operator of spin 3 in an interacting theory, $Q_{\nu\alpha}$ cannot be conserved in an interacting theory.

Conservation of Q_α, on the other hand, is permitted. Since it has spin 1/2, its anticommutator has spin 1, and there is a conserved spin-1 charge: P_μ. This gives rise to the N=1 d=4 *supersymmetry algebra*

$$\begin{aligned} \{Q, \overline{Q}\} &= -2i\gamma^\mu P_\mu, \\ [Q, P_\mu] &= 0, \end{aligned} \tag{1.12}$$

[1] Our spinor and Dirac matrix conventions are those of [1] and will be reviewed in section 1.3 below.

extending the usual Poincaré algebra.

The above arguments are too fast. Proper arguments involve the machinery of spinor representations of the Lorentz group and an analysis of the associativity constraints of algebras like (1.12) to determine the precise right hand sides—see [4] and [1, sections 25.2 and 32.1]. The more general result is that a supersymmetry algebra in any dimension has the form

$$\begin{aligned} \{Q_n, Q_m\} &= \Gamma^\mu_{nm} P_\mu + Z_{nm}, \\ [Q_n, P_\mu] &= 0, \end{aligned} \tag{1.13}$$

where the Q_n label all the spinor supercharges and their adjoints, Γ^μ_{nm} are some c-number coefficients, and Z_{nm} are some scalar conserved charges. The Z_{nm} are called *central charges* of the algebra and can be shown to commute with all other charges; this implies in particular that they can only be the conserved charges of $U(1)$ internal symmetry groups (*e.g.* electric charge or baryon number). If there is more than one independent spinor charge (and its conjugate) the algebra is called an *extended supersymmetry algebra*.

To avoid confusion, it is worth noting that algebras involving both commutators and anticommutators are sometimes called *superalgebras* (or *graded Lie algebras*) even if they do not have the form (1.13). We will reserve the term supersymmetry algebra only for those superalgebras some of whose bosonic generators have the interpretation as energy-momentum.

Finally, it is worth pointing out a few situations where the assumptions of the Coleman-Mandula theorem break down, thus allowing more general symmetry algebras than the supersymmetry algebras outlined above. One case is field theories in d=2 space-time dimensions. Here it is the assumption of analyticity in the scattering angle which fails, since with one spatial dimension there can only be forward or backward scattering. This allows a much wider variety of space-time symmetry algebras (*e.g.* those underlying completely integrable systems), though d=2 supersymmetric theories still exist and play an important role.

Another case are theories of objects extended in p spatial directions called p-branes (1-branes are strings, 2-branes are membranes, *etc.*) which can carry p-form conserved charges (that is, charges $Q_{[\mu_1 \cdots \mu_p]}$ antisymmetric on p indices). For these extended objects it is the assumption of locality which is violated in the Coleman-Mandula theorem. It turns out that in this case the most general space-time symmetry algebra with spinors is still of the form (1.13) but with the p-form charges appearing along with the scalar central charges on the right hand side of the supersymmetry algebra.

1.2 Supersymmetric Quantum Mechanics

In this lecture we examine a toy model of supersymmetric quantum field theory—supersymmetric quantum mechanics. Our aim is to present the main qualitative features of supersymmetric theories and techniques without having to deal with the mathematical, notational and conceptual difficulties associated with four-dimensional quantum field theory. Much of this lecture follows [5].

(Supersymmetric quantum mechanics is interesting in its own right, and not just as a toy model. The dynamics of higher dimensional supersymmetric quantum field theories in finite volume reduce to that of supersymmetric quantum mechanics in the infrared (low energy) limit; this has been used to extract non-perturbative data on supersymmetry breaking in superYang-Mills theories [6]. A supersymmetric quantum mechanics with 16 supercharges appears in the matrix theory description of M theory [7]. And, in mathematics supersymmetric quantum mechanics has proved to be an effective and intuitive tool in proving index theorems about diferential operators and related subjects [8, 9, 10].)

Quantum mechanics can be thought of as quantum field theory in $0+1$ dimensions (*i.e.* no space and one time dimension). The Poincaré algebra reduces simply to time translations, generated by the energy operator H. Field operators $\phi(t)$ in $0+1$ dimensions are just the time-dependent Heisenberg picture operators of quantum mechanics. Thus the quantum mechanics of a spinless particle on the x-axis, described by a position operator x and its conjugate momentum p, can be interpreted as a $0+1$-dimensional scalar quantum theory with the real field ϕ canonically conjugate field π playing the roles of x and p. The ground state of quantum mechanics is the vacuum state of the field theory.

1.2.1 Supersymmetry algebra in 0+1 dimensions

By analogy with the supersymmetry algebra in 3+1 dimensions, we define the supersymmetry algebra in quantum mechanics to be

$$\{Q^\dagger, Q\} = 2H, \qquad \{Q, Q\} = 0, \qquad [Q, H] = 0, \tag{1.14}$$

where Q is the supercharge and H the Hamiltonian.

The superalgebra implies that the energy spectrum is positive. One way to see this is by taking expectation values in any state $|\Omega\rangle$:

$$\langle\Omega|\{Q^\dagger, Q\}|\Omega\rangle = \langle\Omega|Q^\dagger Q|\Omega\rangle + \langle\Omega|QQ^\dagger|\Omega\rangle = |Q|\Omega\rangle|^2 + \left|Q^\dagger|\Omega\rangle\right|^2 \geq 0. \tag{1.15}$$

It follows from (1.14) that $\langle\Omega|H|\Omega\rangle \geq 0$ for all $|\Omega\rangle$ and therefore that $H \geq 0$. This lower bound on the energy spectrum does not mean, of course, that the minimum energy state saturates it, or even (in field theory) that there is *any* minimum energy state. For example, the potential $V(\phi)$ may slope off to infinity as in figure 1.1, so that the energy never attains its minimum and there is no vacuum.

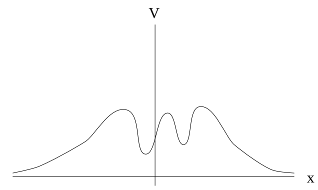

Figure 1.1: A bounded potential may never attain its minimum.

If we diagonalize H by $H|n\rangle = E_n|n\rangle$, then on a given eigenspace $\{Q^\dagger, Q\} = 2E_n$. If $E_n > 0$ we can define

$$a \equiv \frac{1}{\sqrt{2E_n}} Q, \qquad a^\dagger \equiv \frac{1}{\sqrt{2E_n}} Q^\dagger, \tag{1.16}$$

and the supersymmetry algebra becomes

$$\{a^\dagger, a\} = 1, \qquad \{a, a\} = 0, \tag{1.17}$$

which is the algebra of a fermionic creation and annihilation operator (a 2-dimensional Clifford algebra). Its representation theory is very simple: its single irreducible representation is 2-dimensional, and can be represented on a basis of states $|\pm\rangle$ as

$$\begin{aligned} a|-\rangle &= 0 \qquad & a|+\rangle &= |-\rangle \\ a^\dagger|+\rangle &= 0 \qquad & a^\dagger|-\rangle &= |+\rangle. \end{aligned} \tag{1.18}$$

Thinking of $a^\dagger$ as a fermion creation operator, we can assign fermion number $F = 1$ to $|+\rangle$ and fermion number $F = 0$ to $|-\rangle$. Of course, calling $|+\rangle$ and $|-\rangle$ fermionic and bosonic states, respectively, makes no sense in quantum mechanics, but is the reduction of what happens in higher dimensional theories where fermion number is well-defined. (In $2+1$ or more dimensions there is an independent definition of $(-)^F$ as the operator implementing a 2π rotation: $(-)^F = e^{2\pi i J_z}$.)

When there is a state (the vacuum) with $E = 0$, however, the supersymmetry algebra in this energy sector becomes simply

$$\{Q^\dagger, Q\} = 0. \tag{1.19}$$

There is only the trivial (one-dimensional) irreducible representation $Q|0\rangle = Q^\dagger|0\rangle = 0$. These states can be assigned either fermion number, $(-)^F|0\rangle = \pm|0\rangle$, and is a matter of convention.

These properties are also true (qualitatively) of the supersymmetry algebra in other space-time dimensions. Thus the spectrum of a supersymmetric theory will have an equal in number boson and fermion states degenerate in energy (mass) at all positive energies. But, there need not be such a degeneracy among the zero energy states (vacua).

When there exists an $E = 0$ state, we will say that it is a "supersymmetric vacuum". This is because such a state is annihilated by the supersymmetry generators. If there is no $E = 0$ state, then the vacuum is not annihilated by the supersymmetry generators (*i.e.* it transforms under supersymmetry) and we say that supersymmetry is (spontaneously) broken.

1.2.2 Quantum mechanics of a particle with spin

The supersymmetry algebra can be realized in quantum mechanics by a one-dimensional particle with two states (spin). Normally, we would describe such a system by two-component wave functions of x, the particle position. Viewed as a $0+1$-dimensional field theory, we replace x by the field value ϕ, and write the wave function as

$$\Omega = \begin{pmatrix} \omega_+(\phi) \\ \omega_-(\phi) \end{pmatrix}. \tag{1.20}$$

The conjugate momentum operator to ϕ is then

$$\pi = -i\hbar\frac{\partial}{\partial\phi}. \tag{1.21}$$

Define the operators

$$\begin{aligned} Q &\equiv \sigma^-\left(f'(\phi) + i\pi\right), & \sigma^- &\equiv \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ Q^\dagger &\equiv \sigma^+\left(f'(\phi) - i\pi\right), & \sigma^+ &\equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{1.22}$$

where $f(\phi)$ is a real function and $f' = \partial f/\partial\phi$. It is then easy to compute

$$\{Q^\dagger, Q\} = \pi^2 + (f')^2 - \hbar f''\sigma^3 \equiv 2H, \qquad \sigma^3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.23}$$

where we have simply defined the right hand side to be twice the Hamiltonian. It is easy to check that the rest of the supersymmetry algebra, $Q^2 = [Q, H] = 0$, is satisfied by these operators. Furthermore, going back to the quantum mechanical interpretation of this system, H indeed has the form expected for the Hamiltonian for a particle with spin: $\frac{1}{2}\pi^2$ is the usual kinetic energy, $\frac{1}{2}(f')^2$ is a potential, and $-\frac{1}{2}\hbar\sigma^3 f''$ has the form of a the interaction of the spin with a magnetic field $\propto f''$. The only odd thing is the special relation between the form of the potential and the magnetic field that was required by the supersymmetry.

An interesting question for supersymmetric field theories is whether supersymmetry is spontaneously broken. The analogous question here is whether or not there is a supersymmetric (*i.e.* zero energy) vacuum. When we are looking for exact zero-energy states $H|\Omega\rangle = 0$, then, by the supersymmetry algebra (1.14), $Q|\Omega\rangle = Q^\dagger|\Omega\rangle = 0$. Thus we need only look for solutions to the *first order* equations:

$$\begin{aligned} \Omega &= \begin{pmatrix} \omega_+ \\ 0 \end{pmatrix} \quad \Rightarrow \quad (f' + i\pi)\omega_+ = 0 \quad \Rightarrow \quad \omega_+ \propto e^{-f/\hbar}, \\ \Omega &= \begin{pmatrix} 0 \\ \omega_- \end{pmatrix} \quad \Rightarrow \quad (f' - i\pi)\omega_- = 0 \quad \Rightarrow \quad \omega_- \propto e^{+f/\hbar}. \end{aligned} \tag{1.24}$$

For these solutions to correspond to vacua, they must be normalizable. There are three cases:

$$\begin{array}{lccl} (1) & \lim_{\phi\to\pm\infty} f \to +\infty & \Rightarrow & \omega_+ \text{ normalizable, } \omega_- \text{ not;} \\ (2) & \lim_{\phi\to\pm\infty} f \to -\infty & \Rightarrow & \omega_- \text{ normalizable, } \omega_+ \text{ not;} \\ (3) & \lim_{\phi\to+\infty} f = -\lim_{\phi\to-\infty} f & \Rightarrow & \text{neither normalizable.} \end{array} \tag{1.25}$$

So in cases (1) and (2) we find a unique supersymmetric vacuum, while in case (3) we learn that supersymmetry is broken, for there is no supersymmetric vacuum state. With this, we have solved for the supersymmetric vacua of supersymmetric quantum mechanics. The simplification due to the supersymmetry algebra reducing a second order equation (in this case, the Schrödinger equation) to first order equations will be a recurring theme in supersymmetric field theory.

1.2.3 Superspace and superfields

We will now rewrite the supersymmetric quantum mechanics (1.23) using anticommuting (or Grassmann) numbers. These are classical analogs of fermionic operators, which are the $0+1$-dimensional version of fermionic fields in higher dimensional field theory. We will then define a *superspace* by formally extending space-time to include anticommuting coordinates. This will prove to be helpful for developing a representation theory for the supersymmetry algebra like that of ordinary symmetries, and in particular gives a powerful tool for quickly and compactly writing supersymmetric actions.

Define the (Schrödinger picture) quantum mechanical operators ψ and $\psi^\dagger$ by

$$\psi = \sqrt{\hbar}\sigma^+, \qquad \psi^\dagger = \sqrt{\hbar}\sigma^-, \tag{1.26}$$

so that they satisfy the algebra

$$\{\psi, \psi\} = \{\psi^\dagger, \psi^\dagger\} = 0, \qquad \{\psi^\dagger, \psi\} = \hbar. \tag{1.27}$$

In the Heisenberg picture of quantum mechanics, these operators become 0+1-dimensional field operators $\psi(t)$. The point of this renaming is that now our supersymmetric quantum mechanics becomes

$$\begin{aligned} Q &= \psi(f' + i\pi)/\sqrt{\hbar} \\ Q^\dagger &= \psi^\dagger(f' - i\pi)/\sqrt{\hbar} \\ H &= \frac{1}{2}\pi^2 + \frac{1}{2}(f')^2 - \frac{1}{2}[\psi^\dagger, \psi]\, f'', \end{aligned} \tag{1.28}$$

that is, without any explicit factors of $\hbar$ in the Hamiltonian. This allows us to identify a "classical" analog of the ψ fields, and so develop classical methods for treating fermions. The only novelty is that the classical limit of the algebra (1.27) of ψ's is $\{\psi, \psi\} = \{\psi^*, \psi^*\} = \{\psi^*, \psi\} = 0$ (where we have simply set $\hbar \to 0$ and traded Hermitian conjugation for complex conjugation). This is the algebra of anticommuting (or Grassmann) numbers, so we see that classical ψ fields take anti-commuting number values.

Now we can derive the above Hamiltonian from a classical Lagrangian by the canonical method, treating ψ as an independent field. The action that does the job is

$$S = \int dt \left\{ \frac{1}{2}\dot{\phi}^2 + i\psi^*\dot{\psi} - \frac{1}{2}(f')^2 + \frac{1}{2}f''\,[\psi^*, \psi] \right\}, \tag{1.29}$$

where a dot denotes a time derivative. Note that we have adopted the *convention* that the complex conjugate of a product of anticommuting numbers reverses their order without introducing an extra sign:

$$(\theta_1\theta_2)^* = \theta_2^*\theta_1^*. \tag{1.30}$$

It is easy to check that with this convention S is real. Canonical quantization of this action gives $\pi = \dot{\phi}$ as the momentum conjugate to ϕ, and shows that ψ and ψ^* are canonically conjugate fields, with anticommuting Poisson brackets which reproduce (1.27) upon quantization.

We define the supersymmetry variation of any field χ by

$$\delta\chi = [\epsilon^* Q + \epsilon Q^\dagger, \chi], \tag{1.31}$$

where ϵ is an (infinitesimal) constant anticommuting parameter. Because of our conjugation convention (1.30), $\epsilon^*\mathcal{Q}+\epsilon\mathcal{Q}^\dagger$ is anti-Hermitian, so $(\delta\chi)^\dagger = \delta\chi^\dagger$. The infinitesimal anticommuting parameter ϵ is just a bookkeeping device. From (1.28) the action of the supersymmetry generators on the fields ϕ, ψ, ψ^*, is (from now on we set $\hbar = 1$)

$$\begin{aligned} [Q,\phi] &= \psi, & [Q^\dagger,\phi] &= \psi^\dagger, \\ \{Q,\psi\} &= 0, & \{Q^\dagger,\psi\} &= f' - i\pi, \\ \{Q,\psi^\dagger\} &= f' + i\pi, & \{Q^\dagger,\psi^\dagger\} &= 0, \end{aligned} \tag{1.32}$$

so we find the supersymmetry variations

$$\begin{aligned} \delta\phi &= \epsilon^*\psi - \epsilon\psi^\dagger, \\ \delta\psi &= \epsilon(f' - i\pi). \end{aligned} \tag{1.33}$$

We now introduce superspace by extending space-time (in quantum mechanics this is only t) to include an anticommuting parameter θ for each supercharge Q: $t \to (t, \theta, \theta^*)$. A *superfield* Φ is then simply a general function on superspace:

$$\Phi(t,\theta,\theta^*) = \phi(t) + \theta\psi(t) - \theta^*\psi^*(t) + \theta\theta^* F(t). \tag{1.34}$$

Here we have chosen Φ to be a commuting superfield, implying that ϕ and F are commuting fields, while ψ is anticommuting. With our conjugation convention (1.30), if ϕ and F are real, then $\Phi^* = \Phi$. Note that there can be no higher order terms in θ or θ^* since $\theta^2 = (\theta^*)^2 = 0$. We speak of the coefficient fields ϕ, ψ, and F as the *components* of the superfield.

The reason superspace is useful is that, just as the Hermitian generator of time translations,

$$\mathcal{H} = i\frac{\partial}{\partial t}, \tag{1.35}$$

gives a geometrical realization of the Hamiltonian operator H acting on fields, so we can find similar translation operators acting on superfields which obey the supersymmetry algebra. To see this, recall that anticommuting differentiation is defined by

$$\left\{\frac{\partial}{\partial\theta},\theta\right\} = \left\{\frac{\partial}{\partial\theta^*},\theta^*\right\} = 1, \qquad \left\{\frac{\partial}{\partial\theta},\theta^*\right\} = \left\{\frac{\partial}{\partial\theta^*},\theta\right\} = 0, \tag{1.36}$$

and that anticommuting integration is the same as differentiation so that integration of a single Grassmann variable θ is

$$\int d\theta\,\theta = 1, \qquad \int d\theta\, 1 = 0, \tag{1.37}$$

and for multiple integration $d\theta_1 d\theta_2 = -d\theta_2 d\theta_1$. Note that with these definitions, the Hermitian conjugate of the the anticommuting derivative satisfies

$$\left(\frac{\partial}{\partial\theta}\right)^\dagger = \frac{\partial}{\partial\theta^*}. \tag{1.38}$$

This has a minus sign relative to the adjoint of derivatives of commuting parameters (*e.g.* $[\partial/\partial t]^\dagger = -\partial/\partial t$ by integration by parts) because the conjugation convention (1.30) implies that complex conjugation reverses the sign of anticommuting derivatives on commuting superfields:

$$\left(\frac{\partial}{\partial\theta}A\right)^* = -(-)^A \frac{\partial}{\partial\theta^*}A^* \tag{1.39}$$

for a general superfield A, where $(-)^A = +1$ if A is a commuting superfield and $(-)^A = -1$ if A is anticommuting.[2]

Now we can define differential operators on superspace

$$\mathcal{Q} = \frac{\partial}{\partial\theta} + i\theta^*\frac{\partial}{\partial t}, \qquad \mathcal{Q}^\dagger = \frac{\partial}{\partial\theta^*} + i\theta\frac{\partial}{\partial t}, \tag{1.40}$$

which satisfy the supersymmetry algebra:

$$\{\mathcal{Q}, \mathcal{Q}^\dagger\} = 2\mathcal{H}, \qquad \{\mathcal{Q}, \mathcal{Q}\} = 0. \tag{1.41}$$

Thus any superfield automatically provides a representation of the supersymmetry algebra, since we can define the supersymmetry variation of any superfield Φ by

$$\delta\Phi = [\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger, \Phi] = (\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger)\Phi, \tag{1.42}$$

where ϵ is an (infinitesimal) constant anticommuting parameter. It is straight forward to calculate

$$[\mathcal{Q}, \Phi] = \left(\frac{\partial}{\partial\theta} + i\theta^*\frac{\partial}{\partial t}\right)\Phi = \psi + \theta^*(F + i\dot{\phi}) - i\theta\theta^*\dot{\psi}, \tag{1.43}$$

so expanding out both sides of (1.42) in components gives

$$\begin{aligned} \delta\phi &= \epsilon^*\psi - \epsilon\psi^*, \\ \delta\psi &= \epsilon(F - i\dot{\phi}), \\ \delta F &= -i(\epsilon^*\dot{\psi} - \epsilon\dot{\psi}^*). \end{aligned} \tag{1.44}$$

Note that this supersymmetry variation applies to any product of superfields as well. (By the definition of the supersymmetry variation of fields, (1.31), the rule for the variation of a product of superfields is $\delta(\Phi_1\Phi_2) = [\epsilon^* Q + \epsilon Q^\dagger, \Phi_1\Phi_2] = [\epsilon^* Q + \epsilon Q^\dagger, \Phi_1]\Phi_2 = \Phi_1[\epsilon^* Q + \epsilon Q^\dagger, \Phi_2] = \delta\Phi_1\Phi_2 + \Phi_1\delta\Phi_2$; but this is satisfied as well by the superspace definition of the variation of superfields (1.42): $\delta(\Phi_1\Phi_2) = (\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger)(\Phi_1\Phi_2) =$

[2]More properly, one shows that $\int d\theta d\theta^*\, A^*(\partial/\partial\theta)B = \int d\theta d\theta^*\, [(\partial/\partial\theta^*)A]^* B$ for arbitrary superfields A and B.

$[(\epsilon^*\mathcal{Q}+\epsilon\mathcal{Q}^\dagger)\Phi_1]\Phi_2+\Phi_1[(\epsilon^*\mathcal{Q}+\epsilon\mathcal{Q}^\dagger)\Phi_2]$, by the Leibniz rule.) Since the sum or product of superfields is also a superfield, we can apply the above supersymmetry variation to a general polynomial function $\mathcal{L}(\Phi)$ of superfields.

The last equation in (1.44) shows that the supersymmetry variation of the $\theta\theta^*$ ("highest") component of any superfield is a total time derivative. Thus an action of the form $S=\int dt\, F$, where F is the highest component of any superfield, will automatically be supersymmetry invariant. By the rules for integration of anticommuting parameters, this can be written as

$$\int dt d\theta d\theta^*\, \mathcal{L}(\Phi), \tag{1.45}$$

since the $d\theta d\theta^*$ integration picks out only the highest component of $\mathcal{L}$.

We want not only to allow products of superfields in the Lagrangian $\mathcal{L}$, but also derivatives of superfields. To this end it is also convenient to define *covariant derivatives* or *superderivatives* on superspace,

$$\begin{aligned} \mathcal{D} &= \frac{\partial}{\partial\theta} - i\theta^*\frac{\partial}{\partial t}, \\ \mathcal{D}^\dagger &= \frac{\partial}{\partial\theta^*} - i\theta\frac{\partial}{\partial t}. \end{aligned} \tag{1.46}$$

which differ from $\mathcal{Q}$ and $\mathcal{Q}^\dagger$ just by taking $t\to -t$. They satisfy

$$\{\mathcal{D},\mathcal{Q}\} = \{\mathcal{D},\mathcal{Q}^\dagger\} = \{\mathcal{D},\mathcal{D}\} = 0, \qquad \{\mathcal{D},\mathcal{D}^\dagger\} = -2\mathcal{H}, \tag{1.47}$$

which is to say, they anticommute with $\mathcal{Q}$ and $\mathcal{Q}^\dagger$, and satisfy the supersymmetry algebra with a wrong sign. Their action on superfields is just like that of $\mathcal{Q}$ but with $t\to -t$, *e.g.*

$$\begin{aligned} \mathcal{D}\Phi &= \psi + \theta^*(F - i\dot\phi) + i\theta\theta^*\dot\psi, \\ \mathcal{D}^\dagger\Phi &= -\psi^* - \theta(F + i\dot\phi) + i\theta\theta^*\dot\psi^*. \end{aligned} \tag{1.48}$$

The utility of the covariant derivatives is that the covariant derivative of a superfield also transforms like a superfield under supersymmetry transformations:

$$\begin{aligned} \delta\mathcal{D}\Phi &= [\epsilon^* Q + \epsilon Q^\dagger, \mathcal{D}\Phi] = \mathcal{D}[\epsilon^* Q + \epsilon Q^\dagger, \Phi] = \mathcal{D}\delta\Phi \\ &= \mathcal{D}(\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger)\Phi = (\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger)\mathcal{D}\Phi, \end{aligned} \tag{1.49}$$

where the second equality ($\mathcal{D}$ commuting with $\epsilon^* Q + \epsilon Q^\dagger$) follows because $\mathcal{D}$ acts on the superspace coordinates while Q acts on the fields, and the last equality ($\mathcal{D}$ commuting with $\epsilon^*\mathcal{Q} + \epsilon\mathcal{Q}^\dagger$) follows from (1.47). Therefore we have shown that an arbitrary polynomial function of superfields and their superderivatives transforms in the same way under supersymmetry variations as a superfield does (1.44).

For this formalism to be useful, we would like to be able to write our supersymmetric quantum mechanics (1.29) in terms of superfields and covariant derivatives. An apparent problem with this is that the superfield Φ which we would like to associate with the fields ϕ and ψ has an extra component F which does not appear in (1.29); also the supersymmetry variation of the Φ components (1.44) does not match that of the quantum mechanics (1.33). Nevertheless, we can usefully make the above association of fields ϕ, ψ with superfield Φ.

Consider the supersymmetry invariant action

$$\begin{aligned} S &= \int dt d\theta d\theta^* \left\{ -\frac{1}{2}\mathcal{D}\Phi\mathcal{D}^\dagger\Phi + f(\Phi) \right\} \\ &= \int dt d\theta d\theta^* \left\{ \frac{1}{2}\left(\psi + \theta^*(F - i\dot{\phi}) + i\theta\theta^*\dot{\psi}\right)\left(\psi^* + \theta(F + i\dot{\phi}) - i\theta\theta^*\dot{\psi}^*\right) \right. \\ &\qquad \left. + f(\phi) + (\theta\psi - \theta^*\psi^* + \theta\theta^* F) f'(\phi) + \frac{1}{2}(\theta\psi - \theta^*\psi^*)^2 f''(\phi) \right\} \\ &= \int dt \left\{ \frac{1}{2}\left(F^2 + \dot{\phi}^2 + i(\psi^*\dot{\psi} - \dot{\psi}^*\psi)\right) - f'F + \frac{1}{2}f'' \cdot (\psi^*\psi - \psi\psi^*) \right\}, \end{aligned} \tag{1.50}$$

where in the third line we have expanded function $f(\Phi)$ in components by Taylor expanding in the anticommuting coordinates. F is an *auxiliary field* since its classical equation of motion,

$$F = f'(\phi), \tag{1.51}$$

is algebraic (it involves no time derivatives). Furthermore, since F appears only quadratically in the action, we can substitute its classical equation of motion even quantum mechanically (*e.g.* in a path integral formulation of quantum mechanics, the F-integral is Gaussian and just gives the classical result). With this substitution, the above action becomes precisely the original action (1.29).

The great utility of the superfield derives from the fact that it realizes the supersymmetry action on the fields linearly (1.44), as compared to the component formalism with only propagating fields (1.33). This linearization is accomplished through the introduction of the auxiliary F field. Although a superspace and superfields can be defined for arbitrary supersymmetry algebras (1.13) and in arbitrary number of dimensions, the hard problem is whether one can find a suitable (and simple enough) set of auxiliary fields for a given field theory for it to be reproducible by superfields (*i.e.* so that the supersymmetry algebra closes without using the classical equations of motion for the fields). In later lectures when we deal with extended supersymmetric theories in $d \geq 4$ dimensions, we will be forced to abandon superspace techniques for lack of a suitable set of auxiliary fields.

An important aspect of superfields is that the field representations of the supersymmetry algebra which they provide are typically not irreducible; smaller representations can be formed by constraining the superfields in some supersymmetry covariant way. Such constrained superfields can provide potentially new types of terms in supersymmetry invariant actions besides the $\int d\theta d\theta^* \mathcal{L}$ terms discussed above. An example illustrating this (which is somewhat artificial in 0+1 dimensions, but plays an important role in 3+1 dimensions) is the *chiral superfield*. This is a complex superfield X satisfying the additional constraint $\mathcal{D}^\dagger X = 0$. Noting that the coordinate combinations θ and $\tau \equiv t - i\theta\theta^*$ are themselves chiral ($\mathcal{D}^\dagger\theta = \mathcal{D}^\dagger\tau = 0$), it is easy to solve this constraint in general:

$$\begin{aligned} X(t,\theta,\theta^*) &= X(\tau,\theta) = \phi(\tau) + \theta\psi(\tau) \\ &= \phi(t) + \theta\psi(t) - i\theta\theta^*\dot{\phi}(t). \end{aligned} \tag{1.52}$$

It is easy to show that a product of chiral superfields is still chiral, and that $\mathcal{D}^\dagger\Phi$ is chiral whether Φ is or not.

Supersymmetric invariants can be formed as an integral of a chiral field over *half* of superspace,

$$S = \int dt d\theta\, X, \tag{1.53}$$

since by (1.44) the supersymmetry variation of any such term is a total derivative (since $F = -i\dot{\phi}$ for chiral fields). If X is chiral but not of the form $\mathcal{D}^\dagger\Phi$ for some superfield Φ, then such a term cannot be expressed as an integral over all of superspace, and can be used to write potentially new supersymmetry invariant terms in the action. Note that because anticommuting differentiation and integration are the same, we can dispense with the integration if we like. For example,

$$\int dt d\theta d\theta^* \mathcal{L} = \int dt d\theta \frac{\partial}{\partial\theta^*}\mathcal{L} = \int dt d\theta \mathcal{D}^\dagger \mathcal{L}, \tag{1.54}$$

where in the last step we have added a total derivative. What we have seen with the above construction of chiral superfield contibutions to the action is that the converse is not true: not every supersymmetry invariant term can be written as an integral over all of superspace.

Problem 1.2.1 Analyze the low-lying spectrum ($E \simeq 0$) of the supersymmetric quantum system (1.23) when $f(\phi)$ is a generic fourth- or third-order polynomial in ϕ. In particular, are there supersymmetric vacua, what is an estimate of the energies of the next lowest states, and what are their degeneracies?

Problem 1.2.2 Develop a superspace formalism for supersymmetric quantum mechanics with just *one* self-adjoint supercharge Q (as opposed to the one with two independent supercharges Q and $Q^\dagger$ that we discussed above). Can you construct any nontrivial quantum systems with this supersymmetry alone (*i.e.* not as a subalgebra of a supersymmetric quantum mechanics with more supersymmetry generators)?

1.3 Representations of the Lorentz Algebra

In this lecture we upgrade to four dimensions. One of the main technical difficulties of four dimensions compared to one dimension is the complication of the representation theory of the Lorentz algebra. This lecture will be a quick review of this representation theory without reference to supersymmetry.

The finite dimensional representations of the Lorentz algebra classify the different Lorentz covariant fields, local symmetry currents, and conserved charges that can arise in field theory. Our discussion of these representations will also serve to set our notation and conventions for spinors. I will not follow the easiest route to constructing the representations of the d=4 Lorentz algebra, but our path will have the virtue of generalizing to any dimension. Also, we will take the opportunity to review (or, perhaps, introduce) some basic notions in the representation theory of Lie algebras which will be useful in other contexts in later lectures.

Recall that the generators of Lorentz transformations $J^{\mu\nu} = -J^{\nu\mu}$ satisfy the Lie algebra

$$i[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} - \eta^{\nu\sigma} J^{\mu\rho} + \eta^{\mu\sigma} J^{\nu\rho}, \tag{1.55}$$

where $\eta^{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$ is the Minkowski metric. The six independent components of $J^{\mu\nu}$ can be organized into the generators of rotations, $J_i \equiv \frac{1}{2}\epsilon_{ijk} J^{jk}$, and boosts, $K_i \equiv J^{i0}$, where $i, j, k = 1, 2, 3$.

Fields are classified by the finite dimensional representations of this algebra, which means that under an infinitesimal Lorentz transformation $\Lambda^{\mu}_{\nu} = \delta^{\mu}_{\nu} + \eta^{\mu\rho}\omega_{\rho\nu}$ (where $\omega_{\rho\nu} = -\omega_{\nu\rho}$ are infinitesimal parameters), the components of a field ϕ_i, $i = 1, \ldots, n$, mix as

$$\delta\phi_i = \frac{i}{2}\omega_{\mu\nu} \left(\mathcal{J}^{\mu\nu}\right)_i^{\,j} \phi_j, \tag{1.56}$$

where the $n \times n$ matrices $\mathcal{J}^{\mu\nu}$ satisfy the commutation relations (1.55):

$$i[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}] = \eta^{\nu\rho} \mathcal{J}^{\mu\sigma} - \eta^{\mu\rho} \mathcal{J}^{\nu\sigma} - \eta^{\nu\sigma} \mathcal{J}^{\mu\rho} + \eta^{\mu\sigma} \mathcal{J}^{\nu\rho}. \tag{1.57}$$

In this matrix notation, the fields can be thought of as n-component column vectors, and the transformation rule (1.56) becomes $\delta\phi_i = \frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}\phi$, where matrix multiplication is understood. These infinitesimal transformations can be exponentiated to give a finite matrix representation of the action of the *Lorentz group*, *i.e.* the transformation rule for finite values of the $\omega_{\mu\nu}$ parameters is $\phi'_i = D(\omega)_i^{\,j}\phi_j$ where

$$D(\omega) = \exp\left\{\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}\right\}. \tag{1.58}$$

Because of the structure of the Lorentz algebra, the $\mathcal{J}^{\mu\nu}$ cannot all be Hermitian matrices, and thus the finite dimensional representations $D(\omega)$ of the Lorentz group are not unitary. This has to do with the non-compactness of the Lorentz group; indeed we will see later that the rotation generators are Hermitian, but the boosts must be anti-Hermitian, giving the reality conditions

$$\left(\mathcal{J}^{ij}\right)^\dagger = \mathcal{J}^{ij}, \qquad \left(\mathcal{J}^{i0}\right)^\dagger = -\mathcal{J}^{i0}, \qquad i,j = 1,2,3. \tag{1.59}$$

A change of basis of the components of ϕ to $\widetilde{\phi} = S\phi$ implies that $\widetilde{\phi}$ transforms by the representation matrices $\widetilde{\mathcal{J}}^{\mu\nu} = S\mathcal{J}^{\mu\nu}S^{-1}$. Conversely, if there exists a similarity transformation S relating two n-dimensional representation $\mathcal{J}^{\mu\nu}$ and $\widetilde{\mathcal{J}}^{\mu\nu}$ for all μ, ν, then the representations are said to be *equivalent*. If a representation is equivalent to one whose matrices all have the same block diagonal form (with more than one block), then the representation is said to be *reducible*, since we can then split the n field components of $\widetilde{\phi}$ into smaller sets which form representations of the Lorentz algebra by themselves. A representation which cannot be reduced any further is called *irreducible*. It is conventional to denote irreducible representations by their dimensions, so that an n-component field ϕ_i may be said to transform in the $\mathbf{n}$ representation. (Two inequivalent representations of the same dimension are differentiated by an appropriate subscript or superscript, *e.g.* $\mathbf{n}$ and $\mathbf{n}'$.) We will denote a field vector transforming in the $\mathbf{n}$ representation by $\phi^{\mathbf{n}}$ and the generator matrices similarly, $\mathcal{J}^{\mathbf{n}}_{\mu\nu}$; thus, if $n \neq m$, $\phi^{\mathbf{n}}$ and $\phi^{\mathbf{m}}$ are different fields even though I've used the same symbol ϕ for both.

The inverse operation to reducing a representation is the *direct sum* of representations, $\mathbf{n} \oplus \mathbf{m}$, in which two fields $\phi^{\mathbf{n}}$ and $\phi^{\mathbf{m}}$ are concatenated into a single field given by

$$\phi^{\mathbf{n}\oplus\mathbf{m}} = \begin{pmatrix} \phi^{\mathbf{n}} \\ \phi^{\mathbf{m}} \end{pmatrix}, \qquad \mathcal{J}^{\mathbf{n}\oplus\mathbf{m}}_{\mu\nu} = \begin{pmatrix} \mathcal{J}^{\mathbf{n}}_{\mu\nu} & 0 \\ 0 & \mathcal{J}^{\mathbf{m}}_{\mu\nu} \end{pmatrix}. \tag{1.60}$$

All finite dimensional representations of Lie algebras can be built out of ireducible representations by taking direct sums, and so it is sufficient to classify all the irreducible representations. Another way of building larger representations is by taking the *tensor product*, $\mathbf{n} \otimes \mathbf{m}$, of representations, defined by

$$\phi^{\mathbf{n}\otimes\mathbf{m}} = \phi^{\mathbf{n}} \otimes \phi^{\mathbf{m}}, \qquad \mathcal{J}^{\mathbf{n}\otimes\mathbf{m}}_{\mu\nu} = \mathcal{J}^{\mathbf{n}}_{\mu\nu} \otimes \mathbb{1}^{\mathbf{m}} + \mathbb{1}^{\mathbf{n}} \otimes \mathcal{J}^{\mathbf{m}}_{\mu\nu}, \tag{1.61}$$

where the $\otimes$ on the right hand sides denotes the usual tensor product of vectors and matrices, and $\mathbb{1}^{\mathbf{n}}$ denotes the $n \times n$ identity matrix. In components this reads

$$\phi^{\mathbf{n}\otimes\mathbf{m}}_{ia} = \phi^{\mathbf{n}}_i \phi^{\mathbf{m}}_a, \qquad \left(\mathcal{J}^{\mathbf{n}\otimes\mathbf{m}}_{\mu\nu}\right)^{ia}{}_{jb} = \left(\mathcal{J}^{\mathbf{n}}_{\mu\nu}\right)^i{}_j \delta^a_b + \delta^i_j \left(\mathcal{J}^{\mathbf{m}}_{\mu\nu}\right)^a{}_b, \tag{1.62}$$

where $i, j = 1, \ldots, n$ and $a, b = 1, \ldots, m$.

Tensor products of representations are typically reducible.[3] Products of two identical representations can immediately be reduced into their symmetric and antisymmetric parts:

$$\mathbf{n}\otimes\mathbf{n} = (\mathbf{n}\otimes_S \mathbf{n}) \oplus (\mathbf{n}\otimes_A \mathbf{n}), \tag{1.63}$$

where

$$\left(\phi^{\mathbf{n}\otimes_A \mathbf{n}}\right)_{ij} = (\phi^{\mathbf{n}})_i(\widetilde{\phi}^{\mathbf{n}})_j - (\phi^{\mathbf{n}})_j(\widetilde{\phi}^{\mathbf{n}})_i, \tag{1.64}$$

and similarly for the symmetrized product. That these form separate representations follows from the form (1.62) of the generators for products identical representations. This generalizes to n-fold products of the same representation which is the sum of representations with specific patterns of symmetrizations and antisymmetrizations of sets of factors. A useful method for further reducing product representations involves the *invariant tensors* of the algebra. To define invariant tensors, we need to know about the simplest representation all, the unique one dimensional *trivial, singlet,* or *scalar* representation,

$$\mathbf{1}: \qquad \phi, \qquad \mathcal{J}^{\mu\nu} = 0, \tag{1.65}$$

for which the generators simply vanish (*i.e.* the field ϕ in this representation does not transform at all under Lorentz transformations). An invariant tensor arises whenever the singlet representation occurs in the direct sum decomposition of a product of representations; the coefficients of the product factors that appear in the singlet piece form the invariant tensor. For example, if the product $\mathbf{m}\otimes\mathbf{n}\otimes\mathbf{p}$ contains a singlet given by $A^{afi}(\phi^{\mathbf{m}})_a(\phi^{\mathbf{n}})_f(\phi^{\mathbf{p}})_i$ for some coefficients A^{afi}, then the numerical tensor A^{afi} is an invariant tensor. Invariant tensors are useful since they give a way of contracting representations; in the above example, any product of representations whose factors include $\mathbf{m}$, $\mathbf{n}$, and $\mathbf{p}$ can be multiplied with A^{afi}, summing over the respective indices, to obtain a new, smaller representation. Will see examples below.

1.3.1 Tensors

With these generalities under our belts, we now turn to constructing the irreducible representations of the Lorentz algebra. A simple non-trivial representation is the four dimensional *vector* representation,

$$\mathbf{4}: \qquad \phi_\mu, \qquad (\mathcal{J}^{\mu\nu})^\rho_\sigma = \delta^\mu_\sigma\eta^{\nu\rho} - \delta^\nu_\sigma\eta^{\mu\rho}, \tag{1.66}$$

[3] A familiar example is the addition of angular momenta in quantum mechanics, where "adding" angular momenta corresponds to taking the tensor product of representations of the angular momentum ($so(3)$) algebra: $\mathbf{n}\otimes\mathbf{m} = (\mathbf{n}-\mathbf{m}+1)\oplus\cdots\oplus(\mathbf{n}+\mathbf{m}-1)$. Conventionally in quantum mechanics these representations are denoted by their spins $j=(n-1)/2$, instead of by their dimensions $\mathbf{n}$ as we are doing.

where we have used the space-time indices μ, ν, *etc.* as labels for the four components of the field (and we will adopt the convention of raising and lowering these indices with the $\eta^{\mu\nu}$ metric, as well).

From the tensor product of n vector representations we can make the *rank n tensor* representations, $\phi_{\mu_1\cdots\mu_n}$. These are not irreducible. They can be reduced by splitting them into tensors with definite symmetrizations or antisymmetrizations of subsets of their indices.[4] Thus, for example, the 16-dimensional rank 2 tensor representation can be split into a 10-dimensional symmetric and a 6-dimensional antisymmetric rank 2 tensor representation, which we denote by curly or square brackets on the indices:

$$\mathbf{4}\otimes_S\mathbf{4}=\phi_{\{\mu\nu\}}, \qquad \mathbf{4}\otimes_A\mathbf{4}=\phi_{[\mu\nu]}. \tag{1.67}$$

However, such tensors of definite symmetry are not, in general, irreducible, because of the possibility of contracting some of the indices with invariant tensors to form smaller representations. The simplest invariant tensor comes from the product of two vectors, since it is a familiar fact that

$$\phi^\nu\phi'_\nu=\eta^{\mu\nu}\phi_\mu\phi'_\nu \tag{1.68}$$

is a scalar (*i.e.* transforms in the trivial representation). Thus the invariant tensor is the Minkowski metric $\eta^{\mu\nu}$. Treating of $\eta^{\mu\nu}$ as a tensor field, it is easy to show directly from (1.66) and (1.62) that its variation vanishes (hence the name "invariant tensor"). Contracting the symmetric tensor $\phi_{\{\mu\nu\}}$ with $\eta^{\mu\nu}$ gives a singlet (the "trace"), so the remainder forms a 9-dimensional representation, the *traceless symmetric tensor*, which is irreducible:

$$\mathbf{9}: \qquad P^{\rho\sigma}_{\mu\nu}\phi_{\{\rho\sigma\}}=\phi_{\{\mu\nu\}}-\frac{1}{4}\eta_{\mu\nu}\left(\eta^{\rho\sigma}\phi_{\{\rho\sigma\}}\right), \tag{1.69}$$

where

$$P^{\rho\sigma}_{\mu\nu}=\delta^\rho_\mu\delta^\sigma_\nu-\frac{1}{4}\eta_{\mu\nu}\eta^{\rho\sigma} \tag{1.70}$$

is a traceless projection operator: $P^{\alpha\beta}_{\rho\sigma}P^{\rho\sigma}_{\mu\nu}=P^{\alpha\beta}_{\mu\nu}$. Thus we have found the decomposition $\mathbf{4}\otimes_S\mathbf{4}=\mathbf{1}\oplus\mathbf{9}$. The graviton is an example of a field transforming in the traceless antisymmetric representation. Note that contraction of the antisymmetric tensor $\phi_{[\mu\nu]}$ with the symmetric $\eta^{\mu\nu}$ gives zero identically, so does not help in reducing the antisymmetric tensor further.

Another invariant tensor arises in the completely antisymmetric product of four vectors, for such a tensor has only one independent component, and so is a singlet. We can

[4] The combinatorics of splitting a rank n tensor into parts with definite symmetries is the problem of finding the irreducible representations of the permutation group, and is conveniently soved in terms of *Young tableaux*; see [11, section 4.3] for a quick introduction.

write this component as $\epsilon^{\mu\nu\rho\sigma}\phi_\mu\phi_\nu\phi_\rho\phi_\sigma$ where $\epsilon^{\mu\nu\rho\sigma}$ is the unique tensor antisymmetric on four indices with

$$\epsilon^{0123} = +1. \tag{1.71}$$

(Note that, upon lowering indices with the Minkowski metric, we get the opposite sign $\epsilon_{0123} = -1$.) Upon contracting the rank 2 antisymmetric tensor with the ϵ tensor, we get another rank 2 antisymmetric tensor,

$$\phi^*_{[\mu\nu]} = \frac{i}{\sqrt{2}}\epsilon_{\mu\nu\rho\sigma}\phi^{[\rho\sigma]}, \tag{1.72}$$

which we will call the *Hodge dual* tensor. The factor of $i/\sqrt{2}$ is chosen in the definition of the Hodge dual so that (dropping indices) $(\phi^*)^* = \phi$. Given this definition, there are two possible invariant conditions one can put on ϕ, namely

$$\phi^* = \pm\phi. \tag{1.73}$$

An antisymmetric tensor satisfying this condition with the plus sign is said to be a *self dual tensor*, while one with the minus sign is an *anti-self dual tensor*. These are both 3-dimensional irreducible representations, and will be denoted $\mathbf{3}^\pm$. Thus we have found the decomposition $\mathbf{4} \otimes_A \mathbf{4} = \mathbf{3}^+ \oplus \mathbf{3}^-$. Note that because of the i in the definition of the Hodge dual, the $\mathbf{3}^\pm$ representations are necessarily complex, and

$$\left(\mathbf{3}^\pm\right)^* = \mathbf{3}^\mp. \tag{1.74}$$

Thus a real antisymmetric field (like the electromagnetic field strength) must transform in the reducible $\mathbf{3}^+ \oplus \mathbf{3}^-$ representation.

1.3.2 Spinors

In general, for tensor representations of the Lorentz group (and of orthogonal groups in any dimension) the metric and completely antisymmetric tensor are a complete set of invariant tensors, *i.e.* any tensor representation can be completely reduced using them. But not all irreducible representations of the Lorentz algebra are tensor representations; there are also *spinor* representations. A product of an even number of spinors is a tensor representation, while odd numbers of spinors give new representations, which can always be realized in the product of a spinor with some tensor. Lorentz spinors can be constructed (for any dimension) by the following trick. Given representation matrices γ^μ of the *four dimensional Clifford algebra*

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}, \tag{1.75}$$

it is straight forward to show that the matrices

$$\mathcal{J}^{\mu\nu} = -\frac{i}{4}[\gamma^\mu, \gamma^\nu] \tag{1.76}$$

form a representation of the Lorentz algebra. This Clifford algebra has a single, four dimensional irreducible representation (I'm not proving this), whose representation matrices are the *Dirac gamma matrices*, and give rise to the four dimensional *Dirac spinor* representation of the Lorentz algebra. We will denote the associated spinor fields by ψ_α, with spinor indices $\alpha, \beta, \ldots = 1, \ldots, 4$. In what follows we will follow the conventions of [1]; unfortunately there is no set of universally accepted conventions for gamma matrices and spinors.

Products of $\gamma^\rho\gamma^\sigma\cdots$ of gamma matrices can always be chosen to be totally antisymmetric on their $\rho\sigma\cdots$ indices since by the Clifford algebra any symmetric pair in the product can be replaced by the identity matrix. The complete list of independent products is then 1 (the 4×4 identity matrix), γ^μ, $\gamma^{[\mu}\gamma^{\nu]}$, $\gamma^{[\mu}\gamma^\nu\gamma^{\rho]}$, and $\gamma^{[\mu}\gamma^\nu\gamma^\rho\gamma^{\sigma]}$, where the square brackets denote antisymmetrization on n indices without dividing by $n!$. The rank 3 and 4 antisymmetric combinations can be rewritten as

$$\gamma^{[\mu}\gamma^\nu\gamma^{\rho]} = 3!\, i\, \epsilon^{\mu\nu\rho\sigma}\gamma_5\gamma_\sigma, \qquad \gamma^{[\mu}\gamma^\nu\gamma^\rho\gamma^{\sigma]} = 4!\, i\, \epsilon^{\mu\nu\rho\sigma}\gamma_5, \tag{1.77}$$

where we have defined (the phase is a non-universal convention)

$$\gamma_5 \equiv -i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{1.78}$$

which obeys

$$\{\gamma_5, \gamma_\mu\} = 0, \qquad {\gamma_5}^2 = 1, \qquad \gamma_5 = {\gamma_5}^\dagger. \tag{1.79}$$

(The last condition follows from the Hermiticity condition described below.) The set $M = \{1, \gamma_\mu, [\gamma_\mu, \gamma_\nu], \gamma_5\gamma_\mu, \gamma_5\}$ of sixteen matrices forms a (complex) basis for the space of all 4×4 matrices.

Since ${\gamma_5}^2 = 1$, its eigenvalues are all ± 1; since its trace vanishes, it has two $+1$ and two -1 eigenvalues. It is easy to show that $[\gamma_5, \mathcal{J}^{\mu\nu}] = 0$, implying that in the basis which diagonalizes γ_5 all the $\mathcal{J}^{\mu\nu}$ are block diagonal with two 2×2 blocks. Thus we learn that the Dirac spinor is reducible as a representation of the Lorentz algebra. Indeed, we can define two-component *left- and right-handed Weyl spinors* as spinors satisfying

$$\psi_L = \gamma_5\psi_L, \qquad \psi_R = -\gamma_5\psi_R. \tag{1.80}$$

These two 2-dimensional spinor representations of the Lorentz group will be denoted as $\mathbf{2}_L$ and $\mathbf{2}_R$. They can be formed from a general Dirac spinor representation by projecting out the components in the different γ_5 eigenspaces using the *chirality projection operators*

$$\mathcal{P}_\pm \equiv \frac{1}{2}(1 \pm \gamma_5) \tag{1.81}$$

so that

$$\psi_L \equiv \mathcal{P}_+\psi, \qquad \psi_R \equiv \mathcal{P}_-\psi. \tag{1.82}$$

These two representations are inequvalent, though we will see below that they are related by complex conjugation.

One conventionally demands that the γ^μ satisfy definite Hermiticity conditions. Because $-(\gamma^0)^2 = (\gamma^i)^2 = +1$, the eigenvalues of γ^0 are all $\pm i$ while those of the γ^i are ± 1. Thus the γ^i can be taken Hermitian, while γ^0 must be antiHermitian:

$$(\gamma^0)^\dagger = -\gamma^0, \qquad (\gamma^i)^\dagger = +\gamma^i; \tag{1.83}$$

note that this implies the reality conditions (1.59) for the Lorentz generators. Then

$$\beta\gamma_\mu\beta^{-1} = -\gamma_\mu^\dagger, \tag{1.84}$$

where we have defined (the phase is a non-universal convention)

$$\beta \equiv i\gamma^0 \tag{1.85}$$

which obeys

$$\beta^2 = 1, \qquad \beta = \beta^\dagger. \tag{1.86}$$

If a given set of 4×4 matrices γ_μ form a representation of the Clifford algebra, then $\pm\gamma_\mu^T$, $\pm\gamma_\mu^*$, and therefore $\pm\gamma_\mu^\dagger$ also satisfy the Clifford algebra. Since the Clifford algebra has only the one irreducible representation, all these matrices must be related to γ^μ by similarity transformations. γ_5 and β are the similarity matrices for $-\gamma_\mu$ and $-\gamma_\mu^\dagger$. There is also a matrix $\mathcal{C}$, called the *charge conjugation* matrix, such that

$$\mathcal{C}\gamma_\mu\mathcal{C}^{-1} = -\gamma_\mu^T. \tag{1.87}$$

It follows that

$$M^T = \begin{cases} +\mathcal{C}M\mathcal{C}^{-1} & M \;=\; 1,\ \gamma_5\gamma_\mu,\ \gamma_5 \\ -\mathcal{C}M\mathcal{C}^{-1} & M \;=\; \gamma_\mu,\ [\gamma_\mu, \gamma_\nu] \end{cases}. \tag{1.88}$$

One can show that, independent of the representation used for the γ_μ, $\mathcal{C}$ can be chosen to satisfy

$$\mathcal{C}\mathcal{C}^\dagger = 1, \qquad \mathcal{C} = -\mathcal{C}^T. \tag{1.89}$$

In [1] a specific basis for the gamma matrices is chosen for which

$$\gamma_0^T = +\gamma_0,\ \gamma_1^T = -\gamma_1,\ \gamma_2^T = +\gamma_2,\ \gamma_3^T = -\gamma_3. \tag{1.90}$$

and the phase of $\mathcal{C}$ is chosen so that

$$\mathcal{C} = \mathcal{C}^*. \tag{1.91}$$

These are representation-dependent statements, *i.e.* they change under unitary changes of spinor basis. More explicitly, we choose

$$\gamma^0 = -i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma_i = -i\begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \tag{1.92}$$

where 1 is the 2×2 identity matrix and σ_i are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.93}$$

γ_5, β, and $\mathcal{C}$ are then

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathcal{C} = -i\begin{pmatrix} \sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}. \tag{1.94}$$

In summary, in this basis we have

$$\begin{aligned} \gamma_5 = {\gamma_5}^T = {\gamma_5}^* = {\gamma_5}^\dagger = {\gamma_5}^{-1} &= i\gamma_0\gamma_1\gamma_2\gamma_3, \\ \beta = \beta^T = \beta^* = \beta^\dagger = \beta^{-1} &= -i\gamma_0, \\ \mathcal{C} = -\mathcal{C}^T = \mathcal{C}^* = -\mathcal{C}^\dagger = -\mathcal{C}^{-1} &= i\gamma_0\gamma_2, \\ \{\gamma_5, \beta\} = \{\beta, \mathcal{C}\} = [\mathcal{C}, \gamma_5] &= 0. \end{aligned} \tag{1.95}$$

By (1.84), (1.87) and the above relations it follows that

$$M^* = \left\{ \begin{array}{lll} +\beta\mathcal{C}M(\beta\mathcal{C})^{-1} & M = & 1,\ \gamma_\mu,\ [\gamma_\mu, \gamma_\nu] \\ -\beta\mathcal{C}M(\beta\mathcal{C})^{-1} & M = & \gamma_5\gamma_\mu,\ \gamma_5 \end{array} \right. . \tag{1.96}$$

In particular, $\beta\mathcal{C}\mathcal{J}_{\mu\nu}(\beta\mathcal{C})^{-1} = -\mathcal{J}^*_{\mu\nu}$ from which it follows that the spinors ψ^* and $\beta\mathcal{C}\psi$ transform in the same way under the Lorentz group. We'll call $\beta\mathcal{C}$ the *complex conjugation operator*. It gives us another way to reduce a Dirac spinor consistent with the Lorentz algebra: we can impose the reality condition

$$\psi^* = \beta\mathcal{C}\psi \tag{1.97}$$

which is consistent since $(\beta\mathcal{C})^2 = 1$. A spinor satisfying this reality condition is called a *Majorana spinor*. Any Dirac spinor ψ can be decomposed into two Majorana spinors $\psi_\pm$ by the projections

$$\psi_+ = \tfrac{1}{2}(\psi + \beta\mathcal{C}\psi^*), \qquad \psi_- = -i\tfrac{1}{2}(\psi - \beta\mathcal{C}\psi^*). \tag{1.98}$$

Because $\{\gamma_5, \beta\mathcal{C}\} = 0$, it is impossible to impose both a Majorana and a Weyl condition on a spinor. For instance, a spinor satisfying both $\psi^* = \beta\mathcal{C}\psi$ and $\psi = \gamma_5\psi$ vanishes by $\psi^* = \beta\mathcal{C}\psi = \beta\mathcal{C}\gamma_5\psi = -\gamma_5\beta\mathcal{C}\psi = -\gamma_5\psi^* = -(\gamma_5\psi)^* = -\psi^*$. This also shows that the complex conjugation operator interchanges left- and right-handed Weyl spinors. Finally, given a Majorana spinor ψ one can build a Weyl spinor ψ_R, and *vice versa*, by the inverse relations

$$\psi_R = \mathcal{P}_-\psi, \qquad \psi = (\psi_R + \beta\mathcal{C}\psi_R^*). \tag{1.99}$$

Despite this one-to-one map between Majorana and Weyl spinors, they are not equivalent as representations of the Lorentz algebra. In particular, a Majorana spinor can be projected onto both left- and right-handed Weyl spinors (which are complex conjugates of one another) and so transforms under Lorentz rotations as the reducible representation $\mathbf{2}_L \oplus \mathbf{2}_R$. Note that in our specific gamma matrix basis, the chirality projector $\mathcal{P}_-$ simply kills the lower two components of ψ:

$$\psi_R = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix} \qquad \text{if} \qquad \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{1.100}$$

We will be using Majorana spinors in most of these lectures, though it will be convenient at times to work in terms of Weyl spinors.

Now we turn to fermion bilinears, that is to say, the tensor product of two spinor representations. The basic result (which is not too difficult to check) is that for any Dirac spinors $\psi_{1,2}$ the bilinears

$$\psi_1^\dagger\beta M\psi_2, \qquad \text{for } M = 1,\ \gamma_\mu,\ [\gamma_\mu, \gamma_\nu],\ \gamma_5\gamma_\mu,\ \gamma_5, \tag{1.101}$$

transform as tensors under the Lorentz algebra according to the vector indices of M, where matrix multiplication on the spinor indices is understood. Since ψ^* transforms in the same way as $\beta\mathcal{C}\psi$, these tensors could equally well be written as $\psi_1^T(\beta\mathcal{C})^T\beta M\psi_2 = -\psi_1^T\mathcal{C}M\psi_2$, which shows that the matrices $\mathcal{C}M$ are invariant tensors for two spinor and various tensor representations. The form of these invariants motivates the bar notation

$$\overline{\psi} \equiv \psi^\dagger\beta. \tag{1.102}$$

For Majorana spinors, by (1.97), this becomes

$$\overline{\psi} = -\psi^T\mathcal{C}. \tag{1.103}$$

The properties of the Majorana bilinears under transposition and complex conjugation,

$$\begin{aligned} (\overline{\psi}_1 M\psi_2) &= \begin{cases} +(\overline{\psi}_2 M\psi_1) & M = 1,\ \gamma_5\gamma_\mu,\ \gamma_5 \\ -(\overline{\psi}_2 M\psi_1) & M = \gamma_\mu,\ [\gamma_\mu, \gamma_\nu] \end{cases}, \\ (\overline{\psi}_1 M\psi_2)^* &= \begin{cases} +(\overline{\psi}_1 M\psi_2) & M = 1,\ \gamma_\mu,\ [\gamma_\mu, \gamma_\nu] \\ -(\overline{\psi}_1 M\psi_2) & M = \gamma_5\gamma_\mu,\ \gamma_5 \end{cases}, \end{aligned} \tag{1.104}$$

follow from the definition of $\overline{\psi}$ and the properties of β and $\mathcal{C}$. The only subtleties arise from the fermionic nature of spinor fields, *i.e.* that they are values in the anticommuting numbers. In particular, one must remember the minus sign from interchange ψ_1 and ψ_2 in the first equation, our complex conjugation convention which reverses the order of anticommuting factors without introducing a minus sign under complex conjugation, and the extra minus sign in the second equation that comes from undoing the complex conjugation reversal of ψ_1 and ψ_2. It is important to note that these identities apply only to classical (anticommuting number valued) quantities, and *not* to operators (like the supersymmetry generators) which can have non-trivial anticommutation relations.

When $\psi_1 = \psi_2 = \theta$ (with an eye towards superspace) the above relations imply that the $\overline{\theta}\gamma_\mu\theta$ and $\overline{\theta}[\gamma_\mu, \gamma_\nu]\theta$ bilinears vanish identically. Thus the only tensors built from bilinears in a Majorana spinor θ are

$$(\overline{\theta}\theta), (\overline{\theta}\gamma_5\gamma^\mu\theta), (\overline{\theta}\gamma_5\theta). \tag{1.105}$$

The following identities will prove useful in manipulating superfields:

$$\begin{aligned}
\theta_\alpha\theta_\beta &= -\frac{1}{4}\mathcal{C}_{\alpha\beta}(\overline{\theta}\theta) - \frac{1}{4}(\gamma_5\mathcal{C})_{\alpha\beta}(\overline{\theta}\gamma_5\theta) + \frac{1}{4}(\gamma_5\gamma_\mu\mathcal{C})_{\alpha\beta}(\overline{\theta}\gamma_5\gamma^\mu\theta), \\
\theta_\alpha\theta_\beta\theta_\gamma &= -\frac{1}{8}(\overline{\theta}\gamma_5\theta)\left\{\mathcal{C}_{[\alpha\beta}(\gamma_5\theta)_{\gamma]} - (\gamma_5\mathcal{C})_{[\alpha\beta}\theta_{\gamma]}\right\}, \\
\theta_\alpha\theta_\beta\theta_\gamma\theta_\delta &= -\frac{1}{128}(\overline{\theta}\gamma_5\theta)^2\left\{-\mathcal{C}_{[\alpha\beta}\mathcal{C}_{\gamma\delta]} + (\gamma_5\mathcal{C})_{[\alpha\beta}(\gamma_5\mathcal{C})_{\gamma\delta]}\right\},
\end{aligned} \tag{1.106}$$

which, upon contracting appropriately, give

$$\begin{aligned}
\theta_\alpha(\overline{\theta}\theta) &= -(\gamma_5\theta)_\alpha(\overline{\theta}\gamma_5\theta), \\
\theta_\alpha(\overline{\theta}\gamma_5\gamma_\mu\theta) &= -(\gamma_{5\mu})_\alpha(\overline{\theta}\gamma_5\theta), \\
(\overline{\theta}\theta)^2 &= -(\overline{\theta}\gamma_5\theta)^2, \\
(\overline{\theta}\theta)(\overline{\theta}\gamma_5\theta) &= (\overline{\theta}\theta)(\overline{\theta}\gamma_5\gamma_\mu\theta) = (\overline{\theta}\gamma_5\theta)(\overline{\theta}\gamma_5\gamma_\mu\theta) = 0, \\
(\overline{\theta}\gamma_5\gamma_\mu\theta)(\overline{\theta}\gamma_5\gamma_\nu\theta) &= -\eta_{\mu\nu}(\overline{\theta}\gamma_5\theta)^2.
\end{aligned} \tag{1.107}$$

Note that any product of 5 or more θ's vanishes identically since θ has only four independent anticommuting components. These and other useful identities are proven in the appendix to [1, chapter 26].[5]

Finally, I would like to mention a quicker route to the representations of the Lorentz algebra which is special to d=4. One notices that the combinations of rotation and boost generators

$$L_i \equiv \tfrac{1}{2}(J_i + iK_i), \qquad R_i \equiv \tfrac{1}{2}(J_i - iK_i), \tag{1.108}$$

[5] Note that [1] introduces another matrix $\epsilon \equiv -\mathcal{C}\gamma_5$ so that for Majorana spinors $\overline{\theta} = -\theta^T\mathcal{C} = \theta^T\epsilon\gamma_5$.

satisfy two commuting copies of the $so(3)$ (angular momentum) algebra:

$$[L_i, L_j] = i\epsilon_{ijk}L_k, \qquad [R_i, R_j] = i\epsilon_{ijk}R_k, \qquad [L_i, R_j] = 0. \tag{1.109}$$

Thus the finite dimensional representations of the Lorentz group can be classified by pairs of "spins" (j_L, j_R) corresponding to $(2j_L + 1)(2j_R + 1)$-dimensional irreducible representations. The dictionary to our notation is

Name	Field	Dimension	(j_L, j_R)
scalar	ϕ	$\mathbf{1}$	$(0,0)$
left-handed spinor	ψ_L	$\mathbf{2}_L$	$(\frac{1}{2},0)$
right-handed spinor	ψ_R	$\mathbf{2}_R$	$(0,\frac{1}{2})$
vector	ϕ_μ	$\mathbf{4}$	$(\frac{1}{2},\frac{1}{2})$
self dual antisymmetric	$\phi^+_{[\mu\nu]}$	$\mathbf{3}^+$	$(1,0)$
anti-s.d. antisymmetric	$\phi^-_{[\mu\nu]}$	$\mathbf{3}^-$	$(0,1)$
traceless symmetric	$\phi_{\{\mu\nu\}}$	$\mathbf{9}$	$(1,1)$

(1.110)

This classification of Lorentz representations makes reducing tensor products especially easy. For example

$$\begin{aligned} (\tfrac{1}{2},0) \otimes (\tfrac{1}{2},0) &= (0,0) \oplus (1,0), \\ (\tfrac{1}{2},0) \otimes (0,\tfrac{1}{2},) &= (\tfrac{1}{2},\tfrac{1}{2}), \end{aligned} \tag{1.111}$$

follow easily from the usual rules for the addition of angular momenta.

Problem 1.3.1 Show from the d=4 Clifford algebra that the sixteen matrices $M_i = \{1, \gamma_\mu, [\gamma_\mu, \gamma_\nu], \gamma_5\gamma_\mu, \gamma_5\}$ are orthogonal with respect to a trace inner product:

$$\mathrm{tr}(M_i M_j) = \delta^i_j. \tag{1.112}$$

Problem 1.3.2 Show that the d=4 N=1 supersymmetry algebra can be written as

$$\overline{Q}\gamma_\mu Q = -4iP_\mu, \tag{1.113}$$

and that $(\overline{Q}\gamma_\mu Q)^\dagger = -\overline{Q}\gamma_\mu Q$. More generally, compute the traces of $\{Q_r, \overline{Q}_s\}$ with the M_i of the previous problem to determine which tensor operators can appear in the anticommutator by Lorentz invariance; here $r, s = 1, 2, \ldots$, label different supercharges.

Problem 1.3.3 Show that the action

$$S = \int d^4x \left(-\frac{1}{2}\overline{\psi}\gamma^\mu\partial_\mu\psi - \frac{1}{2}m\overline{\psi}\psi - i\frac{1}{2}\widetilde{m}\overline{\psi}\gamma_5\psi \right) \tag{1.114}$$

is real. What is the mass of this free fermion?

1.4 Supermultiplets

Particles states must transform in unitary representations of the 3+1 dimensional Poincaré algebra, generated by the Hermitian generators $J^{\mu\nu}$ of Lorentz transformations and by the energy-momentum four-vector P^μ (generating translations), and satisfying the the Lorentz algebra (1.55) as well as

$$\begin{aligned} i[P^\mu, J^{\rho\sigma}] &= \eta^{\mu\rho}P^\sigma - \eta^{\mu\sigma}P^\rho, \\ i[P^\mu, P^\nu] &= 0. \end{aligned} \tag{1.115}$$

Since the Poincaré algebra is extended by the supersymmetry generators, different particle representations will be related by supersymmetry transformations. These collections of supersymmetry-related particles are called *supermultiplets*.

Particle (scattering) states are labelled by the eigenvalues of the P^μ charges, *i.e.* their four-momenta p^μ (as long as translational invariance is not broken). Since the supercharges commute with the P^μ, supersymmetry transformations will not affect the four-momenta of states. In particular, the different particles in a supermultiplet will all have the same mass.

1.4.1 Poincaré algebra and particle states

Since the group corresponding to the Poincaré algebra is not compact, all its unitary representations (except the trivial representation) are infinite dimensional. This infinite dimensionality is simply the familiar fact that particle states are labelled by the continuous parameters p_μ—their four-momenta. Such representations can be organized by the *little group*, the subgroup of (usually compact) transformations left after fixing some of the non-compact transformations in some conventional way.

In the present case, the non-compact part of the Lorentz group are the boosts and translations. For massive particles, we can boost to a frame in which the particle is at rest

$$p^\mu = (m, 0, 0, 0). \tag{1.116}$$

The little group in this case is just those Lorentz transformations which preserve this four-vector—that is $SO(3)$, the group of rotations. Thus massive particles are in representations of $SO(3)$, labelled by the spin $j \in \frac{1}{2}\mathbb{Z}$ of the $(2j+1)$-dimensional representation

$$|j, j_3\rangle, \qquad -j \leq j_3 \leq j. \tag{1.117}$$

We have derived the familiar fact that a massive particle is described by its four-momentum and spin quantum numbers (as well as any internal quantum numbers).

Massless states are classified similarly. Here we can boost to

$$p^\mu = (E, 0, 0, E), \tag{1.118}$$

(for some conventional value of E) which is preserved by $SO(2)$ rotations around the z-axis.[6] Representations of $SO(2)$ are one dimensional, labelled by a single eigenvalue, the helicity

$$|\lambda\rangle, \tag{1.119}$$

which physically measures the component of angular momentum along the direction of motion. Algebraically λ could be any real number, but there is a topological constraint. Since the helicity is the eigenvalue of the rotation generator around the z-axis, a rotation by an angle θ around that axis produces a phase $e^{i\theta\lambda}$ on wave functions. Now, the Lorentz group is not simply connected: while a 2π rotation cannot be continuously deformed to the identity, a 4π rotation can. This implies that the phase $e^{4\pi i\lambda}$ must be one, giving the quantization of the helicity

$$\lambda \in \frac{1}{2}\mathbb{Z}. \tag{1.120}$$

The student for whom this material is unfamiliar is referred to [12, chapter 2] for a detailed exposition.

1.4.2 Particle representations of the supersymmetry algebra

Recall the d=4 N=1 supersymmetry algebra (1.12) written in terms of a four component Majorana spinor supercharge Q:

$$\{Q, \overline{Q}\} = -2i\gamma^\mu P_\mu, \qquad [Q, P_\mu] = 0. \tag{1.121}$$

This defines the normalization of the supersymmetry generators. The uniqueness of this algebra was discussed in the first lecture. Using the Majorana condition $\overline{Q} = -Q^T\mathcal{C}$ and multiplying by $\mathcal{C}$ gives the algebra in another useful form:

$$\{Q_\alpha, Q_\beta\} = -2i(\gamma^\mu\mathcal{C})_{\alpha\beta}P_\mu. \tag{1.122}$$

[6]Actually, the little group preserving p^μ is isomorphic to the non-compact group of Euclidean motions on the plane—$SO(2)$ plus two "translations" generated by the linear combinations $K_1 + J_2$ and $K_2 - J_1$ of boosts and rotations. However, being a non-compact group itself, this little group's unitary representations are infinite-dimensional, except when the eigenvalues of the "translations" are zero, in which case it effectively reduces to $SO(2)$. The infinite-dimensional representations are considered unphysical because we never see particle states in nature labelled by extra continuous parameters.

For the analysis of supermultiplets that follows, it will be useful to rewrite the supersymmetry algebra in terms of a right-handed Weyl supercharge and its complex conjugate:

$$Q_R = \mathcal{P}_- Q, \qquad Q_R^* = \mathcal{P}_- Q^* = \mathcal{P}_- \beta \mathcal{C} Q, \tag{1.123}$$

where we have used $\gamma_5 = {\gamma_5}^*$ and the Majorana condition in the form $Q^* = \beta\mathcal{C}Q$. Multiplying (1.122) appropriately by $\mathcal{P}_-$ and $\mathcal{P}_-\beta\mathcal{C}$ gives the anticommutators

$$\begin{aligned} \{\mathcal{P}_-Q, \mathcal{P}_-Q\} &= -2i(\mathcal{P}_-\gamma^\mu\mathcal{C}\mathcal{P}_-)P_\mu, \\ \{\mathcal{P}_-Q, \mathcal{P}_-Q^*\} &= -2i(\mathcal{P}_-\gamma^\mu\mathcal{C}(\beta\mathcal{C})^T\mathcal{P}_-)P_\mu. \end{aligned} \tag{1.124}$$

A little gamma matrix algebra shows that $\mathcal{P}_-\gamma^\mu\mathcal{C}\mathcal{P}_- = 0$ and $\mathcal{P}_-\gamma^\mu\mathcal{C}(\beta\mathcal{C})^T\mathcal{P}_- = \mathcal{P}_-\gamma^\mu\beta\mathcal{P}_-$. In our specific gamma matrix basis, recall that $\mathcal{P}_-$ simply annihilates the upper two of the four spinor components, and so the non-zero 2×2 block of $\mathcal{P}_-\gamma^\mu\beta\mathcal{P}_-$ is $i\sigma^\mu$ where we define

$$\sigma^0 = -1, \qquad \sigma^i = \sigma_i, \tag{1.125}$$

the Pauli matrices. Thus in terms of two-component Weyl spinors the supersymmetry algebra becomes

$$\{Q_R, Q_R\} = 0, \qquad \{Q_R, Q_R^*\} = 2\sigma^\mu P_\mu. \tag{1.126}$$

Now we are set to analyze the particle content of supermultiplets. Start with a massive particle state $|\Omega\rangle$ boosted to its rest frame: $p_\mu = (-m, 0, 0, 0)$. Then, acting on this state, the supersymmetry algebra (1.126) becomes

$$\{Q_{Ra}, Q_{Rb}^*\} = 2m\delta_{ab}, \qquad \{Q_{Ra}, Q_{Rb}\} = 0 \tag{1.127}$$

where $a, b = 1, 2$ index the components of the Weyl spinors. The representations of this algebra are easy to construct, since it is the algebra of two fermionic creation and annihiliation operators (up to a rescaling of Q_R by $\sqrt{2m}$). If we assume that Q_{Ra} annihilate a state $|\Omega\rangle$, then we find the four-dimensional representation:

$$|\Omega\rangle, \qquad Q_{Ra}^*|\Omega\rangle, \qquad Q_{R1}^*Q_{R2}^*|\Omega\rangle. \tag{1.128}$$

Suppose $|\Omega\rangle$ is a spin j particle. The Q_{Ra}^* operators transform in the $\mathbf{2}_L$ representation which transforms as spin $\frac{1}{2}$ under rotations. Thus the states $Q_{Ra}^*|\Omega\rangle$, by the rule for the addition of angular momenta, have spins $j+\frac{1}{2}$ and $j-\frac{1}{2}$ if $j\neq 0$, while if $j=0$ they have only spin $\frac{1}{2}$. The operator $Q_{R1}^*Q_{R2}^*$, on the other hand, transforms in the $\mathbf{2}_L \otimes_A \mathbf{2}_L = \mathbf{1}$ representation since the Q^*'s anticommute. In other words it transforms as a singlet ($j=0$) under rotations, and so the state $Q_{R1}^*Q_{R2}^*|\Omega\rangle$ has the same spin j as $|\Omega\rangle$. So, explicitly, the spin content of a massive spinless supersymmetry multiplet is

$$j = 0,\ 0,\ \tfrac{1}{2}, \tag{1.129}$$

while for a massive spinning multiplet, it is

$$j-\tfrac{1}{2},\ j,\ j,\ j+\tfrac{1}{2}. \tag{1.130}$$

You can check that such multiplets have equal numbers of bosonic and fermionic (propagating) degrees of freedom.

For massless particles, we boost to the frame where the four-momentum is $p_\mu = (-E, E, 0, 0)$, and denote the state by $|\Omega\rangle$. The supersymmetry algebra (1.126) is then

$$\{Q_{Ra}, Q^*_{Rb}\} = 4E \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{1.131}$$

This implies that $Q_{R2} = Q^*_{R2} = 0$ on all representations. Thus the massless supersymmetry multiplets are just two-dimensional, for if $|\Omega\rangle$ is annihilated by Q_{R1}, then the only states are

$$|\Omega\rangle, \qquad Q^*_{R1}|\Omega\rangle. \tag{1.132}$$

If $|\Omega\rangle$ has helicity λ, then $Q^*_{R1}|\Omega\rangle$ has helicity $\lambda+\frac{1}{2}$. By CPT invariance, such a multiplet will always appear in a field theory with its opposite helicity multiplet $(-\lambda\,,\ -\lambda-\frac{1}{2})$.

We will only concern ourselves with a few of these representations in these lectures. For massless particles, we will be interested in the *chiral multiplet* with helicities $\lambda = \{-\frac{1}{2}, 0, 0, \frac{1}{2}\}$, corresponding to the degrees of freedom associated with a complex scalar and a Majorana fermion: $\{\phi, \psi_\alpha\}$; and the *vector multiplet* with helicities $\lambda = \{-1, -\frac{1}{2}, \frac{1}{2}, 1\}$, corresponding to the degrees of freedom associated with a Majorana fermion and a vector boson: $\{\lambda_\alpha, A_\mu\}$. Other massless supersymmetry multiplets contain fields with spin 3/2 or greater. The only known consistent (classical) couplings for such fields occur in supergravity and gravity theories. Chiral multiplets are the supersymmetric analog of matter fields, while vector multiplets are the analog of the gauge fields. A terminology we will often use will refer to the fermions in the chiral multiplets as "quarks" and call their scalar superpartners *squarks*; also the fermionic superpartner of the gauge bosons are *gauginos*.

For massive particle multiplets, we have the massive chiral multiplet with spins $j = \{0, 0, \frac{1}{2}\}$, corresponding to a massive complex scalar and a Majorana fermion field: $\{\phi, \psi_\alpha\}$; and a massive vector multiplet with $j = \{0, \frac{1}{2}, \frac{1}{2}, 1\}$ with massive field content $\{h, \psi_\alpha, \lambda_\alpha, A_\mu\}$, where h is a real scalar field. In terms of propagating degrees of freedom, the massive vector multiplet has the same counting as a massless chiral plus a massless vector multiplet. This is indeed the case dynamically: massive vector multiplets arise by the usual Higgs mechanism.

1.4.3 Supersymmetry breaking

The above argument showing that at each mass there are equal numbers of boson and fermion states is valid only when supersymmetry is not spontaneously broken. We made the (hidden) assumption that the supersymmetry algebra charges were finite (*i.e.* that they always mapped states in the Hilbert space to other states in the Hilbert space). Precisely this assumption fails in the case of spontaneous breaking of ordinary (bosonic) internal symmetries: there the generators (charges) of the broken symmetry diverge because a field carrying that charge has formed a condensate in the vacuum. Thus the vacuum carries a finite charge per unit volume, so the total charge of the vacuum diverges. The broken charge does not annihilate the vacuum, and indeed, maps it to a non-normalizable state.

A similar mechanism is at work in spontaneous supersymmetry breaking. Just as in supersymmetric quantum mechanics, supersymmetry is broken if and only if the energy of the vacuum is non-zero. This follows from the supersymmetry algebra, which we can write as (see problem 1.3.2)

$$4P^\mu = Q^\dagger(i\beta\gamma^\mu)Q. \tag{1.133}$$

Since $i\beta\gamma^0 = 1$, it follows that

$$P^0 = \frac{1}{4}Q^\dagger_\alpha Q_\alpha, \tag{1.134}$$

so the energy density of the vacuum, ρ, is given by

$$\rho V = \langle P^0\rangle = \frac{1}{4}\sum_\alpha |Q_\alpha|0\rangle|^2 \geq 0, \tag{1.135}$$

where we are thinking of the system in a finite spatial volume V. Thus a vanishing vacuum energy is equivalent to $Q_\alpha|0\rangle = 0$ and unbroken supersymmetry.

Now, in the infinite volume limit, if supersymmetry is spontaneously broken, the vacuum energy and therefore the norm of $Q_\alpha|0\rangle$ is infinite. In supersymmetric field theory the supercharge is the the integral of a locally conserved *supersymmetry current* S^μ_α: $Q_\alpha = \int d^3\mathbf{x}\, S^0_\alpha$. The volume divergence of $\langle Q^\dagger Q\rangle$ implies that S^μ_α must create from the vacuum a spin-$\frac{1}{2}$ state with vanishing four-momentum and normalization $\langle 0|S^\mu_\alpha|\psi_\beta(p^\nu = 0)\rangle \propto (\gamma^\mu)_{\alpha\beta}/\sqrt{V}$, which is the normalization of a one particle state in finite volume. This massless fermion is called the *goldstino* in analogy to the massless boson associated with a spontaneously broken bosonic global symmetry. We will see goldstinos explicitly in field theory in later lectures. Thus in spontaneously broken supersymmetry the supersymmetry charges Q_α create from any state $|\Omega\rangle$ a non-normalizable partner $|\Omega + \{p_\mu{=}0 \text{ goldstino}\}\rangle$ of opposite statistics (see [1, sections 29.1-2] for a discussion).

The fact that the vacuum energy is the order parameter for supersymmetry breaking means that supersymmetry breaking can be seen even in finite volume regularizations of supersymmetric theories. Indeed, the low energy modes of a finite volume field theory form a quantum mechanical system with a finite number of degrees of freedom, and we saw spontaneous supersymmetry breaking in such a system in section 1.1. (This should be contrasted with bosonic spontaneous symmetry breaking, where in finite volumes quantum tunnelling mixes what would have been degenerate vacua in infinite volume.) Conversely, if supersymmetry is not broken at any finite volume, then it is not broken in the infinite volume limit. Finite volume regularization has been used [6] to show that supersymmetry is not spontaneously broken even non-perturbatively in supersymmetric versions of QCD.

1.5 N=1 Superspace and Chiral Superfields

Fields form representations of the N=1 d=4 supersymmetry algebra which are most conveniently handled in superspace. In this lecture we will introduce the *chiral superfield* describing the chiral supermultiplet. Vector multiplets will be discussed in later lectures.

1.5.1 Superspace

Extend space-time by including one anticommuting spinor coordinate for each supercharge Q_α:

$$x^\mu \rightarrow (x^\mu, \theta_\alpha). \tag{1.136}$$

Because we are working with Majorana supercharges, we also take θ Majorana. Differentiation and integration of θ satisfy the usual rules

$$\left\{ \frac{\partial}{\partial \theta_\alpha}, \theta_\beta \right\} = \int d\theta_\alpha \, \theta_\beta = \delta_{\alpha\beta}. \tag{1.137}$$

The usual chain rule for differentiation implies that if $\theta = M\theta'$ for some matrix M then $(\partial/\partial\theta) = M^{-T}(\partial/\partial\theta')$. Since the Majorana condition for θ, $\overline{\theta} = -\theta^T\mathcal{C} = \mathcal{C}\theta$ implies that

$$\frac{\partial}{\partial \overline{\theta}_\alpha} = \mathcal{C}_{\alpha\beta} \frac{\partial}{\partial \theta_\beta}. \tag{1.138}$$

A compact notation for superspace derivatives is

$$\partial_\alpha \equiv \frac{\partial}{\partial \theta_\alpha}, \qquad \overline{\partial}_\alpha \equiv \frac{\partial}{\partial \overline{\theta}_\alpha}, \qquad \partial_\mu \equiv \frac{\partial}{\partial x^\mu}. \tag{1.139}$$

More writing is saved by dropping spinor indices whenever there is no ambiguity, and by using the usual Dirac slash notation

$$\partial\!\!\!/ = \gamma^\mu \partial_\mu. \tag{1.140}$$

Thus, for example, in this compact notation the N=1 d=4 supersymmetry algebra (1.122) is

$$\{Q, Q\} = -2i \, P\!\!\!/ \mathcal{C}, \tag{1.141}$$

and a useful identity is

$$\overline{\partial}(\overline{\theta} M \theta) = 2M\theta, \tag{1.142}$$

for M any linear combination of the matrices 1, γ_5, and $\gamma_5\gamma^\mu$ (which are the only matrices for which $(\overline{\theta} M \theta)$ does not vanish identically).

We want to realize the supersymmetry algebra in terms of differential operators on superspace, with the four-momentum represented by the usual generator of translations:

$$P_\mu = i\partial_\mu. \tag{1.143}$$

One can check that

$$\mathcal{Q} = -\overline{\partial} + \not{\partial}\theta \tag{1.144}$$

does the job, *i.e.* $\{\mathcal{Q},\mathcal{Q}\} = 2\not{\partial}\mathcal{C}$. (Just to be clear for those having trouble with the indexless notation, the definition of $\mathcal{Q}$ scans as $\mathcal{Q}_\alpha = -\overline{\partial}_\alpha + (\gamma^\mu)_{\alpha\beta}\theta_\beta\partial_\mu$.) We also define the superderivative

$$\mathcal{D} = -\overline{\partial} - \not{\partial}\theta \tag{1.145}$$

which anticommutes with $\mathcal{Q}$ and satisfies the supersymmetry algebra with the wrong sign.

1.5.2 General Superfields

A general superfield $\Phi(x_\mu,\theta)$ is a function on superspace whose supersymmetry variation is given by

$$\delta\Phi = (\overline{\epsilon}\mathcal{Q})\Phi. \tag{1.146}$$

By the Leibniz rule for differentiation on superspace and $\{\mathcal{Q},\mathcal{D}\} = 0$, it follows that an arbitrary polynomial function of superfields and their superderivatives is itself a superfield. Since, as we saw in section 1.3, the product of two θ's is a linear combination of $(\overline{\theta}\theta)$, $(\overline{\theta}\gamma_5\theta)$, and $(\overline{\theta}\gamma_5\gamma_\mu\theta)$; the product of three θ's is proportional to $(\overline{\theta}\gamma_5\theta)\theta$; the product of four θ's is proportional to $(\overline{\theta}\gamma_5\theta)^2$; and the product of five or more θ's vanishes; the most general complex scalar superfield (*i.e.* one whose lowest component is a scalar field) has a component expansion

$$\Phi = \phi + (\overline{\theta}\psi) + (\overline{\theta}\theta)A + (\overline{\theta}\gamma_5\theta)B + (\overline{\theta}\gamma_5\gamma_\mu\theta)V^\mu + (\overline{\theta}\gamma_5\theta)(\overline{\theta}\lambda) - \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2 D, \tag{1.147}$$

where $\phi(x)$, $A(x)$, $B(x)$, and $D(x)$ are complex scalar fields, $\psi(x)$ and $\lambda(x)$ are Dirac spinor fields, and $V^\mu(x)$ is a complex vector field.

It is traditional to always denote the top component of an N=1 superfield by the letter "D", and to write

$$D \equiv [\Phi]_D. \tag{1.148}$$

Note the factor of $-\frac{1}{4}$ implicit in this definition, coming from (1.147). It is a straight forward exercise to show that the supersymmetry variation of the top component of a general superfield is a total space-time derivative,

$$\delta[\Phi]_D = \partial_\mu X \tag{1.149}$$

for some X. (We will write out the supersymmetry variation of a general superfield in components in a later lecture.) This implies that $S = \int d^4x\,[\Phi]_D$ is a supersymmetry invariant. Define

$$d^4\theta \equiv d\theta_1\, d\theta_2\, d\theta_3\, d\theta_4, \tag{1.150}$$

so that $\int d^4\theta\, \theta_4\theta_3\theta_2\theta_1 = 1$. In our spinor basis it is easy to show that $-\frac{1}{4}(\overline{\theta}\gamma_5\theta)^2 = -\frac{1}{2}\theta_4\theta_3\theta_2\theta_1$, so that a supersymmetry invariant action can be written as an integral over all of superspace:

$$S = \int d^4x\,[\Phi]_D = -\frac{1}{2}\int d^4x\, d^4\theta\, \Phi. \tag{1.151}$$

1.5.3 Chiral superfields

The general superfield has many component fields and gives a reducible representation of the supersymmetry algebra. To get an irreducible field representation we must impose a constraint on the superfield which (anti)commutes with the supersymmetry algebra. One such constraint is simply a reality condition, which turns out to lead to a vector multiplet—we will return to this representation in later lectures.

Another constraint we can impose is the so-called *left-chiral superfield* constraint:

$$\mathcal{D}_R\Phi = 0, \tag{1.152}$$

where we have defined the left-handed part of the superderivative by

$$\mathcal{D}_R \equiv \mathcal{P}_-\mathcal{D}, \tag{1.153}$$

and where we recall that $\mathcal{P}_\pm = \frac{1}{2}(1 \pm \gamma_5)$ are the left- and right-handed chirality projection operators. Thus the constraint (1.152) says that the right-handed superderivative annihilates a left-chiral superfield; a *right-chiral superfield* can be similarly defined to be annihilated by $\mathcal{D}_L \equiv \mathcal{P}_+\mathcal{D}$. One can check that

$$[\mathcal{P}_\pm, \mathcal{Q}] = 0, \tag{1.154}$$

implying that $\mathcal{D}_L$ anticommutes with $\mathcal{Q}$. This shows that the left-chiral superfield constraint is consistent: if Φ is a left-chiral superfield, then $Q\Phi$ is too.

This constraint is easy to solve explicitly by noting that

$$\theta_L \equiv \mathcal{P}_+\theta \tag{1.155}$$

and

$$x_+^\mu \equiv x^\mu + \tfrac{1}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta) \tag{1.156}$$

are both annihilated by $\mathcal{D}_R$. Thus the general solution to (1.152) is (assuming the lowest component is a scalar)

$$\begin{aligned}\Phi(x,\theta) &= \phi(x_+) + \sqrt{2}\left(\theta_L^T \mathcal{C}\mathcal{P}_+\psi(x_+)\right) - \left(\theta_L^T\mathcal{C}\theta_L\right)F(x_+)\\ &= \phi - \sqrt{2}(\overline{\theta}\psi_L) + (\overline{\theta}\mathcal{P}_+\theta)F + \frac{1}{2}(\overline{\theta}\gamma_5\gamma_\mu\theta)\partial^\mu\phi\\ &\quad + \frac{1}{\sqrt{2}}(\overline{\theta}\gamma_5\theta)(\overline{\theta}\gamma_5\not{\partial}\psi_L) - \frac{1}{8}(\overline{\theta}\gamma_5\theta)^2\Box\phi, \end{aligned} \tag{1.157}$$

where in the second line $\phi(x)$ and $F(x)$ are complex scalar fields, $\psi(x)$ is a Majorana fermion field, and $\Box \equiv \partial_\mu\partial^\mu$. Actually we see that it is the left-handed Weyl fermion

$$\psi_L \equiv \mathcal{P}_+\psi \tag{1.158}$$

which naturally enters in the left-chiral superfield (hence the name). A right-chiral superfield has the same expansion as in (1.157) but with $\mathcal{P}_+ \to \mathcal{P}_-$.

Note that the complex conjugate of a left-chiral superfield is a right-chiral superfield. This follows from

$$(\mathcal{D}_R\Phi)^* = \mathcal{P}_+^*\mathcal{D}_R^\dagger\Phi^* = \mathcal{P}_+\beta\mathcal{C}\mathcal{D}\Phi^* = \beta\mathcal{C}\mathcal{P}_-\mathcal{D}\Phi^* = \beta\mathcal{C}\mathcal{D}_L\Phi^*. \tag{1.159}$$

We will use only left-chiral superfields from now on, and complex conjugate when we need a right-chiral superfield. If Φ_n are left-chiral superfields, then it is easy to see that $\Phi_1+\Phi_2$ and $\Phi_1\Phi_2$ are also left-chiral superfields. Mixed objects such as $\Phi\Phi^*$ are neither left-chiral superfields nor right-chiral superfields. This can be summarized in the rule that any function of left-chiral superfields but not their complex conjugates is itself a left-chiral superfield. Note, however, that a left-chiral covariant derivative such as $\mathcal{D}_L\Phi$ for a left-chiral superfield Φ is a right-chiral superfield, not a left-chiral superfield. This gives another way to make a left-chiral superfield out of left-chiral superfields, since, for example, $\mathcal{D}_R\Phi^*$ is a left-chiral superfield if Φ is. Also, the space-time derivative of a left-chiral superfield is a left-chiral superfield since $\partial_\mu\Phi \propto \{\mathcal{D}_L,\mathcal{D}_R\}\Phi = \mathcal{D}_R\mathcal{D}_L\Phi$ so $\mathcal{D}_R\partial_\mu\Phi \propto \mathcal{D}_R^2\mathcal{D}_L\Phi = 0$.

The supersymmetry variation of the left-chiral superfield components is

$$\begin{aligned}\delta\phi &= \sqrt{2}\overline{\epsilon}\mathcal{P}_+\psi,\\ \delta\psi &= \sqrt{2}\mathcal{P}_+(\not{\partial}\phi + F)\epsilon + \sqrt{2}\mathcal{P}_-(\not{\partial}\phi^* + F^*)\epsilon,\\ \delta F &= \sqrt{2}\overline{\epsilon}\not{\partial}\mathcal{P}_+\psi, \end{aligned} \tag{1.160}$$

or, in terms of Weyl fermions,

$$\begin{aligned}\delta\phi &= \sqrt{2}\overline{\epsilon}\psi_L,\\ \delta\psi_L &= \sqrt{2}\mathcal{P}_+(\not{\partial}\phi + F)\epsilon,\\ \delta F &= \sqrt{2}\overline{\epsilon}\not{\partial}\psi_L. \end{aligned} \tag{1.161}$$

The $\overline{\theta}\mathcal{P}_+\theta$ component of a left-chiral superfield is traditionally denoted by the letter "F", and we write

$$F \equiv [\Phi]_F. \tag{1.162}$$

From the above supersymmetry variation we see that the F-component of a left-chiral superfield transforms by a total derivative, so its space-time integral is a supersymmetry invariant:

$$S = \int d^4x\,[\Phi]_F = \frac{1}{2}\int d^4x\, d^2\theta_L\, \Phi, \tag{1.163}$$

where we have also expressed it as an intergral over half of superspace by defining

$$d^2\theta_L \equiv d\theta_{L1}\, d\theta_{L2}. \tag{1.164}$$

1.5.4 Chiral superfield action: Kahler potential

Given a collection $\{\Phi^n\}$ of left-chiral superfields labelled by an index n, consider a supersymmetric action of the form

$$S = \frac{1}{2}\int d^4x\,[K(\Phi^n, (\Phi^n)^*)]_D + \int d^4x\,[f(\Phi^n)]_F + \text{c.c.} \tag{1.165}$$

where K, called the *Kahler potential*, is a real function of left-chiral superfields and their complex conjugates with no derivatives, and f, called the *superpotential*, is a complex function of left-chiral superfields alone, and "c.c." denotes the complex conjugate of the superpotential term. In the rest of this lecture we will expand this action in components and verify that it describes propagating and interacting chiral multiplets. Other possible supersymmetric terms, involving derivatives of chiral superfields, could be added to this action, but they turn out to give rise to higher-derivative terms in the component fields. In a later lecture we will explain why the terms with the fewset number of derivatives are generally the most interesting.

Let's start with the Kahler term. Note that a change in the Kahler potential of the form

$$K(\Phi^n, \Phi^{*n}) \to K(\Phi^n, \Phi^{*n}) + f(\Phi^n) + f^*(\Phi^{*n}) \tag{1.166}$$

will not change the component action since the D-component of a chiral superfield like $f(F^n)$ is a total derivative, by (1.160). Thus, expanding K in a power series in the fields Φ^n, the first non-trivial term is quadratic in the fields. These quadratic terms give rise to the free action for massless chiral supermultiplets.

We illustrate this with the simplest case of a single left-chiral superfield Φ:

$$\begin{aligned} S_K &= \frac{1}{2}\int d^4x\,[\Phi^*\Phi]_D \\ &= \int d^4x\,\left[-\partial_\mu\phi^*\partial^\mu\phi - \overline{\psi}_L \partial\!\!\!/ \psi_L + F^*F\right], \end{aligned} \tag{1.167}$$

where I have dropped total derivatives. (It is a good exercise to verify this.) This describes a free complex boson and a free Weyl fermion. This can also be written in terms of a free Majorana fermion using the identity

$$-\int d^4x\, \overline{\psi}_L \not{\partial}\psi_L = -\frac{1}{2}\int d^4x\, \overline{\psi}\not{\partial}\psi, \tag{1.168}$$

which follows since $\overline{\psi}\not{\partial}\gamma_5\psi$ is a total derivative. F has no derivatives and so is an auxiliary field; indeed, it enters the action quadratically, and so, even quantum mechanically, can be substituted by its equation of motion: $F = 0$.

We turn now to a general Kahler potential $K = K(\Phi^n, \Phi^{*n})$. First, we introduce a seemingly redundant, but useful notation in which the index n labelling the different left-chiral superfields Φ^n is barred when taking the complex conjugate. Hence

$$(\Phi^n)^* = \Phi^{*\overline{n}}. \tag{1.169}$$

When we expand in component fields, the coefficients of the expansion will be the Kahler potential, and derivatives of it with respect to its arguments, evaluated on the lowest components of Φ^n,

$$K(\phi^n, \phi^{*\overline{n}}), \qquad \frac{\partial}{\partial\phi^m}K(\phi^n, \phi^{*\overline{n}}), \qquad \textit{etc.}. \tag{1.170}$$

We can now introduce a shorthand notation for the derivatives of K (or any function of the fields) with respect to its arguments:

$$\partial_n \equiv \frac{\partial}{\partial\phi^n}, \qquad \partial_{\overline{n}} \equiv \frac{\partial}{\partial\phi^{*\overline{n}}}. \tag{1.171}$$

It will be useful to think of the complex scalar fields ϕ^n as maps from space-time to a *target space* with complex coordinates $\{\phi^n, \phi^{*\overline{n}}\}$. The Kahler potential is thus a real scalar function on the target space. We can use it to define a metric on target space by

$$g_{m\overline{m}} \equiv \partial_m\partial_{\overline{m}}K(\phi, \phi^*), \tag{1.172}$$

so that the line element reads $ds^2 = g_{m\overline{m}}d\phi^m d\phi^{*\overline{m}}$. (We treat barred and un-barred indices as independent, so that in the line element there are separate summations over m and $\overline{m}$.) Then, in the usual way, one defines a Christoffel symbol

$$\Gamma^k_{\ell m} = g^{k\overline{k}}g_{\ell\overline{k},m}, \qquad \Gamma^{\overline{k}}_{\overline{\ell}\overline{m}} = g^{k\overline{k}}g_{k\overline{\ell},\overline{m}}, \tag{1.173}$$

and a Riemann tensor,

$$R_{k\overline{\ell}m\overline{n}} = g_{k\overline{\ell},m\overline{n}} - \Gamma^p_{km}g_{p\overline{p}}\Gamma^{\overline{p}}_{\overline{\ell}\overline{n}}, \tag{1.174}$$

associated to this metric. Here $g^{m\overline{n}}$ is the inverse to $g_{m\overline{n}}$, and the indices after commas denote derivatives, *e.g.* $g_{k\bar{\ell},\overline{m}} = \partial_{\overline{m}} g_{k\bar{\ell}}$. In terms of these quantities one finds (this is an instructive exercise)

$$\begin{aligned} S_K &= \frac{1}{2}\int d^4x\,[K]_D \qquad (1.175)\\ &= \int d^4x \Big[g_{m\overline{m}} F^m F^{*\overline{m}} - \frac{1}{2} F^m g_{m\overline{m}} \Gamma^{\overline{m}}_{\overline{k\ell}} (\overline{\psi}^{\bar{k}}_L \psi^{\bar{\ell}}_R) - \frac{1}{2} F^{*\overline{m}} g_{m\overline{m}} \Gamma^m_{k\ell} (\overline{\psi}^k_R \psi^\ell_L) \\ &\quad - g_{m\overline{m}} \partial_\mu \phi^m \partial^\mu \phi^{*\overline{m}} - g_{m\overline{m}} \mathrm{Re}(\overline{\psi}^{\overline{m}}_L \not{D} \psi^m_L) + \tfrac{1}{4} g_{k\bar{\ell},m\overline{n}} (\overline{\psi}^k_R \psi^m_L)(\overline{\psi}^{\bar{\ell}}_L \psi^{\overline{n}}_R) \Big], \end{aligned}$$

where

$$D_\mu \psi^m \equiv (\delta^m_\ell \partial_\mu + \Gamma^m_{k\ell} \partial_\mu \phi^k)\psi^\ell. \qquad (1.176)$$

Note that the unbarred indices always adorn the left-handed Weyl fermion fields ψ^m_L (and thus the conjugates $\overline{\psi}^m_R$ of the right-handed fields as well), which fits with their being the superpartners of ϕ^m in left-chiral superfields; likewise right-handed Weyl fermions $\psi^{\overline{m}}_R$ are superpartners of $\phi^{*\overline{m}}$ in a right-chiral superfield. The kinetic terms for the fermions can no longer be written in a simple way in terms of a Majorana field since the $\overline{\psi}^m \not{\partial}\gamma_5 \psi^n$ total derivative can no longer be discarded as the metric $g_{m\overline{n}}$ is not a constant.

The equation of motion found by varying $F^{*\overline{m}}$ is $g_{m\overline{m}} F^m - \frac{1}{2} g_{m\overline{m}} \Gamma^m_{k\ell} (\overline{\psi}^k_R \psi^\ell_L) = 0$. For the kinetic term to have the right sign, $g_{m\overline{m}}$ must be positive definite, and hence invertible, giving $F^m = \frac{1}{2}\Gamma^m_{k\ell}(\overline{\psi}^k_R \psi^\ell_L)$. Substituting gives

$$S_K = \int d^4x \left[-g_{m\overline{m}} \partial_\mu \phi^m \partial^\mu \phi^{*\overline{m}} - g_{m\overline{m}} \mathrm{Re}(\overline{\psi}^{\overline{m}}_L \not{D} \psi^m_L) + \tfrac{1}{4} R_{k\bar{\ell}m\overline{n}} (\overline{\psi}^k_R \psi^m_L)(\overline{\psi}^{\bar{\ell}}_L \psi^{\overline{n}}_R) \right]. \qquad (1.177)$$

This is known as the supersymmetric *non-linear sigma model* for historic reasons. Its interest lies in the fact that it is the second term in the expansion of the (non-renormalizable) low energy effective action of a supersymmetric theory of left-chiral superfields. We will return to this point in later lectures.

It should not be surprising that complex Riemannian geometry has arisen in which the values of the complex scalar fields play the role of complex (or *holomorphic*) coordinates on the target space. Field redefinitions which preserve the chiral nature of the fields, $\phi^n \to f^n(\phi)$, are just complex coordinate transformations on the target space, implying that the target space will naturally have the structure of a manifold. The bosonic kinetic term naturally defines a postive-definite quadratic form on this manifold, thus giving it a metric structure. What is special to supersymmetry is that the target space geometry that occurs is actually *Kahler geometry*—complex geometry in which the metric is locally derived from a Kahler potential as in (1.172).

(Furthermore, the fermion fields ψ^n can be interpreted as a kind of vector in the tangent space to the Kahler manifold, since the space-time covariant derivative (1.176) can be thought of as a covariant derivative on target space pulled back to space-time,

$$D_\mu \psi^m = (\partial_\mu \phi^k) D_k \psi^m, \tag{1.178}$$

where we define the target space covariant derivative in the usual way as

$$D_k \psi^m \equiv (\delta^m_\ell \partial_k + \Gamma^m_{\ell k})\psi^\ell. \tag{1.179}$$

In writing the last two equations, however, we make the assumption that the ψ^n fields are functions of the ϕ^n fields, and depend on the space-time coordinates only implicitly through their dependence on the ϕ^n. This is an unphysical assumption for independently propagating fields.)

1.5.5 Chiral superfield action: Superpotential

The Kahler terms gave us the kinetic terms for massless chiral multiplets when K was quadratic. In the non-linear sigma model for general K there were extra (non-renormalizable) interaction terms, but all included derivatives of some field. To find the non-derivative interaction terms (including mass terms) we need to include superpotential terms in the action, which are the F-components of left-chiral superfields. So consider adding to S_K the terms

$$\begin{aligned} S_f &= \int d^4x \, \{[f(\Phi^n)]_F + \text{c.c.}\} \\ &= \int d^4x \left\{ F^n \partial_n f - \frac{1}{2}(\partial_n \partial_m f)(\overline{\psi}^n_R \psi^m_L) + \text{c.c.} \right\}, \end{aligned} \tag{1.180}$$

where in the second line we are treating f as a function of the lowest components of the left-chiral superfields: $f = f(\phi^n)$. We remove F^n by solving its (linear) equation of motion coming from combining the superpotential term with the non-linear sigma model terms,

$$F^n = \frac{1}{2}\Gamma^n_{k\ell}(\overline{\psi}^k_R \psi^\ell_L) - g^{n\overline{n}} \partial_{\overline{n}} f^*, \tag{1.181}$$

giving the terms

$$S_f = -\int d^4x \left\{ (D_n f) g^{n\overline{n}} (D_{\overline{n}} f^*) + \text{Re}\left[(D_n D_m f)(\overline{\psi}^n_R \psi^m_L) \right] \right\} \tag{1.182}$$

which are to be added to the non-linear sigma model terms (1.177). Here we have used target space covariant derivatives

$$D_m f \equiv \partial_m f, \qquad D_n D_m f \equiv (\delta^\ell_m \partial_n - \Gamma^\ell_{nm}) D_\ell f, \tag{1.183}$$

where the minus sign relative to (1.176) is because $D_m f$ has a lower instead of an upper index.

The first term in (1.182) is a scalar potential, while the second is a generalized Yukawa coupling which includes fermion mass terms. Note that since the metric $g_{n\overline{n}}$ is positive definite (for unitarity), the scalar potential

$$V(\phi) = (\partial_n f) g^{n\overline{n}} (\partial_{\overline{n}} f^*) \tag{1.184}$$

is a sum of squares and so is necessarily non-negative. Thus if there is a solution to the equations

$$\partial_n f = 0 \tag{1.185}$$

for all n, then this solution is a global minimum of the potential. The converse need not be true, however: we will see examples in the next lecture where the global minimum of the potential is not zero. According to our discussion in section 1.4.3, supersymmetry is spontaneously broken if and only if there is a non-zero vacuum energy density. A classical vacuum solution is

$$\phi^n(x^\mu) = \langle \phi^n \rangle, \tag{1.186}$$

i.e. where the scalars take constant values. (The spinor fields do not get vacuum expectation values by Lorentz invariance.) Then the only contribution to the vacuum energy coems from the potential term, and (1.185) is the condition for unbroken supersymmetry. Quantum mechanically, one might worry that vacuum fluctuations (*i.e.* renormalization effects) and fermion bilinear condensates $\langle \overline{\psi}_R \psi_L \rangle$ might affect this conclusion. As we will discuss in later lectures, the superpotential is protected from renormalization by quantum fluctuations in a supersymmetric vacuum, so (1.185) remains valid there. We will also see that in theories of chiral superfields alone, fermion condensates do not form; however in gauge theories such condensates might form (in analogy to chiral symmetry breaking in QCD) giving rise to *dynamical supersymmetry breaking* (= spontaneous supersymmetry breaking through a non-perturbative mechanism). The vacuum structure of strongly coupled supersymmetric gauge theories will be the subject of the second half of this course.

Note that the F^n auxiliary fields were not only auxiliary, but also appeared only quadratically. Thus the classical step of replacing them by their equations of motion is valid quantum mechanically. The fact that the F^n always appear at most quadratically follows simply from the fact that they are the highest components of the chiral superfields. It has become standard terminology to refer to the terms appearing in the scalar potential V coming from the superpotential as F *terms*. (We will later see that there is another contribution to the scalar potential when vector multiplets are included—the D *terms*.)

Although we have not shown it, the action (1.177) and (1.182) is the most general supersymmetric action describing the renormalizable interactions of chiral multiplets alone. This is not to say, however, that left-chiral superfields give the unique superspace way of constructing such actions. For example, another useful superfield, the *linear superfield* satisfying the constraint $\overline{\mathcal{D}}\mathcal{D}\Phi = 0$, can also describe chiral multiplet particle content; but its interactions are no more general than those of chiral superfields. Since left-chiral superfields capture all the relevant interactions, they will suffice for the purposes of these lectures where we use superfields mainly as a convenient way of constructing supersymmetric actions.

1.6 Classical Field Theory of Chiral Multiplets

In this lecture we will explore the classical (tree-level) physics of the chiral superfield actions introduced in the last lecture.

1.6.1 Renormalizable couplings

The general chiral superfield action is

$$S = S_K + S_f \tag{1.187}$$

where S_K is given by (1.177) and S_f by (1.182). If there is a global minimum of the scalar potential at $\langle\phi^n\rangle = \phi_0^n$, then we speak of the vacuum at the point $\{\phi_0^n\}$ in target space. The physics in this vacuum can be deduced by expanding the action around this point in target space

$$\phi^n(x^\mu) = \phi_0^n + \varphi^n(x^\mu). \tag{1.188}$$

Our ability to express the action S in terms of geometrical objects on target space means that the form of the action is invariant under refefinitions of the field variables

$$\phi^n \to \widetilde{\phi}^n(\phi^m) \tag{1.189}$$

accompanied by the usual transformation of the target space metric and by

$$\psi^n \to \widetilde{\psi}^n = (\partial_m\widetilde{\phi}^n)\psi^m. \tag{1.190}$$

By making such a redefinition of the field variables, we can assume that on the target space the coordinates ϕ^n are locally orthogonal, geodesic coordinates at the point ϕ_0^n. In such coordinates (known as *Riemann normal coordinates*) the metric has the expansion

$$g_{m\overline{n}}(\phi) = \delta_{m\overline{n}} + R_{k\overline{\ell}m\overline{n}}(\phi_0)\,\varphi^k\varphi^{*\overline{\ell}} + \mathcal{O}(\varphi^3), \tag{1.191}$$

implying

$$\begin{aligned} g^{n\overline{n}} &= \delta^{n\overline{n}} - R^{n\overline{n}}{}_{k\overline{\ell}}\varphi^k\varphi^{*\overline{\ell}} + \mathcal{O}(\varphi^3), \\ \Gamma^k_{mn} &= R_{m\overline{\ell}n}{}^k\varphi^{*\overline{\ell}} + \mathcal{O}(\varphi^2), \\ K &= \varphi^n\varphi^*_n + \tfrac{1}{4}R_{k\overline{\ell}m\overline{n}}\varphi^k\varphi^m\varphi^{*\overline{\ell}}\varphi^{*\overline{n}} + \mathcal{O}(\varphi^5), \end{aligned} \tag{1.192}$$

where we have dropped reference to the vacuum point ϕ_0 in the coefficients of the expansion in φ, and lower and raise their target space indices with the flat metric $\delta_{m\overline{m}}$ or its inverse $\delta^{m\overline{m}}$. Since the Kahler potential only enters the action with two derivatives (as the metric on target space), the constant and linear terms in the expansion of K

can be dropped without penalty. The reality of K implies the usual reality conditions on the metric and curvature:

$$(g_{m\overline{n}})^* = g_{n\overline{m}}, \qquad (R_{k\overline{\ell}m\overline{n}})^* = R_{\ell\overline{k}n\overline{m}}. \tag{1.193}$$

With the usual scaling of the fields in which the scalars φ have mass dimension 1 and the fermions ψ have mass dimension 3/2 (so their kinetic terms are dimensionless), inspection of (1.177) shows that contributions from the $\mathcal{O}(\varphi^2)$ and higher terms in g give rise to power counting irrelevant (non-renormalizable) terms. However, these terms do contribute to renormalizable terms in S_f. To see this, expand the superpotential about our vacuum ϕ_0. Since f only enters through its derivatives, we can drop any constant piece of f without loss of generality, and write

$$f = v_n\varphi^n + \tfrac{1}{2}\mathcal{M}_{mn}\varphi^m\varphi^n + \tfrac{1}{6}f_{\ell mn}\varphi^\ell\varphi^m\varphi^n + \mathcal{O}(\varphi^4), \tag{1.194}$$

where now

$$v_n = \partial_n f(\phi_0), \qquad \mathcal{M}_{mn} = \partial_m\partial_n f(\phi_0), \qquad f_{\ell mn} = \partial_\ell\partial_m\partial_n f(\phi_0). \tag{1.195}$$

As above we will raise and lower target space indices with $\delta^{n\overline{n}}$, so that, for example,

$$v^{\overline{n}} \equiv \delta^{n\overline{n}} v_n, \qquad v^{*n} \equiv \delta^{n\overline{n}} v^*_{\overline{n}} = \delta^{n\overline{n}}\partial_{\overline{n}} f^*(\phi_0^*). \tag{1.196}$$

It is straight forward to compute

$$\begin{aligned} D_n f &= v_n + \mathcal{M}_{nm}\varphi^m + \tfrac{1}{2}f_{nm\ell}\varphi^m\varphi^\ell + \mathcal{O}(\varphi^3), \\ D_n D_m f &= \mathcal{M}_{nm} + f_{nm\ell}\varphi^\ell - v_k R_{m\bar{\ell}n}{}^k\varphi^{*\overline{\ell}} + \mathcal{O}(\varphi^2). \end{aligned} \tag{1.197}$$

Plugging these expansions into S_f gives mass terms, Yukawa couplings, and φ^3 and φ^4 potential terms. The total action, keeping only renormalizable terms in the expansion around the vacuum ϕ_0 is then

$$\begin{aligned} S &= \int d^4x \Big\{ -\partial_\mu\varphi^*_m\partial^\mu\varphi^m - \tfrac{1}{2}\overline{\psi}_m \partial\!\!\!/ \psi^m - \mathrm{Re}\left[\mathcal{M}^{\overline{n}}_m(\overline{\psi}_{\overline{n}}\mathcal{P}_+\psi^m)\right] \\ &\quad - \mathrm{Re}\left[(f^{\overline{n}}{}_{m\ell}\varphi^\ell - v^{\overline{k}}R^{\overline{n}}{}_{\overline{k}m}{}^\ell\varphi^*_\ell)(\overline{\psi}_{\overline{n}}\mathcal{P}_+\psi^m)\right] - V(\varphi,\varphi^*)\Big\} \end{aligned} \tag{1.198}$$

where the scalar potential is

$$\begin{aligned} V &= v^*_{\overline{n}}v^{\overline{n}} + \left(v^*_{\overline{n}}\mathcal{M}^{\overline{n}}_m\varphi^m + \text{c.c.}\right) + \tfrac{1}{2}\left(v^*_{\overline{n}}f^{\overline{n}}{}_{m\ell}\varphi^m\varphi^\ell + \text{c.c.}\right) \\ &\quad + \varphi^*_m\left(\mathcal{M}^{*m}_{\overline{k}}\mathcal{M}^{\overline{k}}_n - v^*_{\overline{k}}v^{\overline{\ell}}R^{\overline{k}}{}_{\overline{\ell}n}{}^m\right)\varphi^n + \mathcal{O}(\varphi^3,\varphi^4). \end{aligned} \tag{1.199}$$

I have not written out the φ^3 and φ^4 terms since to do that I would have had to keep terms in the expansion of $g_{m\overline{m}}$ out to order φ^4, which is a real mess. I will write them out later for the case where supersymmetry is not spontaneously broken. (Note that in writing the action in terms of Majorana fermions, there is no longer a consistent way to assign barred *versus* unbarred indices to the fermions; I use whichever is convenient.)

The masses of the scalars and spinors can be read off from the terms in (1.198) and (1.199) quadratic in the fields. These terms can be written in an obvious matrix notation on the "flavor" indices k, ℓ, m, and n, as

$$\begin{aligned} S_2 &= \int d^4x \Big\{ -\partial_\mu \varphi^\dagger \partial^\mu \varphi - \varphi^\dagger (\mathcal{M}^\dagger \mathcal{M} - v^\dagger R v) \varphi - \tfrac{1}{2} \varphi^T (v^\dagger f) \varphi - \tfrac{1}{2} \varphi^\dagger (f^* v) \varphi^* \\ &\qquad - \tfrac{1}{2} \overline{\psi} \partial\!\!\!/ \psi - \tfrac{1}{2} \overline{\psi} (\mathcal{M} \mathcal{P}_+ + \mathcal{M}^\dagger \mathcal{P}_-) \psi \Big\}, \end{aligned} \tag{1.200}$$

giving the equations of motion $(\partial\!\!\!/ + \mathcal{M}\mathcal{P}_+ + \mathcal{M}^\dagger \mathcal{P}_-)\psi = 0$ and $(\Box - \mathcal{M}^\dagger \mathcal{M} + v^\dagger R v)\varphi - (f^* v)\varphi^* = 0$. Multiplying the fermion equation by $(\partial\!\!\!/ - \mathcal{M}\mathcal{P}_- - \mathcal{M}^\dagger \mathcal{P}_+)$ and splitting it into left- and right-chiral parts gives

$$\Box \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \mathcal{M}^\dagger \mathcal{M} & 0 \\ 0 & \mathcal{M}\mathcal{M}^\dagger \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \tag{1.201}$$

implying that the masses-squared are given by the eigenvalues of the matrix $\mathcal{M}^\dagger \mathcal{M}$. (Since $\mathcal{M}^\dagger \mathcal{M}$ is Hermitian, it can be diagonalized by a unitary transformation, and its eigenvalues are non-negative real numbers; since $\mathcal{M}$ is symmetric, $\mathcal{M}\mathcal{M}^\dagger = (\mathcal{M}^\dagger \mathcal{M})^*$, so the masses-squared of ψ_R equal those of ψ_L, as had to be the case by CPT invariance.) Likewise, combining the scalar equation with its complex conjugate gives

$$\Box \begin{pmatrix} \varphi \\ \varphi^* \end{pmatrix} = \begin{pmatrix} \mathcal{M}^\dagger \mathcal{M} - v^\dagger R v & f^* v \\ f v^* & \mathcal{M}\mathcal{M}^\dagger - v^\dagger R v \end{pmatrix} \begin{pmatrix} \varphi \\ \varphi^* \end{pmatrix}, \tag{1.202}$$

implying the scalar masses-squared are the eigenvalues of this Hermitian matrix.

Recall from our discussion at the end of last lecture that the condition for supersymmetry to not be spontaneously broken is that the value of the scalar potential vanish in the vacuum:

$$v^{\overline{n}} = 0. \tag{1.203}$$

Thus, in the case of unbroken supersymmetry the scalar mass-squared matrix simplifies dramatically, becoming the same as that of the fermions. This is the expected boson-fermion mass degeneracy. Conversely, when $v^{\overline{n}} \neq 0$ and supersymmetry is spontaneously broken, the scalar masses are split from their erstwhile supersymmetric values, while the fermions are not. Neglecting, for the moment, the target space curvature R

coming from the Kahler potential, the off-diagonal terms in the scalar mass-squared matrix imply that the scalar masses will be split from each other and from the fermions by (schematically)

$$\delta m_\varphi^2 \sim \pm |v_k^* f^k{}_{mn}|. \tag{1.204}$$

This makes sense physically, since v is the order parameter for supersymmetry breaking. We see that the effects of supersymmetry breaking are "transmitted" by fields with $v^* \neq 0$ to the other fields through the dimensionless (Yukawa and φ^4) couplings f_{kmn}. Note that there can be a further overall shift of the scalar masses relative to the fermion masses in cases where the target space curvature R from the Kahler potential cannot be neglected.[7]

Whether supersymmetry is broken or not, the vacuum is at a minimum of the scalar potential V. This implies that the terms linear in φ in (1.199) must vanish:

$$v^\dagger \mathcal{M} \varphi = 0. \tag{1.205}$$

Thus, if supersymmetry is broken, so $v \neq 0$, we learn that $\mathcal{M}^\dagger$, and therefore the fermion mass-squared matrix $\mathcal{M}\mathcal{M}^\dagger$, has a zero eigenvalue. We thus see that whenever supersymmetry is spontaneously broken, there is a massless fermion: the Goldstino.

It is useful to translate the above description of spontaneous supersymmetry breaking in a field theory of left-chiral superfields into more general terms. Supersymmetry is spontaneously broken if and only if the vacuum expectation value of the supersymmetry variation of some field is non-zero. The field in question must be a fermion, since the supersymmetry variation of a boson is a fermion, whose vacuum expectation value vanishes by Lorentz invariance (which we assume to be unbroken). Thus the condition for supersymmetry to be spontaneously broken is that there exists a fermion ψ such that $\langle \delta\psi \rangle = \langle \{Q, \psi\} \rangle \neq 0$. But for left-chiral superfields the supersymmetry variation of the ψ component fields (1.160) is $\delta\psi \sim F + \partial\!\!\!/\phi$, so supersymmetry breaks if and only if $\langle F \rangle \neq 0$, since $\partial_\mu \phi$ can't get a Lorentz invariant vacuum expectation value. Thus, the order parameter for supersymmetry breaking is the expectation value of the F-components of left-chiral superfields. (It may be worth emphasizing that, by contrast, $\langle \phi \rangle \neq 0$ does *not* break supersymmetry.) The fermion field ψ which is not supersymmetry invariant in the ground state is the superpartner of the non-vanishing F-component. This fermion, which shifts under a supersymmetry transformation, is the Goldstino.

Restricting ourselves to the case where supersymmetry is not spontaneously broken,

[7] Often in the literature reference is made to a *supersymmetric mass sum rule* [13] which states that the sum of the scalar masses-squared equals that of the fermion masses-squared. Taking traces of the mass-squared matrices found above, we see that this result only holds when the target space curvature R vanishes. It is also modified when vector multiplets are included.

we write out all the renormalizable terms in the chiral superfield action:

$$\begin{aligned} S &= \int d^4x \Big\{ \varphi^\dagger(\Box - \mathcal{M}^\dagger\mathcal{M})\varphi - \tfrac{1}{2}\overline{\psi}(\partial\!\!\!/ + \mathcal{M}\mathcal{P}_+ + \mathcal{M}^\dagger\mathcal{P}_-)\psi \\ & \quad -\tfrac{1}{2}f_{\ell mn}\varphi^\ell\overline{\psi}^m\mathcal{P}_+\psi^n - \tfrac{1}{2}f^{*\ell mn}\varphi^*_\ell\overline{\psi}_m\mathcal{P}_-\psi_n \\ & \quad -\tfrac{1}{2}\mathcal{M}^{*k}_{\overline{\ell}}f^{\overline{\ell}}_{mn}\varphi^*_k\varphi^m\varphi^n - \tfrac{1}{2}f^{*mn}_{\overline{\ell}}\mathcal{M}^{\overline{\ell}}_k\varphi^k\varphi^*_m\varphi^*_n \\ & \quad -\tfrac{1}{4}f^{*mn}_{\overline{\jmath}}f^{\overline{\jmath}}_{k\ell}\varphi^k\varphi^\ell\varphi^*_m\varphi^*_n \Big\}. \end{aligned} \tag{1.206}$$

1.6.2 Generic superpotentials and R symmetries

Having analyzed the qualititative physics in vacua with and without spontaneously broken supersymmetry, the question remains as to when spontaneous supersymmetry breaking actually takes place in chiral superfield actions. We will review some general statements that can be made about *when* supersymmetry can and cannot be broken in our theories [14]. We have seen that supersymmetry is unbroken if and only if there exists a simultaneous solution to the equations

$$\frac{\partial f(\phi^n)}{\partial \phi^m} = 0 \qquad \text{for } m = 1, \ldots, N. \tag{1.207}$$

These are N complex analytic equations in N complex unknowns (the vacuum expectation values of the chiral fields ϕ^n), and so there will *generically* exist a solution. "Generically" means that by making an arbitrary small change in the couplings (consistent with symmetries) any theory with no solution to (1.207) will be taken to a theory which has a solution. So we learn that for a generic superpotential, supersymmetry is unbroken. It is a general working hypothesis that unless there are some symmetries to restrict the model, there will be generated by quantum corrections all possible terms in the effective action, and thus that the superpotential will be generic.

So, what if the superpotential is constrained by a global internal symmetry? Say it is a $U(1)$ symmetry with charges $Q(\Phi^n) = q_n$. That is to say, the global symmetry transforms the left-chiral superfields (and thus each of their component fields) as

$$\Phi^n \to e^{i\alpha q_n}\Phi^n \tag{1.208}$$

for arbitrary real constant parameter α. Now, the vacuum may or may not spontaneously break this symmetry. If it does not, then (by definition) the vacuum expectation values of all the charged fields are zero. We can then reduce the question of the existence of supersymmetric vacuum to whether there is a solution to $\partial_n f = 0$

restricted to the submanifold of target space where all charged bosons vanish. This submanifold of target space is described by the complex equations

$$\phi^n = 0 \qquad \text{when } q_n \neq 0, \tag{1.209}$$

and thus describes a complex submanifold of target space. Say it is M complex dimensional. Then the restriction of the N complex equations $\partial_n f = 0$ to this submanifold will, by the chain rule, give M independent complex equations. This just takes us back to the previous situation with no global symmetries, and supersymmetry is not broken, generically. If, on the other hand, the $U(1)$ symmetry is spontaneously broken, then at least one of the charged fields will have a non-zero vacuum expectation value. Without loss of generality, we can take it to be Φ_1 and choose the normalization of the $U(1)$ generator so that $q_1 = 1$. For Q to be a symmetry of the action, the superpotential must be neutral, and so it has the form

$$f(\Phi^1, \ldots, \Phi^n) = f(U_2, \ldots, U_N), \text{ where } U_n \equiv \Phi^n/(\Phi^1)^{q_n}. \tag{1.210}$$

(This should be taken as a local statement in target space.) Because $\langle\phi^1\rangle \neq 0$, this change of variables is non-singular, so $\partial_n f = 0$ if and only if $(\partial/\partial U_n) f = 0$, giving $N-1$ equations for $N-1$ variables, and so for a generic superpotential again supersymmetry is unbroken. This argument can be easily generalized to an arbitrary number of $U(1)$ or non-Abelian global symmetries, to show that supersymmetry is generically not broken.

A hidden assumption in the above argument was that under global internal symmetry transformations all components of left-chiral superfields transform in the same way. This is equivalent to the assumption that the global symmetry generators commute with the supercharge Q_α. By the Coleman-Mandula theorem (see section 1.1) all global symmetries must commute with the Poincaré group; however, it is not necessary that they commute with the supersymmetry algebra. Associativity of the superPoincaré algebra implies[8] that there can be at most a single (independent) Hermitian $U(1)$ generator R which does not commute with the supersymmetry generator Q_L, and is conventionally normalized so that

$$[R, Q_L] = -Q_L. \tag{1.211}$$

[8] Say there were a global symmetry algebra with Hermitian generators T^a, $[T^a, T^b] = i f^{ab}_c T^c$ which did not commute with supersymmetry $[T^a, Q_\alpha] = h^a Q_\alpha$. The Jacobi identity $[T^a, [T^b, Q]] + [T^b, [Q, T^a]] + [Q, [T^a, T^b]] = 0$ implies $f^{ab}_c h^c = 0$. Now, by the Coleman-Mandula theorem, any scalar symmetry algebra is a direct sum of a semi-simple algebra $\mathcal{A}_1$ and an Abelian algebra $\mathcal{A}_2$. Since for a semi-simple Lie algebra the Killing form $g^{ab} = f^{ay}_x f^{bx}_y$ is non-degenerate (Cartan's theorem), we can go to a basis in which it is diagonal, and f^{abc} is antisymmetric (we raise and lower indices with g). Then $0 = f_{abc}h^c = f^{bad} f_{abc} h^c \propto h^d$. Thus only the components of h^c in $\mathcal{A}_2$ (the Abelian directions) can be non-zero. But then we can define the linear combination $R = -\sum_a h^*_a T^a / (\sum_b h^*_b h^b)$ to be the $U(1)$ generator with the desired commutation relations. Note that in theories with extended supersymmetries with N supersymmetry spinor charges, non-Abelian $U(N) = U(1) \times SU(N)$ R symmetries are allowed in 4 dimensions.

This single $U(1)$ under which Q_L has charge -1 is called the *R symmetry*. Despite this commutation relation, the R symmetry is *not* part of the supersymmetry algebra. In particular, a given supersymmetric theory may or may not have a conserved R symmetry, and if it does, it may or may not be spontaneously broken.

Since the R symmetry does not commute with supersymmetry, the component fields of a left-chiral superfield do not all carry the same R charge. If, for example, the lowest component ϕ has R charge $R(\phi) = r$, then

$$R(\phi) = r, \qquad R(\psi) = r - 1, \qquad R(F) = r - 2. \tag{1.212}$$

So, if we assign R charge $+1$ to the anticommuting superspace coordinate θ_L (and therefore -1 to θ_R) then the whole left-chiral superfield has the R charge of its lowest component. Since $R(d\theta_L) = -1$, it follows that the R charge of the superpotential

$$R(f) = +2. \tag{1.213}$$

Since this is the only symmetry under which the superpotential is charged, this is, in practice, the simplest way of finding an R symmetry of a given action.

Now suppose we have a theory with an R symmetry with charges $R(\Phi^n) = r_n$. Again, if it is not spontaneously broken, then generically supersymmetry is not either. If it is, we can choose $r_1 = 1$ and $\langle\phi_1\rangle \neq 0$, so that f can be written

$$f = (\Phi^1)^2\, f(U_2, \ldots, U_N), \text{ where } U_n \equiv \Phi^n/(\Phi^1)^{r_n}. \tag{1.214}$$

Then $\partial_n f = 0$ is equivalent to the set of equations $(\partial/\partial U_n)f = 0$ as well as $f = 0$. These are now N equations for $N-1$ unknowns, and so typically have no solution. Thus generically supersymmetry *is* broken in this situation. Our net result is that *if the superpotential is a generic function (constrained only by global symmetries) then supersymmetry is spontaneously broken if and only if there is a spontaneously broken R symmetry.* (We emphasize that the genericity assumption is crucial: it is easy to construct non-generic superpotentials with no R symmetry and broken supersymmetry, and with unbroken supersymmetry and a spontaneously broken R symmetry.)

This would seem to be bad news for supersymmetry phenomenology, for this result implies that along with supersymmetry breaking (which must occur since our vacuum is obviously not supersymmetric) necessarily goes a Goldstone boson (which is not observed) for the spontaneously broken $U(1)_R$ symmetry. This conclusion depends on the effective theory being generic, for there is a set of measure zero in "theory space" which evades this problem. But this way out gives rise to a naturalness problem: it seems unnatural for the effective theory to have exactly the required special couplings with no symmetry reason to enforce them. However, we will learn, when we discuss quantum corrections to supersymmetric actions in later lectures, that the superpotential is *not* generic in this way: there are natural (generic) supersymmetric gauge theories where the quantum fluctuations give rise only to the special "non-generic" superpotentials needed for supersymmetry breaking without an R symmetry. I hope to provide an example of this kind of supersymmetry breaking by the end of the course.

The interrelation of R symmetry and supersymmetry breaking can be illustrated more concretely by writing a specific form for the superpotential. A broad class of theories with an R symmetry have a superpotential which can be put in the general form

$$f = X^n f_n(\Phi^m), \tag{1.215}$$

(with summation over the index n implied) *i.e.* linear in the X^n, where X^n, $n = 1, \ldots, N$ and Φ^m, $m = 1, \ldots, M$ are left-chiral superfields, and f_n are general holomorphic functions. The R charge assignments are thus

$$R(X^n) = +2, \qquad R(\Phi^m) = 0. \tag{1.216}$$

The supersymmetric vacuum equations are

$$\partial_n f = 0 = f_n(\Phi), \qquad \text{and} \qquad \partial_m f = 0 = X^n \partial_m f_n(\Phi). \tag{1.217}$$

If we can find solutions (values of Φ^m) to the first N equations, then the last M equations can be satisfied by choosing $X^n = 0$, and both the R symmetry and supersymmetry are unbroken. But if $N > M$, then generically there are no solutions to the first N equations (because there are more equations than unknowns), so supersymmetry is broken in the vacuum. If we assume the theory has a quadratic Kahler potential (*i.e.* just the canonical kinetic terms and no more), the scalar potential is

$$V = \sum_n |f_n|^2 + \sum_m |X^n \partial_m f_n|^2 \tag{1.218}$$

which can always be minimized by choosing the Φ^m to minimize the first term. The vanishing of the second term then implies M linear constraints on the N X^n, so there is a whole $N-M$ complex dimensional space of (degenerate) vacua of this theory. Except at the special point $X^n = 0$, in the generic such vacuum one of the R charged fields X^n develops an expectation value, and so the $U(1)_R$ is indeed spontaneously broken.

The occurence of a manifold of vacua is characteristic of supersymmetric theories, and will be explored in examples in the next section. In this class of examples, we have a manifold of non-supersymmetric classical vacua. This kind of degeneracy is not stable under quantum corrections, which typically lift the degeneracy and pick out a single vacuum.[9] However, the occurence of a whole manifold of supersymmetric vacua (a complex submanifold in target space) which is stable under quantum corrections is characteristic of supersymmetric theories. We will begin to explore these manifolds of vacua in the next subsections, and in much greater detail later in the course.

[9] Or, perhaps, a single global symmetry orbit in target space.

1.6.3 Moduli space

By the *moduli space* of a theory we will mean its space of supersymmetric vacua. For a theory to have a non-trivial moduli space necessarily means that it has more than one vacuum. In regular (non-supersymmetric) field theories the usual examples of degenerate vacua occur due to broken symmetries, where the broken global symmetry generators relate all the vacua. Any further degeneracies are considered accidental since presumably quantum corrections or small irrelevant operators will lift the non-symmetry-enforced degeneracies. In supersymmetric theories, on the other hand, moduli spaces of degenerate vacua not related by any global symmetry frequently occur and are stable against generic changes of the couplings of the theory which respect the global symmetries. The existence of these moduli spaces of inequivalent vacua is due to the possibility of an R symmetry and the fact that the superpotential depends on the left-chiral superfields holomorphically. For since the superpotential has R charge +2, any field (or combination of fields) Φ with R charge +2 can only enter the superpotential linearly, $f \propto \Phi$, for other R symmetry preserving contributions, like $f \propto \Phi(\Phi\Phi^*)^N$, are not allowed by holomorphy. Vacua then exist for any value of $\langle\Phi\rangle$.

More generally, holomorphy of the superpotential promotes any global symmetry group G of the theory to a *complexified* symmetry group $G_{\mathbb{C}}$ of the superpotential. Suppose the theory has Hermitian global symmetry generators T^a satisfying the Lie algebra of G, $[T^a, T^b] = f^{ab}_c T^c$, and act on the fields by

$$\Phi \to \Phi' = \exp\{i\alpha_a \mathcal{T}^a\}\Phi \tag{1.219}$$

with real α_a, where we are thinking of Φ as an N-component vector of all the left-chiral superfields so that $\mathcal{T}^a$ is an $N \times N$ matrix representation of the lie algebra of G. Then the superpotential obeys

$$f(\Phi') = e^{i\alpha_a q^a} f(\Phi) \tag{1.220}$$

with $q^a = 2$ for the R symmetry generator and zero otherwise. But since f only depends on Φ and not Φ^*, f satisfies (1.220) with *complex* α_a as well. Thus the $G_{\mathbb{C}}$ invariance of the superpotential is found simply by allowing the α_a to be complex in (1.219). In the case of a $U(1)$ symmetry this means simply that $U(1)_{\mathbb{C}}$ invariance is invariance (or covariance) of the superpotential under general complex rescalings of the left-chiral superfields and not just under phase rotations. Since it is the extrema of the superpotential which govern the existence of supersymmetric vacua, this enhanced symmetry of the superpotential is what is responsible for the existence of continuous moduli of inequivalent vacua. We emphasize that the complexified symmetry $G_{\mathbb{C}}$ is a symmetry of the superpotential term only—the Kahler term, and thus the theory as a whole, is invariant only under G.

The extrema of the superpotential satisfy $\partial f(\phi)/\partial\phi^n = 0$ which are a set of holomorphic equations defining the moduli space as a *complex variety* (a complex submanifold with certains kinds of singularities allowed) in target space. The Kahler metric on the target space is pulled back to a Kahler metric on the moduli space. The existence of moduli spaces and the interpretation of their singularities turns out to be a very powerful tool for deriving non-perturbative information about the vacuum structure of supersymmetric field theories.

1.6.4 Examples

We will now illustrate the general properties of classical chiral superfield actions derived in the last three subsections through a series of simple examples. Unless we state otherwise, we will assume that the Kahler potential is quadratic (so we have only canonical kinetic terms). We will also restrict ourselves to superpotentials cubic in the fields so as to have renormalizable interactions.

$\triangleright \qquad f = \mu^2 \Phi$

We start with the simplest example with a single left-chiral superfield. This superpotential is generic given a $U(1)$ R-symmetry under which

$$R(\Phi) = +2. \tag{1.221}$$

There are no extrema of f, so supersymmetry is spontaneously broken in this model. The potential is

$$V = |\partial f|^2 = |\mu|^4, \tag{1.222}$$

showing that there is a whole space of degenerate, non-supersymmetric vacua in this model. Writing out the component action we see that this model describes a free theory of a massless complex scalar and a massive Majorana fermion. In fact, the vacuum degeneracy in this example is a fake since it is lifted by arbitrarily small deformations of the Kahler term. This is true in general of vacuum degeneracies (not related to spontaneously broken global symmetries) of non-supersymmetric vacua, and is the reason why we reserve the term moduli space for manifolds of supersymmetric vacua only. In this example, consider adding to the Kahler term a quartic piece

$$K = \Phi\Phi^* + \frac{C}{4M^2}\Phi^2\Phi^{*2}, \tag{1.223}$$

where C is some real constant and M some mass scale. (This is the type of term we would expect to get from quantum corrections if our theory were an effective description below an energy scale M; we will discuss effective actions in more detail in later lectures.) Then it is easy to compute the inverse Kahler metric on target space and the scalar potential to be

$$V = \frac{M^2|\mu|^4}{M^2 + C|\phi|^2}. \tag{1.224}$$

For $C < 0$ this has a unique minimum at $\langle\phi\rangle = 0$. For $C > 0$ there is no minimum, but the potential slopes off to zero as $|\phi| \to \infty$; of course, in this limit there is no reason to keep just $\mathcal{O}(\phi^4)$ terms in K.

▷ $f = \frac{1}{2}m\Phi^2 + \frac{1}{3}\lambda\Phi^3$

This is the most general renormalizable superpotential for one left-chiral superfield and is known as the *Wess-Zumino model.* Note that we have not written any linear term in Φ since it could always be absorbed by a shift in Φ. On the other hand, the symmetries of this theory become more obvious if we shift $\Phi \to \Phi - m/(2\lambda)$ to get

$$f = \tfrac{1}{3}\lambda\Phi^3 - \tfrac{1}{4}\tfrac{m^2}{\lambda}\Phi + \text{constant} \tag{1.225}$$

which (neglecting the constant) has a $\mathbb{Z}_4$ R symmetry generated by

$$\begin{aligned} \phi &\to -\phi \\ \theta_L &\to i\theta_L. \end{aligned} \tag{1.226}$$

Note that this model is not generic—all odd powers of Φ are allowed by this discrete R symmetry. The supersymmetric vacua determined by $\partial f = 0$ are

$$\phi = \pm\tfrac{1}{2}\tfrac{m}{\lambda}, \tag{1.227}$$

which, though distinct, have equivalent physics since they are related by the R symmetry. Adding higher odd powers of Φ to f will add more pairs of supersymmetric vacua, though now the physics in different pairs will not be related by any symmetry.

▷ $f = \frac{1}{2}\lambda\Phi_1\Phi_2^2$

This superpotential is the generic one with two left-chiral superfields, a $U(1)$ global symmetry (with charge Q), and an R symmetry, under which the fields are charged as

$$\begin{aligned} Q(\Phi_1) &= 1, \qquad & Q(\Phi_2) &= -\tfrac{1}{2}, \\ R(\Phi_1) &= 1, \qquad & R(\Phi_2) &= +\tfrac{1}{2}. \end{aligned} \tag{1.228}$$

Extrema of f are at

$$\Phi_2 = 0, \qquad \Phi_1 = \text{arbitrary}, \tag{1.229}$$

implying a whole moduli space, $\mathcal{M}$, of degenerate but inequivalent classical ground states. We can see that they are inequivalent because their physics is different: the spectrum at any such vacuum is one massless chiral multiplet Φ_1, and one massive chiral multiplet Φ_2 with mass $|\lambda\langle\Phi_1\rangle|$. Since the Kahler potential of this model is $K = \Phi_1\Phi_1^* + \Phi_2\Phi_2^*$, the metric induced on $\mathcal{M}$ is

$$ds^2 = d\Phi_1\, d\Phi_1^*. \tag{1.230}$$

Perturbations (quantum corrections) to the Kahler potential will certainly change this metric on $\mathcal{M}$, but cannot change the topological properties of $\mathcal{M}$ since those were determined by the (generic) superpotential.

(Note that we have fallen into a shorthand notation which uses the same symbol for the left-chiral superfield, its scalar components, and the vacuum expectation value of the scalar component.)

▷ $f = \frac{1}{2}m\Phi_1^2 + \lambda\Phi_1\Phi_2^2$

Adding a mass term to the previous example still leaves a generic renormalizable model with just the R symmetry

$$R(\Phi_1) = 1, \qquad R(\Phi_2) = \tfrac{1}{2}. \tag{1.231}$$

It has a single supersymmetric vacuum at $\Phi_1 = \Phi_2 = 0$ where Φ_1 has mass m and Φ_2 is massless.

Since Φ_1 is massive in the vacuum, at energies low compared to m it should be frozen at its expectation value—there is not enough energy to appreciably excite fluctuations in its field. Thus we can "integrate out" Φ_1 simply by solving its algebraic vacuum equation $\partial_{\Phi_1} f = 0$, and substituting back in f to get an "effective" (low energy) superpotential for Φ_2 alone:

$$f_{\text{eff}} = -\frac{1}{2}\frac{\lambda^2}{m}\Phi_2^4. \tag{1.232}$$

This result is consistent with the global symmetry of the original model; that it is a nonrenormalizable theory is not surprising given it effective status.

Note that there is nothing wrong in principle with leaving Φ_1 in; the only reason to integrate it out is that it plays no dynamical role at low enough energies.

▷ $f = \Phi_1\Phi_2\Phi_3$

This model is generic given a $U(1)_1 \times U(1)_2 \times U(1)_R$ symmetry under which

$$Q_1(\Phi_1) = +1, \qquad Q_1(\Phi_2) = +1, \qquad Q_1(\Phi_3) = -2,$$

$$\begin{array}{lll} Q_2(\Phi_1) = +1, & Q_2(\Phi_2) = -2, & Q_2(\Phi_3) = +1, \\ R(\Phi_1) = \ 0, & R(\Phi_2) = +1, & R(\Phi_3) = +1. \end{array} \tag{1.233}$$

Extrema of f are at

$$\Phi_1\Phi_2 = \Phi_2\Phi_3 = \Phi_3\Phi_1 = 0, \ \Rightarrow \ \{\Phi_1 = \Phi_2 = 0, \ \Phi_3 \text{ arbitrary; \& permutations}\}. \tag{1.234}$$

This example shows that the moduli space of vacua need not be a manifold, but may also have singularities (in this case an intersection point).

$\triangleright$ $\quad f = \mu^2\Phi_1 + m\Phi_2\Phi_3 + g\Phi_1\Phi_2^2$

This superpotential is generic given a $\mathbb{Z}_2$ symmetry (with charge Π) and a $U(1)_R$ symmetry under which

$$\begin{array}{lll} \Pi(\Phi_1) = +, & \Pi(\Phi_2) = -, & \Pi(\Phi_3) = - \\ R(\Phi_1) = 2, & R(\Phi_2) = 0, & R(\Phi_3) = 2. \end{array} \tag{1.235}$$

This is called the *O'Raifeartaigh model.* The extrema of f are at

$$\begin{aligned} 0 = \partial_1 f &= \mu^2 + g\phi_2^2, \\ 0 = \partial_2 f &= m\phi_3 + 2g\phi_1\phi_2, \\ 0 = \partial_3 f &= m\phi_2, \end{aligned} \tag{1.236}$$

which have no solution, implying supersymmetry is broken.

Problem 1.6.1 Compute the potential in the O'Raifeartaigh model and the spectrum of bosons and fermions in the ground states. Where is the vacuum once a $|\Phi_2|^4$ term is added to the Kahler term?

1.7 Vector superfields and superQED

Gauge fields appear in supersymmetric field theories in *vector superfields*. This lecture will focus on classical vector superfields and the effective actions describing their couplings to left-chiral superfields.

1.7.1 Abelian vector superfield

A vector superfield V is a general scalar superfield satisfying a reality condition:

$$\begin{aligned} V = V^* &= C - i(\overline{\theta}\gamma_5\omega) - \tfrac{i}{2}(\overline{\theta}\gamma_5\theta)M - \tfrac{1}{2}(\overline{\theta}\theta)N \\ &\quad + \tfrac{i}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)V_\mu - i(\overline{\theta}\gamma_5\theta)\left(\overline{\theta}\left[\lambda + \tfrac{1}{2}\partial\!\!\!/\omega\right]\right) \\ &\quad - \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2\left(D + \tfrac{1}{2}\Box C\right), \end{aligned} \tag{1.237}$$

where C, M, and N are real scalar fields, V_μ is a real vector field, and ω and λ are Majorana fields. The $\partial\!\!\!/\omega$ and $\Box C$ terms are separated from λ and D in this expansion to make the supersymmetry transformation rules of the components simpler:

$$\begin{aligned} \delta C &= i(\overline{\epsilon}\gamma_5\omega), \\ \delta M &= -(\overline{\epsilon}[\lambda + \partial\!\!\!/\omega]), \\ \delta M &= i(\overline{\epsilon}\gamma_5[\lambda + \partial\!\!\!/\omega]), \\ \delta V_\mu &= (\overline{\epsilon}[\gamma_\mu\lambda + \partial_\mu\omega]), \\ \delta\omega &= (-i\gamma_5\partial\!\!\!/C - M + i\gamma_5 N + V\!\!\!/)\,\epsilon, \\ \delta\lambda &= \left(\tfrac{1}{2}\gamma^\nu\gamma^\mu[\partial_\mu V_\nu - \partial_\nu V_\mu] + i\gamma_5 D\right)\epsilon, \\ \delta D &= i(\overline{\epsilon}\gamma_5\partial\!\!\!/\lambda). \end{aligned} \tag{1.238}$$

As advertised in section 1.5, the variation of the D component is a total derivative.

Since one component is a vector field, V_μ, we expect the interactions of this superfield to have a the usual $U(1)$ gauge invariance

$$V_\mu \rightarrow V_\mu + \partial_\mu\Lambda \tag{1.239}$$

where $\Lambda(x)$ is an arbitrary real scalar field. The only supersymmetry covariant generalization of this gauge invariance is

$$V \rightarrow V + \tfrac{i}{2}(\Omega - \Omega^*), \tag{1.240}$$

where now $\Omega(x,\theta)$ is an arbitrary left-chiral superfield with the usual component expansion

$$\Omega = \Lambda(x_+) - \sqrt{2}\overline{\theta}\mathcal{P}_+ w(x_+) + (\overline{\theta}\mathcal{P}_+\theta)\mathcal{W}(x_+), \tag{1.241}$$

where $x_+^\mu \equiv x^\mu + \frac{1}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)$, Λ and $\mathcal{W}$ are complex scalars and w is a Majorana spinor. Thus the set of *super gauge transformations* (1.240) is much larger that that of ordinary gauge transformations. In components, the effect of a super gauge transformation is to shift the component fields of the vector superfield as

$$\begin{aligned} \delta_\Omega C &= -\mathrm{Im}\Lambda, \\ \delta_\Omega M &= -\mathrm{Re}\mathcal{W}, \\ \delta_\Omega N &= +\mathrm{Im}\mathcal{W}, \\ \delta_\Omega V_\mu &= +\mathrm{Re}\partial_\mu\Lambda, \\ \delta_\Omega \omega &= \tfrac{1}{2}\sqrt{2}w, \\ \delta_\Omega \lambda &= 0, \\ \delta_\Omega D &= 0, \end{aligned} \tag{1.242}$$

where we denote the super gauge transformation by δ_Ω to differentiate it from supersymmetry transformations. We see that this correctly transforms V_μ with gauge parameter $\mathrm{Re}\Lambda$ for a $U(1)$ gauge field. C, M, N, and ω are gauge artifacts since they can be gauge away entirely by an appropriate choice of $\mathrm{Im}\Lambda$, $\mathcal{W}$, and w. λ and D, on the other hand, are gauge invariant. Indeed, we can partially gauge fix to the *Wess-Zumino gauge*

$$C = M = N = \omega = 0 \tag{1.243}$$

at the cost of losing manifest supersymmetry. The Wess-Zumino gauge does not completely fix the gauge—in fact it fixes all of the gauge freedom in (1.240) except for the ordinary $U(1)$ gauge transformations of the vector field V_μ.

A slight relaxation of the Wess-Zumino gauge, which we'll call the *complex gauge* for want of a better name, is to only set

$$M = N = \omega = 0. \tag{1.244}$$

Thus in this gauge there is left unfixed one *complex* scalar field Λ worth of gauge invariance: instead of the usual $U(1)$ group of gauge transformations, in this gauge the vector superfield is covariant under the complexified group $U(1)_{\mathbb{C}}$. This will be a useful observation once we couple to chiral multiplet matter below.

We now want to write down a gauge invariant and supersymmetry invariant action for a single (*i.e.* Abelian) vector superfield. Because the D component of V is gauge invariant by (1.242), an obvious choice is

$$S_{FI} = \int d^4x\, \xi[V]_D = \int d^4x\, \xi D, \tag{1.245}$$

which is known as the *Fayet-Iliopoulos term*. Here ξ is a real constant of dimension of mass-squared, and in the last inequality I have dropped a total derivative. It contains no dynamics, so we clearly need to look further afield. D terms made from polynomials in V won't do the job since, though supersymmetry invariant, they are no longer gauge invariant.

So we look at covariant derivatives. Note that since Ω is a left-chiral superfield, $\mathcal{D}_R\Omega = \mathcal{D}_L\Omega^* = 0$. Also, recall that $\mathcal{D}_{L,R}$ satisfy the supersymmetry algebra with the wrong sign:

$$\{\mathcal{D}_R, \mathcal{D}_R\} = 0, \qquad \{\mathcal{D}_R, \mathcal{D}_L\} = -2\mathcal{P}_- \partial\!\!\!/ \mathcal{C}. \tag{1.246}$$

Thus $\mathcal{D}_L V$ transforms under gauge transformations as $\mathcal{D}_L V \to \mathcal{D}_L V + \frac{i}{2}\mathcal{D}_L\Omega$. To get rid of the dependence on the left-chiral superfield Ω, we should further act with $\mathcal{D}_R$ which annihilates it. But, because of the non-vanishing commutator, $\mathcal{D}_R\mathcal{D}_L\Omega = -2\mathcal{P}_- \partial\!\!\!/ \mathcal{C}\Omega$, so we must act a second time with $\mathcal{D}_R$ so that $\mathcal{D}_R\mathcal{D}_R\mathcal{D}_L\Omega = 0$. Thus we learn that $\mathcal{D}_R\mathcal{D}_R\mathcal{D}_L V$ is gauge invariant. Since $\mathcal{D}_R$ anticommutes with itself, the two $\mathcal{D}_R$'s must appear antisymmetrically on their Weyl spinor indices. But the antisymmetric combination of two two-dimensional spinors representations, $\mathbf{2} \otimes_A \mathbf{2} = \mathbf{1}$, is the singlet representation. Thus (choosing a convenient normalization) the gauge invariant superfield made out of V and superderivatives is

$$W_{La} \equiv \frac{i}{4}(\mathcal{D}_R^T \mathcal{C} \mathcal{D}_R)\mathcal{D}_{La} V. \tag{1.247}$$

This is called the *field strength chiral superfield*. Note that, as indicated by the spinor index a, this superfield is a left-chiral Weyl spinor. It is also a left-chiral superfield since $\mathcal{D}_R W_L \propto \mathcal{D}_R\mathcal{D}_R\mathcal{D}_R X = 0$ because no totally antisymmetric combination of three two-dimensional representations is possible. It is called the field-strength superfield because of its component expansion, which is

$$W_L = \lambda_L(x_+) + \tfrac{1}{2}\gamma^\mu\gamma^\nu\theta_L f_{\mu\nu}(x_+) - i\theta_L D(x_+) - (\theta_L^T\mathcal{C}\theta_L)\,\partial\!\!\!/ \lambda_R(x_+), \tag{1.248}$$

where

$$f_{\mu\nu} \equiv \partial_\mu V_\nu - \partial_\nu V_\mu \tag{1.249}$$

is the gauge field strength, and $x_+^\mu \equiv x^\mu + \frac{1}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)$.

Using the gamma matrix identity

$$\gamma^\mu\gamma^\nu\gamma_5 = -\frac{i}{2}\epsilon^{\mu\nu\rho\sigma}\gamma_\rho\gamma_\sigma, \tag{1.250}$$

it follows that only the self dual part of the $f_{\mu\nu}$ antisymmetric tensor appears in W_L, for

$$2\gamma^\mu\gamma^\nu\theta_L f_{\mu\nu} = (\gamma^\mu\gamma^\nu + \gamma^\mu\gamma^\nu\gamma_5)\theta f_{\mu\nu} = \gamma^\mu\gamma^\nu\theta(f_{\mu\nu} - \tfrac{i}{2}\epsilon_{\mu\nu\rho\sigma}f^{\rho\sigma}). \tag{1.251}$$

Define the *dual field strength* $\widetilde{f}$ by

$$\widetilde{f}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}f_{\rho\sigma}, \tag{1.252}$$

which differs from the Hodge dual defined in section 1.3.1 by a factor of i, so that $\widetilde{\widetilde{f}} = -f$. Then with our definitions, the self dual part of the field strength is $f^+ = f - i\widetilde{f}$, and useful identities are

$$f_{\mu\nu}f^{\mu\nu} = -\widetilde{f}_{\mu\nu}\widetilde{f}^{\mu\nu}, \qquad \tfrac{1}{2}f^+_{\mu\nu}f^{+\mu\nu} = f_{\mu\nu}f^{\mu\nu} - i f_{\mu\nu}\widetilde{f}^{\mu\nu}. \tag{1.253}$$

A right-chiral version, W_R, of the field strength chiral superfield is defined by

$$W_R \equiv -\frac{i}{4}(\mathcal{D}_L^T\mathcal{C}\mathcal{D}_L)\mathcal{D}_R V. \tag{1.254}$$

W_R depends on the field strength in its anti-self dual form $f^- = f + i\widetilde{f}$. W_L and W_R are clearly not simply general chiral superfields; indeed, because of the identity

$$\mathcal{C}_{ab}\mathcal{D}_{La}(\mathcal{D}_R^T\mathcal{C}\mathcal{D}_R)\mathcal{D}_{Lb} = \mathcal{C}_{ab}\mathcal{D}_{Ra}(\mathcal{D}_L^T\mathcal{C}\mathcal{D}_L)\mathcal{D}_{Rb} \tag{1.255}$$

which follows from the anticommutation relations of the superderivatives, W_L and W_R satisfy the constraint

$$(\mathcal{D}_L^T\mathcal{C}W_L) = -(\mathcal{D}_R^T\mathcal{C}W_R). \tag{1.256}$$

This is the superspace version of the Bianchi identity $\epsilon^{\mu\nu\rho\sigma}\partial_\rho f_{\mu\nu} = 0$, as can be checked by going to components.

A gauge, Lorentz, and supersymmetry invariant action term is then

$$S_\tau = \int d^4x\, \frac{\tau}{16\pi i}[W_L^T\mathcal{C}W_L]_F + \text{c.c.} \tag{1.257}$$

where

$$\tau \equiv \frac{\vartheta}{2\pi} + i\frac{4\pi}{g^2} \tag{1.258}$$

is a dimensionless complex constant. We can see the interpretation of g and ϑ by expanding S_τ in components:

$$S_\tau = \int d^4x \left\{ -\frac{1}{2g^2}\overline{\lambda}\partial\!\!\!/\lambda - \frac{1}{4g^2}f_{\mu\nu}f^{\mu\nu} + \frac{\vartheta}{32\pi^2}f_{\mu\nu}\widetilde{f}^{\mu\nu} + \frac{1}{2g^2}D^2 \right\}. \tag{1.259}$$

This shows that in supersymmetric actions V_μ will describe a propagating $U(1)$ gauge boson $f_{\mu\nu}$ (*i.e.* a photon), a massless Majorana fermion λ (the "photino"), and an auxiliary field D. This describes the particle content of the massless vector multiplet. Note that the kinetic terms for λ and $f_{\mu\nu}$ are not canonically normalized—a factor of the constant g has to be absorbed in each. Later, when we discuss the coupling to left-chiral superfields, we will see that g has the interpretation as the gauge coupling constant. ϑ is the *theta angle*, which will play an important role quantum mechanically. As its name suggests, and as we will discuss later, the theta angle is periodic

$$\vartheta = \vartheta + 2\pi, \tag{1.260}$$

implying $\tau = \tau + 1$; it was this simple form of the periodicity relation which prompted our choice of normalization of τ.

The gauge kinetic terms can also be written as an integral over all of superspace (*i.e.* as a D term) using

$$\int d^2\theta_L \, W^2 \sim \int d^2\theta_L \mathcal{D}_R^2 (W\mathcal{D}_L V) \sim \int d^4\theta \, (W\mathcal{D}_L V). \tag{1.261}$$

Note, however, that in this form the integrand is gauge variant. In dealing with effective actions it will be important to have gauge invariant expressions for Lagrangians. The Fayet-Iliopoulos term was written with a gauge variant Lagrangian in (1.245). It can be written in a manifestly gauge invariant way as an integral over a *quarter* of superspace using

$$[V]_D \sim \int d^4\theta V \sim \int d\theta_R^2 d\theta_L^a \mathcal{D}_{La} V \sim \int d\theta_L^a \mathcal{D}_R^2 D_{La} V \sim \int d\theta_L^a W_{La}. \tag{1.262}$$

Thus the FI term can be written as

$$S_{FI} = i\xi \int d^4x \, d\theta_L^a \, W_{La} + \text{c.c.}. \tag{1.263}$$

1.7.2 Coupling to left-chiral superfields: superQED

Recall that a matter field (scalar or spinor) of charge q transforms under an ordinary gauge transformation $V_\mu \to V_m + \partial_\mu \Lambda$ of the gauge field as $\phi \to \exp\{iq\Lambda\}\phi$. The generalization of this to super gauge transformations has the vector superfield and left-chiral superfields Φ^n transforming as

$$\begin{aligned} V &\to V + \tfrac{i}{2}(\Omega - \Omega^*), \\ \Phi^n &\to e^{iq_n\Omega}\Phi^i, \end{aligned} \tag{1.264}$$

where q_n is the charge of the Φ^n left-chiral superfield. In complex gauge where only the (complex) lowest component of Ω is non-zero

$$\Omega(x, \theta) = \Lambda(x), \qquad \text{(complex gauge)} \tag{1.265}$$

this reduces to the usual gauge transformation property on the component fields, all with the same charge q_n, except with a complex gauge transformation parameter. The fact that all the components transform with the same gauge charge means that the gauge symmetry commutes with supersymmetry: the zero-coupling limit of a gauge symmetry reduces to an ordinary global symmetry, not an R symmetry. (It turns out that one is forced to include gravity to write an interacting supersymmetric theory with a gauged R symmetry.)

It is easy to write super gauge and supersymmetry invariant action terms for the left-chiral superfields. For the superpotential terms, gauge invariance simply demands that the superpotential itself be gauge invariant:

$$f(e^{iq_n\Omega}\Phi^n) = f(\Phi^n), \tag{1.266}$$

i.e. that each term in the superpotential have total gauge charge zero. Thus the superpotential action terms

$$S_f = \int d^4x\,[f(\Phi^n)]_F + \text{c.c.} \tag{1.267}$$

remain the same. Since Ω is complex in complex gauge, just as with global symmetries, the superpotential is invariant not just under the $U(1)$ gauge group, but also under its complexification $U(1)_{\mathbb{C}}$.

The Kahler terms need some modification since the basic kinetic terms coming from $K \sim \Phi^*_n\Phi^n$ are not super gauge invariant:

$$\Phi^*_n\Phi^n \to \Phi^*_n e^{-iq_n\Omega^*}e^{iq_n\Omega}\Phi^n. \tag{1.268}$$

This can be fixed by including a factor of e^{-2q_nV} in K for each right-chiral superfield Φ^*_n:

$$S_K = \frac{1}{2}\int d^4x\,\left[\Phi^*_n e^{-2q_nV}\Phi^n\right]_D. \tag{1.269}$$

This can be expanded in components by noting that in complex gauge

$$\begin{aligned} V - C &= -\tfrac{i}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)V_\mu - i(\overline{\theta}\gamma_5\theta)(\overline{\theta}\lambda) - \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2(D + \tfrac{1}{2}\Box C), \\ (V-C)^2 &= -\tfrac{1}{4}(\overline{\theta}\gamma_5\gamma^\mu\theta)(\overline{\theta}\gamma_5\gamma^\nu\theta)V_\mu V_\nu = \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2 V_\mu V^\mu, \\ (V-C)^3 &= 0, \end{aligned} \tag{1.270}$$

so

$$e^V = e^C\left(1 + V - C - \tfrac{1}{8}(\overline{\theta}\gamma_5\theta)^2 V_\mu V^\mu\right). \tag{1.271}$$

Expanding just the Kahler term gives in complex gauge

$$\begin{aligned} S_K = \int d^4x\,e^{-2q_nC}\Big\{ &-(D_\mu\phi^n)^*(D^\mu\phi^n) - \tfrac{1}{2}(\overline{\psi}_n\,\not{D}\psi^n) + F^*_nF^n - q_n\phi^*_n\phi^n(D + \tfrac{1}{2}\Box C) \\ &- 2\sqrt{2}\mathrm{Im}q_n\phi^n(\overline{\psi}_n\mathcal{P}_+\lambda) + 2\sqrt{2}\mathrm{Im}q_n\phi^*_n(\overline{\psi}^n\mathcal{P}_-\lambda)\Big\}, \end{aligned} \tag{1.272}$$

where a sum over n is understood and

$$D_\mu \equiv \partial_\mu - iq_n V_\mu \tag{1.273}$$

is the gauge covariant derivative. Part of the $U(1)_{\mathbb{C}}$ gauge invariance consists of rescalings of the chiral superfields, allowing in particular the factor of e^{-2q_nC} to be absorbed in ϕ, ψ, and F. Also, the gauge coupling g in (1.259) can be absorbed in the gauge field by rescaling $V_\mu \to gV_\mu$ so that the gauge kinetic term has the canonical form $\frac{1}{4}f_{\mu\nu}f^{\mu\nu}$; then the gauge covariant derivatives in (1.272) become $D_\mu = \partial_\mu - igq_nV_\mu$ showing that g is indeed the gauge coupling.

Since D and the F^n appear in the total *supersymmetric QED* action

$$S_{sQED} = S_\tau + S_K + S_f + S_{FI} \tag{1.274}$$

only quadratically, we can replace them by their equations of motion. This gives in complex gauge the scalar potential

$$V(\phi_n^*, \phi^n) = \sum_n e^{2q_nC}|\partial_n f|^2 + \tfrac{1}{2}g^2\left(\xi + \sum_n q_n e^{-q_nC}|\phi^n|^2\right)^2 . \tag{1.275}$$

This immediately implies that supersymmetry is unbroken if and only if the F^n and D fields vanish:

$$\begin{aligned} 0 \;=\; F_n^* \;&=\; -\frac{\partial f}{\partial \phi^n}, \qquad \forall n, \\ 0 \;=\; D \;&=\; -\frac{g^2}{2}\left(\xi + \sum_n q_n e^{-q_nC}|\phi^n|^2\right) . \end{aligned} \tag{1.276}$$

I will refer to these equations as the *vacuum equations*. One usually only sees the vacuum equations written in Wess-Zumino gauge, where $C = 0$. Our more general expression in complex gauge will be useful when we come to solving the D field vacuum equation.

Although the above theory is a consistent classical field theory, quantum mechanically there is another constraint on the couplings coming from *anomalies*. As we will discuss in a later lecture, this theory is inconsistent unless the charges, q_n, of the left-chiral superfields satisfy

$$\sum_n q_n = \sum_n q_n^3 = 0. \tag{1.277}$$

The first constraint is from the mixed gauge-gravitational anomaly, and the second is from the pure gauge anomaly. These constaints can always be satisfied by pairs of left-chiral superfields with opposite charges. Non-trivial solutions exist for five or more charged left-chiral superfields (in fact there is a continuum of solutions). Non-trivial solutions with commensurate charges are harder to find. One such has fifteen chiral fields, corresponding to the hypercharge assignments of one generation of the standard model.

1.7.3 General Abelian gauged sigma model

We can write the most general Abelian gauged sigma model as

$$\begin{aligned} S &= \frac{1}{2}\int d^4x \Big\{ \left[K\left(\Phi^*_n e^{-2q_{An}V_A}, \Phi^n\right) + 2\xi_A V_A \right]_D \\ &\quad + \left[f(\Phi^n) + \frac{1}{16\pi i}\tau_{AB}(\Phi^n)\left(W^T_{LA}\mathcal{C}W_{LB}\right) + \text{c.c.} \right]_F \Big\} \end{aligned} \tag{1.278}$$

where the A, B indices run over different $U(1)$ vector multiplets, q_{An} is the charge of the nth left-chiral superfield under the Ath gauge factor, ξ_A is a Fayet-Iliopoulos constant for each $U(1)$, and $\tau_{AB}(\Phi)$ are generalized gauge couplings and theta angles,

$$\tau_{AB} = \frac{\vartheta_{AB}}{2\pi} + i\frac{4\pi}{(g^2)_{AB}}, \tag{1.279}$$

which can depend only holomorphically on left-chiral superfields since it appears in an F-component. The anomaly cancellation conditions require the charges to satisfy

$$\sum_n q_{An} = \sum_n q_{An}q_{Bn}q_{Cn} = 0, \qquad \text{for all } A, B, C. \tag{1.280}$$

The component expansion of this action gives rise to a gauged version of the nonlinear sigma model (1.177) and (1.182) discussed in section 1.5; the details are left to the reader's imagination except for the resulting scalar potential[10]

$$\begin{aligned} V &= \sum_n e^{2q_{An}C_A}\left|\partial_n f\right|^2 \\ &\quad + \tfrac{1}{2}\sum_{AB}(g^2)_{AB}\left(\xi_A + \sum_n q_{An}e^{-\sum_D q_{Dn}C_D}|\phi^n|^2\right) \\ &\qquad \times \left(\xi_B + \sum_n q_{Bn}e^{-\sum_D q_{Dn}C_D}|\phi^n|^2\right). \end{aligned} \tag{1.281}$$

[10]The left-chiral superfield indices n should really be raised and lowered with the Kahler metric $g_{n\overline{m}}$ and the $|\phi^n|^2$ factors only take that form in a special coordinate system on target space.

Unitarity of the gauge kinetic terms requires the symmetric coupling matrix $(g^2)_{AB}$ to be positive definite, implying the vacuum energy vanishes and supersymmetry is unbroken if and only if the vacuum equations

$$\begin{aligned} 0 &= \partial_n f, && \forall n \quad \text{(F equations)} \\ 0 &= \xi_A + \sum_n q_{An} e^{\sum_B q_{Bn} C_B} |\phi^n|^2, && \forall A \quad \text{(D equations)} \end{aligned} \tag{1.282}$$

are satisfied. Here C_A is the lowest component of the V_A vector superfield. It is gauge-variant; in Wess-Zumino gauge we set $C_A = 0$, giving rise to the "usual" vacuum equations. We will spend the rest of this section discussing the solutions of these equations and the physics they describe.

1.7.4 Higgsing and unitary gauge

When a charged scalar gets a non-zero vacuum expectation value, the gauge symmetry is spontaneously broken, and, by the Higgs mechanism, gauge bosons become massive, "eating" neutral scalars. Precisely the same mechanism works in supersymmetric theories: when a charged left-chiral superfield gets a non-zero vacuum expectation value, the gauge symmetry is spontaneously broken and the gauge bosons in the vector multiplet become massive. If supersymmetry is not also spontaneously broken, then the whole vector multiplet must become massive along with the vector boson.

We can see this in a simple model with two charged left-chiral superfields, $\Phi_\pm$ with charges ± 1, no Fayet-Iliopoulos term, and the superpotential

$$f = -m\Phi_+\Phi_- + \frac{1}{2m}\Phi_+^2\Phi_-^2. \tag{1.283}$$

Then the F equations for a supersymmetric minimum are satisfied by either $\phi_+ = \phi_- = 0$ or $\phi_+\phi_- = m^2$. The first solution does not interest us since it does not break the gauge symmetry. The D equation in Wess-Zumino gauge,

$$|\phi_+|^2 - |\phi_-|^2 = 0, \tag{1.284}$$

implies that $\phi_+ = e^{i\alpha}\phi_-$ with an arbitrary angle α, thus giving the supersymmetric but not gauge invariant vacua

$$\phi_+ = me^{+i\alpha/2}, \qquad \phi_- = me^{-i\alpha/2}. \tag{1.285}$$

This one parameter space of vacua are all really just one vacuum since they are related by the gauge transformation

$$\phi_+ \to \phi_+ e^{+i\beta}, \qquad \phi_- \to \phi_- e^{-i\beta}. \tag{1.286}$$

So, by choosing $\beta = -\alpha/2$, we can gauge transform any of the vacua in (1.285) to the single vacuum

$$\phi_+ = \phi_- = m. \tag{1.287}$$

However, it is hard to see the physical field content in Wess-Zumino gauge. So instead, let us go to *unitary gauge*, in which we fix all of the super gauge invariance by rotating the *whole* left-chiral superfield so that

$$\Phi_+ = m. \tag{1.288}$$

We are free to do this so long as $\langle\Phi_+\rangle \neq 0$. We have chosen m as a convenient value—we could just as well have chosen any non-zero complex number. In physical gauge the action is

$$S = \int d^4x \left\{ \frac{1}{2}\left[e^{-2V}m^2 + \Phi_-^* e^{2V}\Phi_-\right]_D + \left[\frac{\tau}{16\pi i}(W_L)^2 - \frac{1}{2}m^2\Phi_- + \frac{m}{4}\Phi_-^2 + \text{c.c.}\right]_F \right\}, \tag{1.289}$$

where I am writing $(W_L)^2$ as a shorthand for $(W_L^T\mathcal{C}W_L)$. The solution to the F equation is $\phi_- = m$. Expanding about this vacuum, $\Phi_- = m + \delta\Phi_-$, generates the term (using the full component expansion (1.237) of V since we have used up all our gauge freedom)

$$\begin{aligned}\frac{1}{2}m^2\int d^4x\left[e^{2V} + e^{-2V}\right]_D \;\supset\; 2m^2\int d^4x\left[V^2\right]_D \;=\; -2m^2\int d^4x(V_\mu V^\mu + \partial_\mu C\partial^\mu C \\ - \overline{\omega}\partial\!\!\!/\omega - 2CD - M^2 - N^2 - 2\overline{\omega}\lambda)\end{aligned} \tag{1.290}$$

giving the vector boson a mass and making C and ω dynamical. Thus we have as propagating degrees of freedom in the vector superfield a massive scalar C, two massive fermions ω and λ, an a massive vector V_μ, which is the content of the massive vector multiplet. Of the original two left-chiral superfields, one was gauged away ("eaten") and the other remains propagating, and plays the role of the Higgs boson in this example.

1.7.5 Supersymmetry breaking and Fayet-Iliopoulos terms

We saw in section 1.7.2 that the scalar potential was the sum of squares of the D as well as the F terms. We thus have an extra condition to satisfy for there to be a supersymmetric vacuum compared to the case without vector superfields. If the F equations cannot be satisfied by themselves, then, just as in the no vector superfield case, supersymmetry will be spontaneously broken. This kind of breaking is called "F term" or sometimes "O'Raifeartaigh" breaking, and its systematics are just as we discussed in section 1.6.2. It will turn out that if the F term conditions have a solution, then the D term conditions will always also have a solution *if there are no Fayet-Iliopoulos terms.* Thus the Fayet-Iliopoulos terms play a special role in the discussion of supersymmetry breaking. Breaking due to them is called "D term" or sometimes "Fayet-Iliopoulos" breaking. We will see that Fayet-Iliopoulos terms are only allowed for Abelian gauge groups and so only play a role when there are $U(1)$ factors in the gauge gauge group.

In D term breaking, the non-zero vacuum expectation value of a D component is the order parameter for supersymmetry breaking. In particular, the scale of supersymmetry breaking, or equivalently the scale of the mass splittings within multiplets, is given by

$$\delta m^2 \sim g^2 D. \tag{1.291}$$

The factor of the gauge coupling enters since the D term is coupled to the other fields in the theory as part of the gauge multiplet. Furthermore, since the D term breaking only occurs when there is a Fayet-Iliopoulos term, one expects the D vacuum expectation value to be proportional to ξ, the Fayet-Iliopoulos constant. Note that since the supersymmetry variation of the gaugino is $\delta\lambda = \frac{1}{2}\gamma^\nu\gamma^\mu f_{\mu\nu} + i\gamma_5 D$, then in the case of D term breaking the gaugino is shifted, and so is identified with the Goldstino.

We will now determine the conditions under which D term breaking can generically occur. For simplicity, consider the case of a single $U(1)$ vector superfield and N left-chiral superfields, so the vacuum equations are given by (1.276). If the superpotential were generic, then the F equations give N complex analytic equations for N complex unknowns, and so would typically have a solution. However, the superpotential is subject to one constraint—gauge invariance—which may reduce the number of independent equations by one. If the gauge symmetry is not broken (so no charged left-chiral superfields get vacuum expectation values), then the D equation will be satisfied if and only if $\xi = 0$. Thus the Fayet-Iliopoulos term *generically* leads to broken supersymmetry when the F equations do not break the gauge invariance by themselves.

When the solution to the F equations breaks the gauge symmetry, then, as we saw in section 1.6.2, the F equations are equivalent to $N-1$ equations for the $N-1$ unknowns $u_n = \phi_n/\phi_1^{q_n}$, and thus will typically have a one complex dimensional space of solutions. However one real dimension of this space is a gauge artifact: a common phase of the ϕ_n is unobservable by gauge invariance. There remains a one real dimensional space of solutions, plus the single real D equation. But since this last equation is a *real* equation, one can not predict the generic existence of solutions. We will show below the general result that *in the absence of Fayet-Iliopoulos terms and when the F equations have a solution, then there always exists a simultaneous solution to the D vacuum equation.* Thus, in the presence of gauge symmetry breaking, a generic superpotential with no Fayet-Iliopoulos term will lead to a unique supersymmetric vacuum.

Finally, if there is a spontaneously broken $U(1)_R$ symmetry, then, just as in the section 1.6.2, supersymmetry will be generically spontaneously broken—the D equation just adds an additional constraint to the F equations.

Examples

A simple example illustrating D term breaking is the *Fayet-Iliopoulos model*, a theory with two charged left-chiral superfields $\Phi_\pm$ of charges ± 1 with a Fayet-Iliopoulos term ξ and a superpotential

$$f = -m\Phi_+\Phi_-. \tag{1.292}$$

The F equation implies that $\phi_+ = \phi_- = 0$ in the vacuum, but then the D equation (in Wess-Zumino gauge) $|\phi_+|^2 - |\phi_-|^2 + \xi = 0$ has no solution if $\xi \neq 0$.

If we add a term to the superpotential of this model so that it becomes the model of section 1.7.4,

$$f = -m\Phi_+\Phi_- + \frac{1}{2m}\Phi_+^2\Phi_-^2, \tag{1.293}$$

then the F equations have the two solutions $\phi_+ = \phi_- = 0$ and $\phi_- = m^2/\phi_+$. The first solution does not Higgs the gauge symmetry, and cannot satisfy the D equation for $\xi \neq 0$. The second solution, on the other hand gives the D equation (in Wess-Zumino gauge) $|\phi_+|^2 - |m^4/\phi_+|^2 + \xi = 0$ which has a solution

$$2|\phi_+|^2 = \sqrt{\xi^2 + 4|m|^4} - \xi \tag{1.294}$$

for all ξ. Thus this is a supersymmetric vacuum with non-zero ξ and gauge symmetry breaking.

A third example is a theory of $\Phi_\pm$ with superpotential

$$f = -\frac{1}{2m}\Phi_+^2\Phi_-^2. \tag{1.295}$$

The F equations have a moduli space of solutions with two components (intersecting at $\phi_\pm = 0$) given by

$$\begin{aligned} 0 &= \phi_+ \quad \text{and} \quad \phi_- = \text{arbitrary}; \quad \text{or} \\ 0 &= \phi_- \quad \text{and} \quad \phi_+ = \text{arbitrary}. \end{aligned} \tag{1.296}$$

Call the solutions described by the first line the "ϕ_- branch" and those of the second line the "ϕ_+ branch". Then there is a solution to the D equation on the ϕ_- branch only if $\xi > 0$, and on the ϕ_+ branch only if $\xi < 0$.

1.7.6 Solving the D equations

In what follows we will set to zero the Fayet-Iliopoulos terms and the superpotential, thus $\xi_A = f = 0$. We will see that the resulting D equations always have flat directions—whole moduli spaces of solutions. We start by revisiting a simple example.

Example

Consider a $U(1)$ theory with two left-chiral superfields $\Phi_\pm$ of charges ± 1. Then the D equation is in Wess-Zumino gauge

$$0 = |\phi_+|^2 - |\phi_-|^2, \tag{1.297}$$

implying

$$\phi_+ = e^{i\alpha}\phi_- \tag{1.298}$$

for some angle α. The resulting three real dimensional space of vacua $\{\phi_+, \alpha\}$ must be divided by the $U(1)_{\mathbb{R}}$ gauge equivalence which remains in Wess-Zumino gauge:

$$\phi_+ \simeq e^{+i\beta}\phi_+, \qquad \phi_- \simeq e^{-i\beta}\phi_-, \tag{1.299}$$

for β any real angle. We can use this gauge freedom to fix the angle α by choosing $\beta = \alpha/2$ so that $\phi_+ = \phi_-$. Actually, this choice does not completely fix the gauge freedom, since $\beta = \pi + \alpha/2$ would have done just as well. Thus the moduli space can be described as

$$\mathcal{M} = \{\phi_+\}/\{\phi_+ \to -\phi_+\}, \tag{1.300}$$

which means the space of all ϕ_+ quotient the (residual gauge) identification of ϕ_+ with $-\phi_+$.

This space can be conveniently parametrized in terms of the gauge invariant variable

$$M \equiv \phi_+\phi_-. \tag{1.301}$$

M is a good coordinate on $\mathcal{M}$ since every ϕ_+ gives rise to a unique M (since $\phi_+ = \phi_-$), while every value of M determines a ϕ_+ up to a sign. Thus

$$\mathcal{M} = \{M\}. \tag{1.302}$$

So, topologically, the moduli space $\mathcal{M} \simeq \mathbb{C}$.

Metrically, however, it has a singularity. We can compute the induced metric on $\mathcal{M}$ by evaluating the the Kahler potential there:

$$K = \phi_+^*\phi_+ + \phi_-^*\phi_- = 2\phi_+^*\phi_+ = 2\sqrt{MM^*}, \tag{1.303}$$

where in the second equality we have used the Wess-Zumino gauge D equation, and in the third equality the fact that up to a gauge transformation $M = \phi_+\phi_- = \phi_+^2$. So the induced metric is

$$ds^2 = \frac{1}{2}\frac{dM\,dM^*}{\sqrt{MM^*}}. \tag{1.304}$$

Thus there is a metric singularity at $M = 0$, which corresponds to $\phi_\pm = 0$, where the $U(1)$ gauge symmetry is not spontaneously broken, and so the vector superfield is massless there. It is a general rule that singularities in moduli space correspond to vacua with "extra" massless particles. The metric ds^2 is flat everywhere except at the origin, where it has a $\mathbb{Z}_2$ conical singularity (a deficit angle of π corresponding to a delta function curvature). We can see this by going to polar coordinates $M = re^{i\theta}$ where $ds^2 \sim (1/r)dr^2 + rd\theta^2$. Changing radial variables to $u = 2\sqrt{r}$ then gives $ds^2 = du^2 + \frac{1}{4}u^2 d\theta^2$. Thus metrically (as opposed to topologically) the moduli space is the *orbifold* space

$$\mathcal{M} = \mathbb{C}/\mathbb{Z}_2. \tag{1.305}$$

Note that this is only a classical equivalence. Quantum mechanically the Kahler potential gets corrections and changes the metric structure of the moduli space, sharpening the conical singularity into a cusp-like singularity. We will return to the issue of quantum corrections later.

D equations in complex gauge

More generally, the moduli space of a theory with no superpotential is given by the space of all scalar vacuum expectation values satisfying the D equations, modulo gauge equivalences:

$$\mathcal{M} = \{\phi^n | D_A = 0\}/G. \tag{1.306}$$

I claim this space is equivalent to

$$\mathcal{M} = \{\phi^n\}/G_{\mathbb{C}}, \tag{1.307}$$

the space of all scalar vacuum expectation values of the left-chiral superfields modulo *complexified* gauge transformations. Thus we can think of the D equations as just a reflection of the larger $G_{\mathbb{C}}$ gauge invariance that we have seen necessarily appears in a supersymmetric gauge theory with gauge group G. Note that this description also makes it manifest that the moduli space $\mathcal{M}$ is a complex manifold.

We can see this explicitly in the Abelian gauged sigma model (with a single gauge field) as follows. Recall that in complex gauge, in addition to the scalar fields from the left-chiral superfields the real scalar component C of the vector superfield also appears in the potential, giving rise to the D equation

$$0 = \sum_n q_n e^{q_n C} |\phi^n|^2 \tag{1.308}$$

where we are assuming that there is no Fayet-Iliopoulos term. Each term in this equation is postive or negative depending only on the sign of q_n; by anomaly cancellation there will be both positive and negative q_n. Assume that there are non-zero ϕ^n's with both positive and negative charges q_n. Then, since the greatest q_n of a non-zero ϕ^n—call it q_{max}—is postive, as $C \to +\infty$ the right side of (1.308) is dominated by $q_{max} e^{q_{max} C} |\phi^{max}|^2$, and so is positive. Similarly, for $C \to -\infty$ the right side is negative. Therefore, there exists some intermediate value of C for which (1.308) is satisfied. Furthermore, there is a *unique* such value, which follows simply by taking the derivative of (1.308) with respect to C, and noting that it is positive definite. In the case of many Abelian gauge fields, the same argument works simply by applying it to each D_A equation in turn.

Thus we have a unique solution to the D equation for any given set of ϕ^n's, subject only to the constraint that at least one postively and one negatively charged ϕ^n are non-zero; furthermore, these are all the solutions to the D equation, since if all non-zero ϕ^n had the same sign charge, then there is no solution to the D equation (since all its terms are the same sign).[11] However, this is not yet a description of the moduli space of vacua, since we are working in complex gauge and have not fixed the $U(1)_{\mathbb{C}}$ gauge invariance. To do this we must divide out this space by the group of complexified gauge transformations, $G_{\mathbb{C}}$. But this futher division automatically enforces the constraint that vacua with non-zero ϕ^n all of the same sign charge be excluded, thus giving the result (1.307). To see this in the $U(1)_{\mathbb{C}}$ case, recall that a $U(1)_{\mathbb{C}}$ transformation rotates the fields by

$$U(1)_{\mathbb{C}}: \quad \phi^n \to e^{iq_n\Lambda}\phi^n, \qquad \Lambda \in \mathbb{C}. \tag{1.309}$$

Thus a point in target space with, say, ϕ^n non-zero only for $q_n > 0$, can be taken to the origin ($\phi^n = 0$ for all n) by a $U(1)_{\mathbb{C}}$ gauge transformation simply by taking $\mathrm{Im}\Lambda \to +\infty$.

The usefulness of the above result resides not only in showing that solutions to the D equations always exist, but also in providing a relatively simple description of the resulting moduli space. We can see this by trying to find a set of good (non-singular) coordinates which parametrize the quotient of the space of all ϕ^n's (and their complex conjugates) by $U(1)_{\mathbb{C}}$. Say $t(\phi_i, \phi_n^*)$ is one such coordinate function. Then t must be $U(1)_{\mathbb{C}}$-invariant, since if not, a value of t will not specify a submanifold of $\mathcal{M}$ since a $U(1)_{\mathbb{C}}$ transformation changes it. Without loss of generality we can expand t as a sum

[11] This is the step where the vanishing of the Fayet-Iliopoulos parameter is needed.

of monomials $t_{\{\ell,\overline{\ell}\}}$ each of which is of the form

$$t_{\{\ell,\overline{\ell}\}} = \prod_n (\phi^n)^{\ell_n} (\phi^*_n)^{\overline{\ell}_n} \tag{1.310}$$

for some set of exponents $\{\ell_n, \overline{\ell}_n\}$. In order for t to be $U(1)_{\mathbb{C}}$ invariant, each such term must be separately invariant. Thus the set of all such monomials can be taken as a possible basis of coordinate functions on $\mathcal{M}$. In order for $t_{\{\ell,\overline{\ell}\}}$ to be well-defined on the space of ϕ^n's, their exponents must be integers. For $t_{\{\ell,\overline{\ell}\}}$ to be $U(1)_{\mathbb{C}}$ invariant we must have

$$\sum_n q_n \ell_n = \sum_n q_n \overline{\ell}_n = 0. \tag{1.311}$$

This separate cancellation of the ℓ_n and $\overline{\ell}_n$ powers is because Λ is complex in (1.309). Thus each term is a product of two $U(1)_{\mathbb{C}}$ invariant terms—one made only from ϕ^n's and the other from only ϕ^*_n's. So, again, we can take the purely holomorphic terms $t_{\{\ell\}} = t_{\{\ell,\overline{\ell}=0\}}$ as a basis of complex coordinate functions. (The purely anti-holomorphic terms are their complex conjugates.) If there were a negative exponent, say $\ell_n < 0$ in $t_{\{\ell\}}$, then t would not be a good coordinate near points where $\phi^n = 0$, which are certainly points in $\mathcal{M}$. So, finally, we have a basis of good coordinates on $\mathcal{M}$:

$$\{t_{\{\ell\}} | \sum_n q_n \ell_n = 0, \text{ and } \ell_n \geq 0\}. \tag{1.312}$$

However, these coordinates need not all be independent, and will in general satisfy a set of algebraic relations. For example, given any set of monomials $t_{\{\ell\}}$, then any holomorphic polynomials made from them are also good coordinates, but are not algebraically independent.

We have thus developed a description of $\mathcal{M}$ as the space of all holomorphic and gauge invariant polynomials of the ϕ^n modulo algebraic relations among them:

$$\begin{aligned} \mathcal{M} &\equiv \{\phi^n | D \text{ equations}\}/G = \{\phi^n\}/G_{\mathbb{C}} \qquad (1.313) \\ &= \{G\text{-invariant holomorphic monomials of the } \phi^n\}/\{\text{algebraic relations}\}. \end{aligned}$$

We have only shown this for $G = U(1)$, but the extension to an arbitrary number of $U(1)$'s is straightforward. As we will discuss later, this result is also valid for non-Abelian G.

Note, finally, that if we turn on a superpotential in this theory, our analysis of the D term equations is not changed. If there is a moduli space $\mathcal{M}'$ of solutions to the F equations coming from the superpotential, then the branches of this space where not only same-sign charged left-chiral superfields get non-zero vacuum expectation values will also be solutions of the D equations, and so the whole moduli space is

$$\mathcal{M} = \mathcal{M}'/G_{\mathbb{C}} \equiv \{\phi^n | F^m = 0\}/G_{\mathbb{C}}. \tag{1.314}$$

More examples

We will now illustrate this general discussion with a couple of examples. In each case we will see that the (classical) physics we are describing is just the Higgs mechanism. Thus these moduli spaces which appear as solutions to the D equations are often called *Higgs branches.* They are often also called *D-flat directions.*

Consider a $U(1)^N$ theory with $2N+2$ left-chiral superfields with charges

$$\begin{array}{ccccccc} & U(1)_1 & \times & U(1)_2 & \times & \cdots & \times & U(1)_N \\ \Phi_1^\pm & \pm 1 & & 0 & & \cdots & & 0 \\ \Phi_2^\pm & 0 & & \pm 1 & & \cdots & & 0 \\ \vdots & \vdots & & \vdots & & \ddots & & \vdots \\ \Phi_N^\pm & 0 & & 0 & & \cdots & & \pm 1 \\ \Phi_0^\pm & \pm 1 & & \pm 1 & & \cdots & & \pm 1 \end{array} \tag{1.315}$$

A basis of gauge invariant holomorphic monomials is

$$\begin{aligned} M_n &= \phi_n^+ \phi_n^- , \qquad n = 1, \ldots, N, \\ M_0 &= \phi_0^+ \phi_0^- , \\ B &= \phi_0^- \prod_{n=1}^N \phi_n^+ , \\ \widetilde{B} &= \phi_0^+ \prod_{n=1}^N \phi_n^- , \end{aligned} \tag{1.316}$$

which are subject to the single constraint

$$B\widetilde{B} = M_0 \prod_{n=1}^N M_n . \tag{1.317}$$

Thus

$$\mathcal{M} = \{M_0, M_n, B, \widetilde{B}\} / \{B\widetilde{B} = M_0 \textstyle\prod_n M_n\}. \tag{1.318}$$

Counting invariants minus relations, we see that the complex dimension of $\mathcal{M}$ is

$$\dim_{\mathbb{C}}(\mathcal{M}) = N + 2. \tag{1.319}$$

This matches our physical expectation: there are $2(N+1)$ complex fields and N complex gauge invariances which, generically, are all broken, leaving us with $2(N+1) - N = N+2$ complex flat directions.

You might wonder why we don't just use the constraint to solve for one of the gauge invariant fields, say $\widetilde{B}$, in terms of the others, and just say that $\mathcal{M} = \{B, M_0, M_n\}$

without constraints. This is not correct because it misses whole $N+1$ dimensional submanifolds of $\mathcal{M}$ with $B=0$ and one of the M_n or $M_0=0$, with the other coordinates, including $\widetilde{B}$, arbitrary.

At the generic point on $\mathcal{M}$, where the M_0 and all the M_n fields are non-zero, all the charged scalars have non-zero vacuum expectation values and thus completely break (Higgs) the $U(1)^N$ gauge symmetry. At the generic vacuum we therefore expect $N+2$ massless neutral left-chiral superfields corresponding to the $N+2$ complex flat directions. They are necessarily massless for the usual reason that the degeneracy of neighboring vacua in $\mathcal{M}$ means that the energy of a fluctuation of the left-chiral superfields along the flat directions can be made arbitrarily small by making it have a long enough wavelength.

At special points on $\mathcal{M}$, however, there may be extra massless multiplets. These points are typically associated with singularities in $\mathcal{M}$. In the case at hand we expect extra massless vector multiplets whenever one or more of the gauge factors are not Higgsed. We will now find these vacua by examining the singularities of $\mathcal{M}$. Define

$$y \equiv B\widetilde{B} - M_0 \prod_n M_n, \tag{1.320}$$

so that $\mathcal{M}$ is given by the equation $y=0$ in the complex space of $\{M_0, M_n, B, \widetilde{B}\}$. The condition for a singularity on the curve $y=0$ is that

$$dy = 0, \tag{1.321}$$

since this implies that there is a singularity in the tangent space to $\mathcal{M}$, so there are no good local complex coordinates. The dy constraint is

$$0 = dy = Bd\widetilde{B} + \widetilde{B}dB - \sum_{a=0}^{N}\left(\prod_{b\neq a} M_b\right) dM_a, \tag{1.322}$$

implying that singularities occur whenever

$$B = \widetilde{B} = 0, \text{ and at least } \textit{two } M_a = 0. \tag{1.323}$$

Associated to these singularities are points of enhanced gauge symmetry. For instance, when

$$M_1 = M_2 = B = \widetilde{B} = 0 \quad \Rightarrow \quad \phi_1^\pm = \phi_2^\pm = 0, \tag{1.324}$$

so the diagonal $U(1) \subset U(1)_1 \times U(1)_2$ is unbroken (since $\phi_0^\pm \neq 0$). (To deduce that ϕ_1^+ as well as ϕ_1^- vanish, we have to use the D equations.) As another instance, at the singularity

$$M_0 = M_1 = B = \widetilde{B} = 0 \quad \Rightarrow \quad \phi_0^\pm = \phi_1^\pm = 0, \tag{1.325}$$

implying that $U(1)_1$ is unbroken.

As a second example, take the theory of one $U(1)$ gauge field and two left-chiral superfields with charges

$$\begin{array}{cc} & U(1) \\ \Phi_1^\pm & \pm 1 \\ \Phi_2^\pm & \pm 2 \end{array} \tag{1.326}$$

We thus expect $\dim_{\mathbb{C}}\mathcal{M} = 4 - 1 = 3$, and indeed we find it, with the basis of four invariants

$$\begin{array}{ll} M_1 = \phi_1^+\phi_1^- & B = \phi_2^+\phi_1^-\phi_1^- \\ M_2 = \phi_2^+\phi_2^- & \widetilde{B} = \phi_2^-\phi_1^+\phi_1^+ \end{array} \tag{1.327}$$

and the one relation

$$B\widetilde{B} = M_2 M_1^2. \tag{1.328}$$

From

$$\begin{aligned} 0 \;=\; y \;&=\; B\widetilde{B} - M_2 M_1^2, \\ 0 \;=\; dy \;&=\; Bd\widetilde{B} + \widetilde{B}dB - M_1^2 dM_2 - 2M_1M_2dM_1, \end{aligned} \tag{1.329}$$

we find singularities at

$$B = \widetilde{B} = M_1 = 0 \ \ \forall M_2, \quad \Rightarrow \quad \phi_1^\pm = 0 \text{ and } \phi_2^\pm = \text{ arbitrary.} \tag{1.330}$$

This implies that at this singular submanifold for $M_2 \neq 0$ only a discrete $\mathbb{Z}_2$ gauge symmetry is restored. Discrete gauge symmetries have no associated massless gauge bosons. Thus we have found a singularity in the complex structure of $\mathcal{M}$ along a submanifold where there are no extra massless multiplets.

We can compute the metric on $\mathcal{M}$ near this singular submanifold for $M_2 \neq 0$ (and therefore $\phi_2^\pm \neq 0$) by finding the restriction of the Kahler potential

$$K = |\phi_1^+|^2 + |\phi_1^-|^2 + |\phi_2^+|^2 + |\phi_2^-|^2 \tag{1.331}$$

to $\mathcal{M}$. It is easiest to use, say, $\phi_1^\pm$ and ϕ_2^+ as independent coordinates on $\mathcal{M}$ in the region we are interested in. Since the D equation in Wess-Zumino gauge gives $|\phi_2^-|^2 = |\phi_2^+|^2 + \frac{1}{2}|\phi_1^+|^2 + \frac{1}{2}|\phi_1^-|^2$, we can remove ϕ_2^- from the Kahler potential:

$$K = \tfrac{3}{2}|\phi_1^+|^2 + \tfrac{1}{2}|\phi_1^-|^2 + 2|\phi_2^+|^2. \tag{1.332}$$

Furthermore, the phase of ϕ_2^- can also be fixed using the remaining $U(1)$ gauge invariance. Since ϕ_2^- has charge -2 this still leaves unfixed a discrete $\mathbb{Z}_2$ gauge invariance which simultaneously transforms

$$\phi_1^\pm \rightarrow -\phi_1^\pm. \tag{1.333}$$

Thus good gauge invariant variables on $\mathcal{M}$ near the singular submanifold can be taken to be ϕ_2^+, $\omega_1 \equiv (\phi_1^+)^2$, and $\omega_2 \equiv \phi_1^- \phi_1^+$. Rewriting K in terms of these variables then gives the induced Kahler potential on $\mathcal{M}$. The point of this exercize is that the metric has a $\mathbb{Z}_2$ orbifold singularity along the singular submanifold which just comes from the gauge identification (1.333). As we will discuss later, quantum corrections will *not* modify this conical singularity in this case since there are no extra massless particles there. On the other hand, when $M_2 \to 0$ along the singularity (*i.e.* at the origin of $\mathcal{M}$) then a full $U(1)$ is restored and its associated photon becomes massless. In that case we expect quantum corrections to modify the singularity so it becomes cusp-like instead of conical.

The lesson is that cusp-like singularities in the metric on $\mathcal{M}$ correspond to extra massless particles, while orbifold (conical) metric singularities correspond to extra discrete gauge invariances, and that in both cases the complex structure on $\mathcal{M}$ will be singular.

Problem 1.7.1 Solve for the vacua and spectrum of the *Fayet-Iliopoulos model*

$$S = \frac{1}{2}\int d^4x \left[\Phi_-^* e^{2V}\Phi_- + \Phi_+^* e^{-2V}\Phi_+ + 2\xi V\right]_D + \left[\frac{1}{4g^2}(W_L)^2 + m\Phi_+\Phi_- + \text{c.c.}\right]_F \tag{1.334}$$

as a function of its parameters g, ξ, and m.

1.8 Non-Abelian super gauge theory

In this lecture we generalize our construction of supersymmetric $U(1)$ gauge invariant actions to non-Abelian gauge invariance. We begin with a review of ordinary non-Abelian gauge theory.

1.8.1 Review of non-Abelian gauge theory

Compact Lie algebras

The generators t_A of a Lie algebra G satisfy

$$[t_A, t_B] = iC^C{}_{AB} t_C \tag{1.335}$$

where $C^C{}_{AB} = -C^C{}_{BA}$ are real *structure constants* of the algebra, and $A, B, C = 1, \ldots, \dim(G)$, where $\dim(G)$ is the dimension of G. The Jacobi identity for commutators,

$$0 = [[t_A, t_B], t_C] + [[t_C, t_A], t_B] + [[t_B, t_C], t_A], \tag{1.336}$$

implies that the structure constants must satisfy

$$0 = C^D{}_{AB} C^E{}_{DC} + C^D{}_{CA} C^E{}_{DB} + C^D{}_{BC} C^E{}_{DA}. \tag{1.337}$$

An r-dimensional representation of G is a realization of the generators t_A as a set of $r \times r$ matrices satisfying (1.335). Following our notation of section 1.3, we will denote this representation by its dimension as $\mathbf{r}$. If necessary, we will denote the generators in the $\mathbf{r}$ representation by $t_A^{(\mathbf{r})}$.

A compact Lie algebra is one which can be represented by (finite dimensional) Hermitian matrices:

$$t_A^\dagger = t_A. \tag{1.338}$$

Such algebras can be decomposed into simple and $U(1)$ Lie algebras each whose generators commute with all of the generators of the other algebras. A $U(1)$ algebra has a single generator, so satisfies the trivial algebra $[t, t] = 0$, whose irreducible representations are all just 1-dimensional; *i.e.* t is just represented by a real number q, called the charge:

$$t = q \in \mathbb{R}, \qquad (U(1) \text{ algebra}). \tag{1.339}$$

A simple Lie algebra has some structure constant $C^A{}_{BC}$ non-vanishing for each value of the index A. This implies, taking the trace of (1.335) and using the fact that the trace of a commutator vanishes, that the trace of the generators vanish in any representation:

$$\mathrm{tr}(t_A) = 0, \qquad (\text{simple algebra}). \tag{1.340}$$

There is a basis (which we will choose from now on) of generators of any simple compact Lie algebra in which

$$\mathrm{tr}_{\mathbf{r}}(t_A t_B) = C(\mathbf{r})\delta_{AB} \tag{1.341}$$

where $\mathrm{tr}_{\mathbf{r}}$ denotes a trace in the $\mathbf{r}$ representation, and the constant $C(\mathbf{r})$ is called the *quadratic invariant* of the representation $\mathbf{r}$. Note that the quadratic invariants are *not* invariant under rescalings of the generators of the Lie algebra (with a simultaneous rescaling of the structure constants). In a basis in which (1.341) holds, it is easy to check that the structure constant C_{ABC} is totally antisymmetric on its three indices; here we raise and lower the algebra indices A, B, C, with δ_{AB}.

We represent a compact Lie algebra as the *direct product* of its simple or $U(1)$ factors,

$$G = G_1 \times G_2 \times \cdots \times G_n, \tag{1.342}$$

since its irreducible representations are given by the tensor product of irreducible representations of each factor.

The elements g of the Lie group associated to an algebra are generated by the exponential map

$$g(\Lambda) = \exp\{i\Lambda^A t_A\}, \tag{1.343}$$

where Λ^A are real numbers parameterizing the Lie group. In the $\mathbf{r}$ representation, g is an $r \times r$ unitary matrix. A vector is in the $\mathbf{r}$ representation if it is an r-component vector which transforms under the Lie group as

$$\phi \to g(\Lambda)\phi \tag{1.344}$$

where matrix multiplication is understood. For infinitesimal Λ this means that ϕ shifts as

$$\delta_\Lambda \phi_a = i\Lambda^A (t_A^{(\mathbf{r})})_a^{\ b} \phi_b, \tag{1.345}$$

where we have shown the representation indices $a, b = 1, \ldots, r$.

The complex conjugate $\bar{\mathbf{r}}$ of a representation is defined as the representation in which the complex conjugate of a vector in the $\mathbf{r}$ transforms as $\phi^* \to g^*\phi^*$, implying

$$\delta_\Lambda \phi^{*a} = -i\Lambda^A (t_A^{(\mathbf{r})})_a^{*b} \phi^{*b} = -i\Lambda^A (t_A^{(\mathbf{r})})_b^{\ a} \phi^{*b} \tag{1.346}$$

where in the second equality we have used the Hermiticity of t_A, and we have adopted the convention of raising the representation index of the vector upon complex conjugation. Thus, as a matrix statement, we see that the generators in $\bar{\mathbf{r}}$ are related to those in $\mathbf{r}$ by

$$t_A^{(\bar{\mathbf{r}})} = -(t_A^{(\mathbf{r})})^T. \tag{1.347}$$

If the generators are all antisymmetric then they are the same as the generators of the complex conjugate representation, and we say the representation is real.[12]

The *trivial*, *singlet*, or *identity* representation of any simple Lie algebra is the 1-dimensional representation in which $t_A = 0$. In terms of the Lie group this means $g(\Lambda) = 1$ for all Λ.

The *adjoint* representation is a $\dim(G)$-dimensional irreducible representation of G given by

$$(t_A)^B_C = iC^B{}_{AC}, \qquad \text{(adjoint)} \tag{1.348}$$

where we have labelled the rows and columns of the matrix t_A by B and C. This satisfies (1.335) by virtue of the relation (1.337), and is Hermitian by virtue of the antisymmetry of the structure constants. Because of this antisymmetry, we also see that the adjoint representation is real. The quadratic invariant of the adjoint representation, $C(\text{adj})$, is thus given by

$$C(\text{adj}) = \frac{C_{ABC}C^{ABC}}{\dim(G)}, \tag{1.349}$$

and is also called the quadratic Casimir of G, and sometimes denoted C_2.

Finally, the *rank* of a Lie algebra, denoted $\text{rank}(G)$, is the maximal number of mutually commuting generators. The $U(1)^{\text{rank}(G)}$ subalgebra of these generators is called the *Cartan subalgebra* of G.

Classical Lie groups

The classical Lie groups, $SU(N)$, $SO(N)$, and $Sp(2N)$, are defined as the groups of unitary, orthogonal, and symplectic matrices.

$SU(N)$ is the group of $N \times N$ unitary complex matrices with determinant 1. This actually defines the *fundamental*, *defining*, or $\mathbf{N}$, representation of $SU(N)$. In this representation the Lie algebra generators t_A then span the space of traceless Hermitian $N \times N$ matrices. As there are $N^2 - 1$ linearly independent such matrices, the dimension of the algebra is $\dim(SU(N)) = N^2 - 1$. Denoting a vector in the fundamental representation by ϕ_a with $a = 1, \dots, N$, we can form a second irreducible representation, called the *anti-fundamental* by taking its complex conjugate. As above, we will denote it by vectors with raised indices, ϕ^a. We can then form other irreducible representations by taking tensor products and reducing as discussed in section 1.3. For

[12]More generally, a representation might only be equivalent to its complex conjugate, *i.e.* related to it by a similarity transformation: $t_A^{(\overline{\mathbf{r}})} = U t_A^{(\mathbf{r})} U^\dagger$. (The similarity transformation must be unitary to preserve the Hermiticity of the t_A's.) If U is symmetric, then it can be written as $U = V^T V$, and a similarity transformation by V makes $t_A^{(\overline{\mathbf{r}})} = t_A^{(\mathbf{r})}$; in this case the representation is real. Otherwise—if U is not symmetric—the representation is said to be *pseudoreal*.

example, the tensor product of two fundamentals can be reduced into its symmetric and antisymmetric parts, giving

$$\mathbf{N} \otimes \mathbf{N} = [\mathbf{N(N-1)/2}] \oplus [\mathbf{N(N+1)/2}], \tag{1.350}$$

and similarly for the product of anti-fundamentals. All irreducible representations of $SU(N)$ can be found in this way, and by contracting indices using the invariant tensors

$$\delta^a_b, \qquad \epsilon^{a_1 \cdots a_N}, \qquad \epsilon_{b_1 \cdots b_N}. \tag{1.351}$$

Thus, for example,

$$\mathbf{N} \otimes \overline{\mathbf{N}} = [\mathbf{N^2} - \mathbf{1}] \oplus \mathbf{1}, \tag{1.352}$$

where $\mathbf{N^2}-\mathbf{1}$ is the adjoint representation, and the singlet is formed by contracting the $\mathbf{N}$ and $\overline{\mathbf{N}}$ indices with δ^a_b. The invariant epsilon tensors imply relations like $\mathbf{3} \otimes_A \mathbf{3} = \overline{\mathbf{3}}$ for $SU(3)$, or $\overline{\mathbf{2}} = \mathbf{2}$ for $SU(2)$.[13] The rank of $SU(N)$ is $N-1$ since there are $N-1$ independent diagonal traceless $N \times N$ hermitian matrices. In the Dynkin classification of simple Lie algebras, the unitary algebras are the "A series" and are labelled by their rank; thus $SU(r+1) = A_r$. The gauge theory examples we will discuss in these lectures will use $SU(N)$ groups exclusively.

$SO(N)$ is the group of $N \times N$ real orthogonal matrices with determinant $+1$. This actually defines the *fundamental*, *vector*, *defining*, or $\mathbf{N}$, representation of $SO(N)$. In this representation the Lie algebra generators t_A span the space of real antisymmetric $N \times N$ matrices. The dimension of the algebra is thus $\dim(SO(N)) = N(N-1)/2$. The rank of $SO(N)$ is $[N/2]$ (the largest integer part of $N/2$); the $SO(2r+1)$ Lie algebras are denoted B_r, while the $SO(2r)$ ones are denoted D_r in the Dynkin classification. Denote a vector in the fundamental representation by ϕ_a with $a = 1, \ldots, N$. The antisymmetry of the generators means this representation is real. Invariant tensors are

$$\delta_{ab}, \qquad \epsilon_{a_1 \cdots a_N}, \tag{1.353}$$

so, in particular, we can raise and lower indices with δ_{ab}. For example, the tensor product of two fundamentals can be reduced to its symmetric-traceless, antisymmetric and singlet parts, just as in our dicsussion of the tensor representations of the Lorentz group in section 1.3. In this case the antisymmetric representation is the same as the adjoint representation. Indeed, the Lorentz group in d space-time dimensions is just a non-compact form of $SO(d)$, called $SO(d-1,1)$. In addition to the tensor representations, the $SO(N)$ Lie algebra also has spinor representations, which can be constructed from the appropriate Clifford algebra as in section 1.3. (The matrix Lie *group* $SO(N)$ does not admit spinor representations, but its covering space, called $Spin(N)$, does; both $SO(N)$ and $Spin(N)$ have the same Lie algebra.)

[13] The $\mathbf{2}$ of $SU(2)$ is actually pseudoreal.

The compact form of the symplectic group $Sp(2N)$ (or, more properly, $USp(2N)$) is defined as the group of unitary complex $2N \times 2N$ matrices g satisfying the relation

$$gJg^T = J \qquad \text{(with matrix indices: } g_a^c J_{cd} g_b^d = J_{ab} \text{)} \tag{1.354}$$

where the *symplectic form* J is

$$J = \begin{pmatrix} 0 & \mathbb{1}_N \\ -\mathbb{1}_N & 0 \end{pmatrix} \tag{1.355}$$

written in terms of $N \times N$ blocks where $\mathbb{1}_N$ is the $N \times N$ identity matrix. As before, this defines the *fundamental* or $\mathbf{2N}$ representation of $Sp(2N)$. In this representation the Lie algebra generators t_A span the space of Hermitian $2N \times 2N$ matrices satisfying

$$-t_A^T = J t_A J^{-1}. \tag{1.356}$$

Since J is not symmetric (1.356) means that the fundamental representation is pseudoreal. Denote its vectors by ϕ_a, $a = 1, \dots, 2N$. The invariant tensor is J_{ab} which can be used to raise and lower indices. In particular (1.356) can be written as $(Jt_A)^T = (Jt_A)$, implying that the generators with lowered indices are symmetric matrices; thus the dimension of the $Sp(2N)$ algebra is $\dim(Sp(2N)) = N(2N+1)$. All irreducible representations can be formed from tensor products of the fundamentals. For example $\mathbf{2N} \otimes \mathbf{2N} = [\mathbf{N}(\mathbf{2N}+\mathbf{1})] \oplus [(\mathbf{2N}+\mathbf{1})(\mathbf{N}-\mathbf{1})] \oplus \mathbf{1}$: the symmetric (which is the adjoint), the "traceless" antisymmetric, and a singlet. (The traceless antisymmetric is an antisymmetric tensor $\phi_{[ab]}$ satisfying $\mathrm{tr}(\phi J) = 0$.) Finally, the rank of $Sp(2N)$ is N, and in the Dynkin classification $Sp(2r) = C_r$.

So far nothing we have said fixed the normalization of the generators or of the structure constants of our Lie algebras. Conventional normalizations of the generators for the classical groups which are often implicitly used in the physics literature can be summarized by giving the quadratic invariants for their defining representations. We include some additional useful group theory information:

G	rank(G)	dim(G)	C(adj)	dim(fund)	C(fund)	type fund. rep.
$SU(N)$	$N-1$	N^2-1	N	N	1/2	complex
$SO(N)$	$[N/2]$	$N(N-1)/2$	$N-2$	N	1	real
$Sp(2N)$	N	$N(2N+1)$	$N+1$	$2N$	1/2	pseudoreal

A useful formula for computing the quadratic invariants of other representations is

$$C(\mathbf{r_1}) \dim(\mathbf{r_2}) + \dim(\mathbf{r_1}) C(\mathbf{r_2}) = \sum_i C(\mathbf{r_i}), \tag{1.357}$$

where dim($\mathbf{r}$) is the dimension of the representation, and here $\mathbf{r_1} \otimes \mathbf{r_2} = \bigoplus_\mathbf{i} \mathbf{r_i}$. It should be clear, however, that the real invariant quantities are the *ratios* of the quadratic invariants. Also, one should be aware of the following equivalences among Lie algebras:

$$\begin{aligned} SO(3) &\simeq SU(2) \simeq Sp(2) \\ SO(4) &\simeq SU(2)\times SU(2) \simeq Sp(2)\times Sp(2) \\ SO(5) &\simeq Sp(4) \\ SO(6) &\simeq SU(4). \end{aligned} \tag{1.358}$$

There are only five other simple Lie algebras, the exceptional ones, denoted by E_6, E_7, E_8, F_4, and G_2 in the Dynkin classification. A more detailed introduction to the classification and representations of the simple Lie algebras can be found in [15].

Yang-Mills theory

A field $\phi_a(x)$ $a = 1, \ldots, r$ in the $\mathbf{r}$ representation of the gauge algebra G transforms under gauge transformations as

$$\phi(x) \to e^{i\Lambda^A(x)t_A}\phi(x). \tag{1.359}$$

The vector field $V_\mu^A(x)$, $A = 1, \ldots, \dim(G)$, carries an adjoint representation index. From it we can form a matrix in any representation $\mathbf{r}$ by

$$V_\mu \equiv V_\mu^A t_A^{(\mathbf{r})}. \tag{1.360}$$

We then define the covariant derivative of a field in that representation by

$$D_\mu\phi = \partial_\mu\phi - iV_\mu\phi, \tag{1.361}$$

where matrix multiplication on the representation indices is understood. For example,

$$D_\mu V_\nu^A = \partial_\mu V_\nu^A - iV_\mu^B (t_B^{(\mathbf{ad})})^A_C V_\nu^C = \partial_\mu V_\nu^A + V_\mu^B C^A{}_{BC} V_\nu^C \tag{1.362}$$

where in the second equality we have used the definition of the adjoint representation (1.348). Multiplying by t_A then gives

$$D_\mu V_\nu = \partial_\mu V_\nu + C^A{}_{BC} t_A V_\mu^B V_\nu^C = \partial_\mu V_\nu - i[V_\mu, V_\nu], \tag{1.363}$$

where in the second equality we have used the definition of the Lie algebra (1.335). This illustrates the general fact that the action of an adjoint field on an adjoint field can be rewritten as a commutator.

Demanding that the covariant derivative of ϕ also transforms in the $\mathbf{r}$ representation implies that under gauge transformations the vector field transforms as

$$V_\mu \to e^{i\Lambda} V_\mu e^{-i\Lambda} - i(\partial_\mu e^{i\Lambda}) e^{-i\Lambda}, \tag{1.364}$$

where we have defined a Lie algebra-valued gauge transformation parameter Λ by

$$\Lambda = \Lambda^A t_A. \tag{1.365}$$

Note that for infinitesimal Λ this reduces to

$$\delta_\Lambda V_\mu = i[\Lambda, V_\mu] + \partial_\mu \Lambda. \tag{1.366}$$

We recognize the first term as the transformation rule for the adjoint representation, while the second term gives the gauge shift of the potential familiar from Abelian gauge transformations.

The (Lie algebra-valued) field strength tensor is defined as

$$f_{\mu\nu} \equiv i[D_\mu, D_\nu] = \partial_\mu V_\nu - \partial_\nu V_\mu - i[V_\mu, V_\nu], \tag{1.367}$$

which reads in components

$$f^A_{\mu\nu} = \partial_\mu V^A_\nu - \partial_\nu V^A_\mu + C^A{}_{BC} V^B_\mu V^C_\nu. \tag{1.368}$$

The Yang-Mills theory with a simple gauge group is then written

$$S_{YM} = \int d^4x \sum_A \left(-\frac{1}{4g^2} f^A_{\mu\nu} f^{A\mu\nu} + \frac{\vartheta}{32\pi^2} f^A_{\mu\nu} \widetilde{f}^{A\mu\nu} \right). \tag{1.369}$$

(Recall that $\widetilde{f}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} f_{\rho\sigma}$, so that $\widetilde{\widetilde{f}} = -f$.) For a gauge group which is a product of simple and $U(1)$ factors, the action just becomes a sum of such terms with separate field strengths for the generators of each factor group, and separate couplings for each factor group.

Normalizations and the index of a representation

The Yang-Mills action (1.369) implicitly assumes a normalization of the gauge group generators. This follows because under a rescaling $t^A \to \alpha t^A$ of the group generators, one must rescale the structure constants, gauge potential, and field strength as well to keep our definitions the same: $C^A{}_{BC} \to \alpha C^A{}_{BC}$, $V^A_\mu \to \frac{1}{\alpha} V^A_\mu$, and $f^A_{\mu\nu} \to \frac{1}{\alpha} f^A_{\mu\nu}$. This implies, in particular, that the gauge coupling constants are not invariant under these rescalings. To keep the coupling constants in (1.369) invariant under these rescalings, we need to insert a compensating factor. A natural factor is the quadratic "invariant" (1.341) of some representation $\mathbf{r}$, since under the above rescaling, $C(\mathbf{r}) \to \alpha^2 C(\mathbf{r})$.

There remains, however, the arbitrary choice of representation $\mathbf{r}$ to use. This is a matter of convention. For the classical gauge groups (SU, SO, and Sp) it is standard to use the fundamental representations, and so to write (1.369) as

$$S_{YM} = \int d^4x \left(-\frac{1}{4g^2} \mathrm{tr}_f(f^2) + \frac{\vartheta}{32\pi^2} \mathrm{tr}_f(f\widetilde{f}) \right), \tag{1.370}$$

since $\sum_A C(\mathrm{fund}) f^A f^A = \mathrm{tr}_f(f^2)$ where tr_f denotes the trace in the fundamental (or defining) representation. One often sees the gauge kinetic term written with the gauge fields in the adjoint representation of G. In this case the invariant formula is (1.370) with tr_f replaced by Tr (using the conventional notation that a capitalized trace refers to a trace in the adjoint representation) multiplied by an overall factor of $C(\mathrm{fund})/C(\mathrm{adj})$.

It should be clear from this discussion that the real invariant quantities are the *ratios* of the quadratic invariants. From these ratios can be defined the *index* of a representation $T(\mathbf{r})$. Thus only these indices will enter in physical quantities, and not the quadratic invariants. In the case of the classical groups the index is simply

$$T(\mathbf{r}) \equiv C(\mathbf{r})/C(\mathrm{fund}). \tag{1.371}$$

(Mathematically, there is a more general definition, applicable to all simple Lie algebras.) It is a theorem that the index of any representation is an integer. We see that, by definition, the index of the fundamental representation is 1, and, from the above table, that the indices of the adjoint representations are $2N$, $N{-}2$, and $2N{+}2$ for $SU(N)$, $SO(N)$, and $Sp(2N)$, respectively.

1.8.2 Non-Abelian vector superfields

By analogy with the Abelian case, we promote the (Lie algebra-valued) gauge parameter Λ to a left-chiral superfield, Ω. The gauge transformation rule of a left-chiral superfield $\Phi^{(\mathbf{r})}$ in the $\mathbf{r}$ representation of G is

$$\Phi^{(\mathbf{r})} \to e^{i\Omega} \Phi^{(\mathbf{r})}, \tag{1.372}$$

where Ω is also in the $\mathbf{r}$ representation.

A non-Abelian vector superfield $V^A(x,\theta)$, $A = 1, \ldots, \dim(G)$, carries an adjoint representation index. Guessing that the Kahler terms should remain of the same form as in the Abelian case,

$$K = (\Phi^{(\mathbf{r})})^\dagger \exp\{-2V^A t_A^{(\mathbf{r})}\} \Phi^{(\mathbf{r})}, \tag{1.373}$$

we get the gauge transformation rule for the vector superfield

$$e^{-2V} \to e^{-i\Omega^\dagger} e^{-2V} e^{i\Omega} \tag{1.374}$$

where we have defined a Lie algebra-valued vector superfield by

$$V = t_A V^A. \tag{1.375}$$

Expanding this out to leading order gives[14]

$$V \to V + \tfrac{i}{2}(\Omega - \Omega^\dagger) + \ldots, \tag{1.376}$$

implying that an analog of Wess-Zumino gauge exists for non-Abelian vector superfields, in which, just as in the Abelian case,

$$\begin{aligned} V &= -\tfrac{i}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)V_\mu - i(\overline{\theta}\gamma_5\theta)(\overline{\theta}\lambda) - \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2 D, \\ V^2 &= \tfrac{1}{4}(\overline{\theta}\gamma_5\theta)^2 V_\mu V^\mu, \\ V^3 &= 0, \end{aligned} \tag{1.377}$$

but where now all the components are matrix-valued fields in some representation of G. In this gauge the gauge parameter Ω is determined up to a single Hermitian scalar part

$$\Omega(x^\mu, \theta) = \mathrm{Re}\Lambda(x^\mu_+), \tag{1.378}$$

where, as usual, $x^\mu_+ = x^\mu + \frac{1}{2}(\overline{\theta}\gamma_5\gamma^\mu\theta)$. There is also an analog of complex gauge in which both the real and imaginary parts of the lowest component, Λ, of Ω are left unfixed; in this gauge the gauge symmetry is enlarged from G to $G_{\mathbb{C}}$.

The field strength left-chiral superfield, written as a Lie algebra-valued superfield, is defined as

$$W_L \equiv \frac{i}{4}(\mathcal{D}_R^T\mathcal{C}\mathcal{D}_R)e^{-V}\mathcal{D}_L e^V. \tag{1.379}$$

It is straightforward to check that W_L transforms covariantly in the adjoint representation of the gauge group,

$$W_L \to e^{-i\Omega} W_L e^{i\Omega}. \tag{1.380}$$

It is also easy to see that the combination $e^{-V}\mathcal{D}_L e^V$ is a gauge-covariant super derivative. The Bianchi identity is just as in the Abelian case, except that it involves the gauge-covariant derivatives.

[14]There are other terms linear in Ω but with higher powers of V that have been left out: they give rise to the expected adjoint representation transformations of the components of V.

We can write the general supersymmetric gauge theory with simple gauge group as

$$S = \int d^4x \Big\{ \left[\tfrac{1}{2} K\left(\Phi_n^\dagger \exp\{-2V_{(\mathbf{r}_n)}\}, \Phi^n\right)\right]_D + \left[f(\Phi^n) + \frac{1}{16\pi i}\tau(\Phi^n)\,\mathrm{tr}_f(W_L^T \mathcal{C} W_L) + \text{c.c.}\right]_F \Big\} \qquad (1.381)$$

where we have put the representation of the Φ^n left-chiral superfield on V in the Kahler term to emphasize that it should takes values in the $\mathbf{r}_n$ representation. If there were more than one simple factor, we would have a separate field strength superfield W_L and gauge coupling τ for each factor. Note that the left-chiral superfields may be charged under more than one gauge factor—*i.e.* they may transform under a representation which is a tensor product of non-trivial representations for each gauge factor. There is no Fayet-Iliopoulos term in (1.381) since it is not allowed by gauge invariance: the D component of V is not gauge invariant, but transforms in the adjoint representation.

The supersymmetric gauge theory with no matter chiral superfields we will refer to as *superYang-Mills (superYM) theory*. Since there is no matter to couple the different gauge group factors, this theory is just a product of separate superYM theories for each simple gauge group factor separately. In the rest of this course we will be interested in the $SU(N)$ superYM theories. Another set of theories we will be interested in we will call the *superQCD* theories. They are theories with $SU(N_c)$ gauge group and matter (or "quark") chiral fields Φ^n, $n = 1, \ldots, N_f$, in the fundamental ($\mathbf{N_c}$) representation, and an equal number of "anti-quarks" $\widetilde{\Phi}_n$ in the $\overline{\mathbf{N_c}}$. We will call N_c the number of "colors", and N_f the number of "flavors" of superQCD. The simplest superpotential we can write for this theory is

$$f = m_m^n \mathrm{tr}\widetilde{\Phi}_n \Phi^m, \qquad (1.382)$$

where the trace is to remind us that we are to contract the $\mathbf{N_c}$ and $\overline{\mathbf{N_c}}$ indices for gauge invariance. This quark mass term is the most general renormalizable interaction term in this theory (for $N_c > 3$). We will use renormalizablility to restrict ourselves to constant τ (*i.e.* independent of the matter chiral superfields) as well. Near the end of the course we will also consider generalizations of superQCD with *adjoint* matter chiral superfields.

For $N_c > 2$ (so that $\mathbf{N_c}$ and $\overline{\mathbf{N_c}}$ are inequivalent representations) massless superQCD has a $U(N_f)\times U(N_f)$ global symmetry group, with one $U(N_f)$ factor rotating the quark superfields while the other one rotates the anti-quarks. In terms of simple factors $U(N_f) = U(1) \times SU(N_f)$ where the $U(1)$ factor acts by a common phase rotation of all the quark (or anti-quark) superfields. Thus there are two overall $U(1)$ global symmetry factors; the combination under which all the quarks have charge +1 and all the anti-quarks charge -1 is called baryon number, $U(1)_B$, in analogy to ordinary QCD.

The superpotential mass terms explicitly break these symmetries. Separate $SU(N_f)$ field redefinitions of the quarks and antiquarks can always be used to make the mass matrix diagonal, $m_m^n = m_n \delta_m^n$. In this basis it is clear that the global symmetry group is broken down to $U(1)^{N_f}$ with each $U(1)$ factor acting as a separate "baryon number" for each quark-anti-quark flavor pair.

This theory also has a global $U(1)_R$ symmetry. Recalling that the F terms in the action should have total R charge $+2$, we assign R charges to the left-chiral superfields as

$$R(W_L) = R(\Phi^n) = R(\widetilde{\Phi}_n) = +1. \tag{1.383}$$

In terms of component fields this implies that $R(f_{\mu\nu}) = R(D) = R(\psi_L^n) = R(\widetilde{\psi}_{Ln}) = 0$ and $R(\lambda_L) = R(\phi^n) = R(\widetilde{f}_n) = 1$. This symmetry is not broken by the mass terms. These classical global symmetries may suffer from anomalies quantum mechanically, as we will discuss in a later lecture.

The scalar potential from expanding the action (1.381) in components is once again a sum of squares of F and D components. Since the F terms do not involve the vector superfields, they are just the same as in the Abelian case. It is easy to check that the D terms are in Wess-Zumino gauge

$$D^A = \sum_n \mathrm{tr}(\phi_n^\dagger t_A^{(\mathbf{r_n})} \phi^n), \tag{1.384}$$

where I have assumed canonical Kahler terms (otherwise there would be a factor of the Kahler metric lowering the n index on ϕ^n).

The result of section 1.7.6 that the moduli space $\mathcal{M}$ of solutions to the $D^A = 0$ vacuum equations is given by the space of holomorphic G-invariant monomials of the left-chiral superfields modulo algebraic identifications is also valid when G is non-Abelian. See [16, chapter 8] for a discussion of part of this result (the existence of solutions to the D equations) for non-Abelian G. The mathematical procedure defining a Kahler sunmanifold $\mathcal{M}$ of the original Kahler target space by solving the D equations and dividing by the gauge symmetry is known as the *Kahler quotient* construction. For a brief description of various quotient constructions and further references, see [17].

Problem 1.8.1 Show in Wess-Zumino gauge that the non-Abelian gauge transformation (1.374) gives the usual gauge transformation rule for V_μ^A, and gauge-covariant transformation rules for the other components λ^A and D^A.

Problem 1.8.2 Find a basis of holomorphic and gauge invariant monomials in the quark and anti-quark left-chiral superfields, and a complete set of constraints generating the algebraic relations among them, to describe the moduli space of massless $SU(3)$ superQCD with 3 flavors.

Chapter 2

Quantum N=1 Supersymmetry

We now introduce the notion of an *infrared (IR) effective action* which we use to analyze the vacuum structure of four dimensional supersymmetric field theories. The idea is to guess an IR effective field content for the microscopic (UV) theory in question and write down all possible IR effective actions built from these fields consistent with the supersymmetry and other global symmetries of the UV theory. For a "generic" UV theory this would seem to give little advantage for obtaining interesting information about the vacuum structure. However, if the theory has a continuous set of inequivalent vacua, it turns out that selection rules from global symmetries of the UV theory can sometimes constrain the IR effective action sufficiently to deduce exact results.

2.1 Effective Actions

A basic notion in quantum field theory is that of a *low energy* (or *Wilsonian*) *effective action* [18]. This is simply a local action describing a theory's degrees of freedom at energies below a given energy or mass scale μ which we will refer to as the *cutoff scale* or simply the scale of the effective theory. A familiar example is the low energy effective action for QCD: chiral perturbation theory describing the interactions of pions at energies $E < \Lambda_{QCD}$. In such a theory particles heavier than Λ_{QCD} are included in the pion theory as classical sources. The example illustrates the common phenomenon that the degrees of freedom describing the microscopic physics (for QCD quarks and gluons) may be very different from the low energy degrees of freedom (pions). Other examples are the various ten and eleven-dimensional supergravity theories, which appear as effective actions for string/M theory at energies below their Planck scales.

The effective action is obtained by averaging over (integrating out) the short distance fluctuations of the theory down to the scale μ. By locality of the underlying theory the effective action will be local on length scales larger than $1/\mu$, and it will describe in a unitary way physical processes involving energy-momentum transfers less than μ. For processes at energies near μ, the effective couplings and masses will be given by the tree level (classical) couplings in the effective action: they will not be renormalized since the effects of all the higher energy degrees of freedom (that would contribute to loops, *etc.*) have already be integrated out. Thus the effective action at the scale μ is one which describes the physics at that scale by its classical couplings.

Physical processes taking place at a scale E substantially lower than μ will involve quantum corrections due to the flucuations of the modes of the fields in the effective action with energies between E and μ. These corrections can be absorbed in the couplings to define a new effective action at the lower scale E. This change in the effective action with scale is the familiar renormalization group (RG) running of the couplings. Denoting the effective action at scale μ by S_μ, the effect of integrating out fluctuations in a small energy band $\mu > E > \mu - d\mu$ is encoded in a defferential equation for the effective action, the Wilson equation

$$\frac{\partial S_\mu}{\partial \mu} = \mathcal{F}(S_\mu) \tag{2.1}$$

where $\mathcal{F}$ is some functional. Thinking of the action as a (potentially infinite) sum

$$S_\mu = \int d^4x \sum_i g_i(\mu)\mathcal{O}_i \tag{2.2}$$

of local operators $\mathcal{O}_i$ with couplings g_i, the Wilson equation is equivalent to a flow in the infinite dimensional coupling space:

$$\mu\frac{\partial g_i}{\partial \mu} = \beta_i(g_k, \mu). \tag{2.3}$$

A *fixed point* of this flow is an effective theory for which the right hand side vanishes, so the effective theory does not change with scale. Such a theory is naturally called a *scale invariant* theory. Linearizing the RG flow around a fixed point, we can classify the couplings according to their eigenvalues. Negative eigenvalues correspond to operators whose couplings are damped along the flow, and are called irrelevant operators. Positive and zero eigenvalues are correspond to relevant and marginal operators respectively. Thus irrelevant operators become less important in the IR while relevant ones become more important at low energies.

If the fixed point is a free theory (or at weak coupling) then the eigenvalues of the operators can be determined by dimensional analysis. Since the fluctuations of a free field are determined by its kinetic term, if we scale all energies and momenta by a factor μ/μ_0 to lower the cutoff scale $\mu_0 \to \mu$, then lengths scale by μ_0/μ, derivatives by μ/μ_0, and for the kinetic terms in the action

$$S_{\text{kin}} = \int d^4x \left\{ -\tfrac{1}{2}(\partial\phi)^2 - \tfrac{1}{2}(\overline{\psi}\partial\!\!\!/\psi) - \tfrac{1}{4}(f^2) \right\} \tag{2.4}$$

to be scale invariant, we must scale scalar, spinor, and vector fields by

$$\begin{aligned} \phi &\to \left(\frac{\mu}{\mu_0}\right)\phi, \\ \psi &\to \left(\frac{\mu}{\mu_0}\right)^{3/2}\psi, \\ V_\mu &\to \left(\frac{\mu}{\mu_0}\right)V_\mu. \end{aligned} \tag{2.5}$$

If a local operator made from these fields and derivatives scales as $\mathcal{O}_i \to s_i^{\Delta}\mathcal{O}_i$, we say it has (mass) dimension Δ_i. The the effect of a given interaction term in the action due to the operator $\mathcal{O}_i$ then scales as

$$\int d^4x\,\mathcal{O}_i \to \left(\frac{\mu}{\mu_0}\right)^{\Delta_i-4}\int d^4x\,\mathcal{O}_i, \tag{2.6}$$

giving the usual result that operators with dimensions $\Delta_i > 4$ are irrelevant, those with $\Delta_i < 4$ are relevant, while those with $\Delta_i = 4$ are marginal. The utility of the effective action stems from the fact that there are only a finite number of relevant and marginal local operators that one can write down.

In the marginal case, where the classical scaling says the operators don't scale at all, one must turn to quantum corrections to see whether the operator is in fact relevant or irrelevant. (If it remains marginal even quantum mechanically, it is sometimes said to be *exactly marginal*.) In general quantum corrections will modify the classical scaling dimensions. To make this concrete, recall the renormalization of a scalar theory (with a $\mathbb{Z}_2$ global symmetry). The relevant and marginal terms in its effective action at an initial scale μ_0 can be written

$$S_{\mu_0} = S_{\text{free}} + \int d^4x \mu_0^{\Delta_i-4}\lambda_i(\mu_0)\mathcal{O}_i, \tag{2.7}$$

where

$$S_{\text{free}} = \tfrac{1}{2}\int d^4x \left\{ -(\partial\phi)^2 - m^2\phi^2 \right\}, \tag{2.8}$$

and we have explicitly pulled out the classical scaling dimension to make the couplings λ_i dimensionless. Upon lowering the scale to μ we get from the scaling relation (2.6)

$$S_\mu = S_{\text{free}} + \int d^4x \mu^{\Delta_i-4}\lambda_i(\mu)\mathcal{O}_i, \tag{2.9}$$

where the couplings $\lambda_i(\mu)$ are written as functions of the scale since they get correction coming from loops of virtual particles with energies in the range $\mu_0 > E > \mu$. These corrections are a result of interactions and so are small for small λ_i, and only depend on the scale μ directly through logarithms in perturbation theory (since we have written the classical scaling explicitly). Note that since the momentum integration region is finite, these quantum corrections can suffer from neither UV nor IR divergences, even if the particles in the theory are massless. The effective coupling at scale μ then satisfies an RG equation of the form

$$\mu \frac{d\lambda_i}{d\mu} = (\Delta_i - 4)\lambda_i + \beta_i(\lambda_j) \tag{2.10}$$

where the beta function vanishes for vanishing λ_j. For example, for the marginal interaction $\mathcal{O} = \phi^4$ a one-loop calculation gives (heuristically)

$$\mu \frac{d\lambda}{d\mu} = +\lambda^2, \tag{2.11}$$

whose solution is

$$\lambda(\mu) = \frac{\lambda(\mu_0)}{1 + \log(\mu_0/\mu)}, \tag{2.12}$$

from which we see that $\lambda(\mu)$ decreases as $\mu \to 0$. Thus we see that the ϕ^4 interaction is actually irrelevant, and that the ϕ^4 theory is IR free. Note that this conclusion is only good provided $\lambda(\mu_0)$ is small enough to start with so that perturbation theory is reliable. At larger scales ($\mu \to \infty$) λ grows, so eventually the perturbative description breaks down and this is taken as an indication that some new degrees of freedom are needed to describe the physics at scales above the scale where $\lambda \sim 1$.

This analysis of the RG flow of effective theories in terms of classical scaling dimensions of operators and logarithmic quantum corrections is only valid near a free fixed point of the RG flow. There can also be interacting fixed point theories characterized (in part) by scaling dimensions which differ from the classical ones. Such differences of scaling dimensions from their classical values are called *anomalous dimensions*; we will see examples of theories with anomalous dimensions in later lectures. A conventional picture of a quantum field theory is as an RG flow between UV and IR fixed point theories; indeed all the quantum field theories for which we have a precise, non-perturbative definition (*e.g.* on the lattice) are defined in terms of the deformations by relevant operators of a *free* UV fixed point theory (*e.g.*Yang-Mills theory at zero coupling in four dimensions). All the theories we will be discussing in these lectures are of this type.[1]

[1]But one should be aware that a theory described by a local effective action below some scale, may

Among the interaction operators $\mathcal{O}_i$ generated in the RG flow of effective actions such as (2.7) can be terms proportional to the kinetic and mass terms of the fixed point (free) action itself. These terms play a special role. Corrections to the mass term are relevant, and so tend to grow in the IR; unless protected by some symmetry masses are unstable to quantum corrections. In such a case, as we flow to the IR we eventually reach a point where the effective mass of ϕ exceeds the scale μ of our effective action. But the reponse of ϕ to sources on energy scales below its mass are exponentially suppressed (it has no propagating modes) and so it decouples from the low energy physics. Essentially, its mass term dominates its kinetic term and fixes ϕ to be a constant at distances greater than its inverse mass. Thus ϕ acts like a constant in the low energy effective action: we can drop its kinetic term, "integrating out" all the ϕ degrees of freedom.

Corrections to the kinetic terms give rise to wave fucntion renormalization. By including the mass and kinetic terms corrections, we can rewrite the effective action (2.7) without separating out the free part as

$$S_\mu = \int d^4x \left\{ -\tfrac{1}{2} Z(\mu)(\partial\phi)^2 - \tfrac{1}{2}\mu^2 m^2(\mu)\phi^2 - \lambda(\mu)\phi^4 + \ldots \right\} \tag{2.13}$$

where we have introduced a *dimensionless* mass parameter $m(\mu)$. The wavefunction renormalization $Z(\mu)$ can be absorbed in a redefinition of the field variable,

$$\phi \to \phi_c \equiv \sqrt{Z}\phi, \tag{2.14}$$

which we will call the *canonically normalized* field ϕ_c. Rewriting the action in terms of the canonically normalized field variables gives rise to the canonical couplings $m_c(\mu) = m(\mu)/\sqrt{Z(\mu)}$ and $\lambda_c(\mu) = \lambda(\mu)/Z^2(\mu)$. It is the canonical couplings which we usually think of as the effective couplings governing the physics at the scale μ. Thus the effective mass at the scale μ is $\mu m_c(\mu)$. Note that this mass is *not* the physical mass (the position of the pole in the full propagator). Instead the physical mass should be found as the limit of the effective mass at arbitrarily long distances $\mu \to 0$:

$$m_{\text{phys}} = \lim_{\mu\to 0} \mu m_c(\mu). \tag{2.15}$$

More practically, the efective canonical mass should be the same as the physical mass for any scale $\mu < m_{\text{phys}}$ since the ϕ field decouples and undergoes no further RG

not necessarily have such a description in the UV as well. Even within the set of non-gravitational theories, its microscopic description may violate one or more of the assumed properties of locality, Lorentz invariance, or even of the rules of quantum mechanics itself: these properties could just be effective properties in the IR. For example, string theory arguments provide strong evidence for the the existence of an interesting class of non-gravitational six-dimensional Lorentz invariant "little string" theories which are less local than familiar quantum field theories [19].

running below this mass scale. (Indeed, equating the physical and effective canonical masses at the scale $\mu = m_{\rm phys}$ gives the condition $Z(m_{\rm phys}) = m^2(m_{\rm phys})$ which is the usual condition for the pole in the propagator.) Note that for a free theory, or below the decoupling scale, the condition that the canonical effective mass does not run implies that the wavefunction renormalization and dimensionless mass parameters should be related by

$$\frac{\sqrt{Z(\mu)}}{m(\mu)} = \frac{\mu}{m_{\rm phys}}. \tag{2.16}$$

This unexpected-looking scale dependence in a fixed point theory is just the result of our definition of $m(\mu)$ in (2.13).

The above discussion of wavefunction renormalization and the definition of canonical couplings is a reflection of the more general observation that the form of the effective action is not fixed, but is ambiguous up to general (nonsingular) field redefinitions. Such field redefinitions are just changes of variables and should not affect the physical content of the theory. In calculating an effective action, one chooses a specific regularization and a renormalization scheme to compute a definite effective action whose particular form will depend on the scheme, though the actual physics will not. Thus, in a more invariant formulation, the RG flow on the space of couplings should be replaced by a flow of the space of couplings modulo field redfinitions. This is particularly important when trying to determine the relevant and irrelevant operators in an effective theory, since field redefinitons may mix them.

All this you actually know very well already, as a simple example will show: Consider a scalar field theory with potential $V = -\phi^2 + \phi^{100}$. Though the ϕ^{100} term is very irrelevant by power counting, it is needed to stabilize the vacuum at $\langle\phi\rangle = (1/50)^{1/98}$. Shifting to this vacuum and expanding gives a potential $V \sim 2(100\widetilde{\phi}^2/2! + 100^2\widetilde{\phi}^3/3! + 100^3\widetilde{\phi}^4/4!)$ plus irrelevant terms.

Is example of irrelevant to relevant over flow.

Shows why nlsm action is right, essentially an expansion in the number of derivatives of the low energy fields.

Here the potential V is an arbitrary real function of the ϕ^i which is bounded below (for stability), while the coefficient g_{ij} of the generalized kinetic term is a real, symmetric and positive definite tensor (for unitarity).

Thus the nlsm includes the most relevant (in the colloquial sense!) terms in an effective action for determining the vevs of the scalar fields. Thus, with such an effective action, one can solve for the vacuum and expand about it. In this expansion, it is the kinetic scaling dimension which determines the relevant terms. Since kinetic and vacuum dimensions are different, terms which were relevant for determining the vacuum may no longer be relevant in the low energy physics by power counting.

Note that the derivative expansion that we are doing in getting the IREA effectively treats ϕ^i as dimensionless. In the usual discussions of perturbative quantum field theory, one assigns ϕ^i a scaling dimension of (mass). This is because we are interested in the scaling properties of the fluctuations of ϕ about a given vacuum, which are governed by the kinetic terms. But in determining the vacuum itself it is the potential that is important, and so the constant part of ϕ^i (the vevs) should be treated as dimensionless constants. In particular, taking the scale of the low energy effective action to be an energy E does *not* imply that only vacua with $\langle\phi^i\rangle < E$ should be allowed.

...comment on global an gauge symms in effective actions...

...IR free theories, Coleman-Gross ...

We will use low energy effective actions to analyze four dimensional field theories by taking the limit as the cutoff energy scale μ goes to zero, or equivalently, by just keeping the leading terms (up to two derivatives) in the low energy fields. I will call such $\mu \to 0$ low energy effective actions *IR effective actions*. Since an IR effective action describes physics only for arbitrarily low energies, it is, by definition, scale invariant: we simply take the cutoff scale μ below any finite scale in the theory. Scale invariant theories and therefore IREAs can therefore fall into one of the following categories:

Trivial theories in which all fields are massive, so there are no propagating degrees of freedom in the far IR. footnote about can make nonrel eff action around excitations in n-particle state...

Free theories in which all massless fields are non-interacting in the far IR.

(They can still couple to massive sources, but these sources should not be treated dynamically in the IREA.) An example is QED, in which the IREA describes free photons when the lightest charged particle is massive.

Interacting theories of massless degrees of freedom which are usually assumed to be conformal field theories [20]. have anom dimensions

We generally have no effective description of interacting conformal field theories in four dimensions so we must limit ourselves to free or trivial theories in the IR. A large class of these is given by the Coleman-Gross theorem [21] which states that for small enough couplings any theory of scalars, spinors, and $U(1)$ vectors in four dimensions flows in the IR to a free theory. We will therefore focus on IREAs with this field content. Note that other IR free theories are known, for example non-Abelian gauge theories with sufficiently many massless charged scalars and spinors. They will not play as important a role as the $U(1)$ theories, since even within supersymmetric theories they can be destabilized by adding mass terms.

We thus take the field content of our IREA to be a collection of real scalars ϕ^n, left-chiral Weyl spinors ψ_L^a, and $U(1)$ vector fields V_μ^A. Since this theory is free in the IR, no interesting dynamics involving the spinor fields like the formation of scalar condensates occurs (basically by definition). Thus the vacuum structure of this theory is governed by the scalar potential.

In that case the IREA can be written (excluding the spinors) as

$$\mathcal{L} = -V(\phi) + \frac{1}{2} g_{ij}(\phi) D_\mu \phi^i D^\mu \phi^j - \frac{1}{32\pi} \mathrm{Im}\left[\tau_{IJ}(\phi) f^I_{\mu\nu} f^{J\mu\nu}\right], \tag{2.17}$$

where, τ_{IJ} is a complex (gauge invariant) function of the ϕ^i symmetric in I and J and whose imaginary part is positive definite (for unitarity). Eq. 2.17 is called a gauged sigma model on target space.

Defining the real and imaginary parts of the couplings as

$$\tau_{IJ} = \frac{\theta_{IJ}}{2\pi} + i\frac{4\pi}{(e^2)_{IJ}}, \tag{2.18}$$

the generalized Maxwell term can be expanded to

$$\mathcal{L}_{U(1)} = -\frac{1}{4(e^2)_{IJ}} F^I_{\mu\nu} F^{J\mu\nu} + \frac{\theta_{IJ}}{64\pi^2} \epsilon^{\mu\nu\rho\sigma} F^I_{\mu\nu} F^J_{\rho\sigma}, \tag{2.19}$$

showing that the imaginary part of τ_{IJ} is a matrix of couplings and the real part are theta angles.

Since our IREA is is supposed to be free in the IR, we must comment on the meaning of the couplings τ_{IJ}. There are two kinds of vacua to consider. The first is one where a charged field (scalar or spinor) is massless. In this case the one-loop running of the $U(1)$ coupling implies that in the IR the coupling vanishes (corresponding to $\mathrm{Im}\tau \to +i\infty$). The second case is where all the charged fields are massive, in which case the $U(1)$ couplings stop running at energy scales below the mass of the lightest charged particle (just as the electromagnetic coupling is fixed at $\sim 1/137$ on scales below the electron mass). Thus, in this case the coupling $\mathrm{Im}\tau$ in the IREA is the strength of the gauge coupling to *massive (classical) sources.*

The theta angles are coefficients of topological (total derivative) terms in the action which count the instanton number of a given field configuration. Since this is an integer, the theta angles are indeed angles: $\theta_{IJ} \equiv \theta_{IJ} + 2\pi$, implying $\tau_{IJ} \equiv \tau_{IJ} + 1$. It is often remarked that there are no non-trivial instanton field configurations for $U(1)$ gauge groups in four-dimensional space-time, and thus no physics can depend on the θ_{IJ} for $U(1)$ theories. This is not correct for IREAs, however, since the theta angles are couplings to massive sources not described by the IREA fields. In the presence of such sources, the space-time manifold on which the IREA is defined is not all of $\mathbf{R}^4$, but should have the world-lines of the sources removed. On such manifolds there can be non-trivial $U(1)$ bundles, *i.e.* $U(1)$ gauge field configurations with non-zero instanton number. The basic example of this (realizable semi-classically) is when the microscopic theory is a non-Abelian gauge theory Higgsed down to $U(1)$ factors admitting magnetic monopole solutions, so that there are both electrically and magnetically charged sources in the $U(1)$ IREA. In the presence of such sources the instanton number is proportional to products of electric and magnetic charges present (and is an integer because of the Dirac quantization condition).

First, though, let us see how the addition of the $U(1)$ gauge fields affects the moduli space. Two points of target space which are related by a gauge transformation must be identified.

Note that the vacuum expectation values (vevs) of charged scalars can not parameterize the moduli space, because when a charged scalar gets a nonzero vev it Higgses the $U(1)$ it is charged under and thereby gets a mass. It is therefore not a flat direction—*i.e.* changing its vev takes us off the moduli space $\mathcal{M}$. So, since we are interested only in the extreme IR limit, we only need to keep the *neutral* scalars which parameterize $\mathcal{M}$. In this case the IREA Eq. 2.17 simplifies since $V = 0$ on $\mathcal{M}$ by definition and $D_\mu = \partial_\mu$ on neutral scalars. Thus only the metric $g_{ij}(\phi)$ and couplings $\tau_{IJ}(\phi)$ need to be specified. (If we included the fermions, there would also be the coefficient functions of their kinetic terms as well.)

It will be our mission in the rest of these lectures to determine the metric and $U(1)$ couplings on $\mathcal{M}$. Already in the non-supersymmetric case there is more that can be said about the properties of the coupling matrix τ_{IJ}, and is the topic of the next subsection.

2.2 Non-Renormalization Theorems

We will now determine the constaints on the IR effective action of a theory of left-chiral superfields coming from supersymmetry. Unlike global internal symmetries, a spontaneously broken supersymmetry does not imply a set of degenerate vacua related by supersymmetry transformations. Instead, as we have seen, there is typically a single vacuum in which the masses of the states within each supermultiplet are split by a characteristic amount, the scale of the supersymmetry breaking, μ_s. Below this scale there need be no effective supersymmetric description since the superpartners of the light states will have been integrated out. At scales above μ_s, on the other hand, an effective supersymmetric description of the theory (and of the spontaneous breaking of its supersymmetry) is possible by an appropriate choice of renormalization scheme. For the rest of this lecture we will assume that our effective theory is at a scale $\mu > \mu_s$ and that we are working in such a "supersymmetric renormalization scheme." We will save the discussion of examples of quantum theories which spontaneously break supersymmetry to the end of the course.

The classical scaling of the superfields can be determined by dimension counting as in the non-supersymmetric case. From the supersymmetry algebra, if we assign scaling dimension -1 to x^μ, we must assign dimensions to superspace quantities as follows:

	dimension
x^μ, dx^μ	-1
$\partial/\partial x^\mu$	$+1$
θ	$-\frac{1}{2}$
$d\theta$, $\partial/\partial\theta$	$+\frac{1}{2}$

Thus the classical scaling dimension of a left-chiral superfield Φ is $+1$, implying that its propagating components, ϕ and ψ, have their usual dimensions of $+1$ and $+3/2$, respectively. Recalling that the Kahler term can be written as an integral over all of superspace ($d^4x\, d^4\theta$) it follows that for the action to be scale invariant the Kahler potential must have dimension $+2$. Likewise, since the superpotential term can be written as an integral over "half" of superspace ($d^4x\, d^2\theta_L$) it follows that the superpotential must have dimension $+3$.

Flowing down in scale from $\mu_0 \to \mu$ in the IR free theory of coupled left-chiral superfields gives a new supersymmetric effective theory at the scale μ. The leading (two-derivative or two fermion) terms of such a supersymmetric effective action of left-chiral superfields at a scale μ_0 will be of the form we derived classically in section 1.5:

$$S_{\mu_0} = \int d^4x \left\{ \frac{1}{2} \left[\Phi_n^* \Phi^n\right]_D + \left[f(\Phi^n) + \text{c.c.}\right]_F \right\}, \tag{2.20}$$

where the superpotential can be written in general as a sum of terms

$$f = \sum_r \mu_0^{3-d_r} \lambda_r \mathcal{O}_r \tag{2.21}$$

with each term a product of left-chiral superfields

$$\mathcal{O}_r = \prod_i (\Phi^n)^{r_n} \tag{2.22}$$

for some integers r_n, with classical scaling dimensions

$$d_r = \sum_n r_n. \tag{2.23}$$

Here λ_r are the dimensionless effective couplings at the scale μ_0. (We should write a general Kahler term in (2.20) as well, but will stick with a quadratic one for simplicity.) Therefore the effective action at a scale $\mu < \mu_0$ is

$$S_\mu = \int d^4x \left\{ \frac{1}{2} \left[Z_n(\mu) \Phi_n^* \Phi^n \right]_D + \left[\mu^{3-d_r} \lambda_r(\mu) \mathcal{O}_r + \text{c.c.} \right]_F \right\}, \tag{2.24}$$

where we have included possible wavefunction and coupling renormalizations $Z_n(\mu)$ and $\lambda_r(\mu)$, which will depend on the couplings λ_r as well as the scale μ.

2.2.1 Holomorphy of the superpotential

The supersymmetry of the effective action implies that there is a renormalization scheme where the effective couplings at scale μ do not depend arbitrarily on the couplings at the "UV" scale μ_0, but only holomorphically on them [22]. To see this, think of all the coupling constants λ_r which appear in the superpotential at scale μ_0 as classical background left-chiral superfields (*e.g.* as very massive left-chiral superfields with their own superpotential terms which fix the vacuum expectation values of thier scalar components to the values λ_r). It then follows that these couplings can only appear in the effective superpotential holomorphically: only λ_r and not λ_r^* can appear in any quantum corrections to the superpotential, since the superpotential is a function only of left-chiral superfield's, not right-chiral superfield's.

Let us examine more closely the logic of this argument. In the first step we *assume* that the effective theory at the low energy scale μ is described by a supersymmetric theory with a specified set of left-chiral superfields. This is justified in the present case since the theories we are dealing with are IR free, so if they have a description at a scale μ_0 in terms of a certain set of left-chiral superfields, then at a lower scale they still will since the theory just flows to weaker coupling. Later, in strongly-coupled gauge theory examples, we will not have this argument at our disposal and will have to guess the a low energy field content, and then check that the guess is self consistent.

In either case, the next step is to think of the UV couplings as the lowest components of background left-chiral superfields. This step is just a trick—we are certainly allowed to do so if we like (since the couplings enter in the microscopic theory in the same way a background left-chiral superfield would). The point of this trick is that it makes the restrictions on possible quantum corrections allowed by supersymmetry apparent. These restrictions are just a supersymmetric version of "selection rules" familiar from other symmetries in quantum mechanics.

Perhaps an example from quantum mechanics will make this clear. A constant background electric field perturbs the Hamiltonian of a hydrogen atom by adding a term of the form

$$\delta H = E_1 x_1 + E_2 x_2 + E_3 x_3. \tag{2.25}$$

The resulting perturbed energy levels cannot depend on the perturbing parameters E_i arbitrarily. Indeed, one simply remarks that the electric field transforms as a vector $\mathbf{E}$ under rotational symmetries, thus giving selection rules for which terms in a perturbative expansion in the electric field strength it can contribute to. On the other hand, these selection rules are equally valid without the interpretation of the electic field as a background field transforming in a certain way under a symmetry (which it breaks). Instead, one could think of it as an abstract perturbation, and the selection rules follow simply because it is *consistent* to assign the perturbation transformation rules under the broken rotational symmetry. The holomorphy of the superpotential is the same sort of a selection rule, but this time following from supersymmetry: since the UV parameters enter into the action of the UV theory in the same way as the scalar components of chiral superfields do, it is consistent to *assign* these constants supersymmetry transformation properties as if they were the lowest components of chiral superfields.

We can immediately see the power of this supersymmetry selection rule. For suppose our enlarged theory, where we think of one of the couplings λ as a left-chiral superfield, has a $U(1)$ global symmetry under which λ has charge $Q(\lambda) = 1$, *i.e.* in the UV (scale μ_0) superpotential there is a term

$$f \supset \lambda \mathcal{O}_{-1} \tag{2.26}$$

where $\mathcal{O}_{-1}$ is some charge -1 operator. Say we are interested in the appearance of a given operator $\mathcal{O}_{-2}$ of charge $Q(\mathcal{O}_{-2}) = -2$ among the quantum corrections. Normally, one would say that this operator can appear at second and higher orders in perturbation theory in λ, as well as non-perturbatively:

$$\delta f \sim \lambda^2 \mathcal{O}_{-2} + \lambda^3 \lambda^* \mathcal{O}_{-2} + \ldots + \lambda^2 e^{-1/|\lambda|^2} \mathcal{O}_{-2} + \ldots, \tag{2.27}$$

assuming that there is a regular $\lambda \to 0$ limit, so that no negative powers of λ are allowed. However, by the above argument we learn that *only* the second-order term is allowed, all the higher-order pieces, including the non-perturbative ones, are disallowed since they necessarily depend on λ non-holomorphically.

Even more importantly, any operator of *positive* charge under the $U(1)$ symmetry is completely disallowed, since it would necessarily have to have inverse powers of λ as its coefficient. But since we assumed the $\lambda \to 0$ weak coupling limit was smooth (*i.e.* that the phsyics is under control there), such singular coefficients are disallowed. Note that this is again special to supersymmetry, for if non-holomorphic couplings were allowed, one could always include such operators with positive powers of λ^* instead.

This argument can be summarized prescriptively as follows [23]: The effective superpotential is constrained by

- holomorphy in the UV coupling constants,
- "ordinary" selection rules from symmetries under which the coupling constants may transform, and
- smoothness of the physics in various weak coupling limits.

Most of the rest of this course will consist in the systematic application of the above argument to the interesting and strongly coupled case of supersymmetric gauge theories.

A similar argument gives no such restrictions on the way the superpotential coouplings can enter in quantum corrections to the kahler potential. The reason is simply that the Kahler potential can depend on both left-chiral superfields and their conjugates, so superpotential couplings may enter non-holomorphically.

2.2.2 Nonrenormalization theorem for left-chiral superfields

We start by applying this argument to theories just of left-chiral superfields. The simplest case is the Wess-Zumino model of a single left-chiral superfield Φ with microscopic (scale μ_0) superpotential

$$f_{\mu_0} = \tfrac{1}{2}\mu_0\lambda_2\Phi^2 + \tfrac{1}{3}\lambda_3\Phi^3, \tag{2.28}$$

so that the UV effective action is

$$S_{\mu_0} = \int d^4x \left\{ \tfrac{1}{2}\left[\Phi^*\Phi\right]_D + \left[\tfrac{1}{2}\mu_0\lambda_2\Phi^2 + \tfrac{1}{3}\lambda_3\Phi^3 + \text{c.c.}\right]_F \right\}. \tag{2.29}$$

By holomorphy, the effective superpotential at scale $\mu < \mu_0$ is

$$f_\mu = f_\mu(\Phi, \lambda_2, \lambda_3; \mu, \mu_0), \tag{2.30}$$

that is, a function of Φ, λ_2, and λ_3 and not their complex conjugates. (Note that we have made the assumption that the effective theory is still described in terms of a single chiral superfield Φ at the IR scale μ.) The microscopic superpotential is invariant under a global $U(1) \times U(1)_R$ symmetry if we assign the coupling constants appropriate charges:

	$U(1)$	$\times$	$U(1)_R$
Φ	$+1$		$+1$
λ_2	-2		0
λ_3	-3		-1

This implies selection rules for these "symmetries" constraining the effective superpotential to be neutral under the first $U(1)$ and have charge $+2$ under the $U(1)_R$ (as is usual for an R symmetry). Thus the effective superpotential must have the form

$$f_\mu = \mu\lambda_2\Phi^2 g\left(\frac{\lambda_3\Phi}{\mu\lambda_2}\right) \tag{2.31}$$

where g is an arbitrary holomorphic function, and we have put in the powers of μ according to the classical scaling dimensions of Φ and the superpotential. (There can also be a μ_0 dependence in g.) Now, in the $\lambda_3 \to 0$ limit, keeping λ_2 fixed, the theory is free, so only terms with non-negative integers powers of λ_3 can appear in the expansion of g:

$$f_\mu = \sum_{n\geq 0} g_n\, \mu^{1-n}\lambda_2^{1-n}\lambda_3^n\Phi^{n+2}, \tag{2.32}$$

where the g_n are undetermined constants (which can be functions of μ/μ_0). Furthermore, we can also take the $\lambda_2 \to 0$ limit at the same time to conclude that terms with $n > 1$ are disallowed. So we learn that the effective action at scale μ is given by

$$S_\mu = \int d^4x \left\{\tfrac{1}{2}Z\left[\Phi^*\Phi + \ldots\right]_D + \left[g_0\mu\lambda_2\Phi^2 + g_1\lambda_3\Phi^3 + \text{c.c.}\right]_F\right\} \tag{2.33}$$

where we have included a wavefunction renormalization Z of the Kahler term; note that the Kahler term may receive other corrections (*e.g.* $(\Phi^*\Phi)^2$ terms and so on) which we have not written.

It remains to determine the constants g_0 and g_1 which may depend only on μ/μ_0. In the limit that $\lambda_3 \to 0$ the theory is free, with a mass (found after rescaling to canonically normalized fields) $2g_0\mu\lambda/Z$. Equating this to the mass in the UV action (2.29) gives

$$g_0 = \frac{1}{2}\frac{\mu_0}{\mu}Z. \tag{2.34}$$

Now, as discussed in the last lecture, the wavefunction renormalization Z of a free theory is a matter of a choice of the choice of field variables (or, equivalently, a choice of RG scheme). As in the last lecture, it is convenient to choose a scheme in which the masses, like the other couplings, scale with the cut off scale μ to their classical scaling dimension. In this scheme (at zero coupling) we therefore choose

$$Z = \mu/\mu_0 \qquad (\text{when } \lambda_3 = 0) \tag{2.35}$$

giving $g_0 = \frac{1}{2}$. When $\lambda_3 \neq 0$, Z will in general receive corrections shifting it from the above value.

g_1 can now be determined by matching the results of perturbation theory in λ_3 between (2.29) and (2.33). Since the Φ^3 vertex appears in both proportional to the same coupling λ_3, they must match at tree level (*i.e.* classically) and we find immediately that $g_1 = 1/3$. Our result for the effective superpotential is

$$f_\mu = \tfrac{1}{2}\mu\lambda_2\Phi^2 + \tfrac{1}{3}\lambda_3\Phi^3. \tag{2.36}$$

Thus we have shown that the superpotential is non-perturbatively un-renormalized: the low energy superpotential suffered no quantum corrections relative to the UV superpotential, and only differs from it by the classical scalings (which are a matter of a choice of renormalization scheme).

The step above eqn. (2.33) of taking the $\lambda_2 \to 0$ limit deserves a few words. Taking this limit at finite λ_3 does not lead to a free theory; however, by taking *both* λ_3 and λ_2 to zero such that $\lambda_2/\lambda_3 \to 0$, we achieve the desired result. One may wonder, though, whether the $\lambda_2 \to 0$ limit is really smooth, for though the resulting theory is free, it also has a massless particle, and so the effective theory should have IR divergences—perhaps reflected in divergences of the superpotential? This is not the case, though, since we do not do the RG running all the way down to $\mu = 0$, and so we will not see any IR divergences in our Wilsonian effective action.[2]

Let us generalize non-renormalization result a bit further, to a more complicated example:

$$f_{\mu_0} = \mu_0^2\lambda_1\Phi + \mu_0\lambda_2\Phi^2 + \ldots + \mu^{3-r}\lambda_r\Phi^r + \ldots \tag{2.37}$$

which has the global symmetries

	$U(1)$	$\times$	$U(1)_R$
Φ	$+1$		$+1$
λ_r	$-r$		$2-r$

[2]This should be contrasted with the one particle irreducible (1PI) effective action introduced in standard texts in which the effective scale is zero and IR divergences from massless particles do occur.

implying

$$f_\mu = \mu^2 \lambda_1 \Phi\, g\left(\frac{\lambda_2}{\lambda_1}\frac{\Phi}{\mu}, \frac{\lambda_3}{\lambda_1}\frac{\Phi^2}{\mu^2}, \ldots, \frac{\lambda_r}{\lambda_1}\frac{\Phi^{r-1}}{\mu^{r-1}}, \ldots\right). \tag{2.38}$$

Demanding a smooth limit as all $\lambda_r \to 0$ then implies f_μ is the same as the UV superpotential f_{μ_0} except for the classical scaling from $\mu_0 \to \mu$.

This line of argument can be easily generilzed to an arbitrary superpotential with an arbitrary number of left-chiral superfields as in (2.20). A useful trick for dealing with the general case is to replace the UV superpotential $f(\Phi^n, \mu_0)$ with $Y \cdot f(\Phi^n, \mu_0)$ where Y is an auxiliary left-chiral superfield. Setting the scalar component of Y to 1 (and the others components to zero) gives the original theory back. The superpotential of the enlarged theory is linear in Y, the theory is invariant under a $U(1)_R$ symmetry with charges $R(Y) = +2$ and $R(\Phi^n) = 0$. This symmetry and holomorphy is enough to tell us that the effective superpotential has the form $f_\mu = Y g(\Phi^n, \mu)$. Taking the $Y \to 0$ limit in which the theory becomes free and matching the UV and IR effective actions in perturbation theory as above, then implies that $g(\Phi^n, \mu) = f(\Phi^n, \mu)$. Setting $Y = 1$ then gives the general non-renormalization theorem for theories of left-chiral superfields: the UV action (2.29) at scale μ_0 flows to

$$S_\mu = \int d^4x \left\{ \frac{1}{2}\left[Z_n \Phi^*_n \Phi^n\right]_D + \left[\mu^{3-d_r}\lambda_r \mathcal{O}_r + \text{c.c.}\right]_F \right\}. \tag{2.39}$$

The only difference from the general form (2.24) of the effective action at scale μ expected from general RG group arguments is that the couplings λ_r in (2.39) are the same as those in the microscopic theory (2.20). Note that this non-renormalization result shows no contradiction with our assumption that the low-energy degrees of freedom are described byt the same set of left-chiral superfields as in the microscopic theory, in line with our expectations from the Coleman-Gross theorem that scalar and spinor field theory is IR free.

2.2.3 Kahler term renormalization

In order to compare the couplings of the effective action (2.39) to physical effective couplings that would be measured in, say, a scattering experiment with energy transfer of order μ, we should normalize the kinetic terms to their canonical form. We therefore define the *canonical* left-chiral superfields by rescaling them to absorb the wavefunction renormalization factors

$$\Phi^n \to \Phi^n_{\text{cn}} \equiv \sqrt{Z_n(\mu)}\, \Phi^n. \tag{2.40}$$

Then the rescaled action has the same form as the Wilsonian one, but with the superpotential couplings replaced by canonical ones:

$$\lambda_r^{\text{cn}}(\mu) \equiv \left(\frac{\mu}{\mu_0}\right)^{3-d_r} \left(\prod_n Z_n^{-r_n/2}\right) \lambda_r. \tag{2.41}$$

This immediately implies the exact RG equation for the physical superpotential couplings

$$\mu \frac{d\lambda_r(\mu)}{d\mu} = \lambda_r(\mu) \left(3 - d_r - \frac{1}{2} \sum_n r_n \gamma_n(\mu)\right), \tag{2.42}$$

where we have defined the *anomalous dimension* of the Φ^n left-chiral superfield as

$$\gamma_n(\mu) \equiv \frac{d \log Z_n(\mu)}{d \log \mu}. \tag{2.43}$$

Of course, since we have, in general, no exact method of computing these anomalous dimensions, the usefulness of the exact RG equation (2.42) is limited. Nevertheless, we will see an interesting application of it a few lectures from now. Note that the RG equation for the canonical effective mass $m_n^{\text{cn}}(\mu)$, defined as the coefficient of the term $(\Phi_{\text{cn}}^n)^2$ in the canonically normalized superpotential is $dm_n^{\text{cn}}/d \log \mu = m_n^{\text{cn}}(1 - \gamma_n)$. So γ_n is the anomalous dimension of the mass.

In a weakly coupled theory we can compute the wavefunction renormalization of the Kahler potential in perturbation theory. For example, in a Wess-Zumino model with superpotential $f_{\mu_0} = \lambda \Phi^3$, the one loop diagram renormalizing the fermion kinetic term using the Yukawa coupling derived from the superpotential gives

$$Z = 1 + \lambda\lambda^* \log \left|\frac{\mu_0}{\mu}\right| + \ldots \tag{2.44}$$

where the first term is the tree result, the logarithm is the usual one-loop contribution (determined by the symmetries), and the sign is correct since as $\mu \to 0$, $Z \to +\infty$, so that in the IR the theory becomes weakly coupled in the Φ_{cn} variables.

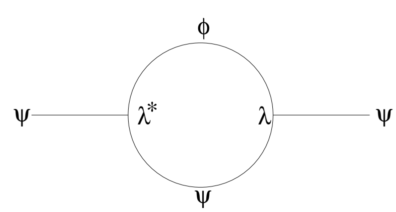

Figure 2.1: One-loop diagram renormalizing the Kahler potential.

A more interesting example arises in a theory with two left-chiral superfields and superpotential

$$f = \lambda \Phi_1 \Phi_2^2, \tag{2.45}$$

in an effective theory at a scale μ_0. This was an example we discussed from a classical point of view in section 1.6.4. There we found, by extremizing f, supersymmetric vacua at

$$\Phi_2 = 0, \qquad \Phi_1 = \text{arbitrary}, \tag{2.46}$$

implying a whole moduli space, $\mathcal{M}$, of degenerate but inequivalent classical ground states. By the non-renormalization theorem of the superpotential, this conclusion does not change once quantum effects are taken into account.

But quantum effects can renormalize the Kahler potential and thus change the metric on $\mathcal{M}$ from its classical value. Since the microscopic Kahler potential of is $K = \Phi_1^*\Phi_1 + \Phi_2^*\Phi_2$, the metric induced on $\mathcal{M}$ is

$$ds^2_{\text{class}} = d\Phi_1\, d\Phi_1^*, \tag{2.47}$$

classically. Now the classical spectrum at any point on $\mathcal{M}$ is one massless chiral multiplet Φ_1 and one massive chiral multiplet Φ_2 with mass proportional to $|\lambda\langle\Phi_1\rangle|$. The Kahler potential will recieve quantum corrections at scales above the mass of Φ_2 coming from virtual Φ_2 states contributing to loops in the Φ_1 propagators. In perturbation theory

$$K = +\Phi_1^*\Phi_1 - \#\Phi_1^*\Phi_1|\lambda|^2 \log\left|\frac{\Phi_1}{\mu_0}\right|^2 + \ldots \tag{2.48}$$

where the first term is the tree level result, and the second comes from the massive Φ_2 at one loop which contributes the usual logarithm of its mass over the cut off mass μ_0. (We are assuming here that the scale of the IR effective action, μ, is smaller than the Φ_2 mass.) The sign of the one loop term leads to a growing K as $\Phi_1 \to 0$, which in turn implies that the canonically normalized effective couplings are going to zero in this limit. Thus the one loop perturbative result becomes exact as we approach the origin of $\mathcal{M}$. Since the Kahler metric is given by $(ds)^2 = g_{1\bar{1}} d\Phi_1 d\Phi_1^*$ with

$$g_{1\bar{1}} = \partial_1\partial_{\bar{1}} K \simeq -|\lambda|^2 \log \Phi_1\Phi_1^* + \text{const}, \tag{2.49}$$

we see that at $\Phi_1 = 0$ there is a metric singularity. It is easy to check that this singularity is at finite distance but that the integrated curvature in a neighborhood of the origin diverges. Thus the moduli space which classically is the complex Φ_1-plane, has a cusp-like singularity quantum mechanicaly. (Note that for Φ_1 large the metric becomes negative indicating a breakdown of unitarity. This is just an artefact of the one loop perturbative expansion: at large Φ_1 the effective coupling is growing and perturbation theory becomes unreliable.) The singularity at $\Phi_1 = 0$ has a physical interpretation: it corresponds to the fact that when $\Phi_1 = 0$ a particle multiplet (Φ_2) is becoming massless. In this case the assumption under which we computed the IR effective action, namely that its scale μ was less than the mass of Φ_2, breaks down. In general, singularities in IR effective actions are a sign of new massless degrees of freedom which must be included in the IR effective action to get a sensible low energy description of the physics there.

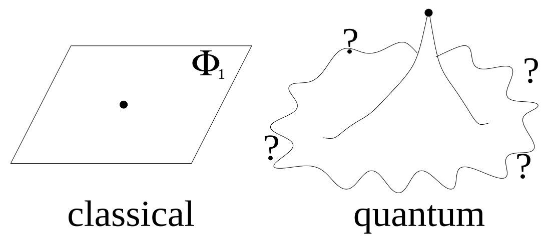

Figure 2.2: The moduli space $\mathcal{M}$ classically and quantum mechanically

2.3 Quantum gauge theories

Before discussing effective actions for supersymmetric gauge theories and the non-renormalization theorems that apply to them, we need to review some basic facts about quantum gauge theories.

2.3.1 Gauge couplings

RG flow

The gauge kinetic term in an effective action at scale μ is

$$\frac{1}{2g^2(\mu)}\mathrm{tr}_f(f^2) \tag{2.50}$$

where $g^2(\mu)$ is the coupling at that scale. The one loop RG implies the coupling "runs" as a function of scale according to

$$\mu\frac{dg}{d\mu} = -\frac{b}{16\pi^2}g^3 + \mathcal{O}(g^5) \tag{2.51}$$

where the first coefficient b of the beta function is given by

$$b = \frac{11}{6}T(\mathrm{adj}) - \frac{1}{3}\sum_a T(\mathbf{r}_a) - \frac{1}{6}\sum_n T(\mathbf{r}_n), \tag{2.52}$$

where the sum on a is over Weyl fermions with the ath fermion in the $\mathbf{r}_a$ representation of the gauge group, and the sum on n is over complex bosons in representations $\mathbf{r}_n$. Recall that $T(\mathbf{r})$ is the index of the representation $\mathbf{r}$; for $SU(N)$, for example, the index of the fundamental representation is 1, and of the adjoint representation is $2N$. Only fields with masses less than the scale μ will contribute to loops, so only these light fields should be included in the sums in (2.52). The solution at one loop order to (2.51) is

$$\frac{1}{g^2(\mu)} = -\frac{b}{8\pi^2}\log\left(\frac{\Lambda}{\mu}\right), \tag{2.53}$$

where we have defined

$$\Lambda \equiv \mu_0 e^{-8\pi^2/bg^2(\mu_0)}, \tag{2.54}$$

the *strong coupling scale* of the gauge group, where μ_0 is any fixed scale and $g^2(\mu_0)$ is the value of the effective coupling there. By (2.51) Λ is independent of the choice of scale μ_0. When the scale of the effective theory approaches Λ, we see that the effective coupling diverges; of course, when this happens the one loop approximation to the

RG running is no longer valid, and higher loop and non-pertrubative effects should be taken into account. One should think of Λ as the approximate scale at which the effective gauge coupling becomes strong (of order one). The trading of the information of a gauge coupling at a given scale for the strong coupling scale of the gauge group

$$\{g^2(\mu_0), \mu_0\} \leftrightarrow \Lambda, \tag{2.55}$$

is known as "dimensional transmutation". In a theory with many gauge groups, $G_1 \times G_2 \times \cdots \times G_n$, there will be correspondingly many gauge group scales $\{\Lambda_1, \cdots, \Lambda_n\}$.

The behavior of the effective coupling has qualitatively different behaviors in the IR depending on the sign of b. For $b > 0$, the coupling is weak in the UV and runs to strong coupling in the IR. Such theories are *asymptotically free* gauge theories. Any theory with a non-Abelian gauge group and no matter (*i.e.* no fermions or scalars in non-trivial representations of the gauge group) will automatically be asymptotically free since only the first term in (2.52) contributes. Adding charged matter can only reduce b. If $b < 0$ then the theory is weakly coupled in the IR and runs to string coupling in the UV. Non-Abelian theories with enough light charged matter will be IR free. Also any Abelian ($U(1)$) theory with a light charged field will be IR free. (Recall that $U(1)$ irreducible representations are all one dimensionsal and are described by their charge, q. The formula (2.52) applies to $U(1)$ gauge factors as well, using $T(q) = q^2$ and recalling that the adjoint representation has $q = 0$.)

If a gauge theory with gauge group G is Higgsed by a charged scalar getting vacuum expectation value ϕ, breaking G to a subgroup H, then the strong coupling scales Λ_G and Λ_H of the two groups are related by matching the RG flows of their couplings (2.53) at the scale ϕ:

$$g^2(\phi)_G = g^2(\phi)_H + \text{const.} \qquad \Rightarrow \qquad \Lambda_H^{b_H} \sim \Lambda_G^{b_G} \phi^{b_H - b_G}, \tag{2.56}$$

where b_G and b_H are the coefficients of the beta function for G and H; see the figure. If $b_G > 0$ so that G is asymptotically free, and $\phi \gg \Lambda_G$, then the Higgsing takes place at weak coupling and the one loop running of the gauge coupling used to do the above matching is a good approximation. Even at one loop the matching in (2.56) is uncertain up to a constant factor due to a scheme dependent one loop *threshold correction* [24].

2.3.2 ϑ angles and instantons

Gauge theories contain another term built solely out of the gauge fields:

$$S_\vartheta = \int d^4x \, \frac{\vartheta}{16\pi^2} \mathrm{tr}_f(f\widetilde{f}) = \int d^4x \, \frac{\vartheta}{8\pi^2} \epsilon^{\mu\nu\rho\sigma} \partial_\mu \mathrm{tr}_f \left(V_\nu \partial_\rho V_\sigma - i\frac{2}{3} V_\nu V_\rho V_\sigma \right) \tag{2.57}$$

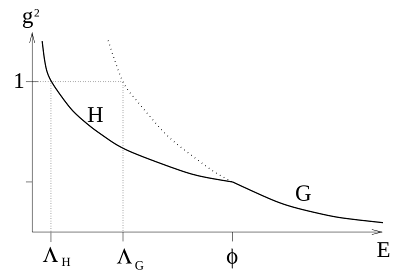

Figure 2.3: Running of the coupling of an asymptotically free gauge theory G Higgsed to H at a scale $\phi \gg \Lambda_G$.

where V_μ is the gauge potential. Since this term is a total derivative it has no effect on the classical equations of motion; however, quantum mechanically it can have an effect since we average over all fluctuations of the gauge fields, not just those satisfying the classical equations of motion. As we will briefly review below, this term is sensitive to *instanton* field configurations; see [25, chapter 7] for a pedagogical introduction to this subject.

Because the ϑ term is a total derivative, it can be expressed as an integral over a surface, which we can think of as a large 3-sphere, S^3, at infinity. For the integral to be finite the field strengths should vanish at infinity, so the gauge potential should be pure gauge there:

$$V_\mu = -i(\partial_\mu g)g^{-1}. \tag{2.58}$$

where $g(x_\mu)$ is an element of the gauge group. Plugging this into (2.57) gives, using the identity $\partial_\mu g^{-1} = -g^{-1}(\partial_\mu g)g^{-1}$,

$$\begin{aligned} S_\vartheta &= -\frac{\vartheta}{24\pi^2}\int d^4x\, \epsilon^{\mu\nu\rho\sigma}\partial_\mu \mathrm{tr}_f\left[g^{-1}(\partial_\nu g)g^{-1}(\partial_\rho g)g^{-1}(\partial_\sigma g)\right], \\ &= -\frac{\vartheta}{24\pi^2}\int_{S^3} d^3\xi\, \epsilon^{abc}\mathrm{tr}_f\left[g^{-1}(\partial_a g)g^{-1}(\partial_b g)g^{-1}(\partial_c g)\right], \end{aligned} \tag{2.59}$$

where in the second line we have written it as a surface integral over the 3-sphere at infinity parameterized by some coordinates ξ_a, $a = 1, 2, 3$. This integral computes the "winding number" of $g(x_\mu)$ around the 3-sphere at infinity and is a topological invariant. More concretely, if we choose g to be in an $SU(2)$ subgroup of the non-Abelian gauge group and of the form

$$g = (g_1)^n, \qquad g_1(x_\mu) = \frac{t + i\mathbf{x}\cdot\boldsymbol{\sigma}}{\sqrt{t^2 + \mathbf{x}^2}}, \tag{2.60}$$

then it is not hard to compute that

$$S_\vartheta[g] = n\vartheta. \tag{2.61}$$

The integer n measuring the winding number of the gauge field configuration is called the *instanton number* of the gauge field configuration. Since it is a topological invariant it will not change under arbitrary continuous deformations of the fields. Since the path integral is over continuous field configurations, and computes $\int \mathcal{D}\phi \ldots e^{iS_\vartheta}$, so

$$\vartheta \to \vartheta + 2\pi \tag{2.62}$$

is an exact symmetry of the theory. More properly, it is an exact equivalence of quantum field theories, since ϑ is a parameter in the action defining the theory. In terms of the complexified gauge coupling $\tau = (\vartheta/2\pi) + i(4\pi/g^2)$, this equivalence reads

$$\tau \simeq \tau + 1. \tag{2.63}$$

This result holds for all simple non-Abelian gauge groups, essentially since they all have $SU(2)$ subgroups.

The physical meaning of the instanton field configurations becomes clearer if we take the surface at infinity to be shaped like a large cylinder as in the figure. Going to temporal gauge, in which $V_t = 0$, it is follows that the contribution to the surface integral for S_ϑ coming from the cylindrical part of the surface vanishes, and S_ϑ reduces to the difference of the integrals over all space of pure gauge configurations on the two "caps" at $t = \pm\infty$. Thus an instanton field configuration interpolates between two vacua of the gauge theory. Both vacua are gauge equivalent (by construction) to the usual vacuum with zero gauge potential, but the gauge transformations which connect them to the usual vacuum cannot be continuously deformed to the identity. For if they could, then we could continuously deform the gauge transformation on the bottom cap to that of the top cap, so the gauge fields in the whole of space-time would be pure gauge and therefore $f_{\mu\nu} = 0$ everywhere. But then S_ϑ would vanish in contradiction to (2.61). Gauge transformations which cannot be continuously connected to the identity are called *large gauge transformations*. The fact that they are weighted differently in the action (by the S_ϑ term) implies that configurations related by large gauge transformations, unlike ones related by gauge transformations connected to the identity, are *not* identified as physically equivalent states. In particular, the different vacua represent a real infinite classical degeneracy of vacua in non-Abelian gauge theory.

The fact that the field strength can not vanish identically for configurations with non-zero instanton number implies that there is necessarily a field energy associated with the gauge field configuration interpolating between the different vacua. Thus there is an energy barrier separating these vacua. Though classically forbidden, there can be quantum mechanical tunnelling between these vacua, lifting their degeneracy. As is usual in quantum mechanics, the tunnelling amplitude is e^{-S_E} where S_E is the Euclidean action of a field configuration which interpolates between the different vacua.

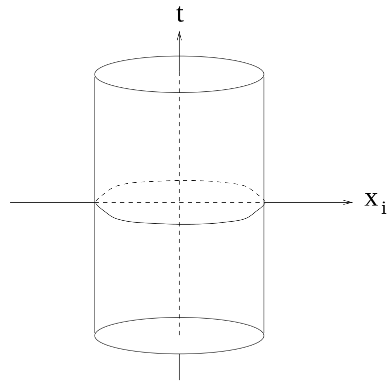

Figure 2.4: Surface "at infinity".

We have seen above that instanton configurations are just such interpolating field configurations. We can put a lower bound on the Euclidean action of an instanton using

$$0 \leq \int d^4x \mathrm{tr}_f (f \pm \widetilde{f})^2 = \int d^4x [2\mathrm{tr}_f(f^2) \pm 2\mathrm{tr}_f(f\widetilde{f})] \tag{2.64}$$

implying that

$$\int d^4x \mathrm{tr}_f(f^2) \geq \left| \int d^4x \mathrm{tr}_f(f\widetilde{f}) \right| = 16\pi^2 |n|. \tag{2.65}$$

This inequality is saturated for self dual or anti-self dual configurations with $f = \pm \widetilde{f}$. (Note that in Euclidean space $\widetilde{\widetilde{f}} = f$, unlike Minkowski space where there is a minus sign in this relation.) The general solution for the self dual one instanton $SU(2)$ Euclidean gauge field configuration in $\partial_\mu V_\mu = 0$ gauge is [26]

$$V_\mu(x + x_0) = \frac{(x \cdot s)s^\dagger_\mu - x_\mu}{x^2 + \rho^2} \tag{2.66}$$

where x_μ is the Euclidean coordinate 4-vector, $s_\mu = (1, i\boldsymbol{\sigma})$ are constant matrices, x_0 is an arbitrary constant 4-vector determining the position of the center of the instanton, while ρ is another arbitrary constant determining the size of the instanton.

Eqn. (2.65) implies that n-instanton contributions to amplitudes will be suppressed by factors of (at least)

$$e^{-S_E} = \left(e^{-8\pi^2/g^2} \right)^{|n|} = \left(\frac{\Lambda}{\mu} \right)^{|n|b}, \tag{2.67}$$

and so are non-perturbative effects going as a power of the gauge group strong coupling scale. Note the appearance of the RG scale μ in this formula. In evaluating the contribution of, say, a one instanton configuration to a given process, one must also integrate over the arbitrary parameters x_0 and ρ appearing in the solution (2.66), giving rise to terms like

$$\int d^4x_0 \int \frac{d\rho}{\rho^5} e^{-8\pi^2/g^2(1/\rho)}, \tag{2.68}$$

where the ρ^{-5} is to get the dimensions right, and the running coupling constant $g^2(\mu)$ is naturally evaluated at the scale $\mu = 1/\rho$ of the size of the instanton. One interprets $\int d^4x_0$ (after exponentiating the one instanton contribution in a "dilute instanton gas approximation") as the space-time integration of a term in the effective Lagrangian induced by the instanton. However, since

$$e^{-8\pi^2/g^2(\rho)} = (\rho\Lambda)^b, \tag{2.69}$$

we see that the ρ integration is IR divergent (*i.e.* as $\rho \to \infty$) for asymptotically free gauge theories ($b > 0$).[3] Of course, the theory becomes strongly coupled in the IR below the scale Λ, and one might expect a semi-classical approximation based on the microscopic (UV) description of the theory to break down. This prevents us from reliably calculating the non-perturbative instanton contributions to amplitudes directly in asymptotically free gauge theories. A situation in which instanton effects *can* be reliably computed is one where the gauge group of an asymptotically free theory is Higgsed ("broken") down to Abelian ($U(1)$) gauge factors (which are IR free) or is completely broken; see the figure. In this case the vacuum expectation value, ϕ, Higgsing the gauge group cuts off the instanton scale integral, $\rho > \phi$, rendering it finite. Furthermore, if $\phi \gg \Lambda$, then the whole instanton contribution is calculated at weak coupling where the semi-classical tunnelling methods are applicable. We will see examples of such instanton effects later in the course, though we will not compute them directly using semi-classical instanton techniques, but indirectly using supersymmetric non-renormalization theorems.

Finally we should note that the instanton configurations we have described depend crucially on the non-Abelian nature of the gauge group and they do not exist in Abelian gauge theories on Minkowski space. For this reason is often remarked that no physics can depend on the theta angles, ϑ_{AB}, in a $U(1)^N$ gauge theory. This is not correct for the theta angles in IR effective actions, however, since there the theta angles are couplings to massive sources not described by the IR effective action fields. In the presence of such sources, the space-time manifold on which the IR effective action is defined is not all of $\mathbb{R}^4$, but should have the world lines of the sources removed. On such manifolds there can be non-trivial $U(1)$ bundles, *i.e.* $U(1)$ gauge field configurations with non-zero instanton number. The basic example of this (realizable semi-classically)

[3] If $0 < b < 4$ the one instanton contribution is finite, but the n instanton contribution for sufficiently large n will still diverge.

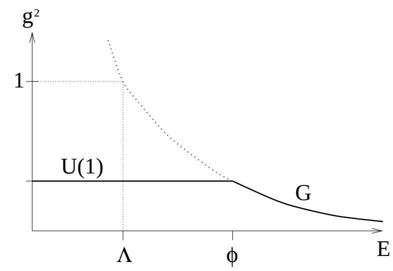

Figure 2.5: Running of the coupling of an asymptotically free gauge theory with gauge group G Higgsed to $U(1)$'s at a scale $\phi \gg \Lambda$. The $U(1)$ couplings do not run below ϕ only because we have assumed there are no charged fields lighter than ϕ; otherwise they would run to even weaker couplings.

is when the microscopic theory is a non-Abelian gauge theory Higgsed down to $U(1)$ factors admitting magnetic monopole solutions, so that there are both electrically and magnetically charged sources in the $U(1)$ IR effective action. In the presence of such sources the instanton number is proportional to products of electric and magnetic charges present and is an integer because of the Dirac quantization condition. Since this is an integer, the theta angles are indeed angles: $\vartheta_{AB} \simeq \vartheta_{AB} + 2\pi$, implying $\tau_{AB} \simeq \tau_{AB} + 1$. This will play an important role in the discussion of electric-magnetic duality in later lectures.

2.3.3 Anomalies

Anomalies refer to classical symmetries which are broken by quantum effects. This means that in the full quantum theory there is no (gauge invariant or covariant) conserved current for an anomalous symmetry. This is important in the case of classical global symmetries, implying as it does that the classical Ward identities are violated, but it does not affect the consistency of the theory. A familiar and important example of an anomalous symmetry is scale invariance: as we saw above, quantum effects in a classically scale invariant Yang-Mills theory make the gauge coupling run with scale. Another kind of anomaly, the *chiral anomaly*, occurs in the conservation of the currents for chiral rotations. If anomalous chiral rotations are gauged, then the resulting theory is *inconsistent*, since we only know how to couple spin-1 fields (gauge fields) in a unitary way to conserved currents. This places restrictions on the allowed gauge group representations of fermions in gauge theories. Chiral anomalies (local or global) arise in four dimensional quantum field theories only in theories where fermions with chiral symmetries are coupled to gauge fields. They can be computed in perturbation theory and only occur at one loop. This is a reflection of the fact that they can also be thought of as IR effects. From this perspective, the existence of anomalies depends only on the field content and charges of the light fermions in the theory, and not on details of the interactions. In what follows I will merely summarize the origin and systematics of anomalies. A discussion of anomalies which is mostly complementary to the approach I'll take here are the introductory sections of [27].

Chiral symmetries are symmetries in which left- and right-handed Weyl fermions transform differently. Consider a symmetry group G with generators t_A, and left-handed Weyl fermions transforming in the $\mathbf{r}$ representation,

$$\delta\mathcal{P}_+\psi^b = i\alpha^A (t_A^{(\mathbf{r})})^b_a \mathcal{P}_+\psi^a, \tag{2.70}$$

where α^A are real constants (the symmetry transformation parameters), $a, b = 1, \ldots, r$ are the $\mathbf{r}$ representation indices, and, as usual in these lectures, ψ is a Majorana fermion, so that $\mathcal{P}_+\psi$ is its left-handed Weyl part. Taking the complex conjugate of (2.70), and using the reality condition defining Majorana fermions, gives

$$\delta\mathcal{P}_-\psi = -i\alpha^A (t_A^{(\mathbf{r})})^T \mathcal{P}_-\psi, \tag{2.71}$$

where we have also used the Hermiticity of the t_A's. We see that the right-handed Weyl component transforms in a representation with generators $-t_A^T$, which is the complex conjugate $\overline{\mathbf{r}}$ representation. Thus chiral symmetries are those in which the fermions transform in complex representations.

In a free field theory (for simplicity) of a Majorana fermion, the conserved currents for the symmetry group G are

$$j_A^\mu = \frac{1}{2}\overline{\psi}\gamma^\mu(\mathcal{P}_+ t_A^{(\mathbf{r})} + \mathcal{P}_- t_A^{(\overline{\mathbf{r}})}\psi = \overline{\psi}\gamma^\mu\mathcal{P}_+ t_A^{(\mathbf{r})}\psi, \tag{2.72}$$

where the second equality follows from interchanging $\overline{\psi}$ and ψ in the second term and using the Reality condition for Majorana fermions. The 3-point function of these currents at one loop will have the form

$$\langle j_A^\mu(x_1) j_B^\nu(x_2) j_C^\rho(x_3)\rangle = \mathrm{tr}_r(t_A t_B t_C) f_{\mu\nu\rho}(x_i) \tag{2.73}$$

for some function $f_{\mu\nu\rho}$, where the trace comes from contracting group generators around the loop. For three global currents this correlator, though having some interesting properties which we will discuss below, does not imply a violation of the conservation law

$$\partial_\mu j_A^\mu = 0 \tag{2.74}$$

for the global symmetry currents.

Now couple these currents (or perhaps only a subset of them) to gauge fields in the usual way,

$$\mathcal{L} = \mathcal{L}_{free} + \sum V_\mu^A j_A^\mu, \tag{2.75}$$

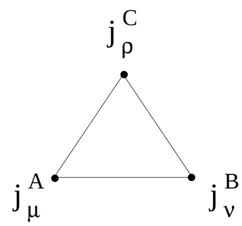

Figure 2.6: One loop diagrom for the 3-point correlator of currents.

and compute $\langle jVV\rangle$, taking care to regulate, impose Bose symmetry on the gauge fields, and covariantize with respect to the gauge group. The resulting one loop diagrams are shown in the figure.[4] Differentiating the result, one finds the *Abelian anomaly* in the conservation law for the global current j^μ_A.

$$\partial_\mu j^\mu_A \propto \mathrm{tr}_r(t_A\{t_B, t_C\}) f^{\mu\nu}_B \widetilde{f}_{C\mu\nu} = \partial_\mu K^\mu_A. \tag{2.76}$$

The symmetrization of the generators in the trace comes from the Bose symmetry between the external gauge fields. The anomaly—the right hand side of (2.76)—is proportional to the Lorentz contraction of the field strength with its dual, which we saw in our discussion of the ϑ term is a total derivative. Thus though (2.76) implies that the j^μ_A current is not conserved, the combination $j^\mu_A - K^\mu_A$ is conserved. But K^μ_A is not gauge invariant, so the classical global symmetry generated by the j^A_μ generators is not a symmetry quantum mechanically. Furthermore, if j^A_μ were itself one of the gauge group generators, we see that the theory would be inconsistent. If there were many fermions in many representations, we would have to sum over the contributions of each of them running in the loop, and so the anomaly would be proportional to $\sum_r \mathrm{tr}_r(t_A\{t_B, t_C\})$.

To summarize, if $\mathrm{tr}_r(t_A\{t_B, t_C\}) \neq 0$ for

- 3 global currents $\Rightarrow$ theory is consistent, and global currents are conserved,
- 1 global and 2 local currents $\Rightarrow$ theory is consistent, but global current is not conserved so global symmetry is anomalous,
- 3 local currents $\Rightarrow$ theory is inconsistent.

[4] Note that gauge covariantizing the result for non-Abelian gauge groups means adding terms with more powers of the external gauge potential. These correspond to one loop, higher point anomalous diagrams in perturbation theory. The "Wess-Zumino consistency conditions" imply, though, that all these higher point amplitudes can be derived from the 3-point amplitude. See, for an explanation with few details, [28, sections 13.3 and 13.4].

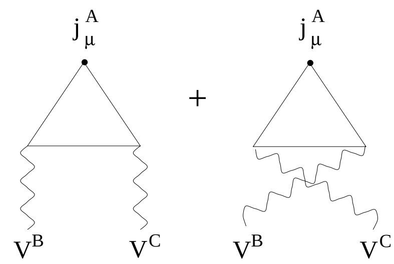

Figure 2.7: One loop diagrams contributing to the $\langle jVV \rangle$ correlator.

As was claimed above, real or pseudoreal representations can give no contribution to the anomaly since for these representations, $t_A = -(t_A)^T$ (up to a unitary similarity transformation), so that

$$\begin{aligned} \mathrm{tr}_r(t_A\{t_B, t_C\}) &= \mathrm{tr}_r\left[(-t_A)^T\left\{(-t_B)^T, (-t_C)^T\right\}\right] = -\mathrm{tr}_r(\{t_C, t_B\}t_A) \\ &= -\mathrm{tr}_r(t_A\{t_B, t_C\}), \end{aligned} \tag{2.77}$$

implying that the trace vanishes. Note that $U(1)$ representations are complex, since the complex conjugate of a charge q representation is a charge $-q$ representation. The vanishing of the anomaly for real representations also implies that only the massless fermions contribute to the anomaly. This follows since a fermion mass term is made from two left-handed Weyl fermions. If one of the fermions is in a complex representation $\mathbf{r}$ of the symmetry group, then the only way the mass term can be invariant under the symmetry is if the other fermion is in $\overline{\mathbf{r}}$. Thus the massive fermions pair up in the real $\mathbf{r} \oplus \overline{\mathbf{r}}$ representations, and thus do not contribute to the anomaly.

As a simple example of these considerations, consider QED with a massless electron, that is to say, a $U(1)_V$ gauge theory with two oppositely charged left-handed Weyl fermions ψ and $\widetilde{\psi}$ (describing the left and right helicity states of the electron, respectively, along with their antiparticles). This theory also has a classical global $U(1)_A$ "axial" symmetry under which both fermions have the same charge:

	$U(1)_V$	$\times$	$U(1)_A$
ψ	-1		$+1$
$\widetilde{\psi}$	$+1$		$+1$

Thus the currents for these two symmetries are

$$j_V^\mu = -\overline{\psi}\gamma^\mu\psi + \overline{\widetilde{\psi}}\gamma^\mu\widetilde{\psi}, \qquad j_A^\mu = +\overline{\psi}\gamma^\mu\psi + \overline{\widetilde{\psi}}\gamma^\mu\widetilde{\psi}. \tag{2.78}$$

Then the gauge current is $j_V = j_R - j_L$ and the global current is $j_A = j_R + j_L$. The anomalies are then

$$\langle j_A j_V j_V \rangle \quad \propto \quad \{1 \cdot (-1)^2 + 1 \cdot (+1)^2\} \neq 0,$$

$$\langle j_V j_V j_V \rangle \quad \propto \quad \{(-1)^3 + (+1)^3\} = 0, \tag{2.79}$$

implying that the global $U(1)_A$ symmetry is anomalous (*i.e.* its current is not conserved), while the gauge symmetry is non-anomalous.

Gauge anomalies

More generally, we can calculate the anomaly conditions for a $U(1)^N$ gauge theory with left-handed Weyl fermions ψ_i with charges q_{iA}, $A = 1, \ldots, N$ under each $U(1)$ factor. The gauge currents are

$$j^\mu_A = \sum_i q_{iA} \overline{\psi}_i \gamma^\mu \psi_i, \tag{2.80}$$

so the gauge anomaly is

$$\langle j_A j_B j_C \rangle \propto \sum_i q_{iA} q_{iB} q_{iC}, \tag{2.81}$$

which must vanish for consistency of the theory. Also, one can insert two $t^{\mu\nu}$ (energy-momentum) tensors in a triangle diagram and couple them to gravity in the usual way ($\int d^4x \sqrt{g} g_{\mu\nu} t^{\mu\nu}$) giving the mixed gauge-gravitational anomaly

$$\langle j^A tt \rangle \propto \sum_i q_{iA}, \tag{2.82}$$

which must also vanish for consistency. This reproduces the anomaly cancellation conditions mentioned in section 1.7.3.

The generalization to non-Abelian gauge anomalies is straightforward. Consider a theory with left-handed Weyl fermions ψ_i in representations $\mathbf{r}_i$ of a gauge group G. Then the gauge anomalies cancel if

$$\sum_i \mathrm{tr}_{r_i}(t_A\{t_B, t_C\}) = 0 \qquad \forall\ A, B, C. \tag{2.83}$$

This is actually a much less restrictive condition for non-Abelian groups than it might seem. For if the anomaly is not zero, it implies that there is a symmetric G invariant tensor $d_{\{ABC\}}$. The only groups which have both complex representations and a symmetric three-index invariant tensor are the $SU(N)$ groups for $N \geq 3$. (Note that $SO(6) \simeq SU(4)$ so $SO(6)$ also belongs to this class).

(There is another kind of anomaly which requires, in our conventions, that the total index of fermions transforming in pseudoreal representations be an even integer. This is relevant only for the $Sp(2n)$ groups. Note that since $SU(2) \simeq Sp(2)$ so it applies to $SU(2)$ as well. This is an anomaly under a kind of large gauge transformation, and so it cannot seen in the perturbative approach we are taking [29].)

Finally, the mixed gauge-gravitational anomalies imply that $\sum_i \mathrm{tr}_{r_i}(t_A) = 0$ for consistency. But for semi-simple groups the generators are automatically traceless. Thus the mixed gauge-gravitational anomalies only constrain the coupling to $U(1)$ gauge factors.

Chiral anomalies (anomalies in global symmetries)

Having satisfied the consistency conditions from the gauge anomalies, we now turn to the physics of anomalous global symmetries. Suppose we have a gauge group G with generators t_A, a global symmetry group $\widetilde{G}$ generated by $\widetilde{t}_A$, and Weyl fermions ψ_i transforming in the $(\mathbf{r_i}, \widetilde{\mathbf{r}}_\mathbf{i})$ representation of $G \times \widetilde{G}$. Then from the triangle diagram with one global current j_A^μ insertion and two gauge insertions we find the anomaly is proportional to

$$\partial_\mu j_A^\mu \propto \sum_i \mathrm{tr}(\widetilde{t}_A^{(\widetilde{\mathbf{r}}_i)} t_B^{(\mathbf{r}_i)} t_C^{(\mathbf{r}_i)}) = \sum_i \mathrm{tr}_{\widetilde{r}_i}(\widetilde{t}_A)\mathrm{tr}_{r_i}(t_B t_C). \tag{2.84}$$

Again, if $\widetilde{G}$ is semi-simple, $\mathrm{tr}(\widetilde{t}_A) = 0$, so there is no anomaly. Thus there are only anomalies in global $U(1)$ symmetries.

So, let us restrict ourselves to the case where the global symmetry is $U(1)$ and the fermions ψ_i have global charge q_i and transform as above in the $\mathbf{r}_i$ representation of some gauge group G. Then the anomaly is (this time including all the factors)

$$\partial^\mu j_\mu = \frac{1}{16\pi^2} \sum_i q_i \mathrm{tr}_{r_i}(f\widetilde{f}) = \frac{\sum_i q_i T(\mathbf{r}_i)}{16\pi^2} \mathrm{tr}_f(f\widetilde{f}). \tag{2.85}$$

This implies that the symmetry is anomalous if $\sum_i q_i T(\mathbf{r}_i) \neq 0$.

The fact that the anomaly is proportional to $\mathrm{tr}_f(f\widetilde{f})$ has some immediate consequences. Most importantly, the efects of the anomaly are equivalent to assigning the ϑ angle transformation properties under global $U(1)$ symmetries according to:

$$\begin{aligned} \psi^i &\rightarrow e^{iq_i\alpha}\psi^i \\ \vartheta &\rightarrow \vartheta + \alpha\left[\sum_i q_i T(\mathbf{r}_i)\right]. \end{aligned} \tag{2.86}$$

This follows since a shift in ϑ generates the right-hand side of the anomalous conservation equation (2.85). In this way we understand the anomalous breaking of the $U(1)$ symmetry as occuring due to an *explicit* breaking: a term $(\vartheta f\widetilde{f})$ in the action is not invariant.

Since the anomaly appears only through the ϑ term, it follows that at most one global $U(1)$ symmetry per gauge factor can be anomalous—by making appropriate linear combinations of their generators, one can choose all others to be non-anomalous.

The global $U(1)$ charge violation due to the anomaly in, say, a scattering process is:

$$\begin{aligned} \Delta Q &= \int_{-\infty}^{+\infty} dt\, \partial_0 Q = \int dt d^3\mathbf{x}\, \partial_0 j_0 \\ &= \int d^4x \left(\boldsymbol{\nabla} \cdot \mathbf{j} + \frac{\sum_i q_i T(r_i)}{16\pi^2} \mathrm{tr}_f(f\widetilde{f}) \right) = \left[\sum_i q_i T(\mathbf{r}_i) \right] n, \end{aligned} \tag{2.87}$$

where we dropped a total derivative of the current, and n in the last line is the (change in the) instanton number. We learned in the last section that processes changing the instanton number are non-perturbative, suppressed at weak coupling by factors of $e^{-8|n|\pi^2/g^2}$. Thus, even though j_μ is not conserved, its charge is conserved to all orders *in perturbation theory*, and at weak coupling the effects of the anomaly are very highly suppressed. For example, baryon number is an anomalous global $U(1)$ in the standard model due to its anomalous 3-point function with the $SU(2)$ gauge bosons; baryon violating amplitudes are therefore suppressed by an instanton factor of about $e^{-8\pi^2/e^2} \sim e^{-2\pi\cdot 137} \sim 10^{-300}$ and so is utterly negligible.

However, the anomalous shift in ϑ (2.86) helps make the physical effects of the ϑ term apparent. If there is a massless Weyl fermion charged under a gauge group G and with no Yukawa interactions, then the theory has an anomalous $U(1)$ symmetry under which only that fermion's phase rotates. By (2.86), by an appropriate such a phase rotation one can the ϑ away, and thus the ϑ angle has no effect in this theory. Alternatively, if there is no such massless fermion without Yukawa couplings, then such a rotation will at the same time give a CP violating phase to the fermion mass or Yukawa term; thus in this case the ϑ angle has observable consequences. For example, this is the origin of the strong CP problem in the standard model due to a possible non-zero value of the ϑ angle in $SU(3)$ QCD; alternatively, if the u quark is massless—if its Yukawa vanishes in the standard model—then the strong CP problem disappears, but we are left with a different naturalness problem of explaining why the u quark Yukawa should vanish.

't Hooft anomaly matching conditions

There is one other property of anomalies that will be important to us. It concerns the triangle diagrams with three global currents which we saw before do not lead to any anomalous symmetry breaking. Nevertheless, the following beautiful argument of 't Hooft [30, chapter 5.3] shows that they compute scale independent information about the theory.

Consider a theory described by a Lagrangian $\mathcal{L}$ at some scale μ, with global (non-anomalous) symmetries generated by currents j^μ_A. Gauge these symmetries by adding in new gauge fields V^A_μ, which I'll call "spectator" gauge fields, thus giving the new theory

$$\mathcal{L}' = \mathcal{L} + \int d^4x \left[\frac{1}{g^2}\mathrm{tr} f^2 + j_A \cdot V^A\right]. \tag{2.88}$$

This may not be a consistent theory, however, due to non-vanishing triangle diagrams for the newly gauged currents j_A. In that case, add in a set of new (spectator) free fermion fields ψ_S in representations to exactly cancel the anomalies and couple them only to the spectator gauge fields. Denoting the currents of the spectator fermions by j^S_A, we then have the enlarged and anomaly-free theory

$$\mathcal{L}'' = \mathcal{L} + \int d^4x \left[\frac{1}{g^2}\mathrm{tr} f^2 + +\overline{\psi}_S \partial\!\!\!/ \psi_S + (j^S_A + j_A)\cdot V^A\right]. \tag{2.89}$$

Since the spectator theory can be made arbitrarily weakly coupled by taking $g \to 0$, the IR dynamics of the enlarged theory are just the IR dynamics of the original theory plus the arbitrarily weakly coupled spectator theory. Thus the anomalies in the spectator theory are just the same as in the UV, and since the whole theory is anomaly-free, the anomaly from the IR currents j_A must also still be the same as in the UV. We can now throw away the spectator theory (take $g = 0$) to learn that the coefficient of the triangle diagram $\mathrm{tr}_r(t_A\{t_B, t_C\})$ for the *global* currents must be the same in the IR as in the UV.

The importance of this result is that the original theory might have been strongly-coupled in the IR in terms of its UV degrees of freedom, so the IR effective action may *a priori* be described by a completely different set of fermionic fields transforming under the global symmetries than appeared in the microscopic description. But 't Hooft's argument gives constraints on the possible IR fermion content, by demanding that their "anomalies" be the same as those of the UV fermions. We will see examples of these constraints in coming lectures.

2.3.4 Phases of gauge theories

One of our aims in exploring the IR effective actions of supersymmetric gauge theories will be to learn what "phase" these theories are in. For example, based on our experience with QCD and with lattice simulations, the vacuum of pure Yang-Mills theories are thought to have a mass gap and confinement. Other gauge theories (*e.g.* the electroweak theory) are known to have a Higgsed vacuum, where a charged scalar forms a condensate. Yet other theories (*e.g.* QED) have unbroken Abelian gauge factors giving rise to long range Coulomb interactions between charges. The question arises as to what are precise characterizations of these different phases of gauge theories, and whether they are qualitatively distinct, or can be continuously transformed into each other by deforming the parameters of the theory.

A way of probing the IR behavior of quantum field theories which gives precise characterizations of these phases is to look at the response of these theories to classical sources—massive particles interacting with the massless fields. For example, the static potential between sources of charge e separated by a distance R in QED with only massive charged fields is the usual Coulomb potential

$$V(R) = \frac{e^2}{R}. \tag{2.90}$$

On the other hand, in massless QED (*i.e.* with massless electrons) the static potential between heavy sources is

$$V(R) = \frac{e^2(R)}{R} \sim \frac{1}{R\log(\Lambda R)}, \tag{2.91}$$

due to the running of the coupling constant to zero in the IR. As another example, in the Abelian Higgs model

$$V(R) \sim \Lambda e^{-\Lambda R} \tag{2.92}$$

due to screening by the charge condensate (or, equivalently, because the gauge bosons now have a mass $\sim \Lambda$).

In pure Yang-Mills theory, the potential is thought to increase linearly with separation

$$V(R) \sim \Lambda^2 R, \tag{2.93}$$

giving charge confinement. This last is often described by the behavior of the expectation value of the *Wilson loop* operator in the limit of large space-time loops:

$$\langle \mathrm{Tr}\mathcal{P}e^{i\oint V}\rangle \sim e^{-\sigma\cdot\mathrm{Area}}. \tag{2.94}$$

The Wilson loop operator tests the response of the theory to the presence of an external source distributed along the loop. If one chooses the loop to be a rectangle of width R and length T (in the time direction), one can interpret the Wilson loop as measuring the action of a process which creates a pair of heavy charges separated by R and holds them there for a time T before annihilating them. The Wilson loop (in Euclidean time) then computes $e^{-TV(R)}$ which for a confining potential then gives the above area law. On the other hand, a Higgs mechanism would be expected to give a perimeter-type law for large Wilson loops, since the energy for separated charges is expected to fall off exponentially due to screening by the Higgs vacuum expectation value. In these last two cases, Higgsing and confinement, Λ is some scale that appears in the full theory. This shows that even though the static potential probes a long-distance aspect of the behavior of these theories, it does not just probe the scale-invariant (arbitrarily low energy) properties of the theory.

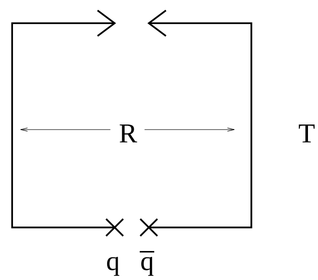

Figure 2.8: The Wilson loop interpreted as probing the response to a heavy quark-antiquark pair.

Let us now discuss some subtleties concerning the distinction between confinement and the Higgs mechanism. These subtleties will be important in interpreting our solutions for the vacuum structure of superQCD. The following discussion copies that of [30, chapter 7.2].

Gauge symmetry is not a really symmetry: it is a redundancy in our description of the physics. Evidence of this fact is that we divide our space of states by gauge transformations, considering two states differing by a gauge transformation as physically equivalent. This is different from what we do in the case of global symmetries: two states connected by a global symmetry transformation are inequivalent states, though they have identical physics.

All physical states in a gauge theory are gauge invariant, by definition. Confinement is sometimes described by saying that only color singlet (*i.e.* gauge invariant) combinations of quarks and gluons are observable as asymptotic states. So isn't confinement trivially a consequence of gauge invariance? Furthermore, if the vacuum is always gauge invariant, there can be no such thing as "spontaneous gauge symmmetry breaking"! Is the Higgs mechanism, in which a field gets a gauge non-invariant vacuum expectation value, in contradiction with gauge invariance?

The answer to both these questions is "no". We can see what is wrong with the above naive descriptions of Higgs and confining behavior in gauge theories through a simple example.

Consider the $SU(2)$ gauge theory with a doublet scalar ϕ (the Higgs), a doublet Weyl spinor ψ (the left-handed electron and neutrino), and a singlet Weyl spinor χ

(the right-handed electron):

$$\mathcal{L} = \frac{1}{g^2} f_{\mu\nu}^2 + D_\mu \phi D^\mu \phi + V(|\phi|) + \psi \not{D} \psi + \overline{\chi} \not{\partial} \chi + y\chi(\psi\phi) + h.c. \tag{2.95}$$

If the minimum of V is at $|\phi| = v$, we usually describe the resulting Higgs mechanism by choosing $\langle\phi\rangle = \binom{v}{0}$ and expand about that vacuum as

$$\phi = \begin{pmatrix} v + h_1 \\ h_2 \end{pmatrix}, \qquad \psi = \begin{pmatrix} e_L \\ \nu \end{pmatrix}, \qquad \chi = e_R, \qquad A_\mu^a = (Z_\mu, W_\mu, \overline{W}_\mu). \tag{2.96}$$

This description seems to break the gauge symmetry. However, dividing by gauge transformations, it is indistinguishable from the following "confined" description where all physical particles are gauge singlets:

$$e_L \sim \frac{1}{v}\psi\overline{\phi}, \quad \nu \sim \frac{1}{v}\psi\phi, \quad \mathrm{Re}(h_1) \sim \frac{1}{v}\phi\overline{\phi} - v, \quad Z_\mu \sim \frac{1}{v^2}\overline{\phi} D_\mu \phi, \quad W_\mu \sim \frac{1}{v^2}\phi D_\mu \phi. \tag{2.97}$$

The reason that we describe the Higgs mechanism in terms of fictitious global quantum numbers (like ν *versus* e_L, *etc.*) is because of our familiarity with global symmetry breaking and the fact that in the weak-coupling limit ($g \to 0$) a theory with a local symmetry looks globally symmetric. In the proper gauge-invariant description, all the physical states are "mesons" or "baryons" of the scalars bound to other fields.

Similarly, in QED (the "Coulomb phase") charges feel a Coulomb potential, $V \sim 1/r$, so charges can be infinitely separated. Thus we can talk about a single electron state $\psi(x)$ even though it is not gauge invariant: it can be thought of as an electron-plus-positron state with a Wilson line running between them, with the positron sent off to infinity. Thus the gauge invariant description of an electron is actually $\psi(x) \exp\{i \int_x^\infty A \cdot dx\}$; indeed, the Wilson line has observable effects when the topology of space-time is not simply connected, *e.g.* the Aharanov-Bohm effect.

The question of whether a gauge theory shows one of these behaviors—Higgs, Coulomb, confining, or something new—is thus a dynamical one. A kinematical question that can be addressed is whether these various long distance behaviors correspond to separate phases, or whether one can smoothly deform, say, Higgs to confining behavior. A Higgs vacuum can arise if there is a scalar condensate which can screen massive charges in the gauge group. One can always imagine the possibility in any strongly coupled theory (whether it has fundamental scalars or not) that the strong coupling dynamics might form such a massless scalar composite, and therefore that a Higgs vacuum might arise. As long as the scalar is in a faithful representation[5] of the gauge group, it can

[5] A faithful representation is one for which for every $g \in G$ with $g \neq 1$, then $R(g) \neq 1$, where $R(g)$ is the representation matrix.

screen all charges, and there can be no invariant distinction between the Higgs vacuum and a confining vacuum since all the asymptotic states are gauge-singlets. The only way a (non-trivial) representation of a simple group can fail to be faithful is if it does not transform under the center of the gauge group.[6] If the microscopic field content of a theory is such that no (composite) scalars in faithful representations can be formed, then there is an invariant distinction between Higgs and confining phases: in the Higgs phase, all but the discrete central charges are screened,[7] whereas in the confining phase all the asymptotic states will be invariant under the center of the gauge group. QCD is such a theory, where the $\mathbb{Z}_3$ center is tied to the electric charges of the fields, so the distinction between confinment and Higgs phases in this case is whether there are charge-1/3 asymptotic states or not.

[6]The center of a group is the subgroup consisting of the set of all elements which commute with all elements of the group. For example, the center of $SU(N)$ is $\mathbb{Z}_N$ realized as overall phase rotations by the Nth roots of unity.

[7]Since the center is a discrete subgroup of the gauge group, and discrete gauge groups have no long-range fields, there can exist asymptotic states charged under the center.

2.4 Non-Renormalization in Super Gauge Theories

We now turn to the quantum mechanical properties of supersymmetric gauge theories. Our aim is to prove non-renormalization theorems for supersymmetric gauge theories, along the lines of the non-renormalization theorem we proved for theories of chiral multiplets.

2.4.1 Supersymmetric selection rules

We start by examining the analog of the holomorphy of the superpotential for an asymptotically free supersymmetric gauge theory. In the action at a scale μ_0, the terms which can be written only as integrals over half of superspace, and therefore must have holomorphic dependence on their fields couplings, are the gauge kinetic and superpotential terms, which we will assemble into a *generalized superpotential* $\widetilde{f}$:

$$\widetilde{f}_{\mu_0} = \frac{\tau(\mu_0)}{8\pi i}\mathrm{tr}_f(W_L^2) + f_{\mu_0}(\Phi^n, \lambda_r, \mu_0). \tag{2.98}$$

Here λ_r are the couplings appearing in the tree-level superpotential

$$f_{\mu_0} = \sum_r \mu_0^{3-d_r}\lambda_r \mathcal{O}_r, \tag{2.99}$$

where $\mathcal{O}_r$ are gauge-invariant composite operators of the Φ^n's of classical scaling dimension d_r.

Recall that at one loop the gauge coupling is

$$\tau(\mu_0) \equiv \frac{\vartheta}{2\pi} + i\frac{4\pi}{g(\mu_0)^2} = \frac{1}{2\pi i}\log\left[\left(\frac{|\Lambda|}{\mu_0}\right)^b e^{i\vartheta}\right], \tag{2.100}$$

where we have used the definition of the strong coupling scale $|\Lambda|$ in the last step. (The absolute value is to remind us that it is a postive real number.) It is thus natural to define a *complex* "scale" in supersymmetric gauge theories by

$$\Lambda \equiv |\Lambda| e^{i\vartheta/b} \tag{2.101}$$

so that

$$\tau(\mu_0) = \frac{b}{2\pi i}\log\left(\frac{\Lambda}{\mu_0}\right). \tag{2.102}$$

Recall that b is the coefficient of the one-loop beta function, given by

$$b = \frac{11}{6}T(\mathrm{adj}) - \frac{1}{3}\sum_a T(\mathbf{r}_a) - \frac{1}{6}\sum_n T(\mathbf{r}_n), \tag{2.103}$$

where the indices a run over Weyl fermions and n run over complex bosons. In a supersymmetric gauge theory, the vector multiplet always includes a Weyl fermion in the adjoint representation (the gaugino), while each chiral multiplet Φ^n has one Weyl fermion and one complex boson, transforming in the same representation of the gauge group $\mathbf{r}_n$. Thus, for supersymmetric gauge theories, b simplifies to

$$b = \tfrac{3}{2}T(\text{adj}) - \tfrac{1}{2}\textstyle\sum_n T(\mathbf{r}_n), \tag{2.104}$$

where the sum on n is over all left-chiral superfields.

Let us assume for definiteness that we are dealing with an asymptotically free theory, so if we take the scale $\mu_0 \gg |\Lambda|$, then the theory is weakly coupled (we might also have to take some of the superpotential couplings to be small). Let us consider how this effective theory will change as we run it down in scale a little to $\mu < \mu_0$. As long as the ratio μ/μ_0 is not too small, the theory should remain weakly coupled, and we expect that the effective theory should be describable in terms of the same degrees of freedom. The effective generalized superpotential will then be

$$\widetilde{f}_\mu = \frac{\tau(\Lambda, \Phi^n, \lambda_r; \mu)}{8\pi i}\text{tr}_f(W_L^2) + f(\Phi^n, \lambda_r, \Lambda; \mu) + \text{irrelevant operators.} \tag{2.105}$$

Here we have written the effective coupling τ and superpotential f as general holomorphic functions of the left-chiral superfields and the bare couplings, as befits terms that appear only as integrals over half of superspace. The irrelevant operators include terms with higher powers of $\text{tr}(W_L^2)$, since W_L has scaling dimension 1.

However, we have to take into account the angular nature of the ϑ angle which implies the identification $\tau(\mu) \simeq \tau(\mu) + 1$. This is true in the effective action at any scale μ since it just follows from the topological quantization of the integral of $\text{tr}_f(f\widetilde{f})$. At the scale μ_0 where we have *defined* ϑ as the phase of Λ^b, this means that as we rotate the phase of Λ^b by 2π, $\Lambda^b \to e^{2\pi i}\Lambda^b$, we have $\tau(\mu_0) \to \tau(\mu_0) + 1$. Now under the RG flow from μ_0 to a slightly lower scale, $\tau(\mu)$ changes continuously with μ. If follows that for any μ, when $\Lambda^b \to e^{2\pi i}\Lambda^b$ we must have $\tau(\mu) \to \tau(\mu) + 1$ by continuity. Thus $\tau(\Lambda, \Phi, \lambda)$ is *not* a general holomorphic function, for as we rotate the phase ϑ of Λ^b by 2π we must have $\tau \to \tau + 1$. This constrains the functional form of τ to be

$$\tau(\mu) = \frac{b}{2\pi i}\log\left(\frac{\Lambda}{\mu}\right) + h(\Lambda^b, \Phi^n, \lambda_r; \mu), \tag{2.106}$$

where h is now an arbitrary holomorphic single valued function of its arguments.

Since we are dealing with an asymptotically free theory, the $\Lambda \to 0$ limit corresponds to the weak coupling limit, in which the effective couplings should not diverge. Thus we have

$$\tau(\mu) = \frac{b}{2\pi i}\log\left(\frac{\Lambda}{\mu}\right) + \sum_{j=0}^{\infty}\Lambda^{bj}h_j(\Phi^n, \lambda_r; \mu), \tag{2.107}$$

in which inverse powers of Λ^b do not appear. By comparing this expression to the perturbative expansion, where $\log\Lambda \sim 1/g^2$ and $\Lambda^{bj} \sim j$-instanton action, we see that the gauge coupling τ in the Wilsonian effective action only gets one loop corrections in perturbation theory, though non-perturbative corrections are allowed.

The superpotential satisfies a similar constraint. If we turned off the gauge coupling ($\Lambda \to 0$) then we would have the our previous non-renormalization theorem for the Wilsonian effective action superpotential which says it does not get renormalized at all: $f_\mu = f_{\mu_0}(\Phi^n, \lambda_r; \mu)$. Turning on the gauge coupling can then only add new terms holomorphic in Λ^b and vanishing as $\Lambda \to 0$, so

$$f_\mu = f_{\mu_0}(\mu) + \sum_{j=1}^{\infty} \Lambda^{bj} g_j(\Phi^n, \lambda_r; \mu), \tag{2.108}$$

for some single valued holomorphic functions g_j, implying no perturbation theory corrections but possible non-perturbative corrections.

In summary, we have learned so far that the effective generalized superpotential has the form

$$\begin{aligned} \widetilde{f}_\mu &= f_{\mu_0}(\Phi^n, \lambda_r; \mu) + \sum_{j=1}^{\infty} \Lambda^{bj} g_j(\Phi^n, \lambda_r; \mu) \\ &\quad + \frac{1}{16\pi^2} \log(\Lambda/\mu)^b \, S + S \sum_{j=0}^{\infty} \Lambda^{bj} h_j(\Phi^n, \lambda_r; \mu) + \mathcal{O}(S^2), \end{aligned} \tag{2.109}$$

where we have defined the dimension 3 composite chiral superfield S to be

$$S \equiv -\mathrm{tr}_f(W_L^2). \tag{2.110}$$

2.4.2 Global symmetries and selection rules

The above non-renormalization theorem can be sharpened in an important way by using the selection rules of other global symmetries in the theory. An important new element is the treatment of the selection rules stemming from anomalous symmetries.

Consider a global symmetry, $U(1)_n$, which rotates only one left-chiral superfield, Φ^n. Thus

$$U(1)_n: \ \Phi^m \to e^{i\alpha\delta_{mn}} \Phi^m \qquad \text{and} \qquad W_L \to W_L. \tag{2.111}$$

As we saw in the last lecture, this symmetry is anomalous, and can be considered as having the effect of rotating the ϑ angle

$$U(1)_n: \ \vartheta \to \vartheta + \alpha T(\mathbf{r}_n), \tag{2.112}$$

where $\mathbf{r}_n$ is the gauge group representation of Φ^n. We can express this by giving Λ an effective charge under $U(1)_n$:

$$U(1)_n: \ \Lambda^b \to e^{i\alpha T(\mathbf{r}_n)}\Lambda^b. \tag{2.113}$$

This gives a selection rule for possible terms appearing in τ and f due to this anomalous symmetry.

Another useful symmetry is the R-symmetry of our theory. Recalling that the R-charge of the superspace Grassmann variable is $R(\theta_L) = 1$, it follows that for $\int d^2\theta_L \text{tr}-f(W_L^2)$ to be R-invariant we must have

$$R(W_L) = 1 \tag{2.114}$$

which implies $R(\lambda_L) = 1$ and $R(f_{\mu\nu}) = R(V_\mu) = 0$ in components. (Note that since $R(V) = 0$ the Kahler terms are automatically R-invariant.) If we define the R-charges of the left-chiral superfields to be zero,

$$R(\Phi^n) = 0, \tag{2.115}$$

then their components wil have charges $R(\phi^n) = 0$ and $R(\psi^n) = -1$. From the charges of the fermions, we see that the anomaly for this symmetry can be compensated by assigning an R charge to Λ of

$$R(\Lambda^b) = T(\text{adj}) - \sum_n T(\mathbf{r}_n). \tag{2.116}$$

We would now like to apply the selection rules following from the $\prod_n U(1)_n \times U(1)_R$ global symmetries (which will typically be explicitly broken by superpotential terms) to constrain the form of the effective generalized superpotential $\widetilde{f}_\mu$. The analysis of these constraints is made simpler by ignoring the constraints that come from gauge-invariance, which we can put back in at the end of our analysis. So, we analyze instead the selection rule constraints on a superpotential where we introduce a separate tree-level coupling for each left-chiral superfield Φ^n:

$$f_{\mu_0} = \sum_m \mu_0^{3-d_n}\lambda_n\Phi^n. \tag{2.117}$$

We emphasize that the λ_n couplings are not gauge invariant; however, transformation properties of any physical coupling λ_r of a gauge-invariant composite operator $\mathcal{O}_r = \prod_i \Phi^{n_i}$ in the microscopic superpotential under the global $U(1)_R \times \prod_n U(1)_n$ symmetries will be the same as those of the corresponding product of λ_n's: $\prod_i \lambda_{n_i}$. The charges of all the fields and couplings can be summarized as:

	$U(1)_n$ $\times$	$U(1)_R$
Φ^m	δ_{nm}	0
λ_m	$-\delta_{nm}$	2
S	0	2
Λ^b	$T(\mathbf{r}_n)$	$T(\text{adj})-\sum_n T(\mathbf{r}_n)$

The general form of any of the unknown non-perturbative terms which may appear in the effective generalized superpotential $\widetilde{f}_\mu$ in (2.109) is

$$\Lambda^{b\alpha} S^\beta \prod_n \lambda_n^{\alpha_n} (\Phi^n)^{\beta_n} \qquad \text{for } \beta = 0, 1 \tag{2.118}$$

for some integers α, α_n, and real numbers β_n. (The powers of the couplings must be integral since the effective generalized superpotential should be single valued in the weak coupling limits $\Lambda, \lambda_n \to 0$.) Terms with $\beta = 0$ give contributions to the g_j functions in the superpotential in (2.109) while the terms with $\beta = 1$ are non-perturbative corrections to the generalized gauge kinetic term, giving contributions to the h_j's in (2.109). The selection rules then imply the conditions

$$\begin{aligned} U(1)_n: \qquad 0 &= \alpha T(\mathbf{r}_n) - \alpha_n + \beta_n \\ U(1)_R: \qquad 2 &= \alpha\left(T(\text{adj}) - \sum_n T(\mathbf{r}_n)\right) + 2\beta + 2\sum_n \alpha_n. \end{aligned} \tag{2.119}$$

Summing the $U(1)_n$ conditions and adding them to the $U(1)_R$ condition implies

$$2 = \alpha T(\text{adj}) + 2\beta + \sum_n (\alpha_n + \beta_n). \tag{2.120}$$

Since, by regularity in the weak coupling limits $\Lambda, \lambda_n \to 0$, $\alpha \geq 1$ and $\alpha_n \geq 0$, and since the smallest value that $T(\text{adj})$ can take is 4 (for $SU(2)$), all solutions to this constraint must have some $\beta_n < 0$. Thus all the possible non-perturbative corrections to the g_j and h_j functions in (2.109) involve inverse powers of the left-chiral superfields. In particular, we find that there are *no non-perturbative corrections to the UV superpotential and one loop gauge couplings.* In other words, the function $\widetilde{f}_\mu$ may contain non-perturbative corrections to the generalized superpotential in the form of new operators involving inverse powers of left-chiral superfields, but none that are functions only of the couplings times the operators S or $\mathcal{O}_r$ present in the UV generalized superpotential $\widetilde{f}_{\mu_0}$. This implies, in particular, that the one loop gauge coupling is exact, and that the λ_r couplings are unrenormalized. (The potential appearance of inverse powers of left-chiral superfields may seem unphysical, but we will meet and interpret such contributions shortly.)

We should emphasize the limitations of this "non-renormalization" theorem: it was only derived for weakly-enough coupled theories where the description in terms of the microscopic degrees of freedom is good. As we run the RG down to the IR, the theory will become strongly coupled, and our description in terms of the Φ^n and W_L fields will break down. More technically, as we run down in scale at some point we can no longer be sure that the "irrelevant operators" in (2.105)—as well as other irrelevant operators appearing elsewhere in the effective action besides the generlized superpotential—are really irrelevant. As we discussed in section 2.1, the characterization of an operator by its scaling dimension (*i.e.* as relevant or irrelevant) only has meaning in the vicinity of an RG fixed point. Upon flowing from a UV weak coupling fixed point to its IR fixed point, operators that were irrelevant in the UV may become relevant in the IR. In short, this non-renormalization theorem in no way solves the essential strong coupling problem of asymptotically free gauge theories by themselves.

2.4.3 IR free gauge theories and Fayet-Iliopoulos terms

So far we have been discussing only asymptotically free (and therefore non-Abelian) gauge theories. Clearly, similar arguments can be applied to IR free gauge theories, as long as we take the scale of our theory *low* enough—$\mu_0 \ll \Lambda$—so that the theory is weakly coupled. Then the RG running to the IR will just make the theory more weakly coupled, so the effective theory should be described by the same degrees of freedom. Thus in IR free theories the weak coupling limit is $\Lambda \to \infty$. We once again find that the gauge coupling is only renormalized at one loop in perturbation theory, and that all non-perturbative corrections must be proportional to powers of Λ^b (since for IR free thoeries $b < 0$). However, since IR free theories are UV strongly coupled, the question of their IR effective couplings is largely moot, unless they are realized as effective theories of some microscopic physics with different degrees of freedom, *e.g.* as an asymptotically free gauge theory whose gauge group is spontaneously broken down to an IR free group. In such a case where the IR free gauge groups are Abelian, there exist techniques relying on the *electric-magnetic duality* of the low energy effective actions which have proved to be strong enough to exactly determine the non-perturbative corrections to the low energy couplings. We will discuss this topic at length in later lectures.

We saw in section 1.7 that Abelian gauge theories admit one extra kind of term, the Fayet-Iliopoulos D-term, which may appear in an effective action at a scale μ_0 as

$$S_{FI,\mu_0} = i \int d^4x d\theta_L^a \, \mu_0^2 \xi_0 W_{La} + \text{c.c.} \tag{2.121}$$

Here ξ_0 is a (dimensionless) real number and not a function of superfields, since for this term to be a supersymmetry invariant, ξ must satisfy $\mathcal{D}\xi = 0$, which implies it must be a constant. We wish to determine the effective Fayet-Iliopoulos coupling as a function $\xi(g, \lambda_r; \mu)$ of the gauge and superpotential couplings g and λ_r upon running down to a scale μ. As ξ is not a function of left-chiral superfields, and since g and λ_r enter the action in the same way as the vacuum expectation values of left-chiral superfields do, there follows the supersymmetric selection rule that g, λ_r, and the (vacuum expectation values of) any left-chiral superfields can *not* enter into quantum corrections of the Fayet-Iliopoulos term. The only other parameters in the theory are the scale μ_0 at which the theory is defined, the RG scale μ of the effective theory, and the gauge charges of the left-chiral superfields q_n. Furthermore, by gauge invariance, the charges and ξ_0 itself can only enter in physical amplitudes in the combinations $g\xi_0$ and gq_n, where g is the $U(1)$ gauge coupling. Thus the effective Fayet-Iliopoulos term coefficient must satisfy

$$g\xi = h(g\xi_0, gq_n, \mu/\mu_0) \tag{2.122}$$

for some real function h, where I have used dimensional analysis to fix some of the μ dependence. Now, for a $U(1)$ gauge theory we can decouple fields by continuously varying the $q_n \to 0$. In this limit the $U(1)$ field is completely decoupled and so the Fayet-Iliopoulos coefficient should have its free value, ξ_0. This implies, in particular, that no inverse powers of gq_n can appear in h. By our supersymmetric selection rule, ξ must be g-independent, implying $\xi = \xi_0\, h_0(\mu/\mu_0) + q_n\, h_n(\mu/\mu_0)$ for some functions h_a; and the $q_n \to 0$ limit implies that $h_0 = 1$. The terms proportional to q_n come by definition from the first order in perturbation theory, namely the tadpole graph with the vector superfield attached to a loop of left-chiral superfields. The sum of all such graphs is proportional to the sum of the charges of the left-chiral superfields, so

$$\xi = \xi_0 + C\left(\sum_n q_n\right)\log\left(\frac{\mu}{\mu_0}\right) \tag{2.123}$$

where C is a numerical constant that can be calculated from perturbation theory. As discussed in section 2.3.3 the requirement of absence of gravitational anomalies requires the sum of the $U(1)$ charges to vanish if the $U(1)$ symmetry is unbroken, and we learn that the Fayet-Iliopoulos term is exactly unrenormalized.

2.4.4 Exact beta functions

We can summarise our non-renormalization theorem for asymptotically free gauge theories by the effective action

$$S_\mu = \int d^4x \Big\{ \frac{1}{2}\left[Z_n(\mu)\Phi_n^* e^V \Phi^n\right]_D + \Big[\sum_r \mu^{3-d_r}\lambda_r \mathcal{O}_r + \text{non-perturbative operators}$$

$$- \frac{\tau(\mu)}{8\pi i} S + \text{c.c.} + \text{irrelevant operators}\Bigg]_F \Bigg\}. \tag{2.124}$$

This is the Wilsonian effective action at the scale μ (and is valid as long as μ is not too much smaller than the UV scale μ_0). Here $\mathcal{O}_r$ are the gauge invariant composite operators appearing in the UV superpotential

$$\mathcal{O}_r = \prod_n (\Phi^n)^{r_n} \tag{2.125}$$

for some integers r_n, with classical scaling dimensions

$$d_r = \sum_n r_n, \tag{2.126}$$

and $2\pi i\tau(\mu) = \log(\Lambda/\mu)^b$ is the Wilsonian effective gauge coupling, and we have included the wave function renormalizations Z_n of the Kahler terms.[8]

In order to compare the couplings of this effective action to physical couplings that would be measured in, say, a scattering experiment with energy transfer of order μ, we must normalize the kinetic terms to their canonical form by rescaling the left-chiral superfields by

$$\Phi^n \to \Phi^n_{\text{cn}} \equiv \sqrt{Z_n(\mu)}\, \Phi^n, \tag{2.127}$$

where a "cn" subscript or superscript denotes canonically normalized fields or couplings, as discussed in section 2.2.3. Then the renormalized Lagrangian has the same form as the Wilsonian one, but with the bare superpotential couplings replaced by canonical ones:

$$\lambda_r^{\text{cn}} \equiv \mu^{3-d_r} \left(\prod_n Z_n^{-r_n/2} \right) \lambda_r, \tag{2.128}$$

which implies the exact RG equation for the canonical superpotential couplings

$$\beta^{\text{cn}}_{\lambda_r} \equiv \frac{d\lambda_r^{\text{cn}}}{d\log\mu} = \lambda_r^{\text{cn}} \left(3 - d_r - \tfrac{1}{2} \sum_n r_n \gamma_n \right), \tag{2.129}$$

where $\gamma_n = d\log Z_n / d\log\mu$ is the anomalous dimension of Φ^n. Of course, we have no "exact" method of computing the γ_n.

In gauge theories, to compare to physical processes we not only need to rescale the left-chiral superfields but also the vector superfield so that the gauge kinetic terms

[8] One may wonder why we haven't written a possible multiplicative renormalization for the vector superfield V: $Z_n \Phi^*{}_n e^{Z_V V} \Phi^n$? This can certainly arise; however, the relation between the field strength superfield W_L and V can also get renormalized: $W_L \sim D^2(e^{-Z_W V} D_\alpha e^{Z_W V})$. Gauge invariance requires that $Z_V = Z_W$, so that these redefinitions can be trivially scaled out of the effective action.

are canonically normalized, and the gauge coupling appears in the covariant derivative acting on the matter fields:

$$S_{\rm cn} = \int d^4x \left\{ \tfrac{1}{2} \left[\Phi^{n*}_{\rm cn} e^{g_{\rm cn} V_{\rm cn}} \Phi^n_{\rm cn}\right]_D + \tfrac{1}{2}\left[\left(1 - \frac{i\vartheta g_c^2}{8\pi^2}\right) {\rm tr}_f (W_L^{\rm cn} \mathcal{C} W_L^{\rm cn}) + {\rm c.c.}\right]_F \right\}. \tag{2.130}$$

To get the action to this form starting from the Wilsonian action in our supersymmetric scheme,

$$S = \int d^4x \left\{ \tfrac{1}{2} \left[Z_n \Phi^{n*} e^V \Phi^n\right]_D + \tfrac{1}{2}\left[\left(\frac{1}{g^2} 1 - \frac{i\vartheta}{8\pi^2}\right) {\rm tr}_f (W_L \mathcal{C} W_L) + {\rm c.c.}\right]_F \right\}, \tag{2.131}$$

we rescale the fields by

$$\begin{aligned} \Phi^n &\rightarrow \Phi^n_{\rm cn} = \sqrt{Z_n}\Phi^n, \\ W_L &\rightarrow W_L^{\rm cn} = \frac{1}{g_{\rm cn}} W_L, \end{aligned} \tag{2.132}$$

which classically gives

$$S^{\rm cl}_{\rm cn} = \int d^4x \left\{ \tfrac{1}{2} \left[\Phi^{n*}_{\rm cn} e^{g_{\rm cn} V_{\rm cn}} \Phi^n_{\rm cn}\right]_D + \tfrac{1}{2}\left[\left(\frac{g^2_{\rm cn}}{g^2} - \frac{i\vartheta g_c^2}{8\pi^2}\right) {\rm tr}_f (W_L^{\rm cn} \mathcal{C} W_L^{\rm cn}) + {\rm c.c.}\right]_F \right\}. \tag{2.133}$$

If this were all there were to the story, then we would have $g = g_c$ to make (2.133) and (2.130) match. However, quantum mechanically the gauge coupling change under the rescaling (2.132) because of anomalies. In particular, we can think of these rescalings as complexified chiral rotations of the superfields

$$\begin{aligned} \Phi^n &\rightarrow e^{i\alpha_n}\Phi^n, \qquad \text{with} \qquad i\alpha_n = \frac{1}{2}\log(Z_n), \\ W_L^{\rm cn} &\rightarrow e^{i\alpha_0} W_L^{\rm cn}, \qquad \text{with} \qquad i\alpha_0 = \frac{1}{2}\log(1/g^2_{\rm cn}). \end{aligned} \tag{2.134}$$

Now, such chiral rotations are anomalous, implying that the ϑ angle is shifted to

$$\vartheta \rightarrow \vartheta + T({\rm adj})\alpha_0 + \sum_n T(\mathbf{r}_n)\alpha_n, \tag{2.135}$$

which when substituted into (2.133) gives rise to a shift in the real part of the ${\rm tr}_f(W_L^{\rm cn}\mathcal{C}W_L^{\rm cn})$ coefficient to

$$\frac{g^2_{\rm cn}}{g^2} + i\frac{\left(T({\rm adj})\alpha_0 + \sum_n T(\mathbf{r}_n)\alpha_n\right) g^2_{\rm cn}}{8\pi^2}. \tag{2.136}$$

Equating this to 1 to match to the canonical action (2.130) determines g_{cn} in terms of g as

$$\frac{1}{g_{\text{cn}}^2} = \frac{1}{g^2} + \frac{1}{16\pi^2}\left(T(\text{adj})\log(1/g_{\text{cn}}^2) + \sum_n T(\mathbf{r}_n)\log(Z_n)\right). \tag{2.137}$$

Taking the logarithmic derivative of this expression with respect to the RG scale μ, and using the definition of the (canonical) beta function as

$$\beta_g^{\text{cn}} \equiv \frac{d(1/g_{\text{cn}}^2)}{d\log\mu}, \tag{2.138}$$

and the exact beta function for the (Wilsonian) coupling, $d(1/g^2)/d\log\mu = b/8\pi^2$, we get the exact expression for the canonical beta function

$$\beta_g^{\text{cn}} = \frac{b + \frac{1}{2}\sum_n T(\mathbf{r}_n)\gamma_n}{8\pi^2 - \frac{1}{2}T(\text{adj})g_{\text{cn}}^2}. \tag{2.139}$$

This result was derived in [31] by different methods. This is the one loop Wilsonian beta function corrected by the anomalous mass dimensions (which involve higher loop effects) as well as by an overall factor which depends on the exact coupling. Conceptually, the anomalous dimensions of the left-chiral superfields enter because these fields enter in the one loop diagrams for the beta function.

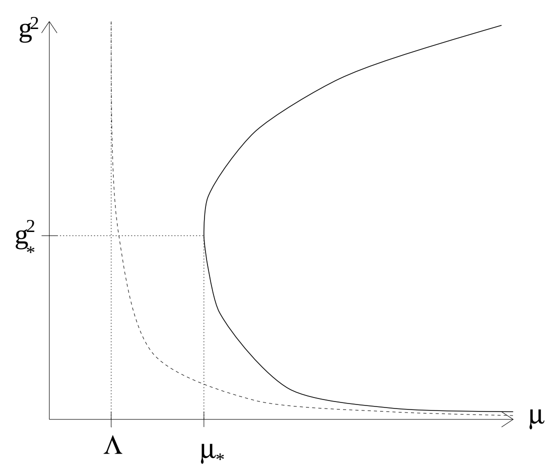

Figure 2.9: Exact RG flow of the canonical gauge coupling in superYang-Mills theory (solid line). The dashed line shows the unmodified one loop running of the coupling for comparison.

The dependence of the exact beta function (2.139) on the anomalous dimensions of the left-chiral superfields can be removed in the case of a theory with no left-chiral superfields, namely in superYang-Mills theory (*i.e.* the theory of a single non-Abelian vector superfield). Then (2.139) gives a differential equation for the coupling as a function of scale:

$$\frac{d(1/g^2)}{d\log\mu} = \frac{\frac{3}{2}T(\text{adj})}{8\pi^2 - \frac{1}{2}T(\text{adj})g^2}, \tag{2.140}$$

(dropping the canonical superscripts) which can be integrated to give the RG flow shown in the figure. This flow equation has two solutions, a physically reasonable lower branch where the coupling goes to zero at large μ, and spurious upper branch where it diverges instead. The two branches meet at a scale $\mu_* \sim \Lambda/T(\text{adj})^{1/3}$ where $g_*^2 \sim 1/T(\text{adj})$ and the beta function has a pole. The fact that the flow does not continue to scales below μ_* reflects the point discussed earlier that the supersymmetric non-renormalization theorems and related exact results do not solve by themselves the strong coupling problem of asymptotically free gauge theory since they are derived under the *assumption* that the effective theory at all scales will be described in terms of the UV fields. Presumably the exact beta function has become invalid by the time it reaches μ_* by new relevant operators (which were irrelevant at weak coupling) being generated during the flow to strong coupling in the IR.

One interesting piece of information about strong coupling which we do extract is that (for, say, $SU(N)$ gauge group) strong coupling occurs at couplings $g_*^2 \sim \mathcal{O}(1/N)$, while the strong coupling scale of the theory is $\mu_* \sim \mathcal{O}(\Lambda/N^{1/3})$, results which may be familiar from the large N expansion of gauge theories.

2.4.5 Scale invariance and finiteness

The exact beta functions for gauge and superpotential couplings we found above can be used to give some interesting information about exactly marginal operators in certain cases. The exact beta functions are proportional to

$$\begin{aligned}\beta_g &\propto b + \tfrac{1}{2}\sum_n T(\mathbf{r}_n)\gamma_n \\ \beta_{\lambda_r} &\propto 3 - d_r - \tfrac{1}{2}\sum_n r_n\gamma_n.\end{aligned} \tag{2.141}$$

If one can arrange the couplings and field content of a theory appropriately so that the conditions for the vanishing of these beta functions are not all lineraly independent, then we can deduce the existence of a submanifold of the g–λ_r coupling space for which the beta functions vanish. Thus changing the couplings in such a way as to stay on this manifold corresponds to turning on an exactly marginal operator. The fixed point theories along this submanifold are all scale invariant (since thier beta functions vanish); the existence of exactly marginal operators then implies that these scale invariant theories have non-trivial interactions.

For example, consider an $SU(N_c)$ gauge theory with $2N_c$ left-chiral superfields Q^i in the $\mathbf{N_c}$ of $SU(N_c)$, and another $2N_c$ left-chiral superfields $\widetilde{Q}_i$ in the $\overline{\mathbf{N_c}}$. (This is superQCD with $N_f = 2N_c$ flavors.) If the theory also has a left-chiral superfield Φ in the adjoint of the gauge group, we can consider adding the operator

$$f = \lambda \sum_n \mathrm{tr} \widetilde{Q}_n \Phi Q^n. \tag{2.142}$$

It is easy to check that with this matter content, the theory is one loop scale invariant: $b = 0$. Since the Q's and $\widetilde{Q}$'s all enter symmetrically (there is an $SU(2N_c)$ flavor symmetry), they will all have the same anomalous dimension γ_Q. We then find

$$\beta_g \propto \beta_\lambda \propto \gamma_\Phi + 2\gamma_Q, \tag{2.143}$$

so there is only one condition on the λ–g parameter plane for scale invariance. Thus there will be a line of fixed points; furthermore, at weak coupling ($g, \lambda \ll 1$) the anomalous dimensions vanish, so this line of fixed points goes through the origin. (In fact, it turns out that the exact curve is $\lambda = g$ and the scale invariant theories along have an N=2 extended supersymmetry.)

As another example, consider a theory with three adjoint left-chiral superfields Φ_i ($i = 1, 2, 3$) and the superpotential

$$f = a \mathrm{tr} \Phi_1 \Phi_2 \Phi_3 + b \mathrm{tr} \Phi_3 \Phi_2 \Phi_1 + c \mathrm{tr}(\Phi_1^3 + \Phi_2^3 + \Phi_3^3). \tag{2.144}$$

Then their anomalous dimensions are all equal $\gamma_i = \gamma$ and

$$\beta_g \propto \beta_a \propto \beta_b \propto \beta_c \propto \gamma, \tag{2.145}$$

so in this case there is actually a three (complex) dimensional space of fixed point theories in the g–a–b–c coupling space, which passes through weak coupling. (A one dimensional submanifold of this set is known to be given exactly by $g = a = -b$ and $c = 0$; the theories on this line have N=4 extended supersymmetry.)

There are many more examples one can construct along these lines; see [32].

This kind of analysis of exactly marginal couplings does not shed light on superQCD theories (whose detailed analysis we will be starting in the next lecture); however perturbative calculations do show the existence of isolated non-trivial fixed points in superQCD with no superpotential couplings. Indeed, one might think then that one could tune the field content of a non-Abelian gauge theory to make the beta function exactly zero, thus acheiving exact scale invariance. Taking $N_f = 3N_c$ in super QCD

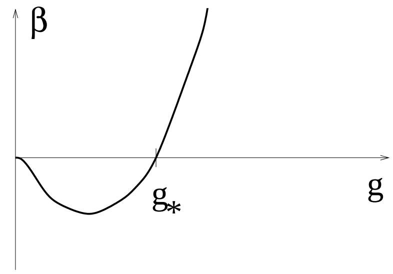

Figure 2.10: pert fp

does not work, though, since that only cancels the one loop beta function; the two loop contributions to the canonical beta function make the theory IR free. However, this observation suggests how to show that interacting scale invariant theories do exist [33]. For $N_f < 3N_c$ a two loop computation gives the β function

$$\beta(g) = -\frac{g^3}{16\pi^2}(3N_c - N_f) + \frac{g^5}{128\pi^4}\left(2N_cN_f - 3{N_c}^2 - \frac{N_f}{N_c}\right) + \mathcal{O}(g^7), \qquad (2.146)$$

which gives an IR fixed point ($\beta = 0$) at a coupling $g_* \sim \sqrt{3N_c - N_f}$. We can trust the existence of this fixed point as long as the coupling g_* is small, so that the higher order terms can be safely neglected. Define $N_f = N_c(3 - \epsilon)$, and take the limit $N_c \to \infty$ (so that $\epsilon \sim 1/N_c$). Then $N_c g_*^2 \sim 4\pi^2\epsilon/3$. Recalling that $N_c g_*^2$ is the expansion parameter for large N_c, we see that there does exist a limit in which the fixed point is at weak coupling. (Note that this argument had nothing to do with supersymmetry—it could have just as well been done in non-supersymmetric QCD). It's implications for superQCD can be summarized in a kind of "phase diagram", see the figure. In the next lectures we will answer the question posed by the question mark.

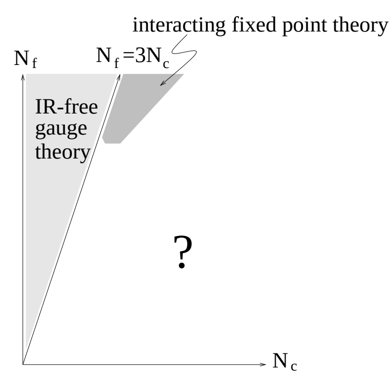

Figure 2.11: Partial "phase diagram" for superQCD theories as a function of the number of colors (N_c) and flavors (N_f).

Chapter 3

The Vacuum Structure of SuperQCD

We define superQCD as the $SU(N_c)$ gauge theory with N_f "quark" left-chiral superfields Q^i_a transforming in the $\mathbf{N_c}$ representation of the gauge group, and N_f "antiquarks" $\widetilde{Q}^a_i$ in the $\overline{\mathbf{N_c}}$. Here $a, b, c = 1, \ldots, N_c$ are color indices, and $i, j, k = 1, \ldots, N_f$ are flavor indices. The most general renormalizable action for superQCD, and the definition of the theory, is

$$S = \int d^4x \left\{ \frac{1}{2} \left[Q^\dagger{}_i e^V Q^i + \widetilde{Q}^{\dagger i} e^{-V} \widetilde{Q}_i \right]_D + \left[\frac{\tau}{8\pi i} \mathrm{tr}_f(W_L^2) + f(Q, \widetilde{Q}) + \text{c.c.} \right]_F \right\}. \quad (3.1)$$

Here the notation in the Kahler terms is meant to imply that e^V is in the fundamental representation, while e^{-V} is in the anti-fundamental. The possible renormalizable superpotential terms are mass terms for the quarks

$$f = m^i_j Q^j \widetilde{Q}_i. \quad (3.2)$$

Our aim is to extract the IR physics of this theory. By IR physics, I mean the physics at arbitrarily low energy scales—*i.e.* the vacuum structure and the massless particles. The reason for concentrating on the IR physics is that this is all that is captured by the Kahler and superpotential terms that we have been keeping in our effective actions. Any finite energy effects will presumably also get contributions from higher derivative terms in the effective action, but the arguments presented here will not be able to determine them.

As we emphasized, the non-renormalization theorems that we proved in section 2.4 for asymptotically free theories only applied to effective actions at scales above the strong coupling scale of the gauge theory, where we were assured of the complete description of the light degrees of freedom in terms of the quark and gluon superfields. In contrast, the non-perturbative description of the superQCD vacua that we are about to develop by definition probes physics well below the strong coupling scale. To do this we will have to make guesses about what the appropriated set of light degrees of freedom are, and then check that those guesses are self-consistent. The result is a compelling picture of the vacuum structure of superQCD; however, the nature of method means that this picture is not immediately generalizable to other gauge theories. In particular, the vacuum structure of supersymmetric gauge theories with chiral matter and/or a complicated enough representation structure has not been fully worked out by these methods.

3.1 Semi-classical superQCD

We start by analyzing the classical vacuum structure of the superQCD, taking into account one loop effects such as the running of gauge couplings and chiral anomalies.

3.1.1 Symmetries and vacuum equations

First, we look at the RG running of the gauge coupling. By our previous formulas, and recalling that for $SU(N_c)$

$$b = \frac{3}{2}T(\text{adj}) - \frac{1}{2}2N_f T(\mathbf{N_c}) = 3N_c - N_f, \tag{3.3}$$

the one-loop running is

$$\frac{8\pi^2}{g^2(\mu)} = (3N_c - N_f)\log\left(\frac{\mu}{\Lambda}\right). \tag{3.4}$$

Thus, the theory is AF for $N_f < 3N_c$ and IR-free for $N_f \geq 3N_c$. (For $N_f = 3N_c$, though there is no one-loop running, typically the two-loop running in the canonical beta function due to the anomalous dimensions of the quarks makes the theory IR-free.)

With zero superpotential, the theory has a non-Abelian global symmetry $U(N_f)$ rotating the quarks, and similarly for the antiquarks; in addition there is a $U(1)_R$ symmetry. We can choose a basis of the $U(1)$ factors to have the following action on the scalar components of the left-chiral superfields (we denote the scalar components by the same symbol as the whole superfield) and on the coupling Λ and superpotential masses m:

	$SU(N_c)$	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_A$	$U(1)_{R'}$	$U(1)_R$
Q^i_a	$\mathbf{N_c}$	$\mathbf{N_f}$	$\mathbf{1}$	1	1	1	$1-\frac{N_c}{N_f}$
$\widetilde{Q}^a_i$	$\overline{\mathbf{N_c}}$	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	-1	1	1	$1-\frac{N_c}{N_f}$
W_L	**adj**	$\mathbf{1}$	$\mathbf{1}$	0	0	1	1
$\Lambda^{3N_c-N_f}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	0	$2N_f$	$2N_c$	0
m^j_i	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	$\mathbf{N_f}$	0	-2	0	$2\frac{N_c}{N_f}$

Here the first column is the gauge symmetry, $U(1)_B$ is the "baryon number", the axial $U(1)_A$ and the $U(1)_{R'}$ as defined are anomalous, and in the last column we have defined a non-anomalous R-symmetry which is a linear combination of these two $U(1)$'s,

$$R = R' - \frac{N_c}{N_f}A. \tag{3.5}$$

(We chose a coefficient of 1 in front of R' to keep the vector superfield having R-charge 1.)

When there are no quarks, $N_f = 0$, the theory is then the superYM theory. The only potential (non-trivial) global symmetry is the anomalous $U(1)_{R'}$. The anomaly shifts the theta angle under a $U(1)_{R'}$ rotation by angle α as

$$\vartheta \to \vartheta + \alpha T(\text{adj}), \tag{3.6}$$

and can be used to shift $\vartheta = 0$. Thus the theta angle can have no observable effect in this theory. But since the theta angle is an angle, we see that a $\mathbb{Z}_{T(\text{adj})}$ discrete subgroup of the $U(1)_{R'}$ is unbroken. The superYM models thus have a discrete chiral symmetry. Aside from the discrete choice of the gauge group, superYM has no free parameters, since the strong coupling scale Λ is dimensionful.

In superQCD with $N_f > 0$ and when there is no superpotential, the scalar potential is just the D terms $V(Q, \widetilde{Q}) \sim \text{tr}D^2$. The condition for a supersymmetric vacuum is then

$$\begin{aligned} 0 = D &= \sum_A \left(Q^{\dagger a}_{\ i} (T_A^{(\mathbf{N_c})})^b_a Q^i_b + \widetilde{Q}^{\dagger i}_{\ a} (T_A^{(\overline{\mathbf{N_c}})})^a_b \widetilde{Q}^b_i \right) \\ &= \sum_A (T_A)^b_a \left(Q^{\dagger a}_{\ i} Q^i_b - \widetilde{Q}^a_i \widetilde{Q}^{\dagger i}_{\ b} \right), \end{aligned} \tag{3.7}$$

where in the first line $T^{(\mathbf{N_c})}$ are the generators in the fundamental and $T^{(\overline{\mathbf{N_c}})}$ are those of the anti-fundamental, and in the second line we have used that fact that the two are related by $(T_A^{(\overline{\mathbf{N_c}})})^a_b = -(T_A^{(\mathbf{N_c})})^a_b \equiv (T_A)^a_b$. Since the T_A are a basis of hermitian traceless matrices, the D term conditions can be written as

$$Q^{\dagger a}_{\ i} Q^i_b - \widetilde{Q}^a_i \widetilde{Q}^{\dagger i}_{\ b} = \frac{1}{N_c}(Q^{\dagger c}_{\ i} Q^i_c - \widetilde{Q}^c_i \widetilde{Q}^{\dagger i}_{\ c})\delta^a_b. \tag{3.8}$$

With non-zero superpotential (*i.e.* non-zero masses) there are also F term equations

$$m^i_j Q^j = m^i_j \widetilde{Q}_i = 0. \tag{3.9}$$

One can always use separate flavor rotations on the quarks and antiquarks to make the mass matrix diagonal. In that case the F equations simply set the vacuum expecatation values of those squarks and antisquarks with non-zero masses to zero. By sending the mass of one flavor to infinity one can decouple that flavor from the rest of the theory, leaving superQCD with one fewer flavor at low energies. By the usual one loop RG matching, the strong coupling scales Λ_{N_f} and Λ_{N_f-1} of the two theories will be related by

$$\Lambda_{N_f-1}^{3N_c-N_f+1} = m\Lambda_{N_f}^{3N_c-N_f} \tag{3.10}$$

where m is the mass of the heavy flavor. This is similar to the RG matching discussed in section 2.3.1, but with the matching done at the scale m of the heavy quark instead of at a vacuum expectation value ϕ controlling the mass of a Higgsed (heavy) gauge boson.

3.1.2 Classical vacua for $N_c > N_f > 0$

The D equations are not hard to solve, using the fact that by appropriate color and flavor rotations we can put an arbitrary Q^i_a into the diagonal form

$$Q = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_{N_f} \\ & & \\ & & \end{pmatrix}, \qquad a_i \in \mathbb{R}^+, \qquad (N_f < N_c). \tag{3.11}$$

Here the columns are labelled by the flavor index and the rows by the color index, and we have shown the result for $N_f < N_c$. Now $\widetilde{Q}$ can simultaneously be put in upper-diagonal form by the remaining $SU(N_f)$ symmetry rotations. Plugging into the D term equations then gives

$$\widetilde{Q}^T = Q. \tag{3.12}$$

Of course, we could have solved the D term conditions in a gauge invariant way by using our result from section 1.7.6 that the D-flat directions are parametrized by the algebraically independent set of holomorphic gauge invariant monomials in the fields. Such a basis (for $N_f < N_c$) is

$$M^i_j = \widetilde{Q}^a_j Q^i_a, \tag{3.13}$$

giving ${N_f}^2$ massless left-chiral superfields whose vacuum expectation values parametrize the moduli space of vacua. (We will discuss shortly how we know these are a basis of gauge invariant states.)

As a check on these results, we can count that the two answers imply the same dimension of our moduli space. From the solution (3.11) and (3.12) we see that at a generic point in the moduli space the gauge symmetry is spontaneously broken from $SU(N_c)$ to $SU(N_c - N_f)$, implying that $({N_c}^2 - 1) - [(N_c - N_f)^2 - 1] = 2N_f N_c - {N_f}^2$ gauge bosons get a mass. But, by the Higgs mechanism, each such massive gauge boson eats a left-chiral superfield, implying that of the original $2N_f N_c$ massless left-chiral superfields, only $2N_f N_c - (2N_f N_c - {N_f}^2) = {N_f}^2$ survive, matching the counting we found from the gauge-invariant solution.

Thus we see that the basic physics occuring here classically is just the Higgs mechanism: the squark vacuum expectation values generically break $SU(N_c) \to SU(N_c - N_f)$. Of course, for non-generic values of the squark vacuum expectation values, the unbroken gauge symmetry can be enhanced, corresponding to points where $\mathrm{rank}(M) < N_f$, or, equivalently, where $\det(M) = 0$.

We can also compute the classical Kahler metric on the moduli space. The Kahler form is $K = {Q^\dagger}^a_i Q^i_a + {\widetilde{Q}^\dagger}{}^i_a \widetilde{Q}^a_i$. The D term equations imply

$${Q^\dagger}^a_i Q^i_b = \widetilde{Q}^a_i {\widetilde{Q}^\dagger}{}^i_b \qquad (N_f < N_c) \tag{3.14}$$

since the trace terms automatically vanish for $N_f < N_c$. Squaring this equation gives $(M\overline{M})^i_j = (\widetilde{Q}^\dagger \widetilde{Q})^i_k (\widetilde{Q}^\dagger \widetilde{Q})^k_j$ which implies that $\widetilde{Q}^\dagger \widetilde{Q} = (M\overline{M})^{1/2}$, and so the Kahler potential is

$$K = 2\mathrm{Tr}(\overline{M}M)^{1/2}. \tag{3.15}$$

This implies the Kahler metric is singular whenever M is not invertible, corresponding to points of enhanced gauge symmetry.

One case that bears special mention is when $N_f = N_c - 1$. Then generically the gauge symmetry is completely broken (there is no $SU(1)$ group). In this case we might expect the IR physics to be under control even quantum mecahnically, since there are no AF gauge groups left.

3.1.3 Classical vacua for $N_f \geq N_c$.

In this case, diagonalizing Q and $\widetilde{Q}$ subject to the D term equations gives the solutions (up to gauge and flavor rotations)

$$\begin{aligned} Q &= \begin{pmatrix} a_1 & & & \\ & \ddots & & \\ & & a_{N_c} & \end{pmatrix}, \qquad a_i \in \mathbb{R}^+, \\ \widetilde{Q} &= \begin{pmatrix} \widetilde{a}_1 & & & \\ & \ddots & & \\ & & \widetilde{a}_{N_c} & \end{pmatrix}, \qquad (N_f \geq N_c), \end{aligned} \tag{3.16}$$

where

$$|\widetilde{a}_i|^2 = a_i^2 + \rho, \qquad \rho \in \mathbb{R}, \tag{3.17}$$

for some constant ρ independent of i. Thus, generically, the gauge symmetry is completely broken on the moduli space.

The gauge invariant description is in terms of the following set of holomorphic invariants:

$$\begin{aligned} M^i_j &= Q^i \widetilde{Q}_j && \text{"mesons"}, \\ B^{i_1 \ldots i_{N_c}} &= Q^{i_1}_{a_1} \cdots Q^{i_{N_c}}_{a_{N_c}} \epsilon^{a_1 \ldots a_{N_c}} && \text{"baryons"}, \\ \widetilde{B}_{i_1 \ldots i_{N_c}} &= \widetilde{Q}^{a_1}_{i_1} \cdots \widetilde{Q}^{a_{N_c}}_{i_{N_c}} \epsilon_{a_1 \ldots a_{N_c}} && \text{"anti-baryons"}. \end{aligned} \tag{3.18}$$

The baryons and anti-baryons vanished identically for $N_f < N_c$ because of the antisymmetrization of the squarks. It is clear that M, B, and $\widetilde{B}$ form a basis of gauge invariants, since they are all that can be made from the $SU(N_c)$ invariant tensors δ^a_b and $\epsilon^{a_1 \ldots a_{N_c}}$.

However, they are an overcomplete basis. One way of seeing this is to note that there are $2\binom{N_f}{N_c} + {N_f}^2$ meson and baryon fields, but by the Higgs mechanism there are only $2N_f N_c - ({N_c}^2 - 1)$ massless left-chiral superfields. Thus there must be relations among the baryons and mesons.

These constraints are easy to find. Since the product of two color epsilon tensors is the antisymmetrized sum of Kronecker deltas, it follows that

$$B^{i_1 \ldots i_{N_c}} \widetilde{B}_{j_1 \ldots j_{N_c}} = M^{[i_1}_{j_1} \cdots M^{i_{N_c}]}_{j_{N_c}}, \tag{3.19}$$

where the square brackets denote antisymmetrization. Also, since any expression antisymmetrized on N_c+1 color indices must vanish, it follows that any product of M's, B's, and $\widetilde{B}$'s antisymmetrized on N_c+1 upper or lower flavor indices must vanish.

A convenient notation for writing these constraints is to denote the contraction of an upper with a lower flavor index by a "$\cdot$", and the contraction of all flavor indices with the totally antisymmetric tensor on N_f indices by a "$*$". For example

$$(*B)_{i_{N_c+1}\dots i_{N_f}} = \epsilon_{i_1\dots i_{N_f}} B^{i_1\dots i_{N_c}}. \tag{3.20}$$

Then (3.19) can be rewritten in this notation as

$$(*B)\widetilde{B} = *(M^{N_c}). \tag{3.21}$$

The constraint coming from antisymmetrizing the N_c+1 flavor indices in the product of one M with a baryon is written

$$M \cdot *B = M \cdot *\widetilde{B} = 0. \tag{3.22}$$

As long as both B and $\widetilde{B}$ are non-zero, an induction argument shows that the above two constraints imply all the other D term constraints: A constraint with, say, k B's and an arbitrary number of M's antisymmetrized on N_c+1 upper indices can be replaced by a constraint with $k-1$ B fields by (3.21). Repeating this process reduces all constraints to (3.21) plus the single constraint with no B fields $*(M^{N_c+1}) = 0$. But this latter constraint is implied by (3.21) and (3.22): $0 = \widetilde{B}(M \cdot *B) = M \cdot *(M^{N_c}) = *(M^{N_c+1})$. When only one of B or $\widetilde{B}$ vanishes, the above arguments fail, and extra constraints would seem to be needed beyond (3.21) and (3.22). I do not know of a simple set of constraints in this case.

It will be useful to write out the simplest cases explicitly. The first case is when $N_f = N_c$. Then we expect ${N_f}^2$+1 massless left-chiral superfields, but we have ${N_f}^2$+2 invariants. The single constraint is just (3.21), which can be written more simply in terms of $*B$ and $*\widetilde{B}$ which have no flavor indices, as

$$y \equiv \det M - (*B)(*\widetilde{B}) = 0, \tag{3.23}$$

where we have used the definition of the determinant which amounts to $\det M = **(M^{N_c})$. Since

$$dy = (\det M)(M^{-1})^i_j dM^j_i - (*B)d(*\widetilde{B}) - (*\widetilde{B})d(*B), \tag{3.24}$$

singularities of the moduli space $y = dy = 0$ occur at

$$B = \widetilde{B} = *(M^{N_c-1}) = 0. \tag{3.25}$$

This last constraint implies rank$(M) < N_c - 1$, so (by referring back to our explicit solutions for Q and $\widetilde{Q}$) we see that there will be at least an unbroken $SU(2)$ gauge group.

In the case $N_f = N_c{+}1$, there are ${N_f}^2$ massless left-chiral superfields, and ${N_f}^2{+}2N_f$ invariants. (3.21) and (3.22) give ${N_f}^2{+}2N_f$ constraints. Therefore, the constraints are not independent in this case. Nevertheless, there is not a smaller set of holomorphic flavor covariant constraints.

When there are quark masses m^i_j the F equations simply amount to the constaraints that any contraction of the mass matrix with a meson or baryon flavor index vanishes:

$$m \cdot M = m \cdot B = m \cdot \widetilde{B} = 0. \tag{3.26}$$

Finally, from the charges of the elementary fields and the definition of the meson and baryon fields, we find their symmetry transformation properties:

	$SU(N_c)$	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_A$	$U(1)_{R'}$	$U(1)_R$
M	$\mathbf{1}$	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	2	2	$2{-}2\frac{N_c}{N_f}$
B	$\mathbf{1}$	$\binom{N_f}{N_c}$	$\mathbf{1}$	N_c	N_c	N_c	$N_c{-}\frac{{N_c}^2}{N_f}$
$\widetilde{B}$	$\mathbf{1}$	$\mathbf{1}$	$\binom{N_f}{N_c}$	$-N_c$	N_c	N_c	$N_c{-}\frac{{N_c}^2}{N_f}$
$\Lambda^{3N_c-N_f}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	0	$2N_f$	$2N_c$	0
m^j_i	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	$\mathbf{N_f}$	0	-2	0	$2\frac{N_c}{N_f}$

where we have included the transformation properties of the gauge coupling and masses as well.

3.2 Quantum superQCD: $N_f < N_c$

In the next five lectures we will study the IR physics of superQCD and will deduce how the classical vacuum structure of superQCD described above is modified by possible non-perturbative quantum corrections. We will work up in the number of flavors; see the review of K. Intriligator and N. Seiberg, hepth/9509066, for another presentation of this material.

Our starting point is therefore superYM theory, which we have seen has a global $\mathbb{Z}_{T(\text{adj})}$ chiral symmetry, a strong coupling scale Λ^{3N_c}, and no elementary scalar fields. What happens in this theory at strong coupling, *i.e.* at scales below Λ? Witten (Nucl. Phys. **B185** (1982) 253) showed that in these theories

$$\text{Tr}(-)^F = \tfrac{1}{2}T(\text{adj}). \tag{3.27}$$

Since this is non-zero, this implies that supersymmetry is not broken. Furthermore, there are at least $\frac{1}{2}T(\text{adj})$ discrete, degenerate vacua. It is a natural guess that the discrete chiral $\mathbb{Z}_{T(\text{adj})}$ symmetry is spontaneously broken, $\mathbb{Z}_{T(\text{adj})} \to \mathbb{Z}_2$, by gaugino condensation, $\langle\lambda\lambda\rangle \sim \Lambda^3 \exp\{4\pi i n/T(\text{adj})\}$ with $n = 1, \dots, \frac{1}{2}T(\text{adj})$, and that each of these $\frac{1}{2}T(\text{adj})$ vacua are "gapped"—*i.e.* they have no massless particles—since there are no flat directions emanating from them and there are no continuous chiral symmetries which might disallow mass terms. We will confirm this picture of the superYM vacua by indirect methods using the non-renormalization theorems and flowing down from $N_f > 0$ superQCD theories by turning on quark masses.[1]

Along a generic classical flat direction in superQCD we have seen that we Higgs the gauge group as $SU(N_c) \to SU(N_c - N_f)$ for $N_f < N_c - 1$, and completely break it otherwise. Far out along a classical flat direction where the scale of the vacuum expectation value of the squark is large compared to Λ, and the gauge group is completely Higgsed, the Higgsing takes place at arbitrarily weak coupling, and so the classical picture of the IR physics as just a nonlinear sigma model (*i.e.* general non-gauge theory) for the light fields M, B and $\widetilde{B}$ (subject to constraints) is reliable. In the cases where there is an unbroken gauge group, because the massless baryon and meson degrees of freedom along the flat directions are gauge neutral, all their couplings to the gauge fields are non-renormalizable. Thus in the IR we expect the theory to decouple into a nonlinear sigma model for the light M fields and an $SU(N_c - N_f)$ superYM theory. We expect the latter factor to have a gap, though, leaving only the nonlinear sigma model.

Assuming, then, that the meson and baryon left-chiral superfields are the correct IR degrees of freedom (at least at generic points on moduli space and with "large enough" vacuum expectation values), the next question to ask is: are the classical flat directions lifted quantum mechanically? The only possible term that could appear in the superpotential consistent with the symmetries and the selection rule for the anomalous symmetry is:

$$f \sim \left[\frac{\Lambda^{3N_c - N_f}}{\det M} \right]^{\frac{1}{N_c - N_f}} . \tag{3.28}$$

This follows since $\det M$ or one of its powers is the only $SU(N_f) \times SU(N_f)$ invariant, and the above powers are fixed by the $U(1)_A$ and $U(1)_R$ symmetries. In the weak coupling limit, $\Lambda \to 0$, we expect this contribution to vanish, and so it can only appear for $N_f < N_c$ because of the sign of the exponent. We thus learn that for $N_f \geq N_c$, no superpotential can be generated, and therefore that the classical flat directions are not lifted. We will explore the $N_f \geq N_c$ theories in later lectures.

If the superpotential term (3.28) *were* generated dynamically for $N_f < N_c$, what would its implications be? Since, qualitatively, $\det M \sim M^{N_f}$, the scalar potential

[1] Another expectation, coming from experience with real and lattice nonsupersymmetric QCD is that superYM will be color confining. We will find evidence of this only when we study the embedding of superYM theory in N=2 extended supersymmetric theories at the end of the course.

derived from the superpotential goes as

$$V = |f'|^2 \sim |M|^{-2N_c/(N_c-N_f)}. \tag{3.29}$$

This potential has no minimum and slopes to zero as $M \to \infty$. This implies that the resulting theory has no ground state: not only are there no supersymmetric vacua, but all potential "vacua" are unstable to moving to large $|M|$! The potential is, however, small at large vacuum expectation value of M, and so it is consistent with our assumptions.

The conclusion that the theory has no ground state is unfamiliar enough that it deserves some discussion. There are a number of ways one might think to avoid this conclusion. First, might quantum effects make the $M = \infty$ "point" in moduli space actually be at finite distance in field space? No, because for large M, we expect perturbation theory to be good, so the classical Kahler potential $K \sim |M|$ should be valid, implying that $M \to \infty$ really is infinitely far away in field space. Could quantum corrections to the Kahler potential for small M lead to new minima of the scalar potential? The reason this is a possibility is that the scalar potential is actually $V \sim g^{n\overline{n}} \partial_n f \partial_{\overline{n}} f^*$, where $g^{n\overline{n}}$ is the inverse Kahler metric. As long as this metric is not degenerate, then the only zeroes of the potential are when $\partial_n f = 0$, which we saw occurs only at $M = \infty$. Modifications to the Kahler metric could create *local* minima in the scalar potential at finite values of M, but these would always be metastable, since there would be lower energy states for large enough M. A final possibility is that quantum corrections actually make the Kahler metric singular at finite M. This is a breakdown of unitarity, implying that other massless fields would need to be added to our IR description. I do not see how to rule out such a possibility, however it seems unlikely that new massless degrees of freedom would enter and yet there would be no sign of them in the superpotential (recall that when we integrated out massless fields, we typically found singularities in the superpotential).

3.2.1 $N_f = N_c - 1$

How can we tell whether this superpotential really is generated dynamically? In the case $N_f = N_c - 1$, the gauge symmetry is completely broken, and instanton techniques can then reliably compute terms in the effective action. (This is because the IR divergence we mentioned before in the the instanton calculation in the superYM case is cut off by the scalar meson vacuum expectation value M.) Furthermore, in this case the superpotential goes as Λ^b, which is just what we expect from a one instanton effect. Such a one instanton calculation has indeed been carried out in the $N_c = 2$, $N_f = 1$ case, finding a non-zero result; for a summary and discussion of this computation, see D. Finnell and P. Pouliot, hepth/9503115. This result not only implies that the superpotential term is generated in the $SU(2)$ theory with one flavor, but also in all superQCD theories with $N_f = N_c - 1$.

To see this, assume the $SU(N_c)$ with $N_c - 1$ flavors generates the superpotential

$$f = c \cdot \frac{\Lambda^{2N_c+1}}{\det M}. \tag{3.30}$$

We will determine c by looking at a convenient flat direction, namely

$$\langle Q \rangle = \langle \widetilde{Q}^T \rangle = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_{N_f} \\ 0 & \cdots & 0 \end{pmatrix}, \qquad \text{with} \quad a_1, \ldots, a_{N_f-1} \sim \mu \gg a_{N_f} \gg \Lambda. \tag{3.31}$$

Along this direction, the large vacuum expectation values break $SU(N_c)$ with $N_f = N_c - 1$ massless flavors down to $SU(2)$ with one massless flavor. In that case the effective superpotential is

$$f = \hat{\Lambda}^5 / a_{N_f}^2 \tag{3.32}$$

where $\hat{\Lambda}$ is the strong coupling scale of the $SU(2)$ and the coefficient $c = 1$ by the instanton calculation. By RG matching, the $SU(2)$ scale is given by[2]

$$\hat{\Lambda}^5 = \Lambda^{2N_c+1}\mu^{4-2N_c} = \Lambda^{2N_c+1}\mu^{2-2N_f} = \Lambda^{2N_c+1}/(a_1^2 \cdots a_{N_f-1}^2). \tag{3.33}$$

Thus, comparing to (3.30), we see that $c = 1$.

3.2.2 $N_f \leq N_c - 1$: effects of tree-level masses

We can extend this result to other numbers of flavors by considering the effect of a tree level superpotential giving masses to the squarks. In the end, by matching to the Witten index result in the pure superYM theory in the infinite mass limit, we will give a separate argument for the appearance of the dynamical superpotential.

Consider adding to our $N_f = N_c - 1$ theory a tree-level superpotential term

$$f_t = m^i_j M^j_i \equiv \mathrm{Tr}(mM). \tag{3.34}$$

The selection rules from the broken flavor symmetries then imply that m can enter into the effective superpotential as

$$f = \mathrm{Tr}(mM) \cdot g\left(\frac{Mm}{\mathrm{Tr}(mM)}, \frac{\Lambda^{2N_c+1}}{(\det M)\mathrm{Tr}(mM)}\right) \tag{3.35}$$

[2]Actually, in this matching there can also be a threshold factor which enters as a possible multiplicative factor between the two sides. This factor is scheme-dependent (since it can be absorbed in a redefinition of what we mean by the strong coupling scales). There exists a scheme, however, in which the threshold factors are always 1, the "$\overline{\text{DR}}$-scheme".

for some holomorphic function g. When $m=0$ and $\Lambda=0$ we should recover the results

$$\begin{aligned} f(m{=}0) &= \frac{\Lambda^{2N_c+1}}{\det M}, \\ f(\Lambda{=}0) &= \mathrm{Tr}(mM). \end{aligned} \tag{3.36}$$

Taking the limits $m \to 0$ and $\Lambda \to 0$ in (3.35) in various ways, one can show by holomorphy that

$$f = \mathrm{Tr}(mM) + \frac{\Lambda^{2N_c+1}}{\det M} \qquad \text{for } N_f = N_c{-}1. \tag{3.37}$$

The F equation for M is then

$$0 = \frac{\partial f}{\partial M^i_j} = m^j_i - (M^{-1})^j_i \frac{\Lambda^{2N_c+1}}{\det M}. \tag{3.38}$$

This implies that

$$\det M = \Lambda^{(2N_c+1)(N_c-1)/N_c} (\det m)^{-1/N_c}, \tag{3.39}$$

and plugging back into (3.38) gives a supersymmetric vacuum at

$$\langle M^i_j \rangle = (m^{-1})^i_j (\det m)^{1/N_c} \Lambda^{(2N_c+1)/N_c}. \tag{3.40}$$

So, as long as we turn on *any* non-degenerate masses for the quarks, we find N_c discrete supersymmetric vacua because of the N_c-th root in (3.40). This is precisely what we expected physically, since after giving masses to the quarks the low energy theory should be pure superYM, which has N_c vacua according to the Witten index.

Instead of turning on a non-degenerate mass matrix, we can turn on a degenerate one so as to integrate out only some of the flavors:

$$m = \begin{pmatrix} 0 & 0 \\ 0 & \widehat{m} \end{pmatrix} \tag{3.41}$$

where the upper left-hand block is $N_f \times N_f$ for some $N_f < N_c{-}1$. Then, in the same block decomposition,

$$M = \begin{pmatrix} \widehat{M} & X \\ Y & Z \end{pmatrix} \tag{3.42}$$

where the fields in the X, Y, and Z blocks all get masses, and so can be integrated out. Letting $i, j = 1, \ldots, N_f$, and $I, J = 1, \ldots, N_c{-}1$, the equations of motion for the X and Y blocks are $0 = \partial f / \partial M^I_j = \partial f / \partial M^i_J$, giving $0 = (M^{-1})^I_j = (M^{-1})^i_J$, which imply

$$M^{-1} = \begin{pmatrix} \widehat{M}^{-1} & 0 \\ 0 & Z^{-1} \end{pmatrix} \quad \Rightarrow \quad M = \begin{pmatrix} \widehat{M} & 0 \\ 0 & Z \end{pmatrix} \tag{3.43}$$

so that the X and Y blocks vanish. The equation of motion for the Z block is

$$0 = \frac{\partial f}{\partial M^I_J} = \widehat{m}^J_I - (Z^{-1})^J_I \frac{\Lambda^{2N_c+1}}{\det \widehat{M} \det Z}, \tag{3.44}$$

implying that

$$\begin{aligned} \det Z &= \frac{\Lambda^{2N_c+1}}{\det \widehat{M}} \left(\frac{\det \widehat{M}}{\Lambda^{2N_c+1} \det \widehat{m}} \right)^{1/(N_c-N_f)}, \\ \langle Z^I_J \rangle &= (\widehat{m}^{-1})^I_J \left(\frac{\Lambda^{2N_c+1} \det \widehat{m}}{\det \widehat{M}} \right)^{1/(N_c-N_f)}, \end{aligned} \tag{3.45}$$

by a similar calculation as in (3.39) and (3.40). Plugging into (3.37) gives

$$\begin{aligned} f &= \mathrm{Tr}(\widehat{m}Z) + \frac{\Lambda^{2N_c+1}}{\det \widehat{M} \det Z} \\ &= (N_c - N_f) \left(\frac{\widehat{\Lambda}^{3N_c-N_f}}{\det \widehat{M}} \right)^{1/(N_c-N_f)}, \end{aligned} \tag{3.46}$$

where we have defined

$$\widehat{\Lambda}^{3N_c-N_f} \equiv (\det \widehat{m}) \Lambda^{2N_c+1}. \tag{3.47}$$

By the usual RG matching, we recognize this as the strong coupling scale of the $SU(N_c)$ theory with N_f flavors. Dropping the hats, (3.46) implies that the superpotential term is dynamically generated for the theories with $N_f < N_c - 1$ with coefficient $N_c - N_f$.

There are a few interesting points to note about this superpotential. A one instanton contribution goes as $\Lambda^b = \Lambda^{3N_c-N_f}$, but (3.46) goes as $\Lambda^{b/(N_c-N_f)}$. The interpretation of this is not clear: does it mean that there are semi-classical field configurations with fractional instanton number which compute this effect? In any case, the usual instanton contribution to the effective action is not well defined in this case due to the IR divergences from the unbroken $SU(N_c - N_f)$ gauge group.

A second interesting point is that these fractional powers imply the superpotential is multivalued as a function of $\langle M \rangle$. We can understand the meaning of this by considering the limit where $\det M \gg \Lambda^{N_f}$, so the theory is broken at a large scale down to $SU(N_c - N_f)$ classically. Since this occurs at weak coupling, we take as light degrees of freedom the M^i_j meson left-chiral superfields and the $SU(N_c - N_f)$ vsf W_α. From our non-renormalization theorem of section 2.4.2, the effective action for the W_α fields will be

$$S_{SU(N_c-N_f)} = \int d^4x \left[\frac{-\hat{b}}{16\pi^2} \log\left(\frac{\widehat{\Lambda}}{\mu} \right) \mathrm{tr} W_L^2 + \text{c.c.} \right]_F \tag{3.48}$$

where $\widehat{\Lambda}$ is the scale of the $SU(N_c{-}N_f)$ superYM theory and $\hat{b} = 3(N_c{-}N_f)$ is its beta function. By the usual RG matching

$$\widehat{\Lambda}^{3(N_c-N_f)} \det M = \Lambda^{3N_c-N_f} \tag{3.49}$$

where Λ is the scale of the original $SU(N_c)$ theory with N_f flavors. We argued earlier that the two sectors corresponding to the $SU(N_c{-}N_f)$ superYM theory and the M nonlinear sigma model decouple in the IR. However, the two sectors are coupled by irrelevant terms through the above dependence of $\widehat{\Lambda}$ on M. In particular, from (3.48) and (3.49) we see that $S_{SU(N_c-N_f)}$ will contain a term

$$S_{SU(N_c-N_f)} \supset \int d^4x \left[\frac{1}{64\pi^2} \mathrm{Tr}(F_M M^{-1})(\overline{\lambda}\mathcal{P}_+\lambda) + \text{c.c.} \right], \tag{3.50}$$

where we have expanded in components: M stands for the lowest component of the M left-chiral superfield, and F_M is its F component, while λ is the gaugino. On the other hand, the dynamically generated superpotential (3.46), from the M-sector of the theory gives rise to the terms

$$S_M \supset -\int d^4x \left[\mathrm{Tr}(F_M M^{-1}) \left(\frac{\Lambda^{3N_c-N_f}}{\det M} \right)^{1/(N_c-N_f)} + \text{c.c.} \right]. \tag{3.51}$$

Solving the F_M equations of motion from $S_M + S_{SU(N_c-N_f)}$ then gives

$$\langle \overline{\lambda}\mathcal{P}_+\lambda \rangle = 64\pi^2 \left(\frac{\Lambda^{3N_c-N_f}}{\det M} \right)^{1/(N_c-N_f)} = 64\pi^2 \hat{\Lambda}^3, \tag{3.52}$$

confirming the expected gaugino condensation in the pure superYM theory. Thus the $N_c{-}N_f$ branches in the superpotential (3.46) correspond to the $N_c{-}N_f$ vacua of the $SU(N_c{-}N_f)$ superYM theory.

Just as we did for the $N_f = N_c{-}1$ theory, we can add in tree-level masses m^i_j. The usual argument using holomorphy, symmetry, and weak coupling limits implies an effective superpotential and vacuum expectation value of the meson field

$$\begin{aligned} f &= \mathrm{Tr}(mM) + (N_c - N_f) \left(\frac{\Lambda^{3N_c-N_f}}{\det M} \right)^{1/(N_c-N_f)} \\ \langle M^i_j \rangle &= (m^{-1})^i_j \left(\Lambda^{3N_c-N_f} \det m \right)^{1/N_c} . \end{aligned} \tag{3.53}$$

This result was first obtained by A. Davis, M. Dine, and N. Seiberg, Phys. Lett. **B125** (1983) 487.

3.2.3 Integrating out and in

The technique of adding masses and integrating out massive degrees of freedom can be generalized, and in many cases is a useful tool for determining exact superpotentials;[3] I will present the basic idea in a somewhat simplified form explained in more detail in K. Intriligator and N. Seiberg, hepth/9509066, section 2.3.

Consider a gauge theory with scale Λ whose D-flat directions are parametrized by a set of gauge invariant composite left-chiral superfields $\mathcal{O}^i$. Then, the dynamics may generate an effective superpotential

$$f_{dyn} = f(\mathcal{O}^i, \Lambda^{b_0}). \tag{3.54}$$

We could probe this theory by adding tree-level couplings

$$f_{tree} = \sum_i J_i \mathcal{O}^i \tag{3.55}$$

to the theory, and then use holomorphy, symmetries and weak coupling limits to constrain the resulting effective superpotential, as we have done above. However, there are many cases in which this can be done more simply:

Think of the couplings J_i as sources for each light degree of freedom. *Assuming the dynamics is trivial (gaussian) in the IR* (so there are no IR divergences to keep the 1PI effective action from existing), we can compute the resulting effective superpotential as a 1PI effective superpotential, $\langle f\rangle$, by the usual Legendre transform[4]

$$\langle f\rangle(J_i, \Lambda^{b_0}) \equiv f_{dyn}(\mathcal{O}^i, \Lambda^{b_0}) + \sum_i J_i \mathcal{O}^i, \tag{3.56}$$

where we replace $\mathcal{O}^i$ on the right-hand side by inverting $J_i = -\partial f_{dyn}/\partial \mathcal{O}^i$, so that $\langle \mathcal{O}^i\rangle = \partial\langle f\rangle/\partial J_i$. In this case the effective superpotential is automatically linear in the sources:

$$f_{eff} = f_{dyn}(\mathcal{O}^i, \Lambda^{b_0}) + \sum_i J_i \mathcal{O}^i \tag{3.57}$$

and the Legendre transform just corresponds to integrating out the left-chiral superfields coupled to the sources.

This can be extended to the gaugino condensate left-chiral superfield

$$S \equiv -\frac{1}{16\pi^2} \mathrm{tr}_f(W_L^2) \tag{3.58}$$

[3] K. Intriligator, R. Leigh, and N.Seiberg, hepth/9403198; K. Intriligator, hepth/9407106.

[4] One may wonder why we can apply the 1PI effective *action* technology to the *superpotential*. This follows simply from the fact that we add sources to left-chiral superfields in the Lagrangian as $\mathcal{L} = \cdots + \int d^2\theta\, J_i \mathcal{O}^i$. So, to compute $\langle \mathcal{O}^i\rangle$ we must differentiate with respect to the F-component of the source left-chiral superfield: $\langle \mathcal{O}^i\rangle = (\partial/\partial F_{J_i}) \int d^2\theta\, \langle f\rangle = (\partial/\partial F_{J_i})[\sum_j (\partial\langle f\rangle/\partial J_j) F_{J_j}] = \partial\langle f\rangle/\partial J_j$.

as well, by treating $\log \Lambda^b$ as its source:

$$f_{eff} = f_{dyn} + \sum_i J_i \mathcal{O}^i + \log \Lambda^b \, S. \tag{3.59}$$

Since the Legendre transform is invertible, we can reverse this procedure and "integrate in" fields as well. As an example, consider the pure $SU(N_c)$ superYM theory, where we have

$$\langle S \rangle = \left(\Lambda^{3N_c}\right)^{1/N_c} = \frac{\partial \langle f \rangle(\Lambda)}{\partial(\log \Lambda^{3N_c})}. \tag{3.60}$$

Solving for $\langle f \rangle$ gives $\langle f \rangle = N_c \Lambda^3$, and taking the (inverse) Legendre transform with respect to the source $\log \Lambda^{3N_c}$ then gives

$$f_{dyn}(S) = \langle f \rangle - \log \Lambda^{3N_c} \cdot S = N_c S(1 - \log S). \tag{3.61}$$

Thus

$$f_{eff} = f_{dyn} + \log \Lambda^{3N_c} \cdot S = S \left[\log \left(\frac{\Lambda^{3N_c}}{S^{N_c}} \right) + N_c \right], \tag{3.62}$$

a result first obtained by G. Veneziano and S. Yankielowicz, Phys. Lett. **B113** (1982) 321, and T. Taylor, G. Veneziano, and S. Yankielowicz, Nucl. Phys. **B218** (1993) 493. However, the meaning of this effective superpotential is not clear, since it implies the left-chiral superfield S is always massive.

3.3 Quantum superQCD: $N_f \geq N_c$

We now move up in the number of flavors to the $N_f = N_c$ and $N_f = N_c+1$ cases. These were first solved by N. Seiberg in hep-th/9402044.

In the last lecture we found for $N_f < N_c$ that

$$\langle M^i_j \rangle = (m^{-1})^i_j (\Lambda^{3N_c - N_f} \det m)^{1/N_c}, \tag{3.63}$$

for an arbitrary mass matrix m. It is not hard to see from symmetries and holomorphy that this expression is the only one allowed, even for $N_f \geq N_c$, though this does not fix its coefficient. But if we consider a theory with $N_f > N_c$ and take masses such that

$$m_1, \ldots, m_{N_c - 1}, \Lambda \ll m_{N_c}, \ldots, m_{N_f}, \tag{3.64}$$

and integrate out the heavy masses, we arrive at an effective $SU(N_c)$ theory with N_c-1 light flavors with a strong coupling scale

$$\hat{\Lambda}^{3N_c - (N_c - 1)} = m_{N_c} \cdots m_{N_c} \Lambda^{3N_c - N_f}, \tag{3.65}$$

by the usual RG matching. Plugging into (3.63) then implies that (3.63) must also hold for all N_c and N_f.

Now, consider taking the limit in (3.63) as $m \to 0$. For $N_f < N_c$ this limit always implied $M \to \infty$, and so there was no vacuum. But for $N_f \geq N_c$, we can take the limit in such a way that M remains finite. By taking $m \to 0$ in different ways, we can "map out" the space of vacua of the $N_f \geq N_c$ theories. The fact that flat directions survive in these theories accords with the fact that no superpotential is dynamically generated.

3.3.1 $N_f = N_c$

In this case, recall that the classical moduli space was parameterized by the "meson" and "baryon" composite left-chiral superfields

$$\begin{aligned} M^i_j &= Q^i \widetilde{Q}_j, \\ B &= Q^{i_1}_{a_1} \cdots Q^{i_{N_c}}_{a_{N_c}} \epsilon^{a_1 \ldots a_{N_c}} \epsilon_{i_1 \ldots i_{N_c}}, \\ \widetilde{B} &= \widetilde{Q}^{a_1}_{i_1} \cdots \widetilde{Q}^{a_{N_c}}_{i_{N_c}} \epsilon_{a_1 \ldots a_{N_c}} \epsilon^{i_1 \ldots i_{N_c}}, \end{aligned} \tag{3.66}$$

which satisfy the constraint

$$\det M - (*B)(*\widetilde{B}) = 0. \tag{3.67}$$

Now, turning on meson masses, by (3.63) gives

$$M = m^{-1} (\det m)^{1/N_c} \Lambda^2, \qquad \Rightarrow \qquad \det M = \Lambda^{2N_c}. \tag{3.68}$$

Since this last formula is independent of m, it will be true in the $m \to 0$ limit. Also, note that when $\det m \neq 0$, that the baryon expectation values must vanish,

$$*B = *\widetilde{B} = 0 \qquad \text{if } \det m \neq 0, \tag{3.69}$$

since the vacuum must transform trivially under $U(1)_B$ because all the fields carrying baryon number are integrated out if $\det m \neq 0$. Taking the limit $m \to 0$, we conclude that $*B = *\widetilde{B} = 0$. These conclusions, (3.68) and (3.69), are not consistent with the classical constraint (3.67). Therefore, the *classical constraints are modified quantum mechanically*, even though no superpotential is dynamically generated.

To see what the quantum modified constraints are, we need to do a little more work, since so far we have only probed the vacua by adding a source for M. By the symmetries, holomorphy, the fact that $\det M = \Lambda^{2N_c}$ when $*B = *\widetilde{B} = 0$, and from demanding that in the weak-coupling limit $\Lambda \to 0$ the quantum constraint reduce to the classical one, the general form of the quantum constraint must be

$$\det M - (*B)(*\widetilde{B}) = \Lambda^{2N_c} \left(1 + \sum_{\alpha,\beta>0} c_{\alpha,\beta} \frac{(\Lambda^{2N_c})^\alpha (*B*\widetilde{B})^\beta}{(\det M)^{\alpha+\beta}} \right). \tag{3.70}$$

Classically we have vacua with arbitrary values of $\langle *B \rangle$ and $\langle *\widetilde{B} \rangle$, so by going far enough out on the classical moduli space where the physics is the Higgs mechanism taking place at arbitrarily weak coupling, we are assured that there will be vacua of the full quantum theory with non-zero baryon vevs with associated meson vacuum expectation values which statisfy the classical constraint arbitrarily well. But, fixing $*B*\widetilde{B}$ at some large (compared to Λ^{2N_c}) constant value, we see that in addition to the asymptotically classical solution with $\det M \sim *B*\widetilde{B}$, any non-zero $c_{\alpha,\beta}$ gives rise to additional solutions going as $\det M \sim (*B*\widetilde{B})^{(\beta-1)/(\beta+\alpha)}$. Such solutions extend out to the perturbative regime of large meson and baryon vacuum expectation values; since no vacua like this are seen in perturbation theory, we must have all the $c_{\alpha,\beta} = 0$, giving the quantum constraint:

$$0 = y = \det M - (*B)(*\widetilde{B}) - \Lambda^{2N_c}. \tag{3.71}$$

What is the physics of these vacua? First, note that $dy = 0$ only at $*B = *\widetilde{B} = \det M = 0$, which does not lie on the constraint surface $y = 0$. Thus we do not expect any enhanced gauge symmetries on this moduli space. Furthermore, while the $*B$, $*\widetilde{B}$ and M vacuum expectation values at typical points on the moduli space spontaneously break all the global symmetries, there are special submanifolds where the global symmetry can be enhanced. For example, at the point $M^i_j = \Lambda^2 \delta^i_j$, $*B = *\widetilde{B} = 0$, the global $SU(N_f) \times SU(N_f) \times U(1)_B \times U(1)_R$ symmetry is only broken to $SU(N_f)_{diag} \times U(1)_B \times U(1)_R$, and the light degrees of freedom are the $*B$ and $*\widetilde{B}$ baryons, as well as the Golstone bosons of the diagonal breaking of the flavor symmetry. This, then, is a supersymmetric version of a vacuum with chiral symmetry breaking, and massless pions and baryons. Another enhanced symmetry point is $M = 0$ and $*B = *\widetilde{B} = i\Lambda^{N_c}$, where only the $U(1)_B$ of the global symmetry is broken. There is no chiral symmetry breaking, and the light fields are the mesons M, as well as a $B\widetilde{B}$ composite (the Goldstone boson of the baryon number).

The difference between the classical and quantum moduli spaces can be summarized by the following cartoon: Classically, the physics is the Higgs mechanism, and at the singularity at the origin, the gauge symmetry is unbroken so there are massless quarks and gluons. In the quantum theory, on the other hand, there is no vacuum with massless gluons, it being replaced by the circle of theories at the neck of the hyperboloid which

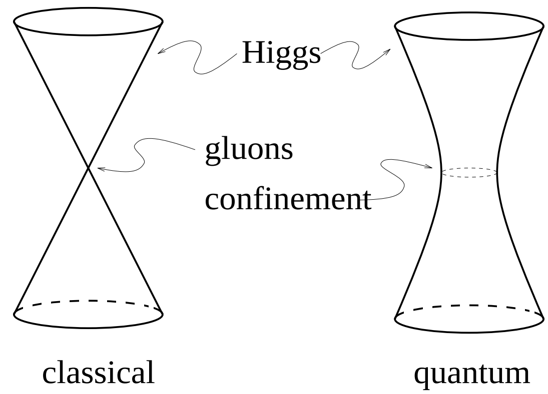

Figure 3.1: $N_f = N_c$ cartoon

have chiral symmetry breaking. This is the expected physics of a confining vacuum. We see that in this theory there is no phase transition separating a "Higgs phase" from a "confining phase". This is in accord with the fact that we have squarks in the fundamental representation which can screen any sources in a Wilson loop.

One may wonder how one can generate a superpotential by integrating out some quarks from this theory, when it doesn't have a superpotential to start with. The point, however, is that the *fluctuations* (and not just the vacuum expectation values) of the meson and baryon fields are constrained by (3.71). One way of seeing this is to note that even after turning on meson masses—which should enable us to probe possible vacua off the constraint surface if they exist—the meson vacuum expectation value still satisfies the constraint (3.68). One can not just naively integrate out the meson fields without taking into account the constraint which couples the meson and baryon fluctuations. To impose this constraint in the action, we add a Lagrange multiplier left-chiral superfield, A, to enforce the constraint. The Lagrange multiplier can be thought of as a left-chiral superfield with no kinetic (Kahler) terms, and therefore no fluctuations. The superpotential (with mass term for the squarks) becomes

$$f = \mathrm{Tr}(mM) + A\left[\det M - (*B)(*\widetilde{B}) - \Lambda^{2N_c}\right]. \tag{3.72}$$

Taking the mass matrix and meson field to be of the form

$$m = \begin{pmatrix} 0 & 0 \\ 0 & \hat{m} \end{pmatrix}, \qquad M = \begin{pmatrix} \hat{M} & X \\ Y & Z \end{pmatrix}, \tag{3.73}$$

where the upper left-hand block is $N_f \times N_f$, one can then use the F equations of motion for the M, $*B$, $*\widetilde{B}$, and A fields to show that $*B = *\widetilde{B} = X = Y = 0$, and solve for the others (just as in the last lecture) giving

$$f = (N_c - N_f)\left(\frac{\hat{\Lambda}^{3N_c-N_f}}{\det \hat{M}}\right)^{1/(N_c-N_f)} \qquad \text{with} \quad \hat{\Lambda}^{3N_c-N_f} \equiv \Lambda^{2N_c}\det \hat{m}. \tag{3.74}$$

3.3.2 $N_f = N_c{+}1$

Recall from the discussion in section 3.1.1 that a basis of composite left-chiral superfields in this case is M^i_j, $*B_i$, and $*\widetilde{B}^i$, satisfying the classical constraints:

$$\begin{aligned} 0 &= (M^{-1})^i_j \det M - (*B)_j(*\widetilde{B})^i, \\ 0 &= (*B)_i M^i_j = M^i_j(*\widetilde{B})^j. \end{aligned} \tag{3.75}$$

We probe the quantum moduli space by turning on quark masses m^i_j. As before, when $\det m \neq 0$, $*B_i = *\widetilde{B}^i = 0$ and

$$M^i_j = (m^{-1})^i_j(\Lambda^{2N_c-1}\det m)^{1/N_c}, \tag{3.76}$$

which imply, in particular, that

$$(*B\cdot M)_i = (M\cdot *\widetilde{B})^i = 0, \quad \text{and} \quad (M^{-1})^i_j \det M = \Lambda^{2N_c-1}m^i_j. \tag{3.77}$$

Unlike the $N_f = N_c$ case where turning on masses kept $\langle M\rangle$ on the constraint surface ($\det M = \Lambda^{2N_c}$), turning on masses in this case allows M to take any value off the constraint surface. This implies we will not be able to implement the quantum constraints with Lagrange mulitpliers in the superpotential—they will have to arise as equations of motion. Also unlike the $N_f = N_c$ case, in the limit $m \to 0$, $\langle M\rangle$ is on the *classical* constraint surface. The possible corrections to the classical constraints consistent with this data involve positive powers of $*B\cdot M\cdot *\widetilde{B}/\det M$ by the symmetries; assuming that turning on baryon sources can probe vacua with arbitrary baryon vacuum expectation values, all these terms must vanish by taking appropriate $M \to 0$ limits. In this way we see that the classical constraints (3.67) remain valid in the full quantum theory.

To see how these can arise as equations motion, we write down the most general dynamical superpotential (consistent with the symmetries):

$$f = \frac{1}{\Lambda^{2N_c-1}}\left[\alpha(*B\cdot M\cdot *\widetilde{B}) + \beta\det M + \det M\, f\left(\det M/{*B}{\cdot}M{\cdot}{*\widetilde{B}}\right)\right]. \tag{3.78}$$

We normally would not allow such a term since it does not vanish in the weak coupling limit $\Lambda \to 0$; however, in this case we will see that it reproduces the classical constraints, so it can be kept. The arbitrary function f must vanish in order to have a smooth $M \to 0$ limit; alternatively, only $f = 0$ will reproduce the classical constraints. The F equations of motion are:

$$\begin{aligned} \frac{\partial f_{eff}}{\partial M} \quad &\Rightarrow \quad 0 = \alpha(*B)_j(*\widetilde{B})^i + \beta(M^{-1})^i_j \det M = 0, \\ \frac{\partial f_{eff}}{\partial(*B, *\widetilde{B})} \quad &\Rightarrow \quad 0 = (*B \cdot M)_i = (M \cdot *\widetilde{B})^i. \end{aligned} \tag{3.79}$$

These are the classical constraints if $\alpha = -\beta$.

Adding in a single mass to integrate out one flavor matches to the $N_f = N_c$ case when $\alpha = 1$. The algebra is as follows. In the superpotential

$$f = \frac{1}{\Lambda^{2N_c-1}} \left(*B \cdot M \cdot *\widetilde{B} - \det M \right) + \mathrm{Tr}(mM), \tag{3.80}$$

let

$$m = \begin{pmatrix} 0 & 0 \\ 0 & \widehat{m} \end{pmatrix}, \quad M = \begin{pmatrix} \widehat{M} & X \\ Y & Z \end{pmatrix}, \quad *B = \begin{pmatrix} W \\ *\widehat{B} \end{pmatrix}, \quad *\widetilde{B} = \begin{pmatrix} \widetilde{W} \\ *\widehat{\widetilde{B}} \end{pmatrix}, \tag{3.81}$$

where the upper left-hand blocks are $N_c \times N_c$. We integrate out X, Y, W, and $\widetilde{W}$ using the equations of motion, leaving us with the equations $\det M = Z \det \widehat{M}$, $*B \cdot M \cdot *\widetilde{B} = Z * \widehat{B} * \widehat{\widetilde{B}}$, and $\mathrm{Tr} mM = Z\Lambda^{2N_c-1}\widehat{m}$, which, when plugged back into f give

$$f = \frac{Z}{\Lambda^{2N_c-1}} \left(*\widehat{B} * \widehat{\widetilde{B}} - \det \widehat{M} + \widehat{\Lambda}^{2N_c} \right), \quad \text{where } \widehat{\Lambda}^{2N_c} \equiv \widehat{m}\Lambda^{2N_c-1}. \tag{3.82}$$

Dropping the hats and identifying Z/Λ^{2N_c-1} with the Lagrange multiplier field A, we indeed recover the $N_f = N_c$ case. Note also that with $\widehat{\Lambda}$ fixed, $\widehat{m} \to \infty$ implies $\Lambda \to 0$. Thus the kinetic terms for A go as $(\partial Z)^2 \sim \Lambda^{4N_c-2}(\partial A)^2 \to 0$, showing that A is indeed a Lagrange multiplier and not a fluctuating field.

We have just shown that the classical and quantum moduli spaces of the $N_f = N_c+1$ theories are the same. In particular, unlike the $N_f = N_c$ case, the singular point at $M = *B = *\widetilde{B} = 0$ remains in the moduli space. Classically this was the point with unbroken $SU(N_c)$ gauge group and massless quarks and gluons. Quantum mechanically, it seems to be a point with massless meson and baryon composites, confinement (no gluons), and no chiral symmetry breaking. On the other hand, we have seen before that singularities in the holomorphic coordinate description of the moduli space are often (though not necessarily) associated with new light degrees of freedom that were not included in our original effective action. How can we tell if that is what actually occurs in this case?

While there is no proof that there cannot be new light degrees of freedom at the origin, the following argument suggests that there are not. We can test the consistency of assuming that only the composite meson and baryon fields are the light degrees of freedom at the origin through the 't Hooft anomaly matching conditions. At the origin, the full global symmetry group is unbroken, under which the microscopic quark left-chiral superfields and the macroscopic meson and baryon left-chiral superfields have charges:

	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_R$
W_L	$\mathbf{N_c{}^2{-}1}$	1	0	1
Q	$\mathbf{N_f}$	1	$+1$	$\frac{1}{N_f}$
$\widetilde{Q}$	1	$\overline{\mathbf{N_f}}$	-1	$\frac{1}{N_f}$
M	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$\frac{2}{N_f}$
B	$\overline{\mathbf{N_f}}$	$\mathbf{1}$	$N_f{-}1$	$1{-}\frac{1}{N_f}$
$\widetilde{B}$	$\mathbf{1}$	$\mathbf{N_f}$	$1{-}N_f$	$1{-}\frac{1}{N_f}$

In terms of the microscopic and macroscopic fermion fields this gives

	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_R$
λ	$\mathbf{N_c{}^2{-}1}$	1	0	1
ψ_Q	$\mathbf{N_f}$	1	$+1$	$\frac{1}{N_f}{-}1$
$\psi_{\widetilde{Q}}$	1	$\overline{\mathbf{N_f}}$	-1	$\frac{1}{N_f}{-}1$
ψ_M	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$\frac{2}{N_f}{-}1$
ψ_B	$\overline{\mathbf{N_f}}$	$\mathbf{1}$	$N_f{-}1$	$-\frac{1}{N_f}$
$\psi_{\widetilde{B}}$	$\mathbf{1}$	$\mathbf{N_f}$	$1{-}N_f$	$-\frac{1}{N_f}$

One can then check that all the anomalies match. For example:

$$\begin{aligned}\mathrm{Tr}R &= 2N_fN_c\cdot\left(\frac{1}{N_f}-1\right)+\left(N_c{}^2-1\right)\cdot 1=-N_f{}^2+2N_f-2 \qquad \text{(micro)}\\ &= N_f{}^2\cdot\left(\frac{2}{N_f}-1\right)+2N_f\cdot\left(-\frac{1}{N_f}\right)=-N_f{}^2+2N_f-2 \qquad \text{(macro)}\end{aligned} \tag{3.83}$$

and

$$\begin{aligned}\mathrm{Tr}R^3 &= 2N_fN_c\cdot\left(\frac{1}{N_f}-1\right)^3+\left(N_c{}^2-1\right)\cdot 1^3=-N_f{}^2+6N_f-12+\frac{8}{N_f}-\frac{2}{N_f{}^2} \qquad \text{(micro)}\\ &= N_f{}^2\cdot\left(\frac{2}{N_f}-1\right)^3+2N_f\cdot\left(-\frac{1}{N_f}\right)^3=-N_f{}^2+6N_f-12+\frac{8}{N_f}-\frac{2}{N_f{}^2} \qquad \text{(macro)}\end{aligned} \tag{3.84}$$

etc.. Because we compute the anomaly by counting states only if their kinetic terms are non-singular, the matching of the anomalies can be taken as evidence for the Kahler potential being smooth at the origin.

In summary, for the $N_f = N_c{+}1$ theories, we have seen that the quantum and classical moduli spaces are the same. The classical moduli space was described by constraints which arose "trivially" from the definition of the composite left-chiral superfields in terms of the microscopic quark left-chiral superfields; while those same constraints in the quantum theory arose as equations of motion.

3.3.3 $N_f \geq N_c{+}2$

Just as in the $N_f = N_c{+}1$ case, we can probe the quantum moduli space by turning on masses m^i_j and using

$$\langle M^i_j \rangle = (m^{-1})^i_j (\Lambda^{3N_c - N_f} \det m)^{1/N_c}. \tag{3.85}$$

For $m \neq 0$ all values of M can be obtained, and by taking $m \to 0$ in various ways we again find that we can arrive at any point on the classical moduli space with vanishing baryon vacuum expectation values. This then implies, using the symmetries and weak coupling limits, that the quantum moduli space must coincide with the classical one.

This immmediately raises the question of the interpretation of the singularity at $M = B = \widetilde{B} = 0$. Unlike the $N_f = N_c$ case, the superpotential (which gives rise to the constraints as equations of motion, and goes as $(\det M)^{1/(N_f - N_c)}$) in this case is singular at the origin, which is a sign that there are extra light degrees of freedom there. Also, the 't Hooft anomaly-matching conditions are not satisfied if one assumes that only M, B, and $\widetilde{B}$, are light there. We thus have a kind of "phase diagram" for the vacua at the origin of moduli space of superQCD; see the figure. We will spend the next two lectures answering the question posed by the question mark.

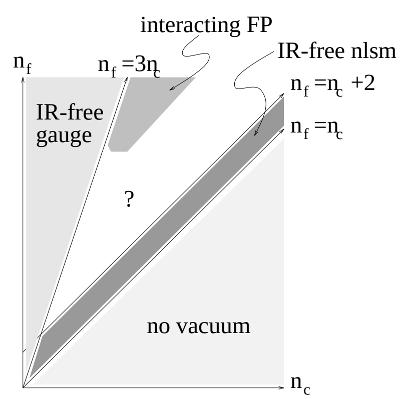

Figure 3.2: superQCD phase diag

3.4 Superconformal invariance

How can we tell the difference between a free scale invariant theory and an interacting one? In a scale invariant field theory the 2-point correlator of a (Lorentz scalar for simplicity) field is

$$\langle\phi(x_1)\phi(x_2)\rangle \propto \frac{1}{|x_1 - x_2|^{2d}} \tag{3.86}$$

where d is the *scaling dimension* of ϕ. For example, for a free scalar field $d = 1$, for a free fermion $d = 3/2$, and for a free $U(1)$ field strength tensor $F^{\mu\nu}$ $d = 2$. At the level of 2-point functions, we might expect the scaling dimensions of interacting fields to be different from their canonical values, as interactions generically give rise to anomalous scaling dimensions. With an extra assumption, we can prove this; we can also derive restrictions on the allowed ranges of scaling dimensions following from unitarity. The extra assumption is conformal invariance, and we will explore its implications in this lecture.

A scale invariant theory (one with vanishing beta functions) is one which is invariant not only under the Poincaré algebra,

$$\begin{aligned} i[J^{\mu\nu}, J^{\rho\sigma}] &= \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\rho} + \eta^{\mu\sigma}J^{\nu\rho}, \\ i[P^\mu, J^{\rho\sigma}] &= \eta^{\mu\rho}P^\sigma - \eta^{\mu\sigma}P^\rho, \\ i[P^\mu, P^\nu] &= 0, \end{aligned} \tag{3.87}$$

but also under dilatations (scale transformations), which we take to be generated by an operator D. We define an operator $\mathcal{O}$ to have *scaling dimension* d if it satisfies the commutation relation

$$i[D, \mathcal{O}] = d\,\mathcal{O}. \tag{3.88}$$

The only non-zero commutator of dilatations with the generators of the Poincaré algebra is

$$i[D, P^\mu] = P^\mu, \tag{3.89}$$

since energy momentum has scaling dimension 1, while the generators $J^{\mu\nu}$ of Lorentz rotations have dimension 0,

$$i[D, J^{\mu\nu}] = 0, \tag{3.90}$$

(*i.e.* D is a Lorentz scalar), and, of course,

$$i[D, D] = 0. \tag{3.91}$$

These commutation relations have a geometrical realization on space-time corrdinates through the differential operators

$$J^{\mu\nu} = -i[x^\mu\partial^\nu - x^\nu\partial^\mu],$$

$$\begin{aligned} P^\mu &= -i\partial^\mu, \\ D &= ix^\mu\partial_\mu, \end{aligned} \tag{3.92}$$

showing that D does indeed generate scale transformations.

This algebra has a unique extension to the larger algebra of *conformal transformations*. It is thought to be only a mild assumption that scale invariant quantum field theories are actually conformally invariant. In particular, there is no known example (I think) of an interacting scale invariant but not conformally invariant 4-dimensional quantum field theory [J. Polchinksi, Nucl. Phys. **B303** (1988) 226].

The conformal algebra has in addition to the Poincaré and dilatation generators, a vector of *conformal generators* K^μ, satisfying the algebra

$$\begin{aligned} i[K^\mu, J^{\rho\sigma}] &= \eta^{\mu\rho}K^\sigma - \eta^{\mu\sigma}K^\rho, \\ i[K^\mu, K^\nu] &= 0, \\ i[D, K^\mu] &= -K^\mu, \\ i[P^\mu, K^\nu] &= 2\eta^{\mu\nu}D + 2J^{\mu\nu}. \end{aligned} \tag{3.93}$$

The first commutator implies that K^μ transforms as a Lorentz vector, and the third implies it has scaling dimension -1.

3.4.1 Representations of the conformal algebra

We are interested in the representations of this symmetry on the Hilbert space of states in a quantum field theory. Thus we are interested in unitary representations of the conformal algebra, that is, representations in which the generators all act as Hermitian operators. As with the unitary of the Poincaré group discussed in section 1.4.1, the conformal group is not compact and so all its unitary representations are infinite dimensional (except the trivial representation). In fact, the conformal algebra is equivalent to $SO(2,4)$, the algebra of rotations and boosts in a six dimensional space with two time-like directions. This can be seen by defining the combinations of generators $\hat{J}^{pq}$ with $p,q = -1,0,1,2,3,4$ and

$$\begin{aligned} \hat{J}^{\mu\nu} &= J^{\mu\nu}, \\ \hat{J}^{-1,4} &= D, \\ \hat{J}^{\mu,-1} &= \tfrac{1}{2}[P^\mu + K^\mu], \\ \hat{J}^{\mu,4} &= \tfrac{1}{2}[P^\mu - K^\mu]. \end{aligned} \tag{3.94}$$

One can then check that the $\hat{J}^{pq}$ satisfies the analog of the Lorentz algebra with a metric $\eta^{pq} = \mathrm{diag}\{-1,-1,1,1,1,1\}$.

As with our discussion of the little group of the Poicaré group, unitary representations of the conformal group are classified by the irreducible representations of the maximal compact part $SO(2) \times SO(4) \subset SO(2,4)$ of the conformal group. Recalling that $SO(4) \simeq SU(2)_L \times SU(2)_R$ (see the discussion at the end of section 1.3.2) we can label the unitary representations of $SO(4)$ by the left and right $SU(2)$ spins (j_l, j_R). Together with the $SO(2)$ eigenvalue d, these spins label the unitary representations of the conformal group. The non-compact generators act as raising and lowering operators, taking us between different states in a given representation.

Define, then, new generators by

$$\begin{aligned} \hat{D} &= \hat{J}^{-1,0}, \\ \hat{P}^a &= \hat{J}^{a,-1} + i\hat{J}^{a,0}, \\ \hat{K}^a &= \hat{J}^{a,-1} - i\hat{J}^{a,0}, \end{aligned} \tag{3.95}$$

where $a, b = 1, 2, 3, 4$. Then $\hat{D}$ generates the $SO(2)$, and the Hermitian conjugate $\hat{P}^a$ and $\hat{K}^a$ are the raising and lowering operators. The remaining $\hat{J}^{ab}$ generate $SO(4)$ rotations. These generators satisfy the same commutation relations as the unhatted generators in (3.87–3.93), but with $\eta^{\mu\nu}$ replaced by δ^{ab}.

It will be useful to replace the four a, b indices with $\alpha, \beta = 1, 2$ indices for $SU(2)_L$ and $\dot{\alpha}, \dot{\beta} = 1, 2$ indices for $SU(2)_R$. (This will be convenient for the supersymmetric generalization.) Thus $\hat{D}$, $\hat{P}^{\alpha\dot{\alpha}}$, $\hat{K}^{\alpha\dot{\alpha}}$, $M_L^{\alpha\beta}$, and $M_R^{\dot{\alpha}\dot{\beta}}$, respectively generate dilatations, translations, special conformal transformations, and the $SU(2)_L \times SU(2)_R \in SO(4)$ rotations. The $SO(4)$ algebra and charges become in this notation

$$\begin{aligned} [M_L^{\alpha\beta}, M_L^{\gamma\delta}] &= i(M_L^{\alpha\delta}\epsilon^{\beta\gamma} + M_L^{\alpha\gamma}\epsilon^{\beta\delta} + M_L^{\beta\delta}\epsilon^{\alpha\gamma} + M_L^{\beta\gamma}\epsilon^{\alpha\delta}) \\ [M_L^{\alpha\beta}, X^\gamma] &= i(X^\alpha\epsilon^{\beta\gamma} + X^\beta\epsilon^{\alpha\gamma}), \end{aligned} \tag{3.96}$$

where X is any generator with a single undotted index. The same formulas hold for $SU(2)_R$ (*i.e.* with dotted indices). Here $\epsilon^{\alpha\beta}$ and $\epsilon^{\dot{\alpha}\dot{\beta}}$ are antisymmetric 2-index tensors with $\epsilon^{12} = \epsilon^{\dot{1}\dot{2}} = +1$. Defining

$$J_L^3 \equiv \tfrac{1}{2}M_L^{12}, \qquad J_L^+ \equiv \tfrac{1}{2}M_L^{11}, \qquad J_L^- \equiv \tfrac{1}{2}M_L^{22}, \tag{3.97}$$

puts the algebra into the familiar $SO(3)$ rotation group form

$$\begin{aligned} &[J_L^3, J_L^\pm] = \pm J_L^\pm, &\qquad &[J_L^+, J_L^-] = 2J_L^3, \\ &[J_L^3, X^1] = \tfrac{1}{2}X^1, &\qquad &[J_L^3, X^2] = -\tfrac{1}{2}X^2. \end{aligned} \tag{3.98}$$

The quadratic casimir $J_L^3(J_L^3 + 1) + J_L^- J_L^+ = j_L(j_L + 1)$ measures the spin j_L of a representation. The casimir can be written in terms of the $M_L^{\alpha\beta}$ as

$$j_L(j_L + 1) = \frac{1}{8}M_L^{\alpha\beta}M_L^{\gamma\delta}\epsilon^{\alpha\gamma}\epsilon^{\beta\delta}, \tag{3.99}$$

where summation over repeated indices is implied. An analogous definition exists for the other spin j_R. The non-zero dimensions of the generators are given by

$$[\hat{D}, \hat{P}^{\alpha\dot{\alpha}}] = +\hat{P}^{\alpha\dot{\alpha}}, \qquad [\hat{D}, \hat{K}^{\alpha\dot{\alpha}}] = -\hat{K}^{\alpha\dot{\alpha}}. \tag{3.100}$$

The special conformal generators and their superpartners satisfy

$$[\hat{P}^{\alpha\dot{\alpha}}, \hat{K}^{\beta\dot{\beta}}] = \frac{i}{2}(M_L^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}} + M_R^{\dot{\alpha}\dot{\beta}}\epsilon^{\alpha\beta}) + \hat{D}\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}. \tag{3.101}$$

Summation on repeated $SU(2)$ indices is implied.

Hermitian conjugation properties of the hatted generators follow from their definition. In particular, $\hat{D}$ is Hermitiean, while Hermitien conjugation exchanges $SU(2)_L$ and $SU(2)_R$, and exchanges $\hat{P}$ with $\hat{K}$:

$$\begin{aligned} \hat{D}^\dagger &= \hat{D}, \\ (M_L^{\alpha\beta})^\dagger = +\epsilon^{\alpha\gamma}\epsilon^{\beta\delta}M_L^{\gamma\delta}, &\qquad (M_R^{\dot{\alpha}\dot{\beta}})^\dagger = +\epsilon^{\dot{\alpha}\dot{\gamma}}\epsilon^{\dot{\beta}\dot{\delta}}M_R^{\dot{\gamma}\dot{\delta}}, \\ (\hat{P}^{\alpha\dot{\alpha}})^\dagger = -\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\hat{K}^{\beta\dot{\beta}}, &\qquad (\hat{K}^{\alpha\dot{\alpha}})^\dagger = -\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\hat{P}^{\beta\dot{\beta}}. \end{aligned} \tag{3.102}$$

We can now find the finite-dimensional representations of the conformal algebra by looking at the *primary* or *highest-weight* states which are those annihilated by the $\hat{K}$'s

$$\hat{K}^{\alpha\dot{\alpha}}|j_L, j_R, d\rangle = 0. \tag{3.103}$$

Here we have labelled the highest-weight state by its eigenvalues under the $SU(2)_L \times SU(2)_R$ rotations and $\hat{D}$. The rest of the states in the representation of which $|j_L, j_R, d\rangle$ is the highest-weight state are formed by acting on it with the "lowering" operator $\hat{P}^{\alpha\dot{\alpha}}$:

$$(\prod P)|j_L, j_R, d\rangle, \tag{3.104}$$

and are called *descendant* states.

All the unitary irreducible representations of the conformal algebra can be classified as follows [Mack, Commun. Math. Phys. **55** (1977) 1]:

identity	$j_L = j_R = 0$	$d = 0$
free chiral	$j_L j_R = 0$	$d = j_L + j_R + 1$
chiral	$j_L j_R = 0$	$d > j_L + j_R + 1$
free general	$j_L j_R \neq 0$	$d = j_L + j_R + 2$
general	$j_L j_R \neq 0$	$d > j_L + j_R + 2$

(3.105)

It is not hard to derive these constraints from the conformal algebra. (The hard part is showing that they are sufficient.) States with Lorentz spins j_L, j_R have internal rotational quantum numbers which we denote by associating a "field" ϕ to the state

$$\phi^{\alpha_1 \ldots \alpha_{2j_L} \dot{\alpha}_1 \ldots \dot{\alpha}_{2j_R}}, \leftrightarrow |j_L, j_R, d\rangle \tag{3.106}$$

where the dotted and undotted indices are separately symmetrized. (From now on the various $SU(2)$ indices of a single field will always be understood to be symmetrized.) All descendant states are generated by applying $\hat{P}^{\alpha\dot{\alpha}}$ to ϕ. The dimension of ϕ is $\hat{D}(\phi) = d$.

The field ϕ is just a notational device to make the $SO(4)$ representation structure clear as far as we are concerned; however there is a way of associating local quantum fields to states in conformal field theories. This can be done by Wick rotating our conformal quantum field theory to Euclidean space and performing *radial quantization* in which we choose the Euclidean dilatation generator D_E as our Hamiltonian. This Hamiltonian generates radial scalings instead of time translations, and so radial quantization foliates the Euclidean $\mathbb{R}^4$ by constant radius S^3's centered around a given point x_0. In the Euclidean theory the translation and special conformal generators P_E^a and K_E^a satisfy the same conjugation relations as in (3.102), and in general the classification of highest weight states of the Euclidean algebra is identical to that of our hatted Minkowski generators $\hat{P}^a$, $\hat{K}^a$, *etc.*. This relation between the Euclidean and Minkowski formulations of the conformal group is important because in the Euclidean formulation dilatations can be used to related any state at a given radius to one localized at x_0 (by scaling the radius to zero). This gives a one to one correspondence of states with local operators (fields) at x_0 in Euclidean conformal field theory. Below we will find in some cases that polynomials in the momenta annihilate a state; this translates to a differential constraint on the associated field through the identification of $P_E^a \leftrightarrow \partial^a$. In the cases that concern us these constraints will imply that the field is free. By Wick rotating back to Minkowski space free fields remain free. Thus will be able to intpret polynomial relations involving the Minkowski $\hat{P}^a$ generators (which, recall, are *not* the translation generators, but are some linear combination of translations, special conformal transformations, dilatations, and boosts) as differential relations on local fields.

In the scalar case, j_L=j_R=0,

$$\begin{aligned}
||\hat{P}^{\alpha\dot{\alpha}}\phi||^2 &= ||\hat{P}^{\alpha\dot{\alpha}}|0,0,d\rangle||^2 \\
&= \langle 0,0,d|(-\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\hat{K}^{\beta\dot{\beta}}\hat{P}^{\alpha\dot{\alpha}})|0,0,d\rangle \\
&= \epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\langle 0,0,d|[\hat{P}^{\alpha\dot{\alpha}},\hat{K}^{\beta\dot{\beta}}]|0,0,d\rangle \\
&= \epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\langle 0,0,d|\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\hat{D}|0,0,d\rangle \\
&= 4\langle 0,0,d|\hat{D}|0,0,d\rangle = 4d, \qquad (3.107)
\end{aligned}$$

implying $d > 0$, and a null state when $d = 0$. Here in the second line we have used the conjugation relation (3.102), in the third the primary state condition (3.103), and in the fourth the commutator (3.101) and the fact that a $j_l = j_r = 0$ state is annihilated by $M_L^{\alpha\beta}$ and $M_R^{\dot\alpha\dot\beta}$. At the next level a similar computation shows

$$||\hat{P}^2\phi||^2 = 8d(d-1), \tag{3.108}$$

implying $d \geq 1$ and a null state at $d = 1$. The field associated to the null state at $d = 1$ therefore satisfies the (Euclidean) free massless wave equation $\partial^2\phi = 0$.

In the chiral case where $j_L{\neq}0$ and $j_R{=}0$, a similar calculation gives

$$||\epsilon^{\alpha\beta_1}\hat{P}^{\alpha\dot\alpha}\phi^{\beta_1\dots\beta_{2j_L}}||^2 = 2(d{-}j_L{-}1), \tag{3.109}$$

implying $d \geq j_L{+}1$, and a null state when $d = j_L{+}1$. This null state gives the free massless wave equation since $\hat{P}^{\beta_1}_{\dot\alpha}\hat{P}^{\dot\alpha}_{\alpha}\phi^{\alpha\beta_2\dots\beta_{2j_L}} = -\frac{1}{2}\hat{P}^2\phi^{\beta_1\dots\beta_{2j_L}}$.

Finally, in the general case $j_L j_R \neq 0$,

$$||\epsilon^{\alpha\beta_1}\epsilon^{\dot\alpha\dot\beta_1}\hat{P}^{\alpha\dot\alpha}\phi^{\beta_1\dots\beta_{2j_L}\,\dot\beta_1\dots\dot\beta_{2j_R}}||^2 = d{-}j_L{-}j_R{-}2, \tag{3.110}$$

implying $d \geq j_L{+}j_R{+}2$, with a null state when the inequality is saturated. This reproduces the classification of unitary conformal representations given in (3.105).

An interesting consequence of this classification is that a $U(1)$ gauge conformal field theory in four dimensions is interacting if and only if it has both electrically *and* magnetically charged conformal fields in its spectrum [P. Argyres, R. Plesser, N. Seiberg, and E. Witten, Nucl. Phys. **B461** (1996) 71]. To see this, recall that an Abelian field strength field $F^{\mu\nu}$ is decomposed into the chiral $(j_L, j_R) = (1, 0)$ and $(0, 1)$ representations $F^\pm$. Then, if $\hat{D}(F^\pm) = 2$ it is a free field so $dF^\pm = 0$, which implies the free Maxwell equations and the Bianchi identities $dF = d * F = 0$. On the other hand, if the field strength is interacting, then $\hat{D}(F^\pm) > 2$, implying $J^\pm \equiv dF^\pm \neq 0$. Since F^+ and F^- are independent representations of the conformal algebra, we learn from the equations of motion

$$dF = J^+ - J^- \equiv J_e \neq 0, \quad \text{and} \quad d * F = J^+ + J^- \equiv J_m \neq 0, \tag{3.111}$$

that the electric and magnetic currents J_e and J_m cannot vanich as quantum fields in this theory.

A related point is that all Abelian gauge *charges* will vanish in a fixed point theory (though they may still couple to massive degrees of freedom). In the case of the interacting $U(1)$ field strength F, though we have seen that its conserved electric and magnetic currents do not vanish, there is no charge at infinity associated with them, because of the rapid decay of correlation functions of F due to its anomalous dimension. This is true even if we include massive or background sources, since the long distance behavior of the fields is governed by the conformal field theory. If, on the other hand, F were free, then we have seen that its associated conserved currents, and thus the charges, vanish. Note, however, massive sources can have long range fields in this case since F has its canonical dimension. (We do not reach a contradiction by taking the mass of a charged source to zero since its $U(1)$ couplings flow to zero in the IR.) Non-Abelian gauge charges need not vanish in the conformal field theory since the above arguments only apply to gauge invariant fields or states.

3.4.2 N=1 superconformal algebra and representations

When we extend the conformal algebra by including the supersymmetry generators Q_L^α, $Q_R^{\dot\alpha}$, we are forced by associativity to include three additional generators: the fermionic *superconformal generators* S_L^α and $S_R^{\dot\alpha}$, and a scalar bosonic R generating $U(1)_R$ rotations. This is in contrast to the usual (non-conformal) supersymmetry algebra where inclusion of the R generator was not mandatory.

By a similar change of basis to hatted operators as in the previous subsection, we can write the $N=1$ superconformal algebra generators and their hermiticity relations as (dropping the hats on all generators from now on)

$$\begin{aligned} R^\dagger &= R, \\ (Q_L^\alpha)^\dagger = +\epsilon^{\alpha\beta}S_L^\beta, &\qquad (S_L^\alpha)^\dagger = -\epsilon^{\alpha\beta}Q_L^\beta, \\ (Q_R^{\dot\alpha})^\dagger = +\epsilon^{\dot\alpha\dot\beta}S_R^{\dot\beta}, &\qquad (S_R^{\dot\alpha})^\dagger = -\epsilon^{\dot\alpha\dot\beta}Q_R^{\dot\beta}. \end{aligned} \tag{3.112}$$

The non-zero dimensions of the generators are given by

$$\begin{aligned} [D, Q_L^\alpha] = +\tfrac{1}{2}Q_L^\alpha, &\qquad [D, S_L^\alpha] = -\tfrac{1}{2}S_L^\alpha, \\ [D, Q_R^{\dot\alpha}] = +\tfrac{1}{2}Q_R^{\dot\alpha}, &\qquad [D, S_R^{\dot\alpha}] = -\tfrac{1}{2}S_R^{\dot\alpha}, \end{aligned} \tag{3.113}$$

and likewise for the $U(1)_R$ charges

$$\begin{aligned} [R, Q_L^\alpha] = +Q_L^\alpha, &\qquad [R, S_L^\alpha] = -S_L^\alpha, \\ [R, Q_R^{\dot\alpha}] = -Q_R^{\dot\alpha}, &\qquad [R, S_R^{\dot\alpha}] = +S_R^{\dot\alpha}. \end{aligned} \tag{3.114}$$

The conformal generators and their superpartners satisfy

$$\begin{aligned} [K^{\alpha\dot\alpha}, Q_L^\beta] = iS_R^{\dot\alpha}\epsilon^{\alpha\beta}, &\qquad [P^{\alpha\dot\alpha}, S_L^\beta] = iQ_R^{\dot\alpha}\epsilon^{\alpha\beta}, \\ [K^{\alpha\dot\alpha}, Q_R^{\dot\beta}] = iS_L^{\alpha}\epsilon^{\dot\alpha\dot\beta}, &\qquad [P^{\alpha\dot\alpha}, S_R^{\dot\beta}] = iQ_L^{\alpha}\epsilon^{\dot\alpha\dot\beta}, \end{aligned} \tag{3.115}$$

while the supersymmetry algebra and its conformal extension are given by:

$$\begin{aligned} \{Q_L^\alpha, Q_R^{\dot\alpha}\} = 2P^{\alpha\dot\alpha}, &\qquad \{S_L^\alpha, S_R^{\dot\alpha}\} = 2K^{\alpha\dot\alpha}, \\ \{Q_L^\alpha, S_L^\beta\} &= M_L^{\alpha\beta} - i(D - \frac{3}{2}R)\epsilon^{\alpha\beta}, \\ \{Q_R^{\dot\alpha}, S_R^{\dot\beta}\} &= M_R^{\dot\alpha\dot\beta} - i(D + \frac{3}{2}R)\epsilon^{\dot\alpha\dot\beta}. \end{aligned} \tag{3.116}$$

In radial quantization, there is again a one-to-one map between states and local operators at the origin. Primary states $|j_L, j_R, d, r\rangle$ are in a representation of $SU(2)_L \times SU(2)_R \times SO(2)_D \times U(1)_R$ and are annihilated by K^μ, S_L^α and $S_R^{\dot\alpha}$. Descendants are formed from the primary states by acting on them with the Q_L and Q_R operators (since P can be expressed as an anticommutator of Q_L and Q_R). The classification of unitary irreducible representations is then [Dobrev and Petkova Phys. Lett. **B162** (1985) 127]:

identity	$j_L = j_R = 0$	$d = 0$	$r = 0$
free left-chiral	$j_R = 0$	$d = +\frac{3}{2}r$	$\frac{3}{2}r = j_L + 1$
left-chiral	$j_R = 0$	$d = +\frac{3}{2}r$	$\frac{3}{2}r > j_L + 1$
free right-chiral	$j_L = 0$	$d = -\frac{3}{2}r$	$\frac{3}{2}r = -j_R - 1$
right-chiral	$j_L = 0$	$d = -\frac{3}{2}r$	$\frac{3}{2}r < -j_R - 1$
free general	$j_L j_R \neq 0$	$d = j_L + j_R + 2$	$\frac{3}{2}r = j_L - j_R$
general	$j_L j_R \neq 0$	$d > \lvert\frac{3}{2}r - j_L + j_R\rvert + j_L + j_R + 2$	

(3.117)

Thus, in general, $d \geq |\frac{3}{2}r|$, with equality only for the left-chiral or right-chiral fields. In the above classification, the left-chiral fields are defined as those with $j_R = 0$. It is easy to see from the superconformal algebra that this implies the usual condition for left-chiral superfields: $Q_R^{\dot\alpha}\phi = 0$; similarly for the right-chiral fields.

3.5 N=1 duality

Let us apply this representation theory of the superconformal algebra to the singularity at the origin of the $N_f \geq N_c{+}2$ moduli space. Recall that the global symmetry group and charges of the superQCD theory for $N_f \leq 3N_c$ is

	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_R$
Q	$\mathbf{N_f}$	$\mathbf{1}$	1	$\frac{N_f-N_c}{N_f}$
$\widetilde{Q}$	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	-1	$\frac{N_f-N_c}{N_f}$
M	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$2\frac{N_f-N_c}{N_f}$
B	$\binom{\mathbf{N_f}}{\mathbf{N_c}}$	$\mathbf{1}$	N_c	$N_c\frac{N_f-N_c}{N_f}$
$\widetilde{B}$	$\mathbf{1}$	$\overline{\binom{\mathbf{N_f}}{\mathbf{N_c}}}$	$-N_c$	$N_c\frac{N_f-N_c}{N_f}$

(3.118)

For sufficiently large N_c and N_f close to (but less than) $3N_c$, then we have seen that the fixed point is close to zero coupling. The zero coupling theory is free so is conformally invariant, and the $U(1)_R$ symmetry in the superconformal algebra is just the microscopic $U(1)_R$ shown above. So for the fixed point at small value of the coupling, it is reasonable to assume that the $U(1)_R$ symmetry in its superconformal algebra is the same, since there is not enough “time” for relevant operators at the zero coupling point to flow to irrelevant operators at the fixed point, and so make a new, “accidental”, $U(1)_R$ symmetry in the IR.

Actually, since there is also the $U(1)_B$ symmetry, the $U(1)_R$ symmetry appearing in the superconformal algebra at the fixed point could be a combination of the $U(1)_R$ and $U(1)_B$ defined above. Notice, however, that this will not affect the R-charge of the meson field, since its baryon number vanishes. We thus read off the scaling dimension of M:

$$D(M) = \frac{3}{2}R(M) = 3\frac{N_f - N_c}{N_f}. \tag{3.119}$$

This implies that for $\frac{3}{3}N_c < N_f < 3N_c$, $1 < D(M) < 2$ and so M is an interacting conformal field.

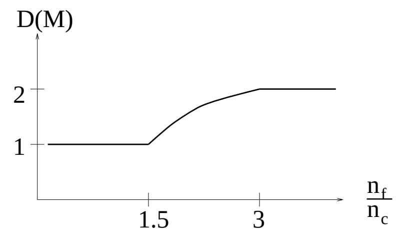

Figure 3.3: dimension of M plot

For $N_f > 3N_c$ this formula implies $D(M) > 2$; however, we know that in this range the IR theory is free, so the quark left-chiral superfields have their canonical dimension of 1, and thus the meson left-chiral superfield must have dimension 2. The reason the above formula fails in this case is that the IR free theory (being free) has an unbroken $U(1)_A$ in the IR which can mix with the $U(1)_R$ defined above.

The relation (3.119) also implies that $D(M) \leq 1$ for $N_f \leq \frac{3}{2}N_c$. Since dimensions less than 1 are not allowed by unitarity, it must be that a new accidental R symmetry arises in this range. It is suggestive that right at $N_f = \frac{3}{2}N_c$, $D(M) = 1$, implying that M is free. This led [N. Seiberg, hep-th/9411149] to guess that $D(M) = 1$ for $N_f \leq \frac{3}{2}N_c$, and so should be treated as an elementary field in an IR free theory in this range.

Since the global symmetry must be the same as in the microscopic theory, one wants this IR free theory to have N_f fundamental flavors in an $SU(\widetilde{N_c})$ gauge theory. In order to be IR free we need $\widetilde{N_c} < \frac{1}{3}N_f$ when $N_f \leq \frac{3}{2}N_c$. A simple choice that works is

$$\widetilde{N_c} \equiv N_f - N_c. \tag{3.120}$$

We will refer to this theory as the “dual theory”, while we will call the original $SU(N_c)$ theory the “direct theory”. (It is sometimes also referred to as the “magnetic theory”, while the direct theory is called the "electric theory"; the reasons for these names will only become clear a few lectures from now.)

We assign the quantum numbers to the fundamental fields in the dual theory as

	$SU(\widetilde{N_c})$	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_R$
M	$\mathbf{1}$	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$2\frac{N_f-N_c}{N_f}$
q	$\widetilde{\mathbf{N_c}}$	$\overline{\mathbf{N_f}}$	$\mathbf{1}$	$\frac{N_c}{N_f-N_c}$	$\frac{N_c}{N_f}$
$\widetilde{q}$	$\overline{\widetilde{\mathbf{N_c}}}$	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	$\frac{-N_c}{N_f-N_c}$	$\frac{N_c}{N_f}$

(3.121)

Here $SU(\widetilde{N_c})$ column are the gauge charges, while the rest are the (non-anomalous) global symmetries. The R-charges of the dual quark left-chiral superfields are fixed by anomaly cancellation. The normalization of their baryon number is chosen so that the dual baryons, $b \equiv q^{\widetilde{N_c}}$ and $\widetilde{b} \equiv \widetilde{q}^{\widetilde{N_c}}$, will have the same baryon number as the direct baryon fields B and $\widetilde{B}$. Indeed, with these assignments, we find the global charges of the gauge-invariant composite left-chiral superfields in the dual model to be

	$SU(N_f)$	$SU(N_f)$	$U(1)_B$	$U(1)_R$
$\widetilde{m}$	$\overline{\mathbf{N_f}}$	$\mathbf{N_f}$	0	$2\frac{N_c}{N_f}$
b	$\overline{\binom{\mathbf{N_f}}{\mathbf{N_f-N_c}}}$	$\mathbf{1}$	N_c	$(N_f-N_c)\frac{N_c}{N_f}$
$\widetilde{b}$	$\mathbf{1}$	$\binom{\mathbf{N_f}}{\mathbf{N_f-N_c}}$	$-N_c$	$(N_f-N_c)\frac{N_c}{N_f}$

(3.122)

where we have defined the dual meson to be $\widetilde{m} \equiv q\widetilde{q}$. Comparing with the global charges of the baryons in the direct theory, we see that they are the same, since as flavor representations $\overline{\binom{N_f}{N_c}} = \binom{N_f}{N_f-N_c}$.

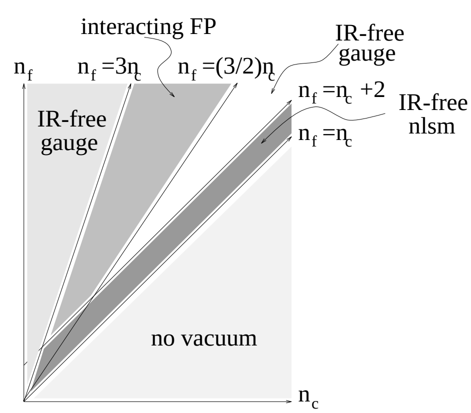

Figure 3.4: complete phase diagram

This educated guess for an alternative (IR equivalent) description of the vacuum physics at the origin of moduli space implies the following "phase diagram" answering the question posed in the last lecture (compare the phase diagram of that lecture). The nature of this proposed solution is quite surprising: the AF direct gauge theory, at least for some range of N_f, is IR-equivalent to an IR-free gauge theory! This naturally raises the question of what is the relation between the IR-free gauge bosons and the direct (microscopic) gauge fields? No precise answer to this question is known.

3.5.1 Checks

Is there any way of checking this proposal?

The first thing to note is that the global symmetries of the direct and dual theories are the same. One can check that the 't Hooft anomaly-matching conditions all work.

The next thing to check is whether these two theories have the same moduli space of vacua: do they have the same light gauge-singlet left-chiral superfields? In the direct theory away from the origin, we have M, B, and $\widetilde{B}$. In the dual theory, the elementary M, and the composite b and $\widetilde{b}$ fields have the same symmetry properties, and so can plausibly be identified. However, the dual theory also has the composite dual meson $\widetilde{m}$. To remove this operator from the dual theory, we must add some superpotential interaction. There is only one term allowed by the symmetries:

$$f_{dual} = \Lambda M q \widetilde{q}, \tag{3.123}$$

where Λ is a dimensionless coupling. Such a coupling is just what we need to remove $\widetilde{m}$ as an independent degree of freedom in the IR, since the F-term equation for M implies that $q\widetilde{q} \equiv \widetilde{m} = 0$. Thus, for non-zero λ we at least have the right counting of light degrees of freedom away from the origin of moduli space.

The superpotential in the dual theory raises a new question, however: what is the correct value of λ? Actually, this is the wrong question, since the superpotential coupling is not exactly marginal.

For example, at the fixed point (the vacuum at the origin of moduli space) when $N_c{+}2 < N_f < \frac{3}{2}N_c$, the dual theory is IR-free, so the gauge-coupling, g_{dual} flows to zero. In a free theory, a Yukawa coupling like (3.123) is irrelevant, so λ also flows to zero. Thus the origin of the λ–g_{dual} plane is the fixed point. In the regime when $\frac{3}{2}N_c < N_f < 3N_c$, the λ and g_{dual} couplings are still irrelevant for large couplings, but $g_{dual} = 0$ is an UV fixed point, since there is supposed to be an IR fixed point at $g_{dual} = g_* > 0$ when $\lambda = 0$. (Recall the form of the 2-loop beta function found at the beginning of last lecture.) However, at this IR fixed point $D(M) = 1$ since it is free (it has no couplings), and $D(q) = D(\widetilde{q}) = 3N_c/(2N_f)$ from their R-charges, implying $D(Mq\widetilde{q}) < 3$, and so is a relevant operator. Thus the superpotential will cause the theory to flow to a fixed point at non-zero $\lambda = \lambda_*$. These RG flows can be illustrated as: Thus we expect the superpotential term to be irrelevant everywhere in the vicinity of the fixed point, except at the fixed point itself. (This situation is often described by saying that the operator in the superpotential is marginal but not exactly marginal.)

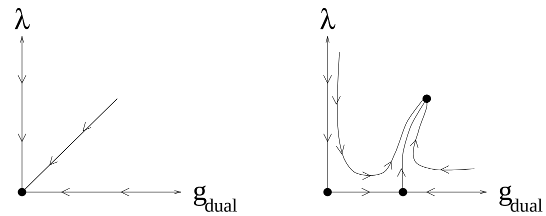

Figure 3.5: RGflow of yukawa-dual

We can therefore trade λ for a scale in the dual theory, and so (as long as it is not zero) its value can have no effect on the scale-invariant far-IR physics. In the case where both the direct theory and the dual theory are AF, each has a gauge strong-coupling scale, Λ and Λ_{dual} respectively. However, the dual theory also has a second scale, which we can define as $\mu \sim \lambda\Lambda_{dual}$. The statement that these two theories are "dual" just means that they flow to the same theory at mass scales well below the smallest of Λ, Λ_{dual} and μ. We can trade λ for μ in the superpotential by noting that in the microscopic theory, M is a composite operator of canonical dimension 2 (in the UV), while in the dual theory it is a fundamental field of dimension 1 (in the UV). Then, if we define a new meson field by

$$M = M_{direct} \equiv \mu M_{dual}, \tag{3.124}$$

the dual superpotential becomes

$$f_{dual} = \frac{1}{\mu} M q\widetilde{q}. \tag{3.125}$$

By the symmetries and holomorphy, the relation between the direct and dual strong-coupling scales must be

$$\Lambda_{direct}^{3N_c-N_f}\Lambda_{dual}^{3(N_f-N_c)-N_f} = (-)^{N_f-N_c}\mu^{N_f}. \tag{3.126}$$

The factor of $(-)^{N_f-N_c}$ can be determined by considering the dual of the dual theory. In this case we expect to regain the original theory with gauge group $SU(N_c)$ and quarks Q (since $N_f - \widetilde{N_c} = N_c$):

$$\begin{array}{ccccc} & \text{dual} & & \text{dual} & \\ Q & \longrightarrow & q\,,\, M & \longrightarrow & Q, N, M \\ & & & & \\ f=0 & & f=\frac{1}{\mu}Mq\widetilde{q} & & f=\frac{1}{\mu}MN+\frac{1}{\widetilde{\mu}}NQ\widetilde{Q} \end{array} \tag{3.127}$$

where in the original theory M is the composite meson $M = Q\widetilde{Q}$, and similarly in the first dual $N = q\widetilde{q}$. From the superpotential of the (dual)^2 theory, we see that the (now fundamental) N mesons are massive and can be integrated-out, giving the required $M = Q\widetilde{Q}$ only if $\widetilde{\mu} = -\mu$. Then (3.126) implies that $\Lambda_{(dual)^2} = \Lambda_{direct}$ with the factor of $(-)^{N_f - N_c}$.

3.5.2 Matching flat directions

We will now analyze the moduli space of deformations of the two theories and show they are the same. We will do somewhat less than this, mainly because (as mentioned in lecture 20) we do not have a convenient description of this moduli space for general N_f and N_c. So we will outline what happens when we turn on vacuum expectation values for the meson field in the two theories. The equivalence of the baryonic directions in moduli space are, as far as I know, less well understood.

Recall that the moduli space of the direct theory is the same as its classical moduli space, and that in the classical moduli space there are flat directions with arbitrary meson vacuum expectation values with rank$(M) < N_c$; see eq. (20.10). Suppose we turn on a vacuum expectation value with rank$(M) = 1$:

$$\langle M \rangle = \begin{pmatrix} a^2 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}, \tag{3.128}$$

corresponding to giving only one component of the squarks a vacuum expectation value. The effect of this on the direct theory for large a is to Higgs the theory from $SU(N_c)$ with N_f flavors down to $SU(N_c{-}1)$ with $N_f{-}1$ flavors. On the other hand, turning on this vacuum expectation value in the dual theory gives rise through the dual superpotential,

$$f_{dual} = \frac{1}{\mu}\mathrm{Tr}Mq\widetilde{q} = \frac{a^2}{\mu}q_1\dot{\gamma}ot\widetilde{q}^1, \tag{3.129}$$

to a mass term for the q_1 and $\widetilde{q}^1$ dual quarks (here the 1 is a flavor index). Again for large a, integrating-out the massive quarks takes the dual theory from $SU(N_f{-}N_c)$ with

N_f flavors to the $SU(N_f - N_c)$ theory with $N_f - 1$ flavors. The equivalence of the direct and dual theories implies that the resulting theories after Higgsing or integrating-out should again be realted by our dual map, which indeed they are:

$$\begin{array}{ccc} \underline{\text{direct}} & & \underline{\text{dual}} \\ & & \\ SU(N_c),\, N_f & \longleftrightarrow & SU(N_f - N_c),\, N_f \\ \downarrow & & \downarrow \\ \text{Higgs} & & \text{mass} \\ \downarrow & & \downarrow \\ SU(N_c - 1),\, N_f - 1 & \longleftrightarrow & SU(N_f - N_c),\, N_f - 1 \end{array} \tag{3.130}$$

Doing the more general case of higher-rank M is equivalent to simply repreating this procedure. Nothing new happens until we take $\text{rank}(M) = N_c$, in which case it can be shown that the resulting direct and dual theories coincide, giving identical non-singular moduli spaces of meson and baryon vacuum expectation values.

Alternatively to turning on vacuum expectation values in the direct theory, we can turn on masses to the fundamental quarks. The corresponding deformation of the dual theory should again give rise to an equivalent theory. Suppose we turn on a mass for just the Q_1 and $\widetilde{Q}^1$ quarks:

$$f_{direct} = m Q_1 \cdot \widetilde{Q}^1. \tag{3.131}$$

For large m, integrating-out this quark then takes the $SU(N_c)$ theory with N_f flavors to an $SU(N_c)$ theory with $N_f - 1$ flavors. In the dual theory, on the other hand, turning on this mass corresponds to the superpotential

$$f_{dual} = \frac{1}{\mu} \text{Tr}(M q \widetilde{q}) + m M_1^1, \tag{3.132}$$

which, upon integrating-out the M_1^1 component (by its F-term equation), gives rise to

$$\langle q_1 \cdot \widetilde{q}^1 \rangle = -m\mu. \tag{3.133}$$

For large m this is just Higgses the dual theory from $SU(N_c - N_f)$ with N_f flavors down to $SU(N_c - N_f - 1)$ with $N_f - 1$ flavors. This is again dual to the corresponding direct theory:

$$\begin{array}{ccc} \underline{\text{direct}} & & \underline{\text{dual}} \\ & & \\ SU(N_c),\, N_f & \longleftrightarrow & SU(N_f - N_c),\, N_f \\ \downarrow & & \downarrow \\ \text{mass} & & \text{Higgs} \\ \downarrow & & \downarrow \\ SU(N_c),\, N_f - 1 & \longleftrightarrow & SU(N_f - N_c - 1),\, N_f - 1 \end{array} \tag{3.134}$$

Again, one can extend this to turning on mass matrices of arbtrary rank by repeating this procedure. This procedure ends with turning on a mass matrix with $\text{rank}(m) = N_f - N_c$, where again the resulting direct and dual theories can be shown to be the same.

So far we have presented strong evidence for Seiberg's duality conjecture for $SU(N_c)$ superQCD. This conjecture posits the IR equivalence of two quite different-looking gauge theories. This is equivalent to saying that the two theories are in the same *universality class*. Using similar arguments there has been developed a fairly rich "phenomenology" of dual sets of theories for other gauge groups and matter content. Perhaps especially interesting among these dualities are chiral/non-chiral dual pairs. No simple constructive rules for predicting other dual pairs has been given. Also, the question of IR equivalences among theories with product gauge groups has not been systematically explored.

Finally, the question of what general lessons can be derived from the existence and systematics of these gauge universality classes has not been answered. There are many suggestions that these dualities are related to a different kind of duality among quantum field theories called *S-duality*. S-duality is the *exact* quantum equivalence of theories with an exactly marginal operator at different values of the coefficient of this operator. (In a few lectures we will discuss the simplest example of such an S-duality: electric-magnetic duality in Abelian gauge theories.) However, there is as yet no clear statement of the relation between N=1 and S-dualities.

Chapter 4

Extended Supersymmetry

In this chapter we will explore gauge theories with extended supersymmetry. With extended supersymmetry there are scalar fields in the same supermultiplet as the gauge bosons. They thus transform under the adjoint representation of the gauge group. The part of moduli space where only these adjoint scalars get vacuum expectation values is called the *Coulomb branch* of the moduli space. This is because a vacuum expectation value for an adjoint scalar can at most Higgs the gauge group to $U(1)$ gauge factors, implying that the Coulomb branch vacua have long distance electromagnetic-like forces (Free photons). In what follows we will discuss the physics peculiar to Coulomb branches first (without refence to extended supersymmetry), and only later will we develop the algebra of extended supersymmetry representations, construct the classical actions, and derive nonrenormalization theorems.

In detail, the first two section deal with two (nonsupersymmetric) aspects of physics in vacua with unbroken $U(1)$ gauge groups: magnetic monopoles and electric-magnetic duality. Section 4.3 will then analyze the simplest example of a Coulomb branch, which occurs in the $SU(2)$ superYM theory with and adjoint left-chiral superfield Φ. This example with a special superpotential interaction is actually N=2 supersymmetric, and was first solved in [42]. The N=1 treatment which we give follows the discussion of [43]. The later sections then analyze extended supersymmetric theories along the lines of the development we gave in chapters 1 and 2 for N=1 theories.

4.1 Monopoles

The first ingredient we need to be aware of is monopoles [64].

First, following Dirac, we ask whether it is possible to add megnetic charges without disturbing the quantum consistency of the electromagnetic coupling. A static magnetic field

$$\vec{B} = \frac{Q_m \hat{r}}{4\pi r^2} \tag{4.1}$$

corresponds to a magnetic charge $\int_{S^2_\infty} \vec{B} \cdot d\vec{S} = Q_m$ at $r = 0$. To couple a charged particle to a background field quantumly we need the vector potential A_μ. There is no solution for A_μ which is regular away from $r = 0$; however we can write the solution as one which is regular in two "patches". Divide a two-sphere S^2 of fixed radius r into a northern half N with $0 \leq \theta \leq \pi/2$, a southern half S with $\pi/2 \leq \theta \leq \pi$ and the overlap region which is the equator at $\theta = \pi/2$. The vector potential on the two halves is then taken to be

$$\vec{A}_N = \frac{Q_m}{4\pi r} \frac{(1 - \cos\theta)}{\sin\theta} \hat{e}_\phi, \qquad \vec{A}_S = -\frac{Q_m}{4\pi r} \frac{(1 + \cos\theta)}{\sin\theta} \hat{e}_\phi. \tag{4.2}$$

Note that on the two halves of the two-sphere the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$ agrees with (4.1). This construction only makes sense if the difference between A_N and A_S on the overlap region is a gauge transformation. At $\theta = \pi/2$

$$\vec{A}_N - \vec{A}_S = \vec{\nabla}\chi, \qquad \chi = \frac{Q_m}{2\pi}\phi, \tag{4.3}$$

so that the difference is a gauge transformation; however, the gauge function χ is not continuous. In quantum mechanics, a gauge transformation acts on wave functions carrying (electric) charge Q_e as $\psi \to e^{-iQ_e\chi}\psi$ so physical quantities will be continuous as long as $e^{-iQ_e\chi}$ is continuous. This then gives us the condition $e^{-iQ_eQ_m} = 1$ or

$$Q_e Q_m = 2\pi n, \qquad n \in \mathbb{Z} \tag{4.4}$$

which is the famous Dirac quantization condition.

Monopoles can be constructed as finite-energy classical solutions of non-Abelian gauge theories spontaneously broken down to Abelian factors [65]. In particular they will occur in the $N = 2$ $SU(2)$ Yang-Mills theory. We illustrate this for simplicity in a (non-supersymmetric) $SU(2)$ theory broken down to $U(1)$ by a real adjoint Higgs:

$$\mathcal{L} = -\frac{1}{4g^2} F^a_{\mu\nu} F^{a\mu\nu} + \frac{1}{2} D^\mu \Phi^a D_\mu \Phi^a - V(\Phi) \tag{4.5}$$

where V has a minimum on the sphere in field space $\sum_a \Phi^a \Phi^a = v^2$. Different directions on this sphere are gauge-equivalent. In the vacuum $\langle \Phi^a \rangle$ lies on this sphere, Higgsing $SU(2) \to U(1)$ and giving a mass $m_W = gv$ to the $W^\pm$ gauge bosons. The unbroken $U(1)$ has coupling g, so satisfies Gauss's law $\vec{D} \cdot \vec{E} = g^2 j^0_e$, where j^μ_e is the electric current density. Thus the electric charge is computed as $Q_e = \frac{1}{g^2} \int_{S^2_\infty} \vec{E} \cdot d\vec{S}$. In the vacuum, the unbroken $U(1)$ is picked out by the direction of the Higgs vev, so $\vec{E} = \frac{1}{v}\Phi^a\vec{E}^a$. With this normalization of the electic charge, we find that the $W^\pm$ bosons have $Q_e = \pm 1$.

Static, finite-energy configurations must approach the vacuum at spatial infinity. Thus for a finite energy configuration the Higgs field Φ^a, evaluated as $r \to \infty$, provides a map from the S^2 at spatial infinity into the S^2 of the Higgs vacuum. Such maps are characterized by an integer, n_m, which measures the winding of one S^2 around the other. Mathematically, the second homotopy group of S^2 is the integers, $\pi_2(S^2) = \mathbb{Z}$. The winding, n_m, is the magnetic charge of the field configuration. To see this, the total energy from the Higgs field configuration:

$$\text{Energy} = \int d^3x \frac{1}{2} D_\mu\Phi^a D^\mu\Phi^a + V(\Phi) \geq \int d^3x \frac{1}{2} D_\mu\Phi^a D^\mu\Phi^a. \tag{4.6}$$

To have finite energy configurations we must therefore ensure that the covariant derivative of Φ^a falls off faster than $1/r$ at infinity. The general solution for the gauge field consistent with this behavior is

$$A^a_\mu \sim -\frac{1}{v^2}\epsilon^{abc}\Phi^b\partial_\mu\Phi^c + \frac{1}{v}\Phi^a A_\mu \tag{4.7}$$

with A_μ arbitrary. The leading-order behavior of the field strength is then

$$F^{a\mu\nu} = \frac{1}{v}\Phi^a F^{\mu\nu} \tag{4.8}$$

with

$$F^{\mu\nu} = -\frac{1}{v^3}\epsilon^{abc}\Phi^a\partial^\mu\Phi^b\partial^\nu\Phi^c + \partial^\mu A^\nu - \partial^\nu A^\mu \tag{4.9}$$

and the equations of motion imply $\partial_\mu F^{\mu\nu} = \partial_\mu * F^{\mu\nu} = 0$. Thus we learn that outside the core of the monopole the non-Abelian gauge field is purely in the direction of Φ^a, that is the direction of the unbroken $U(1)$. The magnetic charge of this field configuration is then computed to be

$$Q_m = \int_{S^2_\infty} \vec{B}\cdot d\vec{S} = \frac{1}{2v^3}\int_{S^2_\infty} \epsilon^{ijk}\epsilon^{abc}\Phi^a\partial^j\Phi^b\partial^k\Phi^c dS^i = 4\pi n_m \tag{4.10}$$

where n_m is the winding number of the Higgs field configuration, recovering the Dirac quantization condition.[1]

[1]The reason this is the Dirac quantization condition (4.4) only for even values of n is that in this theory we could add fields in the fundamental **2** representation of $SU(2)$, which would carry electric charge $Q_e = \pm 1/2$.

Note that for such non-singular field configurations, the electric and magnetic charges can be rewritten as

$$\begin{aligned} Q_e &= \frac{1}{g^2}\int_{S^2_\infty} \vec{E}\cdot d\vec{S} = \frac{1}{g^2 v}\int_{S^2_\infty} \Phi^a \vec{E}^a \cdot d\vec{S} = \frac{1}{g^2 v}\int d^3x \vec{E}^a \cdot (\vec{D}\Phi)^a \\ Q_m &= \int_{S^2_\infty} \vec{B}\cdot d\vec{S} = \frac{1}{v}\int_{S^2_\infty} \Phi^a \vec{B}^a \cdot d\vec{S} = \frac{1}{v}\int d^3x \vec{B}^a \cdot (\vec{D}\Phi)^a \end{aligned} \tag{4.11}$$

using the vacuum equation of motion and the Bianchi identity $\vec{D}\cdot\vec{E}^a = \vec{D}\cdot\vec{B}^a = 0$ and integration by parts.

If we consider a static configuration with vanishing electric field the energy (mass) of the configuration is given by

$$\begin{aligned} m_M &= \int d^3x \left(\frac{1}{2g^2}\vec{B}^a\cdot\vec{B}^a + \frac{1}{2}\vec{D}\Phi^a\cdot\vec{D}\Phi^a + V(\Phi)\right) \geq \int d^3x \left(\frac{1}{2g^2}\vec{B}^a\cdot\vec{B}^a + \frac{1}{2}\vec{D}\Phi^a\cdot\vec{D}\Phi^a\right) \\ &= \frac{1}{2}\int d^3x \left(\frac{1}{g}\vec{B}^a - \vec{D}\Phi^a\right)^2 + \frac{vQ_m}{g}, \end{aligned} \tag{4.12}$$

giving the BPS bound

$$m_M \geq \left|\frac{vQ_m}{g}\right|. \tag{4.13}$$

This semi-classical bound can be extended to *dyons* (solitonic states carrying both electric and magnetic charges):

$$m_D \geq gv\left|Q_e + i\frac{Q_m}{g^2}\right|. \tag{4.14}$$

A theta angle has a non-trivial effect in the presence of magnetic monopoles: it shifts the allowed values of electric charge in the monopole sector of the theory [66]. To see this, consider gauge transformations, constant at infinity, which are rotations in the $U(1)$ subgroup of $SU(2)$ picked out by the Higgs vev, that is, rotations in $SU(2)$ about the axis $\hat{\Phi}^a = \Phi^a/|\Phi^a|$. The action of such an infinitesimal gauge transformation on the field is

$$\delta A^a_\mu = \frac{1}{v}(D_\mu\Phi)^a \tag{4.15}$$

with Φ the background monopole Higgs field. Let $\mathcal{N}$ denote the generator of this gauge transformation. Then if we rotate by 2π about the $\hat{\Phi}$ axis we must get the identity

$$e^{2\pi i\mathcal{N}} = 1. \tag{4.16}$$

Including the θ term, it is straightforward to compute $\mathcal{N}$ using the Noether method,

$$\mathcal{N} = \frac{\partial \mathcal{L}}{\partial \partial_0 A^a_\mu} \delta A^a_\mu = Q_e - \frac{\theta Q_m}{8\pi^2}, \tag{4.17}$$

where we have used the definitions (4.11) of the electric and magnetic charge operators. This result implies

$$Q_e = n_e + n_m \frac{\theta}{2\pi} \tag{4.18}$$

where n_e is an arbitrary integer and $n_m = Q_m/4\pi$ determines the magnetic charge of the monopole. We will henceforth label dyons by the integers (n_e, n_m). Note that the BPS bound becomes

$$M_D \geq gv \left| \left(n_e + n_m \frac{\theta}{2\pi} \right) + i n_m \frac{4\pi}{g^2} \right| = gv |n_e + \tau n_m|. \tag{4.19}$$

This result is classical; quantum mechanically, the coupling τ runs, and gv and $g\tau$ will be replaced by functions of the strong coupling scale Λ and the vevs. In theories with extended supersymmetry the (quantum-corrected) BPS bound can be computed exactly, and states saturating the bound can be identified [57]. For example, in the $N = 2$ $SU(2)$ theory the BPS mass formula becomes [42]

$$M_D = |a(U) n_e + b(U) n_m|, \tag{4.20}$$

where a and b are holomorphic functions of U and Λ^4 satisfying

$$\frac{\partial b(U)}{\partial a(U)} = \tau(U), \tag{4.21}$$

with $a(U)$ a function we will determine in section 4.3 below.[2]

4.2 Electric-magnetic duality

Maxwell's vacuum equations are invariant under the substitution

$$\vec{E} \to \vec{B}, \qquad \vec{B} \to -\vec{E}. \tag{4.22}$$

Covariantly this substitution takes the form $f_{\mu\nu} \to *f_{\mu\nu}$, where recall that the Hodge dual $*f_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} f^{\rho\sigma}$. This substitution invariance of the free Maxwell equations is broken by the presence of electric source terms

$$\partial_\mu f^{\mu\nu} = j^\nu_e, \qquad \partial_\mu * f^{\mu\nu} = 0, \tag{4.23}$$

[2] $b(U)$ is often called $a_D(U)$ in the literature.

For this reason it is of no practical interest in everyday applications of electromagnetism. However, if we include magnetic source terms so that

$$\partial_\mu * F^{\mu\nu} = j^\nu_m, \tag{4.24}$$

with j^ν_m the magnetic four-current, we make Maxwell's equations symmetric under the substitution (4.22) and simultaneous interchange of electric and magnetic currents. This substitution invariance of Maxwell's equations is called *electric-magnetic duality*. As it involves an action on the couplings (exchanging electric and magnetic sources) it is not a symmetry of electromagnetism. Rather, it is simply an ambiguity in the description of low energy electromagnetism: what you call electric versus magnetic is a matter of choice.

Let us generalize this to many $U(1)$ gauge factors. It is convenient to discuss the $U(1)$ gauge fields in the language of forms. Thus we define the one-form potentials and their 2-form field strengths by

$$\begin{aligned} V^I &= V^I_\mu dx^\mu \\ f^I &= dV^I = \frac{1}{2} f^I_{\mu\nu} dx^\mu \wedge dx^\nu, \end{aligned} \tag{4.25}$$

and the Hodge dual of a p-form $C = C_{\mu_1 \cdots \mu_p} dx^{\mu_1} \cdots dx^{\mu_p}$ to be the $(4-p)$-form

$$*C \equiv \frac{1}{p!} \epsilon_{\mu_1 \cdots \mu_4} C^{\mu_1 \cdots \mu_p} dx^{\mu_{p+1}} \wedge \cdots \wedge dx^{\mu_4}, \tag{4.26}$$

so that $**C = (-)^{p+1} C$. It is also convenient to introduce a kind of "self dual" field strength defined by

$$\mathcal{F}^I = f^I - i * f^I \tag{4.27}$$

so that the $U(1)^n$ Maxwell action becomes

$$S = -\frac{1}{16\pi} \int \mathrm{Im} \left[\tau_{IJ} \mathcal{F} \wedge * \mathcal{F} \right], \tag{4.28}$$

where τ_{IJ} is a complex matrix of couplings,

$$\tau = \frac{\vartheta_{IJ}}{2\pi} + i \frac{4\pi}{(g^2)^{IJ}}. \tag{4.29}$$

The classical Maxwell's equations with electric and magnetic sources follow from the action

$$S = \int \left(-\frac{1}{2e^2} f \wedge * f + V \wedge * j_e + \widetilde{V} \wedge * j_m \right), \tag{4.30}$$

where, away from any electric sources $\widetilde{V}$ is defined through $*f = d\widetilde{V}$. The Dirac quantization condition [39] implies that if there are electric sources of unit strength, so that a stationary point source at the origin would have $*j_e = \delta^{(3)}(\mathbf{x})dx^1 \wedge dx^2 \wedge dx^3$, then the strength g^2 of a magnetic source (*i.e.* $*j_m = g^2\delta^{(3)}dx^1 \wedge dx^2 \wedge dx^3$) obeys $g^2 = 4\pi n_m/e^2$ for n_m an integer. With these normalizations, we call the (integer) strength of the electric source, n_e, the electric charge, and n_m the magnetic charge. The equations of motion following from Eq. 4.30 are

$$\frac{1}{e}d{*f} = en_e\delta^{(3)}, \qquad \frac{1}{e}df = \frac{4\pi}{e}n_m\delta^{(3)}, \tag{4.31}$$

which are invariant under the electric-magnetic duality transformation

$$\begin{aligned} (f/e) \to *(f/e), &\qquad *(f/e) \to -(f/e), \\ n_m \to n_e, &\qquad n_e \to -n_m, \\ e \;\leftrightarrow &\;\; 4\pi/e. \end{aligned} \tag{4.32}$$

The minus signs are because $**f = -f$ in Minkowski space.

We can show that this duality of the classical equations of motion holds quantum mechanically as well, though this should be obvious since we are just talking about a free theory. We will also take this opportunity to generalize the above discussion to n $U(1)$ factors and include the theta angles. We compute physical quantities in the quantum theory as a path integral over all gauge potential configurations $\int \mathcal{D}V^I e^{iS}$. This can be rewritten as a path integral over field strength configurations as long as we insert the Bianchi identity as a constraint: $\int \mathcal{D}f^I \mathcal{D}\widetilde{V}_J e^{iS'}$, where $4\pi S' = 4\pi S + \int \widetilde{V}_I \wedge df^I$. Here $\widetilde{V}_I$ is a (one-form) Lagrange multiplier enforcing the Bianchi identity, and whose normalization is chosen so that it couples to monopoles with strength one. Performing the Gaussian functional integral over f^I using $\int \widetilde{V}_I \wedge df^I = \int \widetilde{f}_I \wedge f^I = \frac{1}{2}\int \mathrm{Im}[\widetilde{\mathcal{F}}_I \wedge *\mathcal{F}^I]$ where $\widetilde{\mathcal{F}}_I$ is related to $\widetilde{f}_I = d\widetilde{V}_I$ as in (4.27), we find an equivalent action, $\widetilde{S}$, for $\widetilde{V}_I$:

$$\widetilde{S} = -\frac{1}{16\pi}\int \mathrm{Im}\left[(-\tau^{IJ})\widetilde{\mathcal{F}}_I \wedge *\widetilde{\mathcal{F}}_J\right], \tag{4.33}$$

where τ^{IJ} is the matrix inverse of τ_{IJ}:

$$\tau^{IJ}\tau_{JK} = \delta^I_K. \tag{4.34}$$

Thus the free $U(1)$ gauge theory with couplings τ_{IJ} is quantum mechanically equivalent to another such theory with couplings $-\tau^{IJ}$. This is the electric-magnetic duality "symmetry". It is not really a symmetry since it acts on the couplings—it is an equivalence between two descriptions of the physics.

The electric-magnetic duality transformation

$$S:\ \tau_{IJ} \rightarrow -\tau^{IJ}, \tag{4.35}$$

together with the invariance of the physics under 2π shifts of the theta angles (integer shifts of $\mathrm{Re}\tau_{IJ}$)

$$T^{(KL)}:\ \tau_{IJ} \rightarrow \tau_{IJ} + \delta_I^K \delta_J^L + \delta_I^L \delta_J^K, \tag{4.36}$$

generate a discrete group of duality transformations:

$$\tau_{IJ} \rightarrow (A_I{}^L \tau_{LM} + B_{IM})(C^{JN}\tau_{NM} + D^J{}_M)^{-1}, \tag{4.37}$$

where

$$M \equiv \begin{pmatrix} A_I{}^K & B_{IL} \\ C^{JK} & D^J{}_L \end{pmatrix} \in Sp(2n, \mathbb{Z}). \tag{4.38}$$

The conditions on the $n \times n$ integer matrices A, B, C, and D for M to be in $Sp(2n, \mathbb{Z})$ are (in an obvious matrix notation)

$$\begin{aligned} AB^T = B^T A, &\qquad B^T D = D^T B, \\ A^T C = C^T A, &\qquad D^T C = CD^T, \\ A^T D - C^T B &= AD^T - BC^T = 1, \end{aligned} \tag{4.39}$$

and imply that

$$M^{-1} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix}. \tag{4.40}$$

We have seen that under an electric-magnetic duality transformation, a massive (classical) dyonic source with magnetic and electric charges $(n_m^I, n_{e,J})$ in the original description couples to the dual $U(1)$'s with charges $(n_{e,I}, -n_m^J)$. The effect of a $T^{(KL)}$ theta angle rotation on the charges is $(n_m^I, n_{e,J}) \rightarrow (n_m^I, n_{e,J} - n_m^K \delta_J^L - n_m^L \delta_J^K)$, as follows from the generalization of the Witten effect [40] to n $U(1)$ factors. Together these generate the action

$$(n_m\ n_e) \rightarrow (n_m\ n_e) \cdot M^{-1} \tag{4.41}$$

of the $Sp(2n, \mathbb{Z})$ electric-magnetic duality group on the $2n$-component row vector of magnetic and electric charges.

In the case of a single $U(1)$ factor the coupling matrix becomes a single complex number

$$\tau = \frac{\vartheta}{2\pi} + i\frac{4\pi}{g^2}, \tag{4.42}$$

and the electric-magnetic duality transformations simplify to the group $SL(2, \mathbb{Z}) = Sp(2, \mathbb{Z})$ of duality transformations:

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d}, \quad \begin{pmatrix} n_e \\ n_m \end{pmatrix} \rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} n_e \\ n_m \end{pmatrix}, \qquad a, b, c, d \in \mathbb{Z}, \quad ad - bc = 1, \tag{4.43}$$

which is generated by the electric-magnetic duality transformation (S) and the 2π shift of the theta angle (T),

$$\begin{aligned} S: \quad \tau &\rightarrow \frac{-1}{\tau}, \qquad (n_e, n_m) \rightarrow (n_m, -n_e), \\ T: \quad \tau &\rightarrow \tau{+}1, \qquad (n_e, n_m) \rightarrow (n_e{+}n_m, n_m). \end{aligned} \tag{4.44}$$

Thus electric-magnetic duality simply expresses the equivalence of free $U(1)$ field theories coupled to classical (massive) sources under $Sp(2n, \mathbf{Z})$ redefinitions of electric and magnetic charges. The importance of this redundancy in the Lagrangian description of IR effective actions becomes apparent when there is a moduli space $\mathcal{M}$ of inequivalent vacua. In that case, upon traversing a closed loop in $\mathcal{M}$ the physics must, by definition, be the same at the beginning and end of the loop, but the Lagrangian description need not—it may have suffered an electric-magnetic duality transformation. This possibility is often expressed by saying that the coupling matrix τ_{IJ}, in addition to being symmetric and having positive definite imaginary part, is also a section of a (flat) $Sp(2n, \mathbf{Z})$ bundle with action given by Eq. 4.38.

Electric-magnetic duality can be generalized to other free theories with $U(1)$ gauge invariances. For example, in four dimensions we can also consistently couple a two-form field $B = \frac{1}{2} B_{\mu\nu} dx^\mu \wedge dx^\nu$ if it is invariant under the gauge transformation $\delta B = d\Lambda$ for an arbitrary one-form Λ. Then the gauge-invariant field strength is the three-form $H = dB$, and the IR free Lagrangian is $\mathcal{L} \sim H \wedge {*H}$. We can define a dual "magnetic" field strength one-form by $\widetilde{H} \equiv {*H}$, and, away from sources, its gauge potential (zero-form) Φ by $\widetilde{H} = d\Phi$. In this case the gauge transformations are shifts of Φ by constants, and the Lagrangian becomes $\mathcal{L} \sim d\Phi \wedge {*d\Phi}$. Thus electric-magnetic duality implies that the two-form potential theory is equivalent to that of a derivatively-coupled real scalar field. In particular, we lost no generality by not including two-form potentials in our free IR effective actions. In a general space-time dimension d, electric-magnetic duality relates IR free $U(1)$ theories of p-form potentials to those of $(d-p-2)$-form potentials; the resulting discrete duality groups (including theta angle rotations) have been worked out.[41]

Finally, electric-magnetic duality extends trivially to supersymmetric theories as well. For example, treating the field strength left-chiral superfield W_L in $\int d^2\theta_L (\tau/32\pi i) \cdot (W_L^2)$ as an independent field in the path integral, the Bianchi identity, $\mathcal{D}_L W_L = -\mathcal{D}_R W_R$, can be implemented by a real vector superfield V Lagrange multiplier. We add to the action

$$\frac{i}{16\pi} \int d^4x\, d^4\theta\, V \mathcal{D}_L W_L = \frac{-i}{16\pi} \int d^4x\, d^2\theta_L\, \mathcal{D}_L V W_L = \frac{1}{16\pi} \int d^4x\, d^2\theta_L\, \widetilde{W}_L W_L, \tag{4.45}$$

plus its complex conjugate. Performing the Gaussian integral over W_L gives an equivalent action

$$\widetilde{S} = \int d^4x\, d^2\theta_L\, \frac{1}{32\pi i} \left(\frac{-1}{\tau}\right) (\widetilde{W}_L^2) + \text{c.c.} \tag{4.46}$$

4.3 An $SU(2)$ Coulomb branch

Let us consider an N=1 supersymmetric $SU(2)$ gauge theory with an adjoint left-chiral superfield Φ^a_b and no superpotential. (This theory is actually N=2 supersymmetric, as we will discuss later.) The adjoint representation can be thought of as the set of hermitian traceless 2×2 matrices, acted on by gauge transformations as $\Phi \to g\Phi g^{-1}$ where $g \in SU(2)$. Since the scalar component of Φ is complex, it takes values in the set of complex traceless 2×2 matrices (no hermiticity condition). The classical moduli space is parametrized by the singlet composite left-chiral superfield

$$U = \text{Tr}\Phi^2. \tag{4.47}$$

Higher powers of Φ in the color trace are just polynomials in U. Thus the classical moduli space is the complex U-plane. The classical Kahler potential is $K \sim (\overline{U}U)^{1/2}$, so there is a $\mathbb{Z}_2$ conical singularity at the origin, corresponding to the vacuum where the full $SU(2)$ symmetry is restored.

This moduli space is actually in a Coulomb phase. One way of seeing this is to note that Φ has left-chiral superfield degrees of freedom, U has one, so only two were given mass. Thus only two of the three $SU(2)$ are Higgsed, so it must be that $SU(2) \to U(1)$ on the U-plane. This can be seen more directly by noting that the D-term equations imply

$$[\Phi^\dagger, \Phi] = 0, \tag{4.48}$$

which implies that Φ can be diagonalized by color rotations:

$$\Phi = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}, \tag{4.49}$$

and there is a discrete gauge identification $a \simeq -a$. It is easy to see that (4.49) leaves the diagonal $U(1) \subset SU(2)$ unbroken, and the light field U is neutral under this $U(1)$.

The light degrees of freedom are thus a $U(1)$ vector superfield W_L and the left-chiral superfield U. This theory has an anomaly-free $U(1)_R$ symmetry under which $R(W_L) = 1$ and $R(U) = 0$. There is also an anomalous $U(1)_A$ under which W_L is neutral, $U \to e^{i\alpha}U$, and $\vartheta \to \vartheta + 2\alpha$. By the usual arguments, there is no dynamically generated superpotential for U, so the classical flat directions are not lifted.

This is not the whole story, though, since there is also the kinetic term for the vector superfield:

$$S_{\text{coul}} = \text{Kahler} + \int d^4x \left[\frac{1}{16\pi i}\tau(U,\lambda^4)\text{tr}(W_L^2) + c.c.\right]_F \tag{4.50}$$

where $\tau(U) = \frac{\vartheta(U)}{2\pi} + i\frac{4\pi}{g^2(U)}$ is the low energy coupling, which can depend on U. Our goal will be to detemine the coupling function $\tau(U)$. What is the meaning of this $U(1)$ coupling in an IR effective theory? Classically, the AF $SU(2)$ theory is being Higgsed at the scale $U^{1/2}$ down to $U(1)$; since the fields charged under the $U(1)$ (*e.g.* the $W^\pm$ bosons) get masses of order $U^{1/2}$, they decouple at smaller scales, and the $U(1)$ coupling does not run. Thus the IR coupling τ just measures the $SU(2)$ coupling at the scale $U^{1/2}$. By asymptotic freedom, for $\langle U\rangle \gg \Lambda^2$ this one-loop description of the physics should be accurate. The question is what happens for $\langle U\rangle < \Lambda^2$.

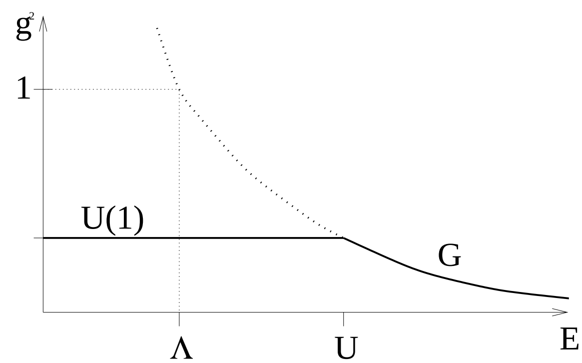

Figure 4.1: Running of the coupling of an asymptotically free gauge theory with gauge group G Higgsed to $U(1)$'s at a scale $U \gg \Lambda^2$. The $U(1)$ couplings do not run below $\langle\Phi\rangle$ because there are no charged fields lighter than ϕ.

Recall that under the anomalous $U(1)_A$ rotations the theta angle and therefore τ shifts. By the angular nature of the theta angle, the shift $\tau \to \tau+1$ is a symmetry. This, plus holomorphy and matching to the one-loop β-function at weak coupling (large U) implies

$$\tau(U) = \frac{1}{2\pi i}\log\left(\frac{\Lambda^4}{U^2}\right) + \sum_{n=0}^{\infty} c_n \left(\frac{\Lambda^4}{U^2}\right)^n . \tag{4.51}$$

The first term is just the one-loop $SU(2)$ β-function. The non-perturbative term with coefficient c_n corresponds to an n-instanton contribution. (Since the model is Higgsed, the instantons have an effective IR cutoff at the scale U, so these instanton effects are calculable; the first two coefficients have been calculated.)

As we make a large circle in the U-plane, the effective coupling shifts, $\tau \to \tau-2$, corresponding to an unobservable theta angle shift $\vartheta \to \vartheta-4\pi$. Note that there is a global discrete symmetry of this model which acts on the U-plane as

$$\mathbb{Z}_2 : \qquad U \to -U, \tag{4.52}$$

and so takes $\tau \to \tau - 1$ (a 2π shift in the theta angle). This $\mathbb{Z}_2$ is part of the anomaly-free $\mathbb{Z}_4$ subgroup of the anomalous $U(1)_A$.

Solving for the vacuum structure of the $SU(2)$ theory is thus reduced to determining this function $\tau(U)$. It is worth examining the formula (4.51) in some detail. The first, logarithm, term came from matching to the one-loop running of the microscopic coupling for $U \gg \Lambda^2$. Any other terms containing multiple logarithms, or any non-constant coefficient of the single logarithm term are not allowed, since they would necessarily imply $\tau(U)$ transformations under theta angle rotations which are U-dependent, and therefore not in $SL(2,\mathbb{Z})$ since $SL(2,\mathbb{Z})$ is a discrete group of transformations. The absence of these higer logarithm terms is equivalent to the absence of all higher-loop corrections to the running of the microscopic coupling.

The terms proportional to Λ^{4n} correspond to a non-perturbative n-instanton contribution. Since the model is Higgsed for large U, the instantons have an effective IR cutoff at the scale U, so these instanton effects are calculable; the first two coefficients have been calculated [63]. In principle one could compute $\tau(U)$ by calculating all the n-instanton contributions, and then analytically continuing (4.51) to the whole U-plane; in practice this is too hard. Instead, we follow N. Seiberg and E. Witten's more physical approach to determining $\tau(U)$ [42].

So far we have been doing the "standard" analysis of the low energy effective action for this theory. But there are two puzzles which indicate that we are missing some basic physics:

- (1.) The effective coupling $\tau(U)$ is holomorphic, implying that $\mathrm{Re}\tau$ and $\mathrm{Im}\tau$ are harmonic functions on the U-plane. Since they are not constant functions, they therefore must be unbounded both above and below. In particular this implies that $\mathrm{Im}\tau = \frac{1}{g^2}$ will be negative for some U, and the effective theory will be non-unitary!

- (2.) If we add a tree level mass $f_{tree} = m\mathrm{tr}\Phi^2 = mU$, then, for $m \gg \Lambda$, Φ can be integrated out leaving a low energy pure $SU(2)$ superYM theory with scale $\widehat{\Lambda}^6 = m^2\Lambda^4$. This theory has a gap, confinement, and two vacua with gaugino condensates $\langle\lambda\lambda\rangle = \pm m\Lambda^2$. But, in our low energy theory on the U-plane, there are no light charged degrees of freedom to Higgs the photon.

The remainder of this lecture presents the physical ingredients which resolve these puzzles. In the next lecture we return to this $SU(2)$ theory and solve for $\tau(U)$.

Now, we learned last lecture that this theory can have magnetic monopoles. Indeed, one can show that there are semi-classically stable solitons with charges $(n_e, n_m) = (0, \pm 1)$ in this theory, and they turn out to lie in chiral multiplets of the supersymmetry algebra. Furthermore, from (4.51) we see that changing the phase of U shifts the effective theta angle. In particular under the global $\mathbb{Z}_2$: $U \to e^{i\pi}U$, $\tau \to \tau - 1$. From the associated duality transformation on the charges of any massive states (4.18), we see that there will be $(\mp 1, \pm 1)$ dyons in the spectrum. Repeating this procedure, we find there must be a whole tower of semi-classically stable dyons of charges $(n, \pm 1)$ for arbitrary integers n.

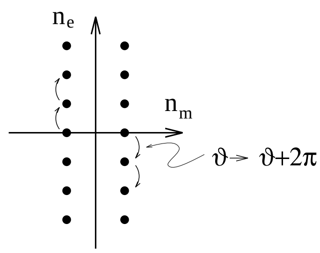

Figure 4.2: dyon spectrum

The existence of these dyon states suggests a possible resolution to one of our puzzles: perhaps at some strong coupling point on the moduli space, *e.g.* $U = U_0$ with

$$U_0 \sim \Lambda^2, \tag{4.53}$$

one of these dyons becomes massless, thereby providing the light charged left-chiral superfield needed to Higgs the $U(1)$.

4.3.1 Physics near U_0

Making this assumption, let us check that it gives rise to the desired physics. We add to our low energy effective theory two left-chiral superfields M and $\widetilde{M}$ oppositely charged under the $U(1)$. (We need two for anomaly-cancellation in the $U(1)$.) Since we are supposing that these states become massless at $U = U_0$, we can expand the effective superpotential around this point as

$$f = (U - U_0)M\widetilde{M} + \mathcal{O}((U - U_0)^2). \tag{4.54}$$

The D equations from the coupling to the $U(1)$ gauge field imply

$$|M| = |\widetilde{M}|, \tag{4.55}$$

while the F equations from (4.54) are

$$\begin{aligned} 0 &= \frac{\partial f}{\partial U} = M\widetilde{M}, \\ 0 &= \frac{\partial f}{\partial M} = (U - U_0)\widetilde{M}. \end{aligned} \tag{4.56}$$

The solutions are $M = \widetilde{M} = 0$ with U arbitrary, which is just the U-plane Coulomb branch.

Now we add a bare mass term for the adjoint Φ, and see if we lift the flat directions and obtain a discrete set of gapped vacua. The bare mass term is $f_{tree} = m\mathrm{tr}\Phi^2 = mU$. By the selection rule for the anomalous $U(1)_A$ under which m has charge -1 and the $U(1)_R$ under which m is assigned charge 2, and the usual non-renormalization argument, the low energy effective superpotential must be of the form

$$f = (U - U_0)M\widetilde{M} + mU + \mathcal{O}((U - U_0)^2). \tag{4.57}$$

The D and F equations become

$$\begin{aligned} 0 &= |M| - |\widetilde{M}|, \\ 0 &= M\widetilde{M} + m, \\ 0 &= (U - U_0)\widetilde{M}, \end{aligned} \tag{4.58}$$

whose solutions are $|M| = |\widetilde{M}| = m^{1/2}$ and $U = U_0$. Thus the Coulomb branch is indeed lifted, and there is only a single vacuum. This vacuum has a gap, since the charged left-chiral superfields M and $\widetilde{M}$ get non-zero vevs, thereby Higgsing the $U(1)$.

In this analysis, we have implicitly assumed (in writing down the D terms) that M and $\widetilde{M}$ were electrically charged with respect to the $U(1)$ field. But, by electric-magnetic duality, our analysis is valid for any dyonic charges. This is because M and $\widetilde{M}$ are the only light charged fields in the theory near U_0, so we can by an electric-magnetic duality transformation rotate any (n_e, n_m) to a description in which they are proportional to $(1, 0)$.[3] Then in this description the above analysis is valid.

Now, for $m \gg \Lambda$ we expect to recover the two gapped vacua of the pure $SU(2)$ superYM theory. Recalling the $\mathbb{Z}_2$ symmetry of the theory, it is natural to assume that there are two points on the U-plane where charged left-chiral superfields become massless in the $m = 0$ theory, and they are at $U = \pm U_0$. Since Λ is the only scale in the theory, we take $U_0 = \Lambda^2$. (We can take this as the definition of Λ, if we like.)

[3] More precisely, there is an $SL(2, \mathbb{Z})$ transformation which takes them to $(q, 0)$ where q is the greatest common divisor of n_e and n_m.

4.3.2 Monodromies

Can this assumption be checked? Yes, by examining the behavior of τ as a function of U. Recall the other puzzle we had about the physics on the Coulomb branch: since $\tau(U)$ is holomorphic, $1/g^2 \sim \mathrm{Im}\tau$ is harmonic and therefore unbounded from below, violating unitarity.

This puzzle is resolved by noting that τ is not, in fact, a holomorphic function of U. In particular, by electric-magnetic duality, as we traverse closed loops in the U-plane, τ need not come back to the same value, only one related to it by an $SL(2,\mathbb{Z})$ transformation. Mathematically, this is described by saying that τ is a section of a flat $SL(2,\mathbb{Z})$ bundle. This multi-valuedness of τ can be described by saying that τ is a holomorphic function on a cut U-plane, with cuts emanating from some singularities, and with the jump in τ across the cuts being an element of $SL(2,\mathbb{Z})$. The two points $U = \pm\Lambda^2$ at which we are assuming there are massless charged left-chiral superfields are the natural candidates for the branch points as shown in the figure. The presence of these cuts allows us to avoid the conclusion that $\mathrm{Im}\tau$ is unbounded.

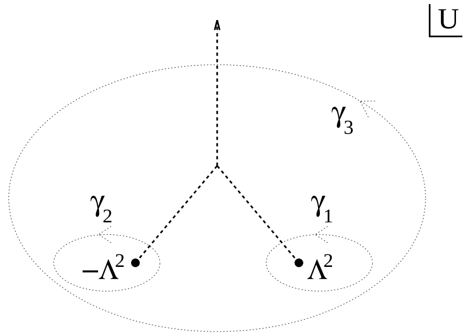

Figure 4.3: Cut U-plane with three loops. The cuts have been placed in an arbitrary manner connecting the two possible strong-coupling singularities, and a possible singularity at weak coupling ($U = \infty$).

Upon traversing the various loops γ_i in the above figure, τ will change by the action of an $SL(2,\mathbb{Z})$ element. These elements are called the *monodromies* of τ, and will be denoted $\mathcal{M}_i$.

We first calculate $\mathcal{M}_3$, the monodromy around the weak-coupling singularity at infinity. By taking γ_3 of large enough radius, τ will be accurately given by its one-loop value, the first term in (4.51). Taking $U \to e^{2\pi i}U$ in this formula gives $\tau \to \tau - 2$, giving for the monodromy at infinity[4]

$$\mathcal{M}_3 = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix}. \tag{4.59}$$

In order to calculate the $\mathcal{M}_{1,2}$ monodromies, let us first calculate the monodromy we would expect if the left-chiral superfield becoming massless at the associated singularity had charge (n_e, n_m). By a duality transformation we can change to a basis where this charge is purely electric: $(\widetilde{n}_e, 0)$. In this basis the physics near the $U = U_0$ singularity is just that of QED with the electron becoming massless. This theory is IR free, so the behavior of the low energy effective coupling will be dominated by its one-loop expression, at least sufficiently near U_0 where the mass of the charged left-chiral superfield $\sim U - U_0$ is arbitrarily small:

$$\widetilde{\tau} = \frac{\widetilde{n}_e^{\,2}}{\pi i} \log(U - U_0) + \mathcal{O}(U - U_0)^0. \tag{4.60}$$

By traversing a small loop around U_0, $(U - U_0) \to e^{2\pi i}(U - U_0)$, we find the monodromy

$$\widetilde{\tau} \to \widetilde{\tau} + 2\widetilde{n}_e^{\,2} \qquad \Longrightarrow \qquad \widetilde{\mathcal{M}} = \begin{pmatrix} 1 & 2\widetilde{n}_e^{\,2} \\ 0 & 1 \end{pmatrix}. \tag{4.61}$$

Now let us duality transform this answer back to the basis where the charges are (n_e, n_m). The required $SL(2, \mathbb{Z})$ element will be denoted $\mathcal{N} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, and satisfies

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} n_e \\ n_m \end{pmatrix} = \begin{pmatrix} \widetilde{n}_e \\ 0 \end{pmatrix}, \quad \text{and} \quad ad - bc = 1 \quad \text{with} \quad a, b, c, d \in \mathbb{Z}. \tag{4.62}$$

The transformed monodromy is then

$$\mathcal{M} = \mathcal{N}\widetilde{\mathcal{M}}\mathcal{N}^{-1} = \begin{pmatrix} 1 + 2n_e n_m & 2n_e^2 \\ -2n_m^2 & 1 - 2n_e n_m \end{pmatrix}. \tag{4.63}$$

Now, by deforming the γ_i contours in the U-plane, we find that the three monodromies must be related by

$$\mathcal{M}_3 = \mathcal{M}_1 \mathcal{M}_2. \tag{4.64}$$

Assuming that a left-chiral superfield with charges (n_{e1}, n_{m1}) becomes massless at $U = \Lambda^2$, while one with charges (n_{e2}, n_{m2}) does so at $U = -\Lambda^2$, and substituting into (4.64) using (4.59) and (4.63) gives as solutions

$$(n_{e1}, n_{m1}) = \pm(n, 1), \qquad (n_{e2}, n_{m2}) = \pm(n-1, 1), \qquad \text{for all} \quad n \in \mathbb{Z}. \tag{4.65}$$

[4]This actually only determines the monodromy up to an overall sign. The sign is determined by noting that $U \to e^{2\pi i} U$ has the effect of $\Phi \to -\Phi$ on the elementary Higgs field, so it reverses the sign of the low energy electromagnetic field which in terms of $SU(2)$ variables is proportional to $\mathrm{tr}(\Phi F)$. Thus it reverses the sign of electric and magnetic charges, giving an "extra" factor of $-\mathbb{1} \in SL(2, \mathbb{Z})$.

This set of charges actually represents a single physical solution. This is because taking $U \to e^{i\pi}U$ takes us to an equivalent theory by the $\mathbb{Z}_2$ symmetry; but this corresponds to shifting the low energy theta angle by 2π which in turn shifts all dyon electric charges by their magnetic charges. Repeated applications of this shift can take any of the above solutions to the solution

$$(n_{e1}, n_{m1}) = \pm(0,1), \qquad (n_{e2}, n_{m2}) = \pm(-1,1). \tag{4.66}$$

The plus and minus sign solutions must both be there by anomaly cancellation in the low energy $U(1)$. We thus learn that there is a consistent solution with a monopole becoming massless at $U = \Lambda^2$ and a charge $(-1,1)$ dyon becoming massless at $U = -\Lambda^2$. Some progress has been made in weakening the initial assumption that there are just two strong-coupling singularities [67].

4.3.3 $\tau(U)$

With the monodromies around the singularities in hand, we now turn to finding the low energy coupling τ on the U-plane. The basic idea is that τ is determined by holomorphy and demanding that it match onto the behavior we have deterined above at $U = \infty$ and $U = \pm\Lambda^2$. Seeing how to solve this "analytic continuation" problem analytically is not obvious, however. Seiberg and Witten did it by introducing an auxiliary mathematical object: a family of tori varying over the Coulomb branch.

This is a useful construction because the low energy effective coupling τ has the same properties as the complex structure of a 2-torus. In particular, the complex structure of a torus can be described by its *modulus*, a complex number τ, with $\text{Im}\tau > 0$. In this description, the torus can be thought of as a parallelogram in the complex plane with oposite sides identified, see the figure. Furthermore, the modulus τ of such a torus gives equivalent complex structures modulo $SL(2,\mathbb{Z})$ transformations acting on τ. Therefore, if we associate to each point in the U-plane a holomorphically-varying torus, its modulus will automatically be a holomorphic section of an $SL(2,\mathbb{Z})$ bundle with positive imaginary part, which are just the properties we want for the effective coupling τ.

At $U = \pm\Lambda^2$, magnetically charged states become massless, implying that the effective coupling $\text{Im}\tau \to 0$. (Recall that by $U(1)$ IR freedom, when an electrically charged state becomes massless, the coupling $g \to 0$, implying $\tau \to +i\infty$. Doing the duality transform $\tau \to -1/\tau$ gives the above result for a magnetic charge becoming massless.) From the parallelogram, we see this implies that the torus is degenerating: one of its cycles is vanishing.

Now, a general torus can be described analytically as the Riemann surface which is

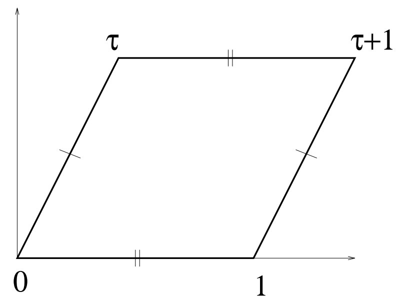

Figure 4.4: A complex torus as a parallelogram in the complex plane with opposite sides identified.

the solution $y(x)$ to the complex cubic equation

$$y^2 = (x - e_1)(x - e_2)(x - e_3). \tag{4.67}$$

We can think of this as a double-sheeted cover of the x-plane, branched over the three points e_i and the point at infinity. We let this torus vary over the U-plane by letting the e_i vary: $e_i = e_i(U, \Lambda)$. By choosing the cuts to run between pairs of these branch points, and "gluing" the two sheets together along these cuts, one sees that the Riemann surface is indeed topologically a torus. Furthermore, the condition for a nontrivial cycle on this torus to vanish is that two of the branch points collide. Since we want this to happen at the two points $U = \pm\Lambda^2$, it is natural to choose $e_1 = \Lambda^2$, $e_2 = -\Lambda^2$, and $e_3 = U$:

$$y^2 = (x - \Lambda^2)(x + \Lambda^2)(x - U). \tag{4.68}$$

Furthermore, note that this choice has a manifest $U \to -U$ symmetry, under which $x \to -x$ and $y \to \pm iy$.

Given this family of tori, one can compute their moduli as a ratio of line integrals:

$$\tau(U) = \frac{\oint_\beta \omega}{\oint_\alpha \omega}, \tag{4.69}$$

where ω is the (unique) holomorphic one-form on the Riemann surface,

$$\omega = \frac{dx}{y} = \frac{dx}{\sqrt{(x^2 - \Lambda^4)(x - U)}}, \tag{4.70}$$

and α and β are any two non-trivial cycles on the torus which intersect once. For example, we might take α to be a cycle on the x-plane which loops around the branch points at $\pm\Lambda^2$, while β is the one which loops around the branch points at Λ^2 and U. If we chose the cuts on the x-plane to run between $\pm\Lambda^2$ and between U and ∞, then the α cycle would lie all on one sheet, while the β cycle would go onto the second sheet as it passes through the cut; see the figure. Since the integrand in (4.70) is a closed one form ($d\omega = 0$), the value of τ does not depend on the exact locations of α and β, but only on how they loop around the branch points.

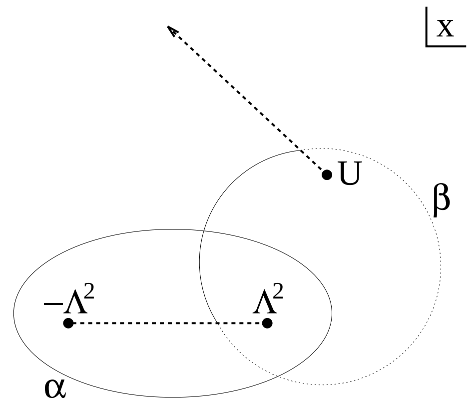

Figure 4.5: Cut x-plane with α and β cycles.

We can now check that our family of tori (4.68) indeed give rise to the correct low energy τ. By taking $U \to \infty$, it is not hard to explicitly evaluate (4.70) to find agreement with the first term in the weak-coupling expansion (4.51).[5] Also, without having to explicitly evaluate the integrals in (4.70), one can check that it reproduces the correct monodromies as U goes around the singularities at $\pm\Lambda^2$ by tracking how the α and β cycles are deformed as U varies. Finally, it turns out that the family of tori (4.68) is the unique one with these properties [42].

4.3.4 Dual Higgs mechanism and confinement

In summary, we have found the solution for $SU(2)$ with a massive adjoint in which, at zero mass, there is a complex U-plane of degenerate vacua in a Coulomb phase. The vacua at $U = \pm\Lambda^2$ are special since a monopole and dyon, respectively, becomes massless there. When we turn on a non-zero mass for the adjoint, all the vacua on the U-plane are lifted, except for the two massless points. At those points, the scalar monopole or dyon fields condense, Higgsing the (appropriate electric-magnetic dual) $U(1)$. This is illustrated in a picture of the combined moduli and parameter space of the model:

One puzzle that may remain concerning this solution is that for an adjoint mass $m \gg \Lambda$ we expected to find two confining vacua of the low energy pure superYM

[5] Though perhaps only up to an $SL(2,\mathbb{Z})$ transformation if I made the wrong choice for my α and β cycles.

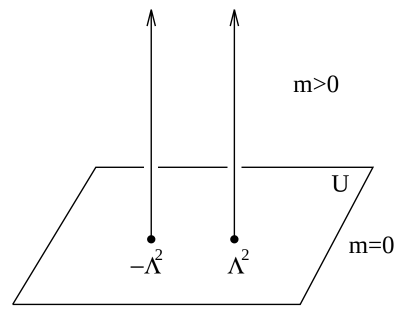

Figure 4.6: Moduli and parameter space for $N{=}1$ $SU(2)$ gauge theory with a massive adjoint left-chiral superfield.

theory, yet we seem to have found instead two Higgs vacua. This is not quite right, though, since the Higgs mechanism taking place is not the usual condensation of an electrically charged scalar field, but of magnetically (and dyonically) charged scalars.

To see what this means, let us recall the basic physics of the Higgs mechanism. When an electric charge condenses, it screens any background electromagnetic fields, damping them exponentially—this is a consequence of the photon aquiring a non-zero mass. This means that electric sources in the theory are essentially free, for their electric fields can be "absorbed" by the electric condensate, and their interaction energy will drop off exponentially. Magnetic charges, on the other hand, behave very differently, because the magnetic field lines have no condensate source to end on. The result is that magnetic field lines tend to be excluded from the vacuum; this is called the Meissner effect in superconductors. The minimum energy configuration is for the magnetic field to be confined to a thin flux tube connecting opposite magnetic charges, leading to confining forces between them. Thus, in the Higgs mechanism, electric charges are screened and magnetic charges are confined.

To see what happens when magnetic charges, instead of electric charges, condense, we simply do an electric-magnetic duality transformation. Thus in the dual Higgs effect, magnetic charges are screened, and electric charges are confined. So we have indeed found confinement in our $SU(2)$ solution at the monopole point. This is a concrete realization of a picture of confinement in non-Abelian gauge theories proposed in the '70's by S. Mandelstam and by G. 't Hooft.

Finally, at the dyonic point, by another duality transformation, it is not hard to see that both electric and magnetic charges are confined, though any dyonic charges proportional to $(-1,1)$ will just be screened. This is a realization of an "oblique confinement" phase of non-Abelian gauge theories proposed by 't Hooft.

Bibliography

[1] S. Weinberg, *The Quantum Theory of Fields III: Supersymmetry* (Cambridge University Press, 2000).

[2] S. Coleman and J. Mandula, Phys. Rev. **D159** (1967) 1251.

[3] E. Witten in *Proceedings of the International School of Subnuclear Physics, Erice 1981*, ed. A. Zichichi (Plenum Press, 1983).

[4] R. Haag, J. Lopuszanski, and M. Sohnius, Nucl. Phys. **B88** (1975) 257.

[5] E. Witten, Nucl. Phys. **B188** (1981) 513.

[6] E. Witten, Nucl. Phys. **B202** (1982) 253.

[7] T. Banks, W. Fischler, S. Shenker, and L. Susskind, Phys. Rev. **D55** (1997) 5112, hep-th/9610043.

[8] E. Witten, J. Diff. Geom. **17** (1982) 661.

[9] L. Alvarez-Gaumé, Comm. Math. Phys. **90** (1983) 161; J. Phys. **A16** (1983) 4177.

[10] D. Friedan and P. Windey, Nucl. Phys. **B235** (1984) 395.

[11] T. Cheng and L. Li, *Gauge Theory of Elementary Particle Physics* (Oxford University Press, 1984).

[12] S. Weinberg, *The Quantum Theory of Fields I: Foundations* (Cambridge University Press, 1995).

[13] S. Ferrara, L. Girardello, and F. Palumbo, Phys. Rev. **D20** (1979) 403.

[14] A. Nelson and N. Seiberg, Nucl. Phys. **B416** (1994) 46, hep-ph/9309299.

[15] R. Cahn, *Semi-Simple Lie Algebras and Their Representations* (Benjamin/Cummings Publishing, 1984).

[16] J. Wess and J. Bagger, *Supersymmetry and Supergravity* (2nd edition, Princeton University Press, 1992).

[17] I. Antoniadis and B. Pioline, Int. J. Mod. Phys. **A12** (1997) 4907, hep-th/9607058.

[18] For a quick review with references see J. Polchinski, in *Recent Directions in Particle Theory: From Superstrings and Black Holes to the Standard Model (TASI - 92)*, eds. J. Harvey and J. Polchinski (World Scientific, 1993), hep-th/9210046.

[19] For a review and discussion of little sting theories from this point of view see, for example, O. Aharony, Class. Quant. Grav. **17** (2000) 929, hep-th/9911147; and A. Kapustin, hep-th/9912044; and references therein.

[20] J. Polchinski, Nucl. Phys. **B303** (1988) 226.

[21] S. Coleman and D. Gross, Phys. Rev. Lett. **31** (1973) 851.

[22] N. Seiberg, Phys. Lett. **B318** (1993) 469, hep-ph/9309335.

[23] N. Seiberg, in *PASCOS 1994*, ed. K.C. Wali (World Scientific, 1995), hep-th/9408013.

[24] S. Weinberg, Phys. Lett. **B91** (1980) 51; L. Hall, Nucl. Phys. **B178** (1981) 75.

[25] S. Coleman, *Aspects of Symmetry*, (Cambridge University Press, 1985).

[26] A. Belavin, A. Polyakov, A. Schwarz and Y. Tyupkin, Phys. Lett. **B59** (1975) 85.

[27] L. Alvarez-Gaumé and E. Witten, in *Current Algebra and Anomalies*, eds. S. Treiman *et. al.* (Princeton University Press, 1985).

[28] M. Green, J. Schwarz and E. Witten, *Superstring Theory*, (Cambridge University Press, 1987).

[29] E. Witten, Phys. Lett. **B117** (1982) 324.

[30] G. 't Hooft, *Under the Spell of the Gauge Principle* (World Scientific, 1994).

[31] V. Novikov, M. Shifman, A. Vainshtein and V. Zakharov, Nucl. Phys. **B229** (1983) 381.

[32] R. Leigh and M. Strassler, Nucl. Phys. **B447** (1995) 95, hep-th/9503121.

[33] T. Banks and A. Zaks, Nucl. Phys. **B196** (1982) 189.

[34] For an introduction to the $N = 1$ and 2 supersymmetric results, see *e.g.* K. Intriligator, N. Seiberg, Nucl. Phys. **BC45** (1996) 1, hep-th/9509066; M. Peskin, hep-th/9702094.

[35] See *e.g.* S. Minwalla, hep-th/9712074.

[36] See *e.g.* P. Argyres, R. Plesser, N. Seiberg, E. Witten, Nucl. Phys. **B461** (1996) 71, hep-th/9511154.

[37] See *e.g.* J. Wess, J. Bagger, *Supersymmetry and Supergravity* (2nd edition, Princeton University Press, 1992); P. West, *Introduction to Supersymmetry and Supergravity* (2nd edition, World Scientific, 1990); M. Sohnius, Phys. Rep. **128** (1985) 39.

[38] See *e.g.* N. Seiberg in *Proceedings of the Trieste Spring School* (1997), hep-th/9705117.

[39] P. Dirac, Proc. R. Soc. **A133** (1931) 60; J. Schwinger, Phys. Rev. **144** (1966) 1087, **173** (1968) 1536; D. Zwanziger, Phys. Rev. **176** (1968) 1480,1489.

[40] E. Witten, Phys. Lett. **B86** (1979) 283.

[41] A. Shapere and F. Wilczek, Nucl. Phys. **B320** (1989) 669.

[42] N. Seiberg and E. Witten, Nucl. Phys. **B426** (1994) 19, hep-th/9407087.

[43] In the $N = 1$ context, K. Intriligator and N. Seiberg, Nucl. Phys. **B431** (1994) 551, hep-th/9408155.

[44] N. Seiberg, Phys. Lett. **B206** (1988) 75.

[45] S. Elitzur, A. Forge, A. Giveon, E. Rabinovici, Phys. Lett. **B353** (1995) 79, hep-th/9504080; Nucl. Phys. **B459** (1995) 160, hep-th/9509130; S. Elitzur, A. Forge, A. Giveon, K. Intriligator, E. Rabinovici, Phys. Lett. **B379** (1996) 121, hep-th/9603051.

[46] V. Novikov, M. Shifman, A. Vainshtein and V. Zakharov, Phys. Lett. **B166** (1986) 329; M. Shifman, A. Vainshtein and V. Zakharov, Phys. Lett. **B166** (1986) 334; M. Shifman and A. Vainshtein, Nucl. Phys. **B277** (1986) 456.

[47] W. Siegel, Phys. Lett. **B84** (1979) 193; I. Jack, D. Jones and K. Roberts, Z. Phys. **C63** (1994) 151, hep-ph/9401349.

[48] For the original and classic examples, see N. Seiberg, Phys. Rev. **D49** (1994) 6857, hep-th/9402044; Nucl. Phys. **B435** (1995) 129, hep-th/9411149.

[49] P. Argyres, R. Plesser and N. Seiberg, Nucl. Phys. **B471** (1996) 159, hep-th/9603042.

[50] P. Argyres, R. Plesser and A. Shapere, Nucl. Phys. **B483** (1997) 172, hep-th/9608129.

[51] N. Seiberg and E. Witten, Nucl. Phys. **B431** (1994) 484, hep-th/9408099.

[52] P. Argyres, to appear in *Proceedings of the Feza Gürsey Institute Summer School*, (Istanbul, 1998).

[53] P. Argyres, *Supersymmetric Effective Actions in Four Dimensions,* to appear in the proceedings of the Trieste Spring School, March 1998.

[54] E. Witten, Nucl. Phys. **B500** (1997) 3, hep-th/9703166.

[55] See, however, J. Maldacena, Adv. Theor. Math. Phys. **2** (1998) 231, hep-th/9711200; S. Gubser, I. Klebanov and A. Polyakov, Phys. Lett. **B428** (1998) 105, hep-th/9802109; E. Witten, Adv. Theor. Math. Phys. **2** (1998) 253, hep-th/9802150.

[56] C. Montonen and D. Olive, Phys. Lett. **B72** (1977) 117; H. Osborn, Phys. Lett. **B83** (1979) 321.

[57] E. Witten and D. Olive, Phys. Lett. **B78** (1978) 97.

[58] A. Sen, Phys. Lett. **B329** (1994) 217, hep-th/9402032.

[59] P. Argyres, R. Plesser and A. Shapere, Phys. Rev. Lett. **75** (1985) 1699, hep-th/9505100.

[60] P. Argyres and A. Shapere, Nucl. Phys. **B461** (1996) 437, hep-th/9509175.

[61] P. Argyres, Adv. Theor. Math. Phys. **2** (1998) 293, hep-th/9706095.

[62] P. Argyres and A. Buchel, Phys. Lett. **B431** (1988) 317, hep-th/9804007.

[63] N. Dorey, V. Khoze and M. Mattis, Phys. Lett. **B388** (1996) 324, hep-th/9607066.

[64] J. Harvey, in *1995 Summer School in High-Energy Physics and Cosmology: Proceedings*, edited by E. Gava *et. al.* (World Scientific, 1997), hep-th/9603086.

[65] G. 't Hooft, Nucl. Phys. **B79** (1974) 276; A. Polyakov, J.E.T.P. Lett. **20** (1974) 194.

[66] E. Witten, Phys. Lett. **B86** (1979) 283.

[67] R. Flume, M. Magro, L. O'Raifeartaigh, I. Sachs and O. Schnetz, Nucl. Phys. **B494** (1997) 331, hep-th/9611123.

[68] For discussions of the various definitions of rigid special Kähler geometry with references, see for example P. Fré, Nucl. Phys. Proc. Suppl. **BC45** (1996) 59, hep-th/9512043; I. Antoniadis and B. Pioline, Int. J. Mod. Phys. **A12** (1997) 4907, hep-th/9607058; B. Craps, F. Roose, W. Troost and A. Van Proeyen, Nucl. Phys. **B503** (1997) 565, hep-th/9703082; D. Freed, hep-th/9712042.

[69] R. Donagi and E. Witten, Nucl. Phys. **B460** (1996) 299, hep-th/9510101.

[70] For the first such papers, see A. Gorsky, I. Krichever, A. Marshakov, A. Mironov and A. Morozov, Phys. Lett. **B355** (1995) 466, hep-th/9505035; E. Martinec and N. Warner, Phys. Lett. **B549** (1996) 97, hep-th/9509161; T. Nakatsu and K. Takasaki, Mod. Phys. Lett. **A11** (1996) 157, hep-th/9509162; E. Martinec, Phys. Lett. **B367** (1996) 91, hep-th/9510204.

[71] For a review of these constructions, see *e.g.* R. Donagi in *Surveys in Differential Geometry*, ed. S.T. Yau, alg-geom/9705010.

[72] For a review of Riemann surfaces, see for example, D. Mumford, *Tata Lectures on Theta I* (Birkhäuser, 1983).

[73] P. Argyres and A. Faraggi, Phys. Rev. Lett. **74** (1995) 3931, hep-th/9411057; A. Klemm, W. Lerche, S. Theisen, and S. Yankielowicz, Phys. Lett. **B344** (1995) 169, hep-th/9411048; U. Danielsson and B. Sundborg, Phys. Lett. **B358** (1995) 273, hep-th/9504102; A. Hanany and Y. Oz, Nucl. Phys. **B452** (1995) 283, hep-th/9505075; A. Brandhuber and K. Landsteiner, Phys. Lett. **B358** (1995) 73, hep-th/9507008.

[74] A. Klemm, W. Lerche, P. Mayr, C. Vafa, N. Warner, Nucl. Phys. **B477** (1996) 746, hep-th/9604034.

[75] For an introduction to the string/M theory properties used in this subsection, see for example Chapter 14 of J. Polchinksi, *String Theory*, vol. II (Cambridge, 1998).

[76] E. Witten, Nucl. Phys. **B460** (1996) 335, hep-th/9510135.