SUPERSYMMETRIC EFFECTIVE ACTIONS IN FOUR DIMENSIONS

PHILIP C. ARGYRES Newman Laboratory, Cornell University, Ithaca NY 14853, USA argyres@mail.lns.cornell.edu

I review the construction of $N=0,\,1,\,{\rm and}\,\,2$ supersymmetric low energy effective actions in four space-time dimensions.

1 Introduction

The aim of these lectures is to introduce some of the arguments that have been used successfully in the last five years to obtain exact information about strongly coupled field theories. I will focus on four-dimensional field theories (without gravity), although the techniques described here have been applied to theories in other dimensions and to string/M theory as well.

The basic notion is that of a low energy (or Wilsonian) effective action. This is simply a local action describing a theory's degrees of freedom at energies below a given scale E. An 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. 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. If there is a sufficiently small ratio E/Λ between the cutoff energy scale E and the energy scale Λ characteristic of the dynamics of the degrees of freedom being averaged over, renormalization group arguments imply that the effective action can be systematically expanded as a power series in E/Λ —essentially an expansion in the number of derivatives of the fields.

We will use low energy effective actions to analyze four dimensional field theories by taking the limit as the cutoff energy scale E goes to zero, or equivalently, by just keeping the leading terms (up to two derivatives) in the low energy fields. I will call such $E \rightarrow 0$ low energy effective actions *infrared effective actions* (IREAs). The idea is to guess an IR effective field content for the microscopic (UV) theory in question and write down all possible IREAs built from these fields consistent with the global symmetries of the UV theory. For a "generic" UV theory this is no better than doing chiral perturbation theory for QCD, and would seem to give little advantage for obtaining exact results. 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 IREA sufficiently to deduce exact results. There are a number of reviews deriving these exact results ¹ assuming the constraints from supersymmetry. The purpose of these lectures will be to deduce and explain these constraints in a relatively non-technical way.

We will start with properties of general IREAs and then progressively specialize to those with N = 1 and then N = 2 supersymmetry. The constraints on the IREAs become progressively more restrictive as the number of supersymmetries is increased; in the N = 2 case they are strong enough to allow quite general and restrictive properties of the moduli space of vacua of gauge theories to be deduced. Important topics omitted include the properties of interacting IREAs—the representation theory of superconformal algebras² and their use in analyzing IREAs;³ instead these lectures concentrate on IR free effective actions. Also missing (as much as possible) are details of supersymmetry algebras and the construction of their representations—many good texts and review articles cover this material⁴—or the application of the ideas presented here to theories in other dimensions.⁵

Since an IREA describes physics only for arbitrarily low energies, it is, by definition, scale invariant: we simply take the cutoff scale E 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.

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.⁶

We generally have no effective description of interacting conformal field theories in four dimensions^{*a*} so we must limit ourselves to free or trivial theories in the IR. A large class of these is given by the Coleman-Gross theorem⁷ 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

 $[^]a \mathrm{See}$ however lectures in this volume on the anti de Sitter/conformal field theory correspondence.

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.

2 IREAs with No Supersymmetry

We thus take the field content of our IREA to be a collection of real scalars ϕ^i , Weyl spinors ψ^a_{α} , and U(1) vector fields A^I_{μ} . Here α and μ are the space-time spinor and vector indices, while *i*, *a*, and *I* label the different field species.

Since this theory is free in the IR, no interesting dynamics involving the spinor fields (like the formation of scalar condensates) can occur (basically by definition). Thus the vacuum structure of this theory is governed by the scalar potential. So, dropping the other fields we write the general Lagrangian with up to two derivatives for a set of real scalars

$$\mathcal{L} = -V(\phi) + \frac{1}{2}g_{ij}(\phi)\partial_{\mu}\phi^{i}\partial^{\mu}\phi^{j}.$$
(1)

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). Let's assume V attains its minimum value, which without loss of generality we take to be V = 0.

Minimizing the generalized kinetic energy term implies that in the vacuum the scalars should all be constant. Denoting these constant values by the same symbols as for the fields, the set of all possible vacua is then seen to naturally have the structure of a Riemannian manifold

$$\mathcal{M}_0 = \{\phi^i\}\tag{2}$$

with metric g_{ij} . Note that the manifold defined by Eq. 2 is independent of the particular choice of scalar fields used in Eq. 1 because an arbitrary nonsingular field redefinition $\phi^i \to \tilde{\phi}^i(\phi)$ transforms g_{ij} in the same way as a metric transforms under a change of coordinates. Eq. 1 is called a sigma model on \mathcal{M}_0 .

If V = 0 identically, then \mathcal{M}_0 would describe a manifold of vacua of this theory. We call such a manifold of vacua the *moduli space* of the theory. Without any extra symmetries to constrain it, generically $V \neq 0$, so \mathcal{M}_0 is not the moduli space, but instead

$$\mathcal{M}_V = \mathcal{M}_0 / \{ V = 0 \} \tag{3}$$

is. At least locally \mathcal{M}_V has the structure of a submanifold of \mathcal{M}_0 .

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.

Now let us incorporate the U(1) gauge fields into our discussion of the moduli space. Some of the scalar fields may be charged under the $U(1)^n$ gauge group of the IREA. The infinitesimal $U(1)^n$ action of the gauge group on the scalars then generates a diffeomorphism of \mathcal{M}_0

$$\phi^i \to \phi^i + \xi^i_I(\phi), \tag{4}$$

where I labels the U(1) generators. For Eq. 4 to be a symmetry of the IREA it is easy to see that it must both leave V invariant and be an isometry of the metric g_{ij} . In that case the IREA can be written (excluding the spinors) as

$$\mathcal{L} = -V(\phi) + \frac{1}{2}g_{ij}(\phi)D_{\mu}\phi^{i}D^{\mu}\phi^{j} - \frac{1}{32\pi}\mathrm{Im}\left[\tau_{IJ}(\phi)\mathcal{F}^{I}_{\mu\nu}\mathcal{F}^{J\mu\nu}\right],\qquad(5)$$

where, treating the ξ_I^i as (Killing) vectors generating the isometry, we have $D_{\mu} = \partial_{\mu} + A_{\mu}^I \xi_I$. The last term in Eq. 5 is a generalized Maxwell term for the U(1) field strengths $F_{\mu\nu}^I = \partial_{\mu} A_{\nu}^I - \partial_{\nu} A_{\mu}^I$, where we have defined

$$\mathcal{F}^{I}_{\mu\nu} = F^{I}_{\mu\nu} - \frac{i}{2} \epsilon_{\mu\nu\rho\sigma} F^{I\rho\sigma}, \qquad (6)$$

and 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. 5 is called a gauged sigma model on \mathcal{M}_0 .

Defining the real and imaginary parts of the couplings as

$$\tau_{IJ} = \frac{\theta_{IJ}}{2\pi} + i \frac{4\pi}{(e^2)_{IJ}},\tag{7}$$

the generalized Maxwell term can be expanded to

$$\mathcal{L}_{U(1)} = -\frac{1}{4(e^2)_{IJ}} F^{I}_{\mu\nu} F^{J\mu\nu} + \frac{\theta_{IJ}}{64\pi^2} \epsilon^{\mu\nu\rho\sigma} F^{I}_{\mu\nu} F^{J}_{\rho\sigma}, \tag{8}$$

showing that the imaginary part of τ_{IJ} is a matrix of couplings and the real part are theta angles. We will discuss the physical interpretation of these couplings momentarily.

First, though, let us see how the addition of the U(1) gauge fields affects the moduli space. Two points of \mathcal{M}_0 which are related by a gauge transformation Eq. 4 must be identified. Thus \mathcal{M}_0 or \mathcal{M}_V (since V is gauge invariant) is replaced by \mathcal{M} , formed by dividing by the action of the gauged isometry group $U(1)^n$:

$$\mathcal{M} = \mathcal{M}_V / U(1)^n. \tag{9}$$

The metric g' induced on \mathcal{M} is not simply the restriction of g to $\mathcal{M} \subset \mathcal{M}_0$, but is instead defined by $g'_{ij}d\phi^i d\phi^j = g_{ij}D\phi^i D\phi^j$, where $D\phi^i \equiv d\phi^i + A^I\xi^i_I$, with $d\phi^i = \partial_\mu \phi^i dx^\mu$ and $A^I = A^I_\mu dx^\mu$ thought of as one-form valued tangent vectors to \mathcal{M}_0 . This construction is known as a *Riemannian quotient*, and is just a geometrical realization of the (classical) Higgs mechanism.

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 $\text{Im}\tau \to +i\infty$). 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 ~ 1/137 on scales below the electron mass). Thus, in this case the coupling $\text{Im}\tau$ 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 spacetime, and thus no physics can depend on the θ_{II} 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).

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. 5 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.

2.1 Electric-Magnetic Duality

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

$$A^{I} = A^{I}_{\mu}dx^{\mu}$$

$$F^{I} = dA^{I} = \frac{1}{2}F^{I}_{\mu\nu}dx^{\mu} \wedge dx^{\nu},$$
(10)

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

$$*C \equiv \frac{1}{p!} \epsilon_{\mu_1 \cdots \mu_4} C^{\mu_1 \cdots \mu_p} dx^{\mu_{p+1}} \wedge \cdots \wedge dx^{\mu_4}, \tag{11}$$

so that $**C = (-)^{p+1}C$. In this language Eq. 6 becomes $\mathcal{F}^I = F^I - i*F^I$ and the Maxwell part of the IREA Eq. 5 becomes

$$S = -\frac{1}{16\pi} \int \operatorname{Im} \left[\tau_{IJ} \mathcal{F} \wedge * \mathcal{F} \right].$$
(12)

The classical Maxwell's equations with electric and magnetic sources follow from the action

$$S = \int \left(-\frac{1}{2e^2} F \wedge *F + A \wedge *j_e + \widetilde{A} \wedge *j_m \right), \tag{13}$$

where, away from any electric sources \tilde{A} is defined through $*F = d\tilde{A}$. The Dirac quantization condition⁸ 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. 13 are

$$\frac{1}{e}d*F = en_e\delta^{(3)}, \qquad \frac{1}{e}dF = \frac{4\pi}{e}n_m\delta^{(3)}, \tag{14}$$

which are invariant under the electric-magnetic duality transformation

$$(F/e) \to *(F/e), \qquad *(F/e) \to -(F/e),$$

$$n_m \to n_e, \qquad n_e \to -n_m,$$

$$e \quad \leftrightarrow \quad 4\pi/e. \tag{15}$$

The minus signs are because **F = -F in four-dimensional 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 $\int \mathcal{D}A^I e^{iS}$. This can be rewritten as a path integral over field strength configurations as long as we insert the Bianchi identity as a constraint: $\int \mathcal{D}F^I \mathcal{D}\tilde{A}_J e^{iS'}$, where $4\pi S' = 4\pi S + \int \tilde{A}_I \wedge dF^I$. Here \tilde{A}_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 $\int \tilde{A}_I \wedge dF^I = \int \tilde{F}_I \wedge$ $F^I = \frac{1}{2} \int \text{Im}[\tilde{\mathcal{F}}_I \wedge *\mathcal{F}^I]$ where $\tilde{\mathcal{F}}_I$ is related to $\tilde{F}_I = d\tilde{A}_I$ as in Eq. 6, we find an equivalent action, \tilde{S} , for \tilde{A}_I :

$$\widetilde{S} = -\frac{1}{16\pi} \int \operatorname{Im} \left[(-\tau^{IJ}) \widetilde{\mathcal{F}}_I \wedge * \widetilde{\mathcal{F}}_J \right], \qquad (16)$$

where τ^{IJ} is the matrix inverse of τ_{IJ} : $\tau^{IJ}\tau_{JK} = \delta^{I}_{K}$. 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

$$S: \ \tau_{IJ} \to -\tau^{IJ}, \tag{17}$$

together with the invariance of the physics under 2π shifts of the theta angles (integer shifts of $\text{Re}\tau_{IJ}$)

$$T^{(KL)}: \ \tau_{IJ} \to \tau_{IJ} + \delta_I^K \delta_J^L + \delta_I^L \delta_J^K, \tag{18}$$

generate a discrete group of duality transformations:

$$\tau_{IJ} \to (A_I{}^L \tau_{LM} + B_{IM})(C^{JN} \tau_{NM} + D^J{}_M)^{-1},$$
 (19)

where

$$M \equiv \begin{pmatrix} A_I^K & B_{IL} \\ C^{JK} & D^J{}_L \end{pmatrix} \in Sp(2n, \mathbf{Z}).$$
(20)

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)

$$AB^{T} = B^{T}A, \qquad B^{T}D = D^{T}B,$$

$$A^{T}C = C^{T}A, \qquad D^{T}C = CD^{T},$$

$$A^{T}D - C^{T}B = AD^{T} - BC^{T} = 1,$$
(21)

and imply that

$$M^{-1} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix}.$$
 (22)

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 ⁹ to n U(1) factors. Together these generate the action

$$(n_m \ n_e) \to (n_m \ n_e) \cdot M^{-1} \tag{23}$$

of the $Sp(2n, \mathbb{Z})$ electric-magnetic duality group on the 2*n*-component row vector of magnetic and electric charges.

Thus electric-magnetic duality simply expresses the equivalence of free U(1) field theories coupled to classical (massive) sources under $Sp(2n, \mathbb{Z})$ redefinitions of electric and magnetic charges. The importance of this redundancy in the Lagrangian description of IREAs 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) $Sp(2n, \mathbb{Z})$ bundle with action given by Eq. 20.

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 $\tilde{H} \equiv *H$, and, away from sources, its gauge potential (zero-form) Φ by $\tilde{H} = d\Phi$. 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 IREAs. 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.¹⁰

2.2 Effective Actions of Asymptotically Free Gauge Theories

As an example of an application of the above considerations, and to introduce the main class of field theories that we will be interested in, we discuss in this subsection effective actions of asymptotically free (AF) gauge theories.

A microscopic (UV) theory is characterized by some parameters (e.g. masses, strong coupling scales, theta angles, dimensionless couplings). We can always take ratios of these parameters to describe them by at most one scale Λ and a set of dimensionless parameters λ_k . The coefficient functions g_{ij} and τ_{IJ} of the IREA will, in general, depend on Λ and the λ_k . Determining this dependence of these IR quantities on UV parameters is the ultimate goal of the techniques reviewed in these lectures.

Consider an AF gauge theory kinetic term

$$\mathcal{L} = \frac{\tau_0}{32\pi i} \operatorname{tr}(\mathcal{F} \cdot \mathcal{F}) \tag{24}$$

Here we are thinking of \mathcal{L} as an effective action at a scale μ_0 , and g_0 is the coupling at that scale. For g_0 small enough we can calculate with arbitrary

accuracy the RG running of the coupling from the one loop result

$$\mu \frac{dg}{d\mu} = -\frac{b_0}{16\pi^2} g^3 + \mathcal{O}(g^5) \quad \Rightarrow \quad \frac{1}{g^2(\mu)} \simeq -\frac{b_0}{8\pi^2} \log\left(\frac{|\Lambda|}{\mu}\right), \qquad (25)$$

where we have defined

$$|\Lambda| \equiv \mu_0 e^{-8\pi^2/b_0 g_0^2},\tag{26}$$

the strong coupling scale of the gauge group. The coefficient of the one-loop beta function is given by

$$b_0 = \frac{11}{6}T(adj) - \frac{1}{3}\sum_a T(R_a) - \frac{1}{12}\sum_i T(R_i)$$
(27)

where the indices a run over Weyl fermions in representations R_a of the gauge group, and *i* runs over real scalars in the representations R_i . T(R) is the index of the representation R; for SU(N), for example, the index of the fundamental representation is 1, and of the adjoint representation is 2N.

Thus the complex gauge coupling is

$$\tau_0 \equiv \frac{\theta}{2\pi} + i\frac{4\pi}{g_0^2} = \frac{1}{2\pi i} \log\left[\left(\frac{|\Lambda|}{\mu_0}\right)^{b_0} e^{i\theta}\right],\tag{28}$$

where we have used the definition of the strong coupling scale $|\Lambda|$ in the last step. It is thus natural to define a complex "scale" by

$$\Lambda = |\Lambda| e^{i\theta/b_0} \qquad \Rightarrow \qquad \tau_0 = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\mu_0}\right). \tag{29}$$

Since we are dealing with an AF theory $(b_0 > 0)$, if we take the scale $\mu_0 \gg |\Lambda|$, then the theory is weakly coupled. 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 gauge coupling τ will then be some function $\tau(\Lambda, \phi^i; \mu)$ of the strong coupling scale Λ , the renormalization scale μ , and any scalar vevs ϕ .

We also have to take into account the angular nature of the theta angle $\theta \simeq \theta + 2\pi$, which means that as we rotate the phase of $\Lambda^{b_0} \to e^{2\pi i} \Lambda^{b_0}$, we must have $\tau \to \tau + 1$. This constrains the functional form of τ to be

$$\tau(\Lambda,\phi;\mu) = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\mu}\right) + f(\Lambda^{b_0},\overline{\Lambda}^{b_0},\phi;\mu),\tag{30}$$



Figure 1: Running of the coupling of an AF 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.

where f is an arbitrary single-valued function of its arguments (no cuts in the complex Λ^{b_0} plane).

Since we are dealing with an AF theory, the $\Lambda \to 0$ limit corresponds to the weak coupling limit, in which the effective couplings should not diverge. This allows non-analytic contributions to f of the form $(\log |\Lambda|)^{-n} \sim g^{2n}$ (an *n*-loop perturbative contribution) as well as analytic terms $\Lambda^{b_0 n} \sim e^{-8n\pi^2/g^2}$ (an *n*-instanton contribution).

This discussion has only applied to weakly coupled AF 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_i and \mathcal{F} fields may break down. However, in a case where a charged scalar gets a large vev Higgsing the AF gauge group down to a $U(1)^n$ subgroup at a large scale (and therefore weak coupling), the analysis of the preceding paragraphs can be applied to the low energy $U(1)^n$ theory; see Fig. 1. Since this theory is IR free, we can then run the RG scale to the far IR, giving for the low energy couplings

$$\tau_{IJ} = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\phi}\right) + f(\Lambda, \overline{\Lambda}, \phi).$$
(31)

Here, since we have run to the far IR, the renormalization scale $\mu = 0$, and so, by dimensional considerations, ϕ has to appear in the combination Λ/ϕ . (More complicated dependences can occur if there is more than one ϕ modulus.)

By itself this does not tell us much, since f is not determined. But an interesting constraint on f comes from the requirement that τ_{IJ} be a section of an $Sp(2n, \mathbb{Z})$ bundle. This allows τ_{IJ} to "jump" by an element of $Sp(2n, \mathbb{Z})$

upon making a circle in the ϕ moduli space. For example, using the matrix notation of Eq. 20, upon making a large circle in a complex ϕ plane τ could "jump" as

$$\tau(e^{2\pi i}\phi) = \frac{A \cdot \tau(\phi) + B}{C \cdot \tau(\phi) + D}.$$
(32)

The element $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2n, \mathbb{Z})$ is called the *monodromy* around the given closed path in the moduli space.

3 N = 1 Supersymmetric IREAs

The preceding discussion of electric-magnetic duality monodromies in the moduli space is largely moot in the case of non-supersymmetric theories, since generically they do not have non-trivial moduli spaces. One familiar example of a manifold of vacua occurs in theories with a spontaneously broken global symmetry. In this example, however, the vacua are related by the broken global symmetry generators and therefore necessarily have equivalent physics. Any further vacuum degeneracy is usually considered "accidental" and has to be engineered by fine-tuning parameters in the UV theory. For example, without any other symmetry, one would expect V to take on its minimum only at a discrete point in \mathcal{M}_0 , so \mathcal{M} is generically just a point. From a low energy perspective the problem is that degenerate vacua involve having exactly flat directions in the potential for some scalar fields.

Supersymmetry is another global symmetry which can constrain the form of the scalar potential by relating the scalars to the Weyl spinor fields in the IREA. Essentially because the way the spinor fields can enter is constrained by Lorentz invariance the form of the the kinetic terms and scalar potential are also constrained, in particular sometimes to have exactly flat directions. Note that since supersymmetry transformations take scalars to spinors, they can not relate different vacua (scalar vevs) and so do not imply equivalent physics along the flat directions. These supersymmetric selection rules will be the subject of succeeding subsections.

We start with N = 1 supersymmetry, the minimum amount of supersymmetry in four dimensions. Just using some basic facts about representations of the supersymmetry and Lorentz algebras (*i.e.*, avoiding detailed constructions needed for existence proofs) we can fairly quickly derive the important selection rules for the IREAs.

Recall that representations of the Lorentz algebra in four dimensions, $so(3,1) \simeq_{\mathbf{C}} su(2)_L \times su(2)_R$, can be labeled by their left and right su(2)"spins" (j_L, j_R) . The smallest representations are scalars (0,0), left handed Weyl spinors $(\frac{1}{2}, 0)$, right handed Weyl spinors $(0, \frac{1}{2})$, vectors $(\frac{1}{2}, \frac{1}{2})$, self-dual antisymmetric tensors (1,0), and anti-self-dual antisymmetric tensors (0,1). Also, complex conjugation reverses left and right spins of representations, so, by CPT invariance, fields with $j_L \neq j_R$ must be complex.

Supersymmetry generators are space-time spinors. The minimum amount of supersymmetry corresponds to a single complex $(\frac{1}{2}, 0)$ Weyl spinor supercharge Q. The basic N = 1 superalgebra is then

$$\{Q,\overline{Q}\} = P, \qquad \{Q,Q\} = \{\overline{Q},\overline{Q}\} = 0, \qquad [Q,P] = [\overline{Q},P] = 0, \qquad (33)$$

where \overline{Q} is the hermitian conjugate of Q and therefore a $(0, \frac{1}{2})$ Weyl spinor, P is the generator of space-time translations—a $(\frac{1}{2}, \frac{1}{2})$ vector of charges which act on fields as ∂_{μ} —and all Lorentz indices and necessary Clebsch-Gordon coefficients (Pauli matrices and the like) have been suppressed.

The irreducible representation of Eq. 33 containing a scalar field ϕ is easy to construct. ^b First, $\psi \equiv Q\phi$ is a $(\frac{1}{2}, 0)$ spinor, and for $Q^2\phi = 0$ we therefore need $Q\psi = 0$. For $\{Q, \overline{Q}\}\phi = P\phi$ we need $\overline{Q}\psi = P\phi$. If ϕ were real then $\overline{Q}\phi = \overline{\psi}$, implying $\overline{Q}^2\psi = P\overline{\psi}$ in contradiction with Eq. 33. Thus we must take ϕ complex and set $Q\overline{\phi} = 0$, giving

$$Q\phi = \psi, \qquad \overline{Q}\phi = 0,$$

$$Q\psi = 0, \qquad \overline{Q}\psi = P\phi,$$
(34)

along with their complex conjugates

$$\overline{Q\phi} = \overline{\psi}, \qquad Q\overline{\phi} = 0,
\overline{Q\psi} = 0, \qquad Q\overline{\psi} = P\overline{\phi}.$$
(35)

It is easy to see that this satisfies Eq. 33; (ϕ, ψ) is called a *chiral multiplet*. Eq. 34 already shows the most important feature of four dimensional supersymmetry for IREAs—complex conjugation of the scalars is tied to the chirality of the spinors. This gives the moduli space \mathcal{M} a complex structure.

To see this, we examine the general scalar kinetic terms in the IREA: $h_{ij}P\phi^iP\phi^j + \overline{h_{ij}}P\overline{\phi}^iP\overline{\phi}^j + g_{i\overline{j}}P\phi^iP\overline{\phi}^{\overline{j}}$, where h and g are functions of the ϕ^i and their complex conjugates and we are using the (somewhat redundant) notation of putting a bar over the species label of complex conjugated fields as well as over the fields themselves. Hermiticity implies that g satisfies the reality condition $\overline{g_{i\overline{j}}} = g_{j\overline{i}}$, while h is symmetric in i and j.

We will now show that $h_{ij} = 0$ (our first supersymmetric selection rule). Consider the term $\mathcal{L}_{\phi^2} = h_{ij} P \phi^i \phi^j$. Its Q variation contains the term $Q \mathcal{L}_{\phi^2} \supset$

 $[^]b{\rm In}$ what follows I am constructing on-shell supersymmetry transformations, assuming that supersymmetry is not spontaneously broken.

 $h_{ij}P\psi^{(i}P\phi^{j)}$. For the Lagrangian to be supersymmetric, this variation should cancel against the Q variation of some other term. But, from Eq. 33 it is easy to see that no other such term is possible and therefore $h_{ij} = 0$. Thus the general N = 1 scalar kinetic term is

$$\mathcal{L} = g_{k\overline{j}} P \phi^k P \overline{\phi}^{\overline{j}}.$$
(36)

Thus the moduli space \mathcal{M} is a complex manifold with metric $g_{i\overline{j}}$; such manifolds are called a Hermitian.

In fact \mathcal{M} satisfies a stronger condition. To see this, consider the term $\mathcal{L}_{\psi^4} = r_{\overline{\imath j} k \ell} \overline{\psi}^{\overline{\imath}} \overline{\psi}^{\overline{\jmath}} \psi^k \psi^\ell$, where r is a function of the ϕ^i symmetric on $\overline{\imath j}$ and $k\ell$ (since $\psi^k \psi^\ell$ have to be combined antisymmetrically on their spinor indices to make a scalar, so, since they anticommute, they are symmetric in their $k\ell$ indices). Then $Q\mathcal{L}_{\psi^4} \supset r_{\overline{\imath j} k \ell} \overline{\psi}^{\overline{\imath}} P \overline{\phi}^{\overline{\jmath}} \psi^k \psi^\ell$. The only term that can cancel this variation is $\mathcal{L}_{\psi^2 \phi} = \Gamma_{\overline{\imath j} k} \overline{\psi}^{\overline{\imath}} P \overline{\phi}^{\overline{\jmath}} \psi^k$ for some function Γ of the ϕ since $Q\mathcal{L}_{\psi^2 \phi} \supset \Gamma_{\overline{\imath j} k, \ell} \overline{\psi}^{\overline{\imath}} P \overline{\phi}^{\overline{\jmath}} \psi^k \psi^\ell$, where $\Gamma_{\overline{\imath j} k, \ell} \equiv (\partial/\partial \phi^\ell) \Gamma_{\overline{\imath j} k}$. For this cancelation to work implies $\Gamma_{\overline{\imath j} k} = \Gamma_{\overline{\imath i} k}$. Now, $\overline{Q} \mathcal{L}_{\psi^2 \phi} \supset \Gamma_{\overline{\imath j} k} \overline{\psi}^{\overline{\imath}} P \overline{\phi}^{\overline{\jmath}} P \phi^k$, and this variation can only be canceled by the kinetic term Eq. 36 since $\overline{Q}\mathcal{L} \supset g_{\overline{k} \overline{\imath}} \overline{\psi}^{\overline{\imath}} P \overline{\phi}^{\overline{\jmath}} P \phi^k$. But then the symmetry of $\Gamma_{\overline{\imath j} k}$ in $\overline{\imath j}$ implies $g_{k\overline{\jmath},\overline{\imath}} = g_{k\overline{\imath},\overline{\jmath}}$. This condition and its complex conjugate have the solution (locally)

$$g_{k\overline{\jmath}} = \mathcal{K}_{,k\overline{\jmath}},\tag{37}$$

where \mathcal{K} is some real function of the ϕ and $\overline{\phi}$, called the *Kähler potential*; manifolds with such metrics are called Kähler manifolds. Thus the moduli space \mathcal{M} of an N = 1 supersymmetric theory is a Kähler manifold.

We can add in the U(1) vector fields as well. The field strength can be written in terms of a self-dual antisymmetric tensor \mathcal{F} defined in Eq. 6 and its anti-self-dual complex conjugate tensor. Since the (1,0) self-dual tensor representation appears in the tensor product of two left-handed Weyl spinor representations, we can make an (on-shell) N = 1 supersymmetry representation from a $(\frac{1}{2}, 0)$ spinor λ and \mathcal{F} much as in Eq. 34

$$Q\lambda = \mathcal{F}, \qquad \overline{Q}\lambda = 0, Q\mathcal{F} = (P\overline{\lambda})_{(\frac{1}{2},0)}, \qquad \overline{Q}\mathcal{F} = (P\lambda)_{(1,\frac{1}{2})},$$
(38)

along with their complex conjugates. In the second line I have included subscripts to emphasize which Lorentz spin components of the right hand sides appear. In general, $P\lambda \sim (\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, 0) = (1, \frac{1}{2}) \oplus (0, \frac{1}{2})$, but one of these irreducible representations is projected out in the second line as follows from the Lorentz spin content of $Q\mathcal{F}$ and $\overline{Q}\mathcal{F}$. (λ, \mathcal{F}) form the field strength chiral multiplet; but we will call it the *vector multiplet* since that is the name of a related multiplet in which the vector potential lives.

Because of the $Q\mathcal{F} = P\overline{\lambda}$ relation in Eq. 38, the closure of the superalgebra on the vector multiplet is a bit more complicated than in the chiral multiplet case. On the one hand, $(P\overline{\lambda})_{(\frac{1}{2},0)} = 0$ by the equations of motion for a massless Weyl spinor, so on shell for massless vector multiplets this term could just be dropped. On the other hand it will be useful keep this term when deriving selection rules for N = 2 supersymmetric actions in the Sec. 4.

The general kinetic term we can write for the vector multiplets in the IREA is

$$\mathcal{L} = \mathrm{Im}\tau_{IJ}\mathcal{F}^I \cdot \mathcal{F}^J, \tag{39}$$

where τ_{IJ} is a section of an $Sp(2n, \mathbb{Z})$ bundle over \mathcal{M} , and so depends on the vevs of the chiral multiplet scalars ϕ^i (the coordinates of \mathcal{M}). $\mathcal{F}^I \cdot \mathcal{F}^J$ is a shorthand for $\mathcal{F}^I \wedge *\mathcal{F}^J = \mathcal{F}^I_{\mu\nu} \mathcal{F}^{J\mu\nu}$. τ_{IJ} is in fact a holomorphic section. This follows from the fact that $Q\mathcal{L} \supset \tau_{IJ,\overline{i}} \overline{\psi}^{\overline{i}} \mathcal{F}^I \cdot \mathcal{F}^J$, but there is no other term that can cancel this supersymmetry variation, so

$$\tau_{IJ,\overline{\imath}} = 0. \tag{40}$$

In summary, the bosonic part of the general N = 1 supersymmetric IREA is given by the Lagrangian

$$\mathcal{L} = \mathcal{K}_{,i\overline{\jmath}}(\phi,\overline{\phi})\partial_{\mu}\phi^{i}\partial^{\mu}\overline{\phi}^{\overline{\jmath}} - \frac{1}{32\pi} \mathrm{Im}\left[\tau_{IJ}(\phi)\mathcal{F}^{I}_{\mu\nu}\mathcal{F}^{J\mu\nu}\right].$$
(41)

3.1 Nonrenormalization Theorems

The holomorphic nature of τ_{IJ} has strong consequences. ^{11,12} Denote the microscopic (UV) parameters of the theory (*e.g.* masses, strong coupling scales, theta angles, dimensionless couplings) by some scale Λ and a set of dimensionless parameters { λ }. The coefficient functions \mathcal{K} , τ_{IJ} of the IREA are functions of Λ and the λ_k which we wish to determine.

Now, it is a fact that these UV parameters enter into the action of N = 1 UV theory in the same way as the scalar components of chiral superfields do. (Showing this fact involves writing down general forms of asymptotically free N = 1 gauge actions, the subject of many standard texts. ⁴) Thus, it is consistent to *assign* these constants supersymmetry transformation properties as if they were the lowest components of chiral superfields. This is often

expressed by saying that we can think of all the UV parameters as classical background chiral superfields. This implies that whatever strong dynamics takes place upon flowing to the IR, these parameters will only enter the IREA in the way chiral multiplet scalars do. In particular, τ_{IJ} will be a holomorphic function of Λ and the λ —*i.e.* if λ is a coupling, then only λ and not $\overline{\lambda}$ can appear in any quantum corrections to τ_{IJ} , since τ_{IJ} is a function only of chiral superfields and not their complex conjugates.^{13,14}

Let us examine more closely the logic of this argument. We are *assuming* that the effective theory (the IREA) will be described by a nonlinear sigma model of some set of light chiral fields which are not necessarily simply a subset of those of the UV theory. We have no derivation of this hypothesis—we can only test it to see if it gives consistent answers. The couplings of the effective theory will be some functions of the couplings of the microscopic theory, which we would like to solve for. The next step of thinking of the couplings in the superpotential as background chiral superfields 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 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 the familiar "selection rules" of quantum mechanics.

Perhaps an example from quantum mechanics will make this clear: Recall the Stark effect, in which one calculates corrections to the hydrogen atom spectrum in a constant background electric field. Thus we perturb the Hamiltonian by adding a term of the form $\delta H = E_1 x_1 + E_2 x_2 + E_3 x_3$. As the E_i are just some constants, this term explicitly breaks rotational invariance. But 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 **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 electric 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 τ_{IJ} is the same sort of a selection rule, but this time following from supersymmetry. The unfamiliar feature of it is that the UV parameters do not explicitly break the supersymmetry.

We can immediately see the power of this supersymmetry selection rule. For suppose our enlarged theory (thinking of the λ as a chiral superfield) has a U(1) global symmetry under which, say, λ has charge $Q(\lambda) = 1$, *i.e.* in the UV gauge coupling there is a term $\tau_{UV} \supset \lambda \mathcal{O}_{-1}$, where \mathcal{O}_{-1} is some charge -1 operator. Say we are interested in the appearance of a given operator \mathcal{O}_{-10} of charge $Q(\mathcal{O}_{-10}) = -10$ among the quantum corrections. Normally, one would say that this operator can appear only at tenth and higher orders in perturbation theory: $\delta \tau \sim \lambda^{10} \mathcal{O}_{-10} + \lambda^{11} \overline{\lambda} \mathcal{O}_{-10} + \ldots + \lambda^{10} e^{-1/|\lambda|^2} \mathcal{O}_{-10} + \ldots$, (assuming that there is a regular $\lambda \to 0$ limit, so that no negative powers of λ are allowed), where I've also indicated potential non-perturbative contributions as well. However, by the above argument we learn that *only* the tenth-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.* the physics 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 $\overline{\lambda}$ instead.

This argument can be summarized prescriptively as follows:¹⁵ The effective (macroscopic) τ_{IJ} is constrained by

- (1) holomorphy in the (microscopic) coupling constants,
- (2) "ordinary" selection rules from symmetries under which the coupling constants may transform or from electric-magnetic duality, and
- (3) smoothness of the physics in various weak-coupling limits.

Much of the progress in understanding the non-perturbative dynamics of supersymmetric gauge theories of the past half decade years has resulted from the systematic application of the above argument.

The most important application of this argument is to AF gauge theories. Consider an AF gauge theory with complex strong coupling scale

$$\Lambda = \mu_0 e^{-8\pi^2/b_0 g_0^2} e^{i\theta/b_0} \tag{42}$$

so that the one loop complex coupling at a scale μ is

$$\tau = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\mu}\right). \tag{43}$$

The coefficient of the one-loop beta function is given by

$$b_0 = \frac{3}{2}T(adj) - \frac{1}{2}\sum_i T(R_i)$$
(44)

in an N = 1 supersymmetric gauge theory, where the sum is over each chiral multiplet (ϕ^i, ψ^i) transforming in the representation R_i of the gauge group. This follows from Eq. 27 if we recall that the vector multiplet includes one Weyl fermion in the adjoint representation, and each chiral multiplet has a complex scalar and Weyl fermion in the representation R_i .

As in Eq. 30, the effective τ at scale μ will have the functional form

$$\tau(\Lambda, \phi^i; \mu) = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\mu}\right) + f(\Lambda^{b_0}, \phi^i; \mu), \tag{45}$$

where f is now an arbitrary holomorphic function of its arguments. The important point is that τ can only depend holomorphically on Λ and ϕ^i . Since we are dealing with an AF theory, the $\Lambda \to 0$ limit corresponds to the weak coupling limit, in which the effective couplings should not diverge. Thus we have

$$\tau = \frac{b_0}{2\pi i} \log\left(\frac{\Lambda}{\mu}\right) + \sum_{n=1}^{\infty} \Lambda^{b_0 n} a_n(\phi^i;\mu), \tag{46}$$

(*i.e.* inverse powers of Λ^{b_0} do not appear). By comparing this expression to the perturbative expansion, where $\log \Lambda \sim 1/g^2$ (a one loop perturbative contribution) while $\Lambda^{b_0n} \sim e^{-8n\pi^2/g^2}$ (an *n*-instanton contribution), 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.

As in the discussion of Sec. 2.2, these considerations also apply to the $U(1)^n$ couplings τ_{IJ} of the IREA of an AF theory Higgsed with a sufficiently large vev. In the case of a single U(1) and a single chiral multiplet vev ϕ , the IREA coupling has the form

$$\tau = \frac{1}{2\pi i} \log\left(\frac{\Lambda_0^b}{\phi^\alpha}\right) + \sum_{n=1}^\infty a_n \left(\frac{\Lambda_0^b}{\phi^\alpha}\right)^n.$$
(47)

Here ϕ is some gauge invariant combination of the scalar Higgs fields in the UV theory, which can be determined classically; the power α with which it appears is determined by dimensional considerations. Now, as we make a large circle in the ϕ plane, τ undergoes the monodromy

$$\tau \to \tau - \alpha,$$
 (48)

which should be an element of $Sp(2, \mathbf{Z})$, implying that α is an integer. In any given AF gauge theory Higgsed to U(1)'s this can indeed be checked to be the case, and is a reflection of the Witten effect.⁹ The combination of the $Sp(2n, \mathbf{Z})$

and holomorphic properties of τ_{IJ} in some cases is sufficient to determine it exactly.^{12,16}

We should emphasize the main limitation of this "not much renormalization" theorem: it is only derived for weakly 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_i and \mathcal{F} fields may break down. For example, 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, and leads to exact non-perturbative expressions for the Wilsonian beta-function in N = 1 theories.¹⁷ However, for low enough scales, these exact beta functions have singular behavior indicative of the breakdown of the description of the low energy physics in terms of the assumed degrees of freedom. (These exact beta functions can nevertheless be used to show the existence of exactly marginal operators in many interesting cases.¹⁸)

Finally, it is important to note that the statements of this and other non-renormalization theorems only hold in a renormalization scheme which preserves the supersymmetric selection rules.¹⁹ For instance, the scalar field strength renormalizations depend on the UV parameters as well as their complex conjugates since the Kähler potential does. So, if one worked in a scheme in which one insisted on the canonical normalization of the scalar kinetic terms, one would have to rescale the holomorphic τ_{IJ} couplings by the nonholomorphic field strength renormalizations, thus invalidating the supersymmetric selection rule.

3.2 N = 1 Supersymmetric Effective Action Potential Terms

So far we have dealt only with the effective action in the far IR limit which only massless neutral scalars and U(1) gauge bosons survive. This begs the question of whether a non-trivial moduli space exists for a given theory. To answer this question we need to examine possible N = 1 supersymmetric potential terms for both neutral and charged scalar fields. We can not do this in as easy and direct manner as we did for the kinetic terms since inclusion of potential terms necessarily takes us off-shell—by definition the potentials vanish on constant scalar field configurations satisfying the equations of motion. (The supersymmetric constraints on possible scalar potential terms can, of course, be deduced from a direct but fairly technical construction of supersymmetric actions.⁴)

We will deduce the supersymmetric form of the potential terms by the

following indirect argument. Denote the (Kähler) manifold of scalar vevs in the IREA of Eq. 41 by $\mathcal{M}_0 = \{\phi^i\}$. Now think of this action as an effective action at some finite scale E and imagine turning on some relevant operators (a potential) at a scale much less than E. The resulting IREA will again be of the form of Eq. 41, though the set of fields will be smaller (*i.e.* just those minimizing the potential). In particular the moduli space \mathcal{M} of the new IREA must also be a Kähler manifold. Since this construction takes place at arbitrarily weak coupling (since the IREA is IR free), the same set of low energy degrees of freedom can be used to describe the effective action at all scales below E, and so \mathcal{M} must be some Kähler submanifold of \mathcal{M}_0 . It now only remains to write the most general potential whose minimization picks out such a submanifold and also preserves the invariance of the effective action holomorphic reparameterizations (field redefinitions).

One way of singling out a Kähler submanifold $\mathcal{M} \subset \mathcal{M}_0$ is by specifying a set of holomorphic conditions $\{F_i(\phi^j) = 0\}$. Then it is straightforward to check that $\mathcal{M} = \mathcal{M}_0/\{F_i = 0\}$ is not only a complex submanifold but also is necessarily istelf Kähler with Kähler potential simply the restriction of the Kähler potential on \mathcal{M}_0 . The potential giving rise to these complex conditions must itself be real, though, suggesting it must be of the form $V = \sum_i F_i \overline{F_i}$. This, however, is not a reparametrization invariant formula. The correct formula must use the Kähler metric on \mathcal{M}_0 to contract the indices of the F_i , implying an F term potential

$$V_F = F_i g^{i\overline{j}} \overline{F}_{\overline{j}},\tag{49}$$

where $g^{i\overline{j}}$ is the inverse of $g_{i\overline{j}}$. Postive definiteness of g (from unitarity of the effective action) implies that V_F takes its minimum value of 0 when the F terms vanish individually:

$$F_i = 0. (50)$$

Note that $V_F = 0$ is also the condition for supersymmetry to be unbroken in the vacuum, since if not the supersymmetry algebra Eq. 33 implies that $\langle \{Q, \overline{Q}\} \rangle \neq 0$ and therefore for some component of Q we would have $Q|0\rangle \neq 0$.

There is a further constraint (besides holomorphicity) on the F_i . Claiming that V_F is reparametrization invariant assumes that F_i transforms as a vector under reparametrizations. The only way (without some extra structure on \mathcal{M}_0) to form such an object out of functions of the scalar vevs ϕ^i is as the derivative of a holomorphic function on \mathcal{M}_0 :

$$F_i = \mathcal{W}_{,i}.\tag{51}$$

 \mathcal{W} is an arbitrary (gauge invariant) holomorphic function on \mathcal{M}_0 called the *superpotential*.

A general \mathcal{W} gives rise to an independent condition $F_i = 0$ for each complex coordinate direction *i*, and thus generically one expects \mathcal{M} to be a single point. It would seem that we are no closer to getting a non-trivial moduli space of vacua with N = 1 supersymmetry than we were without it. But the holomorphicity of \mathcal{W} gives rise to non-renormalization theorems (in the same way that the holomorphicity of τ_{IJ} does) which allow one (in favorable cases) to specify UV couplings which lead to special low energy superpotentials which admit non-trivial moduli spaces.²⁰ Another way of putting it is that even though non-trivial moduli spaces of inequivalent vacua are still "accidental" in N = 1 supersymmetric theories, our knowledge of their RG flows allow us to arrange the necessary accident.

When \mathcal{M}_0 has isometries, there is an additional structure that one can use to construct potential terms. Suppose \mathcal{M}_0 has a $U(1)^n$ group of isometries generated by $\phi^i \to \phi^i + \xi_I^i$ for $I = 1, \ldots, n$. Then we can gauge those isometries with the low energy $U(1)^n$ gauge group; equivalently, and perhaps more descriptively, we turn on an electric charge under the $U(1)_I$ gauge group for the scalars which are shifted under the action of the $U(1)_I$ isometry. (Recall the relation between isometries and low energy gauge groups mentioned in Sec. 2 above.) For example, if the U(1) isometries are realized linearly, then $\xi_I^i = q_{Ii}\phi^i$ (no sum on *i*) where q_{Ii} is the charge of the complex scalar ϕ^i under the $U(1)_I$ gauge group. Appropriately minimally coupling the charged scalars leads to an effective action as in Eq. 5. Letting the charged ϕ^i 's get vevs induces a potential (by the Higgs mechanism) of the form

$$V_D = D_I (\mathrm{Im}\tau)^{IJ} D_J, \tag{52}$$

where $(\text{Im}\tau)^{IJ}$ is the matrix inverse of $\text{Im}\tau_{IJ}$ and the *D* terms are given by

$$D_I = \operatorname{Re}(\xi_I^i \mathcal{K}_{,i}) + \xi_I^0, \tag{53}$$

where ξ_I^0 are real constants called *Fayet-Illiopoulos* terms. Since $\text{Im}\tau_{IJ}$ is postive definite by unitarity of the effective action, V_D takes its minimum value $V_D = 0$ when the D terms vanish individually:

$$D_I = 0, \tag{54}$$

which is also a condition for supersymmetry to be unbroken. From the expression for the D term it follows that when all the fields are neutral any non-zero Fayet-Illiopoulos term spontaneously breaks supersymmetry. Henceforth we ignore Fayet-Illiopoulos terms: they can be shown to obey a stringent non-renormalization theorem which prevents them from being generated in an effective action if they are not generated at one loop in perturbation theory.

One may worry that since the D-term equations are not holomorphic that upon solving them one finds a moduli space $\mathcal{M}' = \mathcal{M}_0/\{D_I = 0\}$ which is not Kähler, in contradiction to our supersymmetric selection rule for the IREA. Actually, since some scalars are charged, to find the moduli space we must also divide by the action of the gauge group: $\mathcal{M} = \mathcal{M}'/U(1)^n$. It turns out that this process always leads to a Kähler \mathcal{M} , and is known²¹ as a Kähler quotient construction, and can be described in a holomorphic way as division of \mathcal{M}_0 by the natural action of the complexified gauge group: $\mathcal{M} = \mathcal{M}_0/U(1)^n_{\mathbf{C}}$. The end result is that the D terms always have a solution which is the Kähler submanifold parametrized by the holomorphic gauge neutral combinations of scalars.

4 N = 2 Supersymmetric IREAs

The basic (no central charges) N = 2 superalgebra is, in the indexless notation of the last section,

$$\{Q_m, \overline{Q}_n\} = \delta_{mn} P, \qquad \{Q_m, Q_n\} = 0, \qquad m, n = 1, 2.$$
 (55)

This is just two copies of the N = 1 algebra, Eq. 33; in particular, it has two N = 1 subalgebras generated by Q_1 and Q_2 . Note that the N = 2algebra has an $SU(2)_R$ group of automorphisms under which Q_m transforms as a doublet. (Global symmetries under which the supercharges transform are called *R* symmetries.)

On shell irreducible representations of Eq. 55 are easy to construct. Because of the Q_1 subalgebra, any N = 2 representation must be made up of N = 1 representations. Suppose that one of these N = 1 representations is a chiral multiplet, as in Eq. 34. Under the action of the Q_2 generators this multiplet must also form an N = 1 representation; but it is easy to see that there is no way to do this consistent with the N = 2 algebra. If we replace the initial N = 1 chiral multiplet by an N = 1 vector multiplet, one also finds no solution. So N = 2 representations are formed by combining at least two N = 1 multiplets. It is not hard to go through the various possibilities to find that the two solutions are the *hypermultiplet*, made from two N = 1 chiral multiplets (ϕ, ψ) and $(\tilde{\phi}, \tilde{\psi})$ which satisfy

$$\begin{array}{lll}
Q_n\phi_m &= \epsilon_{nm}\psi, & \overline{Q}_n\phi_m &= \delta_{nm}\widetilde{\psi}, \\
Q_n\psi &= 0, & \overline{Q}_n\psi &= \epsilon_{nm}P\phi_m, \\
Q_n\overline{\psi} &= \delta_{nm}P\phi_m, & \overline{Q}_n\overline{\psi} &= 0,
\end{array}$$
(56)

where I have defined $\phi_n = (\overline{\phi}, \phi)$; and the vector multiplet, made from one N = 1 chiral multiplet (ϕ, ψ) and one N = 1 vector multiplet (λ, \mathcal{F}) which satisfy

$$\begin{array}{lll}
Q_n\phi &=\lambda_n, & \overline{Q}_n\phi &=0, \\
Q_n\lambda_m &=\epsilon_{nm}\mathcal{F}, & \overline{Q}_n\lambda_m &=\delta_{nm}P\phi, \\
Q_n\mathcal{F} &=\epsilon_{nm}P\overline{\lambda}_m, & \overline{Q}_n\mathcal{F} &=-\epsilon_{nm}P\lambda_m,
\end{array}$$
(57)

where I have defined $\lambda_n = (\psi, \lambda)$.

Important distinguishing features of the hypermultiplet are that its scalars form a complex $SU(2)_R$ doublet and that this $SU(2)_R$ action mixes an N = 1chiral multiplet scalar with an *anti*-chiral multiplet partner. This has the immediate consequence that when coupling hypermultiplets in N = 2 gauge theories, they always transform in a real representation $R_i \oplus \overline{R_i}$ (where R_i is the representation of the N = 1 chiral multiplet and so $\overline{R_i}$ is the representation of its anti-chiral partner). The bosonic degrees of freedom in a vector multiplet, by contrast, are a single complex scalar and a (real) vector field, both transforming in the adjoint of the gauge group, and both singlets under $SU(2)_R$. In particular, in the case of $U(1)^n$ gauge group, which we are interested in for describing IREAs, the vector multiplet scalars are necessarily neutral.

In an N = 2 IREA with gauge group $U(1)^n$ and neutral hypermultiplets, the general action (following from, say, the $Q_1 N = 1$ supersymmetry) would be just as in Eq. 41, where the i, j indices run over all the complex bosons (whether in hyper or vector multiplets). To take into account the N = 2structure, let us now reserve the i, j indices for the complex (doublet) scalars ϕ_n^i of hypermultiplets, and label the complex (singlet) scalars of the vector multiplets ϕ^I .

The first N = 2 selection rule we wish to derive is that no $\partial_{\mu} \phi^{I} \partial^{\mu} \overline{\phi^{i}}$ kinetic terms can occur. To see this, suppose there was a term $\mathcal{L} \sim \mathcal{K}_{,I\overline{i}} P \phi^{I} \cdot P \overline{\phi}_{n}^{i}$. Then $Q_m \mathcal{L} \supset \mathcal{K}_{,I\overline{i}} P \phi^{I} \cdot P \widetilde{\psi} \delta_{nm}$. But it is easy to see, refering to Eq. 57, that there is no term whose Q_m variation can could cancel this, implying that we must have $\mathcal{K}_{,I\overline{i}} = 0$. This in turn implies that the Kähler potential splits into the sum of two pieces depending on the hypermultiplet vevs and the vector multiplet vevs separately:

$$\mathcal{K} = \mathcal{K}_H(\phi_n^i, \overline{\phi}_n^i) + \mathcal{K}_V(\phi^I, \overline{\phi}^I).$$
(58)

Thus the kinetic terms for the scalars also split as

$$\mathcal{L} = g_{i\overline{j}}(\phi^k) \,\partial\phi^i \cdot \partial\overline{\phi}^{\overline{j}} + g_{I\overline{J}}(\phi^K) \,\partial\phi^I \cdot \partial\overline{\phi}^{\overline{J}},\tag{59}$$



Figure 2: Cartoon of a classical N = 2 moduli space. The Higgs and mixed branches intersect along a Higgs submanifold A, while the mixed branch intersects the Coulomb branch along a Coulomb submanifold B.

implying that the moduli space has a natural (local) product structure

$$\mathcal{M} = \mathcal{M}_H \times \mathcal{M}_V; \tag{60}$$

 \mathcal{M}_H is the subspace of \mathcal{M} along which only the hypermultiplet vevs vary while the vector multiplet vevs remain fixed, and *vice versa* for \mathcal{M}_V . In cases where \mathcal{M}_V is trivial (a point), $\mathcal{M} = \mathcal{M}_H$ is called a *Higgs branch* of the moduli space; when \mathcal{M}_H is trivial \mathcal{M}_V is called the *Coulomb branch* (since there are always the massless U(1) vector bosons from the vector multiplets). Cases where both \mathcal{M}_H and \mathcal{M}_V are non-trivial are called *mixed branches*.

In general the total moduli space of a given theory need not be a smooth manifold—it may have "jumps" where submanifolds of different dimensions meet. Classically this occurs as a result of the Higgs mechanism: a charged scalar vev Higgses some vector multiplets, typically lifting them (making them massive). But at the special point where the charged vev is zero, the vector multiplets become massless, leading to extra flat directions and a jump in the dimensionality of the moduli space. Hence, at least classically, the general picture of an N = 2 moduli space is a collection of intersecting manifolds, which can be Higgs, Coulomb, or mixed branches,^{22,23} see Fig. 2.

This classical picture is, of course, modified quantum mechanically. However, a further N = 2 supersymmetric selection rule relating the metric on \mathcal{M}_V to the generalized coupling τ_{IJ} greatly restricts the possible form of these modifications. To see this, consider the $U(1)^n$ kinetic term $\mathcal{L} \sim \tau_{IJ} \mathcal{F}^I \mathcal{F}^J$. Then $Q\mathcal{L} \supset \tau_{IJ} \mathcal{F}^I P \overline{\lambda}^J$. To cancel this variation then requires a fermion kinetic term $\mathcal{L}' \sim \tau_{IJ} \lambda^I P \overline{\lambda}^J$. Then $\overline{Q} \mathcal{L}' \supset \tau_{IJ} P \phi^I P \overline{\lambda}^J$. Finally, to cancel this variation requires a scalar kinetic term $\mathcal{L}'' \sim \tau_{IJ} P \phi^I P \phi^J$. Adding also the complex conjugate terms then implies

$$g_{I\overline{J}} = \mathrm{Im}\tau_{IJ}.\tag{61}$$

In particular, since $g_{I\overline{I}}$ is a function only of the ϕ^{I} , so must τ_{IJ} be.

Now, consider an N = 2 AF gauge theory with dynamically generated scale Λ ($\tau \sim \log \Lambda$). Since Λ appears in τ_{IJ} (at, say, one loop), it appears in the Lagrangian in the same way a scalar vev ϕ^I of an N = 2 vector multiplet would. Therefore, we can think of $\log \Lambda$ as a background U(1) vector superfield. Since the metric on the Higgs branch is independent of vector superfields, it is independent of Λ . Finally, we can use the fact that the classical theory is obtained in the limit $\Lambda \to 0$ to conclude that the Higgs metric is given exactly by the classical answer.²² Note also that any masses for hypermultiplets also enter into the one loop running of the gauge coupling, and so can be promoted to background vector superfields. (These vector superfields correspond to the gauging of global flavor symmetries.) We immediately learn that the metric on the Higgs branch is independent of the masses.

We thus learn that only the Coulomb branch can receive quantum corrections, and that the mixed non-baryonic branch will retain its classical product structure of a hypermultiplet manifold times the vector multiplet manifold corresponding to the subspace of the Coulomb branch along which the nonbaryonic and Coulomb branches intersect; see Fig. 3. A key fact about the Coulomb branch is that though it can be corrected quantum mechanically, it is never wholly lifted in AF N = 2 gauge theories. This is because there is a Coulomb branch for large adjoint scalar vevs where the AF gauge theory is Higgsed to $U(1)^n$ at arbitrarily weak coupling. Quantum corrections in the resulting N = 2 IREA cannot lift these flat directions since the only way (at weak coupling) to give mass to the U(1) photons in the vector multiplets is by the Higgs mechanism; but there are no charged scalars in the vector multiplet. Thus, unlike N = 1 supersymmetric gauge theories, N = 2 supersymmetric gauge theories are always guaranteed to have a non-trivial moduli space of physically inequivalent vacua.

Since the hypermultiplet manifolds can be determined classically in N = 2 supersymmetric gauge theories, we will not consider them further. But one should note that they are constrained by N = 2 supersymmetry to be *hyperKähler* manifolds. This is essentially a manifold which is simultaneously Kähler with respect to three different complex structures. These complex structures transform as a triplet under the $SU(2)_R$ symmetry. Hypermultiplet



Figure 3: Map of Cartoon of a quantum N = 2 moduli space. The Higgs branch and the Higgs (hypermultiplet vev) directions of the mixed branch remain unmodified from their classical geometries, though they may be deformed in the Coulomb (vector multiplet vev) directions. The Coulomb branch is generally different from its classical geometry.

branches can be lifted by the Higgs mechanism, which corresponds geometrically to a *hyperKähler quotient* construction.²¹

The geometry of the vector multiplet manifolds are similiarly constrained by N = 2 supersymmetry. Consider the vector multiplet kinetic term $\mathcal{L} \sim \tau_{IJ} \mathcal{F}^I \mathcal{F}^J$. Then $Q_n \mathcal{L} \supset \tau_{IJ,K} \mathcal{F}^I \mathcal{F}^J \lambda_n^K$. The only term that could cancel this variation is $\mathcal{L}' \sim \tau_{IJ,K} \mathcal{F}^I \lambda_\ell^J \lambda_m^K \epsilon_{\ell m}$. For this term to be a scalar, the two λ 's have to combine to form a (1,0) Lorentz representation, that is, symmetrically on their spinor indices. But since they are antisymmetrized on their $SU(2)_R$ indices and are anticommuting fields, they must therefore be symmetric under interchange of J and K. Thus such a term could only cancel the symmetric part proportional to $\tau_{I(J,K)}$ of the supersymmetry variation of \mathcal{L} , and so N = 2supersymmetry can only be preserved if

$$\tau_{IJ,K} = \tau_{IK,J}.\tag{62}$$

This integrability condition together with Eq. 61 and the $Sp(2n, \mathbb{Z})$ transformation properties of τ_{IJ} can be taken as the definition of a *rigid special Kähler* manifold.²⁴

This rigid special Kähler structure together with certain important physical assumptions^{11,25} have been used to solve for the exact Coulomb branch geometry of many AF N = 2 gauge theories. An exposition of these constructions is the subject of another review.²⁶

Acknowledgments

It is a pleasure to thank the organizers and students of the Trieste Spring School for a stimulating visit. This work is supported in part by NSF grant PHY-9513717 and an A.P. Sloan fellowship

References

- For an introduction to the N = 1 and 2 supersymmetric results, see e.g. K. Intriligator, N. Seiberg, Nucl. Phys. Suppl. BC 45, 1 (1996), hepth/9509066; M. Peskin, hep-th/9702094.
- 2. See *e.g.* S. Minwalla, hep-th/9712074.
- See e.g. P. Argyres, R. Plesser, N. Seiberg, E. Witten, Nucl. Phys. B 461, 71 (1996), hep-th/9511154.
- See e.g. J. Wess, J. Bagger, Supersymmetry and Supergravity (2nd edition, Princeton University Press, Princeton NJ, 1992); P. West, Introduction to Supersymmetry and Supergravity (2nd edition, World Scientific, Singapore, 1990); M. Sohnius, Phys. Rep. 128, 39 (1985).
- See e.g. N. Seiberg in Proceedings of the Trieste Spring School (1997), hep-th/9705117.
- 6. See, however, J. Polchinski, Nucl. Phys. B 303, 226 (1988).
- 7. S. Coleman, D. Gross, *Phys. Rev. Lett.* **31**, 851 (1973).
- P. Dirac, Proc. R. Soc. A 133, 60 (1931); J. Schwinger, Phys. Rev. 144, 1087 (1966), 173, 1536 (1968); D. Zwanziger, Phys. Rev. 176, 1480,1489 (1968).
- 9. E. Witten, *Phys. Lett.* B 86, 283 (1979).
- 10. A. Shapere, F. Wilczek, Nucl. Phys. B **320**, 669 (1989).
- 11. N. Seiberg, E. Witten, Nucl. Phys. B 426, 19 (1994), hep-th/9407087.
- In the N = 1 context, K. Intriligator, N. Seiberg, Nucl. Phys. B 431, 551 (1994), hep-th/9408155.
- 13. N. Seiberg, *Phys. Lett.* B **206**, 75 (1988).
- 14. N. Seiberg, *Phys. Lett.* B **318**, 469 (1993), hep-ph/9309335.
- 15. N. Seiberg, in *PASCOS 1994*, hep-th/9408013.
- S. Elitzur, A. Forge, A. Giveon, E. Rabinovici, *Phys. Lett.* B **353**, 79 (1995), hep-th/9504080; *Nucl. Phys.* B **459**, 160 (1995), hepth/9509130; S. Elitzur, A. Forge, A. Giveon, K. Intriligator, E. Rabinovici, *Phys. Lett.* B **379**, 121 (1996), hep-th/9603051.
- V. Novikov, M. Shifman, A. Vainshtein, V. Zakharov, *Phys. Lett.* B 166, 329 (1986); M. Shifman, A. Vainshtein, V. Zakharov, *Phys. Lett.* B 166, 334 (1986); M. Shifman, A. Vainshtein, *Nucl. Phys.* B 277, 456

(1986).

- 18. R. Leigh, M. Strassler, Nucl. Phys. B 447, 95 (1995), hep-th/9503121.
- W. Siegel, *Phys. Lett.* B 84, 193 (1979); I. Jack, D. Jones, K. Roberts, *Z. Phys.* C 63, 151 (1994), hep-ph/9401349.
- For the original and classic examples, see N. Seiberg, *Phys. Rev.* D 49, 6857 (1994), hep-th/9402044; *Nucl. Phys.* B 435, 129 (1995), hep-th/9411149.
- For a brief description of various quotient constructions and further references, see for example I. Antoniadis, B. Pioline, *Int. J. Mod. Phys.* A 12, 4907 (1997), hep-th/9607058.
- P. Argyres, R. Plesser, N. Seiberg, Nucl. Phys. B 471, 159 (1996), hep-th/9603042.
- P. Argyres, R. Plesser, A. Shapere, Nucl. Phys. B 483, 172 (1997), hep-th/9608129.
- For a discussion of various the various definitions of rigid special Kähler geometry with references, see for example B. Craps, F. Roose, W. Troost, A. Van Proeyen, *Nucl. Phys.* B 503, 565 (1997), hep-th/9703082; D. Freed, hep-th/9712042.
- 25. N. Seiberg, E. Witten, Nucl. Phys. B 431, 484 (1994), hep-th/9408099.
- P. Argyres, to appear in Proceedings of the Feza Gürsey Institute Summer School, (Istanbul, 1998).