SUPERSYMMETRY AND NON-PERTURBATIVE BETA FUNCTIONS

Nathan SEIBERG [1]
Institute for Advanced Study, Princeton, NJ 08540, USA

Received 25 February 1988

We discuss $N=2$ supersymmetric gauge theories in four dimensions. Several non-renormalization theorems are derived. In particular, we give a simple proof that the perturbative beta function is given by the one-loop result and we clarify the use of the "multiplet of anomalies" in this theory. This non-renormalization theorem is violated by non-perturbative effects – instantons contribute to the beta function for the gauge coupling, g, and to the beta function for the vacuum angle, Θ. We also derive a non-renormalization theorem to all perturbative orders around the instanton. $N=4$ supersymmetric gauge theories are shown to be *exactly* finite.

Supersymmetric theories are known to respect certain perturbative non-renormalization theorems (for reviews about supersymmetry and for lists of references see refs. [1,2]. In $N=1$ theories the superpotential is not renormalized in perturbation theory [3], the beta function in $N=2$ theories is not modified beyond one-loop order [1,3] and the $N=4$ theory is finite [4]. These non-renormalization theorems are only perturbative. It is known that the first one is violated by instanton effects. This was shown using two different points of view. One approach is that of ref. [5]. The other approach was started in ref. [6] and is summarized in ref. [7] which also gives an extensive list of references. Here we will use the point of view and the techniques of ref. [5] to study in detail the dynamics of the $N=2$ and the $N=4$ theories. We will give a new proof that the one-loop beta function of the $N=2$ theory is not modified in perturbation theory. However, instantons generate a beta function both for the coupling constant, g, and for the theta parameter, Θ:

$$\beta_g(g,\Theta)=Ag^3+B\cos\Theta\, g^3\exp(-8\pi^2/g^2) +\mathrm{O}(\exp(-16\pi^2/g^2)), \tag{1}$$

$$\beta_\Theta(g,\Theta)=C\sin\Theta\exp(-8\pi^2/g^2) +\mathrm{O}(\exp(-16\pi^2/g^2)), \tag{2}$$

where A, B and C are constants. We will also show that the same phenomenon does not happen in the $N=4$ theory. Furthermore, assuming that this theory exists we prove that it is *exactly finite*. It should be mentioned that our renewed interest in the subject was motivated by the recent exciting work of Witten [8] where a twisted version of $N=2$ supersymmetry was shown to have remarkable properties. Hopefully, our analysis will be useful in future explorations of Witten's theory.

The main physical elements in the analysis of ref. [5] are the following:

(1) These theories have flat directions where the gauge symmetry is broken. In some cases the gauge symmetry is broken with no unbroken non-abelian subgroup and therefore reliable instanton calculations can be performed. In these cases there are no infrared divergences and the approximations are controlled by a small parameter.

(2) The symmetries of the problem help us determine the form of the effective action. The analyticity of the superpotential is extremely powerful in uniquely determining its form.

(3) "Miraculous" relations between the classical action of the instanton, the one-loop beta function

[1] On leave of absence from the Department of Physics, Weizmann Institute of Science, Rehovot 76100, Israel.

0370-2693/88/$ 03.50 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

and the charges of the non-anomalous symmetries lead to a simple and consistent picture.

Here we will use the same elements in the study of the $N=2$ theory. Our analysis is best presented in the superspace formalism of ref. [9]. (We will later show how all our results can be obtained using the more standard $N=1$ superspace.) In this formalism the leading terms in the lagrangian are given in terms of an *analytic function*, $\mathscr{F}$. This is a function of the $N=2$ gauge superfields. Ψ^a, where a is an index labelling the adjoint representation of the gauge group. This superfield is a chiral superfield which is subject to a constraint [9]. In terms of $N=1$ superfields it includes a vector multiplet, V^a, and a chiral multiplet, Φ^a. In components (after eliminating all the auxiliary fields) we find the following fields: a gauge field, A^a_μ, two fermions, ψ^{ai}, $i=1, 2$ ($\psi^{a1}=\lambda^a$ is the gluino in V^a and $\psi^{a2}=\psi^a$ is the fermion in Φ^a) and a complex scalar, ϕ^a. The $N=2$ lagrangian is given by

$$\mathscr{L}=\int \mathrm{d}^4\theta\, \mathscr{F}(\Psi^a)+\text{c.c.} \tag{3}$$

$\mathscr{F}$ has to be analytic because it is integrated only over half of the $N=2$ superspace. In $N=1$ language this becomes

$$\mathscr{L}=\int \mathrm{d}^2\theta\, \mathrm{d}^2\bar\theta\, \tfrac{1}{2}K(\Phi^a, \Phi^{\dagger a}, V^a) + \int \mathrm{d}^2\theta f_{ab}(\Phi^a)\, W^aW^b+\text{c.c.}, \tag{4}$$

where we use standard notation. The Kähler potential, K, and the analytic function, f_{ab}, in this equation are given by

$$K=2\partial_a\mathscr{F}(\mathrm{e}^V)_{ab}\Phi^{\dagger b}+\text{c.c.},\quad f_{ab}=\tfrac{1}{2}\partial_a\partial_b\mathscr{F}. \tag{5}$$

It is now clear that the theory has flat directions and a potential for the scalars ϕ^a can arise only from the D-terms. Furthermore, a superpotential for the chiral superfield Φ^a cannot be generated perturbatively or non-perturbatively because it is not compatible with $N=2$ supersymmetry.

The minimal theory is given by

$$\mathscr{F}=(1/8g^2)\,(\Psi^a)^2. \tag{6}$$

It is easy to see that it leads to the standard renormalizable $N=2$ theory. The classical global symmetry of this theory is $SU(2)\times U(1)$. The $SU(2)$ symmetry rotates the two θ's and the $U(1)$ factor multiplies them by a phase. In components, the $SU(2)$ symmetry rotates only the two fermions. Normalizing the $U(1)$ charge of θ to one, the charge of the superfields Ψ^a is two. This $U(1)$ symmetry is anomalous:

$$\partial_\mu j^\mu=(g^2N_c/8\pi^2)\,F\tilde F \tag{7}$$

[we have taken the gauge group to be $SU(N_c)$]. Therefore, non-perturbative effects explicitly break this $U(1)$ symmetry down to a Z_{4N_c}.

In what follows we first limit ourselves to the case where the gauge symmetry is $SU(2)$. In this case, it is easy to see that the expectation value of ϕ^a in the flat directions breaks the global $U(1)$ symmetry and the gauge symmetry is broken down to $U(1)$. The low-energy theory (after the $N=2$ Higgs mechanism) includes one $N=2$ multiplet, Ψ which includes the gauge boson for the unbroken gauge symmetry. It is related to the gauge invariant superfield $\Psi^2=(\Psi^a)^2$. This low-energy theory includes no non-abelian gauge interactions and therefore, it exhibits smooth behavior in the infrared. Therefore, there is no problem in constructing the effective action for this superfield. Furthermore, the semi-classical analysis which we will perform is free of infrared singularities and is controlled by a small parameter $g(\langle(\phi^a)^2\rangle)$ which can be made arbitrarily small.

The effective action for this supermultiplet has to be $N=2$ supersymmetric. The leading terms are given by

$$\mathscr{L}_{\text{eff}}=\int \mathrm{d}^4\theta\, \mathscr{F}_{\text{eff}}+\text{c.c.} \tag{8}$$

We first analyze $\mathscr{F}_{\text{eff}}$ in perturbation theory. Invariance under the $U(1)$ symmetry restricts its form to

$$\mathscr{F}^{\text{pert.}}_{\text{eff}}=(1/8g^2)\,\Psi^2[A_1+A_2\log(\Psi^2/\Lambda^2)], \tag{9}$$

where A_1 and A_2 are two constants which may depend on g and Λ is the scale at which g is defined. Λ can be redefined by changing the value of the constant A_1. We can pick Λ such that $A_1=1$. Although the effective action is $U(1)$ invariant, for non-zero A_2 the effective lagrangian is not. Under a $U(1)$ transformation with a phase α it is shifted by

$$\delta\mathscr{L}^{\text{pert.}}_{\text{eff}}=\alpha(A_2/g^2)\,F\tilde F. \tag{10}$$

Using eq. (7) we determine $A_2=g^2/4\pi^2$. The second term in eq. (9) is therefore a one-loop effect. Writing the effective action in components, we find

$$f_{\text{eff}}^{\text{pert.}}=(1/8g^2)\,[1+3g^2/4\pi^2 + (g^2/4\pi^2)\ln(\phi^2/\Lambda^2)]\,, \tag{11}$$

and therefore the effective gauge coupling is

$$1/g_{\text{eff}}^2(\langle\phi^2\rangle)=(1/g^2)\,[1+3g^2/4\pi^2 + (g^2/4\pi^2)\ln(\langle\phi^2\rangle/\Lambda^2)] + \text{non-perturbative effects.} \tag{12}$$

From this we can obtain the beta function of the theory:

$$\beta(g)=-(1/4\pi^2)\,g^3 + \text{non-perturbative effects.} \tag{13}$$

We recover the known one-loop result and the fact that the perturbative beta function is not modified by higher order corrections.

An important point must be stressed here. We could try to use a similar argument in any $N=1$ theory and thus restrict the form of the coefficient of W_α^2 which is always an analytic function. This would seem to lead to the same conclusion, namely that the perturbative beta function is given by its one-loop approximation. This fact is related to the well-studied subject of the so-called "multiplet of anomalies" [10]. The point is that the U(1) current is in the same supersymmetry multiplet with the conformal current and therefore, the anomaly in the U(1) current is related to the beta function. Since the anomaly satisfies the Adler–Bardeen theorem, one may be tempted to argue that the beta function is also only a one-loop effect. This argument is known [11] to be wrong. At the quantum level the current which satisfies the Adler–Bardeen theorem is not the supersymmetric partner of the conformal current. In the effective action approach the point is that the field with simple transformation laws under the U(1) symmetry is a complicated function of the field which is related in a simple way to the gauge coupling. More explicitly, the function f in front of W_α^2 can have corrections of the form $[\ln(\Phi^2/\Lambda^2)]^n$ for any n where Φ is a generic chiral superfield. Clearly, the field Φ cannot have simple U(1) transformation laws. However, using a holomorphic (and therefore supersymmetric) field redefinition, we can find a chiral superfield $\tilde{\Phi}$ which transforms simply under the U(1) symmetry and $f=C_1(g)+C_2\ln(\tilde{\Phi}^2/\Lambda^2)$ where $C_1(g)$ depends on the coupling g and C_2 is a constant. The effective coupling constant depends simply on $\tilde{\Phi}$. This does not mean that the beta function is given by the one-loop approximation in the original scheme because the field redefinition is singular at $\Phi=0$.

In view of this discussion one might wonder if our analysis above of the $N=2$ theory is sensible. We assumed that the field Ψ has simple U(1) transformations and we did not allow for the possibility of a field redefinition. However, it is easy to understand why this assumption is correct. Thinking of the theory as an $N=1$ theory we could redefine the chiral superfields Φ in an $N=1$ invariant way. Such a redefinition necessarily breaks $N=2$ if the superfield V is not redefined as well. An $N=1$ supersymmetric redefinition of V is not gauge invariant and therefore it is impossible. The same conclusion can be obtained in $N=2$ superspace. Although a holomorphic redefinition of Ψ preserves the fact that it is a chiral superfield, it is not consistent with the constraint Ψ has to satisfy. We see that in $N=2$ we cannot perform field redefinitions and therefore Ψ must have simple U(1) transformation laws at the quantum level and our analysis above is correct. $N=2$ supersymmetry makes arguments based on the "multiplet of anomalies" simple and reliable.

After finding the perturbative effective action, we move on to discuss the non-perturbative effects. Eq. (9) was derived by imposing invariance under the U(1) symmetry. Non-perturbatively, this symmetry is no longer valid and the effective action should no longer respect it – it should only be invariant under the discrete Z_8 symmetry. At the level of one instanton the U(1) symmetry should be violated by eight units. This fact determines the contribution of one instanton to the function $\mathcal{F}$ to be of the form

$$\mathcal{F}^{\text{instanton}}=a_1\exp(-8\pi^2/g^2)\,\Lambda^4/\Psi^2. \tag{14}$$

We will later show that because of the renormalization group the coefficient a_1 cannot depend on g. By symmetry considerations alone we cannot rule out the possibility that an anti-instanton generates $\mathcal{F}^{\text{anti-instanton}}\sim\Psi^6/\Lambda^4$. However, unlike eq. (14), this term leads to correlation functions which blow up at weak coupling (large $\langle(\phi^a)^2\rangle$) and therefore it cannot be generated. Clearly, if eq. (14) is generated by an instanton, an anti-instanton generates the complex conjugate of this equation.

In order to show that this term is indeed generated and a_1 is non-zero we can perform a detailed instanton calculation. Since it is very similar to the calculations in ref. [5], we will only point out the main features. In the absence of the scalar field expectation value (at the origin in field space) the instanton has eight fermionic zero modes. In terms of components, four of them are gluino zero modes and the other four are ψ zero modes. More precisely, in the conventions we use there are four $\lambda^\dagger$ and four $\psi^\dagger$ zero modes. Four out of the eight zero modes are related to the $N=2$ supersymmetry (there are only four zero modes because four of the supersymmetry generators annihilate the instanton and do not generate zero modes) and the other four are related to the $N=2$ superconformal symmetry of the classical theory. When the scalar expectation value is non-zero the instanton is no longer a solution of the equations of motion. The subtleties associated with this fact were discussed in detail in ref. [5]. The scalar field expectation value has two effects: it makes the instanton calculation infrared finite and it reduces the number of fermionic zero modes. An explicit calculation shows that the four superconformal zero modes are lifted. The wave functions of the two $\lambda^\dagger$ supersymmetry zero modes mix with ψ and the wave functions for the two $\psi^\dagger$ supersymmetry zero modes mix with λ. Notice that this is consistent with the symmetries because the scalar expectation value breaks the discrete Z_8 symmetry to Z_2. Despite this mixing, there are still four zero modes. The zero mode wave functions of ψ and λ behave as $1/x^3$ for large x (x is the distance from the instanton). Therefore, the instanton generates an effective interaction of the form

$$\lambda^2\psi^2. \tag{15}$$

The coefficient of this interaction is proportional to the typical factor of instanton calculations $\exp[-8\pi^2/g^2(\langle\phi^2\rangle)]$ and to $1/\langle\phi^2\rangle$ on dimensional grounds. Collecting these factors, we find that the effective interaction is proportional to

$$\begin{aligned}&(1/\langle\phi^2\rangle)\exp[-8\pi^2/g^2(\langle\phi^2\rangle)]\,\lambda^2\psi^2,\\ &=\exp(-8\pi^2/g^2)\,(\Lambda^4/\langle\phi^6\rangle)\,\lambda^2\psi^2\end{aligned} \tag{16}$$

and the effective action has a term proportional to

$$\exp(-8\pi^2/g^2)\,(\Lambda^4/\phi^6)\,\lambda^2\psi^2. \tag{17}$$

This is one of the terms which are implied by eq. (14). Hence, we have established that the coefficient a_1 in eq. (14) is non-zero. The precise value of the coefficient a_1 depends on the regularization scheme but it cannot be set to zero. [Remember that when the regularization scheme is changed, the value of A_1 in eq. (9) is changed.] Notice that as in the calculations in ref. [5] we found a "miraculous" relation between the instanton classical action and the one-loop beta which determine the exponent of ϕ in eq. (16) and the discrete symmetry which determines the exponent of Ψ in eq. (14).

Higher order perturbative corrections around the instanton would lead to factors of $\ln\phi^2$ in front of eq. (17). Such terms cannot be combined with other terms into the invariant function (14). Therefore, we conclude that our result is not affected by perturbative corrections around the instanton and the coefficient a_1 is independent of g.

The instanton also affects other terms in the effective action. It renormalizes the coefficient of $(F_{\mu\nu})^2$. This can be seen in the instanton calculation by coupling the λ and ψ zero modes to two A_μ's. An easier way to calculate the effect is to expand eq. (14) in components [#1]. Assuming for simplicity that $\langle(\phi^a)^2\rangle$ is real (we will later relax this assumption), we find

$$\begin{aligned}&1/g^2_{\rm eff}(\langle\phi^2\rangle)\\ &\quad=(1/g^2)\,[1+3g^2/4\pi^2+(g^2/4\pi^2)\ln(\phi^2/\Lambda^2)\\ &\quad+24g^2a_1\exp(-8\pi^2/g^2)\,(\Lambda^2/\phi^2)^2\\ &\quad+O(\exp(-16\pi^2/g^2))].\end{aligned} \tag{18}$$

Therefore, the beta function is

$$\begin{aligned}\beta(g)=&-(1/4\pi^2)\,g^3+48a_1g^3\exp(-8\pi^2/g^2)\\ &+O(\exp(-16\pi^2/g^2)).\end{aligned} \tag{19}$$

We see that the instanton changes the beta function

[#1] As for the perturbative non-renormalization theorem, we do not need to use $N=2$ superspace to derive the same results. Using $N=1$ superspace, it is easy to see that since eq. (17) is generated, $f_{\rm eff}$ is affected by the instanton. We could use all the symmetry arguments we used above to restrict the form of $f^{\rm instanton}$ rather than of $\mathscr{F}^{\rm instanton}$. From this point of view $N=2$ is needed only in order to argue that there is no ambiguity associated with field redefinitions.

and thus violates the perturbative non-renormalization theorem about it. Notice that our answer for the beta function disagrees with the conclusion of ref. [6] that the exact beta function in this theory is given by the one-loop answer.

The non-perturbative correction to the beta function is interesting. A well-defined calculation of a non-perturbative beta function makes sense only if the perturbative expansion of the beta function converges (and in particular, if the perturbative beta function vanishes). Clearly, our calculation satisfies this requirement because of the perturbative non-renormalization theorem. Instanton corrections to beta functions were also discussed in refs. [12–14].

We could add to the $N=2$ lagrangian a theta term $(\Theta/32\pi^2)\,F\tilde{F}$ but because of the U(1) symmetry we can redefine it to zero. However, there may still be non-trivial CP violation which arises from the expectation value of the scalar, $\langle(\phi^a)^2\rangle$. Defining Θ to be the coefficient of $(1/32\pi^2)\,F\tilde{F}$ in the low energy effective action in a given vacuum, we find Θ dependence in the beta function of g and renormalization of Θ:

$$\beta_g(g,\Theta) = -(1/4\pi^2)\,g^3+48a_1\cos\Theta g^3\exp(-8\pi^2/g^2) +\mathrm{O}(\exp(-16\pi^2/g^2)), \tag{20}$$

$$\beta_\Theta(g,\Theta)=-768\pi^2 a_1\sin\Theta\exp(-8\pi^2/g^2) +\mathrm{O}(\exp(-16\pi^2/g^2)). \tag{21}$$

Our analysis is easily extended to multi-instantons and to other gauge groups. For SU(N_c) gauge groups we find

$$\mathscr{F}_{\text{eff}}=\Psi^2\left(1/8g^2+(N_c/64\pi^2)\ln(\Psi^2/\Lambda^2) + \sum_{l=1}^{\infty} a_l\exp(-8\pi^2 l/g^2)\,(\Lambda^2/\Psi^2)^{lN_c}\right). \tag{22}$$

The lth term in the sum is the contribution of the configurations with topological charge l. Clearly, there is a non-renormalization theorem in perturbation theory around all these configurations (the coefficients a_l are independent of g). We do not know if the sum in eq. (22) converges. Even if it does, it is not clear whether there are other non-perturbative effects weaker than instantons which should also be taken into account.

In view of the non-perturbative beta function in the $N=2$ theory, it is natural to ask if the $N=4$ theory which is perturbatively finite remains finite non-perturbatively. Working again along the flat directions and repeating the analysis above we find eight fermionic zero modes which are not lifted by the scalar expectation value. These are associated with the action of eight of the sixteen supersymmetry generators on the instanton configuration. Clearly, this does not change the Kähler potential or the coefficient of $F^2_{\mu\nu}$. It leads to an eight-fermion term in the effective action which can be written in $N=1$ superspace as a D-term with covariant derivatives. Since the coefficient of $F^2_{\mu\nu}$ is not modified, instantons do not generate a beta function in this case.

A simple argument shows that the conclusion of the previous paragraph is exact and is not limited to the instanton approximation [#2]. The effective action of the massless fields along the flat direction has to be $N=4$ supersymmetric. This restricts its form. In particular, the flat direction cannot be lifted – a potential along the flat direction is not compatible with $N=4$ supersymmetry. Also, the kinetic terms for the bosons are determined to be trivial. Therefore, there cannot be any scalar dependent function in front of the kinetic term for the gauge bosons. Hence, the effective gauge coupling is independent of the scalar expectation value and the beta function is *exactly* zero. We conclude that the $N=4$ theory is finite both perturbatively and non-perturbatively.

It is a pleasure to thank Tom Banks, Michael Dine and Edward Witten for helpful discussions. This work was supported in part by NSF grant PHY-86-20266.

[#2] We assume here that the $N=4$ theory exists as an interacting quantum field theory and there are no non-perturbative anomalies.

References

[1] S.J. Gates Jr., M.T. Grisaru, M. Roček and W. Siegel, Superspace or one thousand and one lessons in supersymmetry (Benjamin/Cummings, Menlo Park, CA, 1983).

[2] J. Wess and J. Bagger, Supersymmetry and supergravity, Princeton Series in Physics (Princeton U.P., Princeton, NJ, 1983).
[3] M.T. Grisaru, W. Siegel and M. Roček, Nucl. Phys. B 159 (1979) 429.
[4] M.F. Sohnius and P.C. West, Phys. Lett. B 100 (1981) 245; S. Mandelstam, Nucl. Phys. B 213 (1983) 149.
[5] I. Affleck, M. Dine and N. Seiberg, Nucl. Phys. B 241 (1984) 493; B 256 (1985) 557.
[6] V.A. Novikov, M.A. Shifman, A.I. Vainshtein, M.B. Voloshin and V.I. Zakharov, Nucl. Phys. B 229 (1983) 394; V.A. Novikov, M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Nucl. Phys. B 229 (1983) 381, 407; M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Phys. Lett. B 166 (1986) 329.
[7] D. Amati, K. Konishi, Y. Meurice, G.C. Rossi and G. Veneziano, Phys. Rep., to be published.
[8] E. Witten, preprint IASSNS/HEP-87/72.
[9] R. Grimm, M. Sohnius and J. Wess, Nucl. Phys. B 133 (1978) 275; P. Breitenlohner and M.F. Sohnius, Nucl. Phys. B 187 (1981) 409.
[10] S. Ferrara and B. Zumino, Nucl. Phys. B 87 (1975) 207.
[11] See e.g. M. Grisaru and P.C. West, Nucl. Phys. B 254 (1985) 249; M.A. Shifman and A.I. Vainshtein, Nucl. Phys. B 277 (1986) 456.
[12] C.G. Callan, R.F. Dashen and D.J. Gross, Phys. Lett. B 77 (1978) 4684; Phys. Rev. D 17 (1978) 2717; D 18 (1978) 270; D 20 (1979) 3279.
[13] T. Banks and N. Seiberg, Nucl. Phys. B 273 (1986) 109.
[14] M. Dine, N. Seiberg, X.G. Wen and E. Witten, Nucl. Phys. B 278 (1986) 769; B 289 (1987) 319.