Self-duality and modular invariance in gauge theories

1990 J. Phys. A: Math. Gen. 23 3895

(http://iopscience.iop.org/0305-4470/23/17/019)

J. Phys. A: Math. Gen. 23 (1990) 3895-3901. Printed in the UK

Self-duality and modular invariance in gauge theories*

Leah Mizrachi†

Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland

Received 14 November 1989

Abstract. It is shown that self-duality of gauge theories in the large-N limit when g^2N is kept finite extends to an invariance under the full modular group in parameter space when a θ-term is included. We suggest using modular covariant combinations of modular forms to construct the generating functional in the large-N limit. This fixes the phase transitions to be on the boundaries of the horocircles defining the fundamental domain of the modular group.

Duality is a general concept relating physical quantities in different regions of parameter space. So far there has been use of various duality relations in very different fields of physics, where in all cases the parameter multiplying the kinetic term in the Lagrangian gets inverted by the transformation, thus relating a strong to a weak coupling region in parameter space‡. This is the case of Kramers-Wannier duality [1] in lattice models where the role of a coupling constant is played by the temperature, and for self-dual models, i.e. those invariant under the transformation the phase structure in parameter space can be analysed [2]. Another example is electric-magnetic duality in Abelian gauge theory which is invariant under the interchange of electric and magnetic fields provided both electric and magnetic charges appear in the model. Interestingly enough, magnetic charges are proportional to the inverse of the electric ones as dictated by Dirac quantisation condition. Indeed, in both cases the duality transformation can be implemented by a Fourier transform in function space [3], thus leading to the inversion of the coupling parameter (temperature in lattice models or electric charge in gauge theory). In gauge theory, though, an additional SO(4) rotation in Euclidean spacetime has to be performed in order to interchange the roles of electric and magnetic fields§. With the same technique one can analyse non-Abelian gauge theory [5], which strictly speaking is not self-dual, yet for weak coupling [6] or equivalently in the large-N limit it becomes self-dual provided g^2N is finite [7]. Here again it is an electric-magnetic duality [8], as in the non-Abelian models magnetic charges are built in intrinsically due to the existence of the 't Hooft-Polyakov monopoles [9] (or Wu-Yang monopoles

* This paper is based upon the University of Geneva preprint UGVA-DPT 1989/01-603 (January 1989).

† Supported by the Swiss National Science Foundation.

‡ In the more general case the dual coupling parameter is a decreasing function of the original one, and becomes the inverse coupling only within the Gaussian approximation.

§ A similar SO(4) rotation was suggested by 't Hooft [4], in his attempt to define electric-magnetic duality in non-Abelian gauge theory by isolating the contributions of electric and magnetic vortices.

in pure Yang–Mills theory [10]). Other duality relations, e.g. the duality between scalar and antisymmetric tensor field theories [11], or the invariance in compactified string models under the interchange $R \to 1/2R$ [12] (where R is the radius of compactification) can also be studied this way; i.e. by performing the Fourier transform in function space. This last invariance turned out to be very useful in relating and classifying conformal field theories with $C = 1$ [13].

An extension of this technique to models where the generalised kinetic term is not diagonal in function space is almost immediate thus proving the invariance of σ-models under the transformation $(G+B) \to \frac{1}{4}(G+B)^{-1}$ [14], (G and B are the background metric and the antisymmetric tensor field respectively), or getting the duality relation in gauge theories with θ-term. For this last case it turns out to be useful to define a complex parameter $\zeta_{\rm L}$†

$$\zeta_{\rm L} = \frac{\theta}{2\pi} + \mathrm{i}\frac{2\pi}{Ng^2} \tag{1}$$

which gets inverted by the duality transformation both on the lattice [15] and in the continuum [16, 17]. In addition, the invariance of the lattice model under $\theta \to \theta + 2\pi$ turns, in fact, the duality transformation into a modular transformation in parameter space generated by two transformations

$$\zeta \to -\frac{1}{\zeta} \tag{2a}$$

$$\zeta \to \zeta + 1. \tag{2b}$$

Thus the self-dual lattice models containing the θ-term become modular invariant. This modular invariance was emphasised in [15] and was used to suggest the phase structure of this type of models provided there is at least one point of phase transition in the coupling parameter g (temperature for lattice models).

We would like to point out here that this modular invariance is a property of gauge theories in the large-N limit provided g^2N is kept finite. Furthermore, we propose in the following a form for the partition function of large-N Yang–Mills theory when $\theta \neq 0$, which can be used to probe into the phase structure of the theory. This is an extension of a previous proposal [7] for the partition function of large-N Yang–Mills theory when $\theta = 0$. Undoubtedly a knowledge of the phase structure of non-Abelian gauge theory would be very useful in understanding confinement.

We recall that duality transformation in gauge theory is implemented by performing a Fourier transform in function space accompanied by an SO(4) rotation in Euclidean spacetime [3]. Using the radial gauge [18]‡

$$x_\mu A^a_\mu(x) = 0 \tag{3}$$

where $A^a_\mu(x)$ are the gauge potentials, the generating functional can be written as

$$Z_N(g, \theta; J) = \int \mathscr{D}A^a_\mu \delta(x_\mu A^a_\mu) \exp\left[-\int \mathrm{d}^4x\left(\frac{1}{4g^2} F^a_{\mu\nu}F^a_{\mu\nu} - \frac{\mathrm{i}\theta}{32\pi^2} F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} + \mathrm{i}J^a_\mu A^a_\mu\right)\right] \tag{4}$$

† This defines the parameter $\zeta_{\rm L}$ for lattice models. We will later on define the appropriate parameter $\zeta_{\rm G}$ for gauge theory (equation (18)). The difference results from the different normalisations of the kinetic and θ-term in gauge theory compared with the lattice model (see [15, 16]).

‡ The duality transformation can be performed in a gauge independent way [19] (see also [16]), whereas the radial gauge used here simplifies the large-N limit. Renormalisability within this gauge condition was discussed in [20], where its similarity with the axial gauge was emphasised. (For a general review of various gauge conditions and their properties see [21], whereas applications of the radial gauge can be found in [22].)

where

$$F^a_{\mu\nu}=\partial_\mu A^a_\nu-\partial_\nu A^a_\mu+f^{abc}A^b_\mu A^c_\nu \tag{5}$$

is the field strength with f^{abc} as structure constants of the gauge group,

$$\tilde{F}^a_{\mu\nu}=\tfrac{1}{2}\mathscr{E}_{\mu\nu\rho\sigma}F^a_{\rho\sigma} \tag{6}$$

and J^a_μ is the external source. Performing now the duality transformation in the radial gauge (3) it can be shown that the dual model in the large-N limit is the same as the original one provided g^2N is kept finite [7]. The gauge potentials of the dual theory, B^a_μ, also satisfy the radial gauge condition (3), whereas the dual coupling constant, g^θ_{D}, and the dual θ-parameter, θ_{D}, are given by [16, 17]

$$(g^\theta_{\mathrm{D}})^2=\frac{g^2}{4\pi^2N^2}\left[\left(\frac{8\pi^2}{g^2}\right)^2+\theta^2\right] \tag{7a}$$

$$\theta_{\mathrm{D}}=\frac{-4\pi^2N^2\theta}{(8\pi^2/g^2)^2+\theta^2}. \tag{7b}$$

To get this result the identity

$$\exp\left[-\frac{1}{4g^2}\int \mathrm{d}^4x\, F^a_{\mu\nu}F^a_{\mu\nu}\right]=c\int \mathscr{D}K^a_{\mu\nu}\exp\left[-\int \mathrm{d}^4x\left(\frac{\alpha g^2}{4}K^a_{\mu\nu}K^a_{\mu\nu}+\frac{\mathrm{i}\beta}{2}\tilde{K}^a_{\mu\nu}F^a_{\mu\nu}\right)\right] \tag{8}$$

was used [17]. Here c is an (infinite) constant which can be ignored as it will cancel out once correlation functions are studied [7], and α, β are arbitrary non-zero parameters satisfying

$$\beta^2/\alpha=1. \tag{9}$$

The identity (8) is in fact the essence of the duality transformation, which involves a Fourier transform in function space and an SO(4) rotation (coupling of $F^a_{\mu\nu}$ to $\tilde{K}^a_{\mu\nu}$ which is defined as in (6)). We note, in particular, that the coupling parameter is inverted ($g\to 1/g$) in the process of the transformation. Taking $\theta=0$ in (4) the parameters α and β can be fixed once the integration over A^a_μ is carried out. Substituting (8) in (4) and using the inversion formula [18]

$$A^a_\mu(x)=\int_0^1 \mathrm{d}\gamma[y_\rho F^a_{\rho\mu}(y)]_{y=\gamma x} \tag{10}$$

which is valid in the radial gauge (3) we can integrate A^a_μ in (4) and get the dual theory defined in terms of the new variables $K^a_{\mu\nu}$. Studying now the large-N limit when g^2N is finite, we find that $K^a_{\mu\nu}$ are fixed to be the dual field strength, $G^a_{\mu\nu}$, defined in terms of the dual gauge potentials, B^a_μ, as in (5). Other interaction terms, which are not of the gauge type (involving an antisymmetric tensor field $\omega^a_{\mu\nu}$ satisfying the condition $x_\mu\omega^a_{\mu\nu}(x)=0$), disappear in the large-N limit†. There is an additional surface term, though, $\mathrm{i}\beta/2\int \mathrm{d}^4x\, G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}$ which is generated by the duality transformation even when $\theta=0$ in the original theory [6, 16]. Noting that for SU(N) gauge group

† When performed in a gauge independent way the dual fields $K^a_{\mu\nu}$ are constrained to satisfy the Bianchi identity [16, 19], whose general solution can be found in [23] (see also the appendix of [19] for a perturbative approach to this solution). The dual theory turns out to be a gauge theory coupled to two vector potentials satisfying generalised gauge invariance [23].

this surface term can be written as [24]

$$\frac{1}{32\pi^2}\int d^4x\, G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} = n + \frac{\boldsymbol{m}\cdot\boldsymbol{k}}{N} \tag{11}$$

where n is the Pontryagin index (second Chern class index), $\boldsymbol{k} = (k_x, k_y, k_z)$ is the first Chern class index associated with vortices arising from the twisted boundary conditions in the x, y, z directions and $\boldsymbol{m} = (m_x, m_y, m_z)$ are the magnetic charges. Both $\boldsymbol{k}$ and $\boldsymbol{m}$ are integers defined modulo N, and n is an integer. Thus up to a factor $32\pi^2$ the surface term (11) is a rational number. Hence if we take

$$\beta = \frac{N}{4\pi} \tag{12}$$

we find the contribution

$$\exp\left[-\frac{iN}{8\pi}\int d^4x\, G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right] = 1. \tag{13}$$

It is one for all the instanton contributions including those with twisted boundary conditions [24], and it disappears from the dual theory [17]. Note that had we taken $\beta = N/8\pi$, as in [17], (13) would still be valid. However, we will later on see that for the choice (12) for β the theory is manifestly invariant under the full modular group $SL(2, \mathbb{Z})$ once we add the θ-term. Whereas for the choice $\beta = N/8\pi$ only the invariance under the subgroup Γ^1_2 of the modular group would be manifest. With the choice (12) for β we find

$$\alpha = \beta^2 = \frac{N^2}{16\pi^2} \tag{14}$$

and the dual coupling constant constant when $\theta = 0$ becomes

$$g_D = \frac{4\pi}{Ng}. \tag{15}$$

When $\theta \neq 0$ an identity similar to (8) can be used. Here a complex parameter

$$z_\pm = \frac{1}{g^2} \pm \frac{i\theta}{8\pi^2} \tag{16}$$

has to be defined†, then the Euclidean action in (4) gets diagonalised for

$$F^\pm_{\mu\nu} = \tfrac{1}{2}(F_{\mu\nu} \pm \tilde{F}_{\mu\nu}) \tag{17}$$

with coefficients $z_\mp$ in front of the terms $\frac{1}{4}(F^\pm_{\mu\nu})^2$ respectively, thus resulting in the transformation $z_\pm \to \alpha/z_\pm$ when the Fourier transform (8) is used. Continuing as before we choose α and β as in (14) and (12) respectively in order to eliminate the surface term generated by the duality transformation itself. This, though, would not eliminate the θ-term from the dual theory as (13) is not necessarily valid for an arbitrary coefficient θ in front of the surface term. Instead, we find the dual parameters g^θ_D, θ_D given as in (7); they are defined from the real and imaginary parts of z_+ (or z_-) in (16), after its inversion by the duality transformation [16, 17].

† We scaled here the parameters $z_\pm$ of [16] by a factor $8\pi^2$, as the transformation of $z_\pm$ as defined in (16) is simpler than that for the parameters defined in [16].

Since the parameter θ in (4) is proportional to a rational number (see (11)), we find the additional invariance $\theta \to \theta + 2\pi N$. Thus by defining the parameter ζ_G

$$\zeta_G = \frac{iz_-}{\alpha} = \frac{\theta}{2\pi N} + i\frac{4\pi}{g^2 N} \tag{18}$$

we find that the model is invariant under the transformations (2), which generate the full modular group SL(2, $\mathbb{Z}$). Note that had we taken $\beta = N/8\pi$, we would have found the invariance under $\Gamma_2^1 \subset \mathrm{SL}(2, \mathbb{Z})$ which is generated by (2*a*) and $\zeta_G \to \zeta_G + 2$. Thus the invariance under full modular group would not be manifest.

We have thus shown that SU(N) gauge theory is invariant under modular transformations in parameter (g, θ) space in the large-N limit provided g^2N is finite. Note that Re ζ_G can be neglected in this limit when θ is finite. However, when the whole range of θ is considered it has to be kept, thus recovering the full modular group. This modular invariance was emphasised in the analogous lattice model in [15]. The difference, though, is that in gauge theory it is valid in the large-N limit only (when additional interaction terms can be ignored [7]), whereas in the lattice model it is an exact statement.

Once we know that we have modular invariance the partition function can be constructed from modular covariant combinations of modular forms [25], much the same as partition functions in string models are constructed. The difference, though, is that here the modular parameter ζ_G (or ζ_L for lattice models) is not integrated over, whereas the modular parameter τ defining the torus of a closed string model (for genus one), is integrated over the classical fundamental region of the modular group with a modular invariant measure (and modular invariant partition function).

Noting that modular forms are bounded beyond a strip Im $\zeta_G \geq d$ in the complex ζ_G plane where d is some positive constant, we find that the partition function would be a well-behaved function for $4\pi/g^2N > d$. In addition, in the region $0 < \mathrm{Im}\, \zeta_G < d$ modular forms are bounded within the horocircles $|\zeta_G - (p/q) - id_{p,q}| < d_{p,q}$, which are circles of radii $d_{p,q}$ centred at $p/q + id_{p,q}$. Here p/q is a rational number and $d_{p,q}$ is a positive real number. Thus the possible phase transition regions for the large-N gauge theory would be the boundaries of the horocircles (see figure 1) defining the fundamental domain of the modular group [25]†. This would be correct provided g^2N is kept finite. Adding the large N corrections (which are damped by factors of e^{-N} [7]), cannot change the phase structure as long as the large N expansion is valid. This is due to the fact that the boundaries between different phases are regions of singularities, which cannot be shifted by a viable expansion within its radius of convergence. The question is, of course, how far one can push N and still have the same results as in the large-N region. Here, a knowledge of the radius of convergence of the expansion is needed.

The modular invariance of gauge theories in the large-N limit when g^2N is finite, is linked with the existence of the twisted configurations (11). Of particular importance are the vortices which have long been conjectured to be responsible for confinement in gauge theory, even though it has never been really proven. With the self-duality we found [7], it may be possible to write a strong coupling expansion which could be used to study this question more quantitatively. The inversion of the effective coupling parameter may be useful for this end.

† A similar phase structure was found in [15] for the lattice Z_N model.

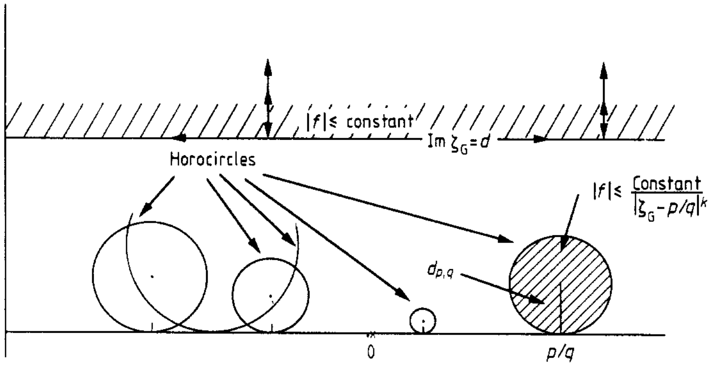

Figure 1. The horocircles in the complex ζ_G plane centred at $p/q+\mathrm{i}d_{p,q}$ with radii $d_{p,q}$ where p/q is a rational number and $d_{p,q}$ is a real positive number. These are the regions of possible phase transitions in large-N gauge theory provided g^2N is kept finite. For Im $\zeta_G > d$ (where d is a positive integer) the modular forms defining the partition function in this limit are bounded functions of ζ_G.

It is tempting to end this paper with a few speculative remarks. Modular invariance is an essential property of closed string models. It is related to the conformal invariance of the two dimensional model. This conformal invariance yields on the one hand the Virasoro algebra (or more generally the Kac-Moody algebra, when there is in addition an invariance under a gauge group G), and the fusion rules on the other hand. These fusion rules were conjectured [26] and proven [27] to be diagonalised by the generators of the modular transformations (2a). They also generate the dual diagrams of the Veneziano amplitude (which is yet another concept of duality, here related to the fusion rules). Alternatively, the Kac-Moody algebra was found to be responsible for the integrability of various models [28]. (A general review on integrability of quantum theories can be found in [29].) The question we raise is: does that mean that similar properties may be found in gauge theories in the large-N limit? It is, of course, important to note that in string theory modular invariance is found in τ, which is not a parameter of the theory, as it is a variable of integration. Yet the invariance $R \to 1/2R$ in the compactified string models [12] (or $G+B) \to \frac{1}{4}(G+B)^{-1}$ [14] in the σ-model approach to string theory) is essentially of the type we are talking about (Kramers-Wannier duality [1]), as these (R, G, B) can be thought of as parameters of the string theory. Indeed, it was pointed out in [30] (see also [28]) that Kramers-Wannier duality leads in general to the existence of an infinite set of conserved charges, here the infinite set of generators of gauge theory in the large-N limit, thus making it closed to an integrable system [29]. Furthermore, string theory is a theory of extended objects (vortices); is this a coincidence only or can that be used to get a better understanding of confinement which is so tightly related to the existence of vortices.

Acknowledgments

I wish to thank all the people who arranged my recent visit to the USA, where this work was conceived. In particular I am grateful to L Brown, P Freund, N Khuri, P van Nieuwenhuizen, C Rebbi, J Schwarz, S Weinberg and T-M Yan for the kind invitations and for the hospitality. I am also indebted to the Department of Theoretical

Physics in the University of Geneva for the support. The encouragement of Professors R Gatto and H Ruegg is particularly appreciated.

Note added. After completing this paper we became aware of [31] where it has been suggested that classical Yang-Mills theory in the large-N limit possesses an additional gauge invariance; that of area preserving coordinate transformations of an internal sphere attached to every spacetime point. It is based on the isomorphism between the algebra of SU(N) gauge theory when $N \to \infty$ and the infinite dimensional algebra of area preserving diffeomorphisms of a relativistic spherical membrane (after gauge fixing), which was proven in [32]. This seems to us related to the ideas presented in this paper.

References

[1] Kramers H A and Wannier C H 1941 *Phys. Rev.* D **60** 252
[2] Savit R 1980 *Rev. Mod. Phys.* **52** 453
[3] Sugamoto A 1979 *Phys. Rev.* D **19** 1820
[4] 't Hooft G 1979 *Nucl. Phys.* B **153** 141
[5] Seo K, Okawa M and Sugamoto A 1979 *Phys. Rev.* D **19** 3744
Seo K and Okawa M 1980 *Phys. Rev.* D **21** 1614
[6] Itabashi K 1981 *Prog. Theor. Phys.* **65** 1423
[7] Mizrachi L 1984 *Phys. Lett.* **139B** 374
[8] Mizrachi L 1984 *Ann. Phys.* **153** 417
[9] 't Hooft G 1974 *Nucl. Phys.* B **79** 276
Polyakov A M 1974 *Zh. Eksp. Teor. Fiz. Pis. Red.* **20** 430 (1974 *JETP Lett.* **20** 194)
[10] Wu T T and Yang C N 1975 *Phys. Rev.* D **12** 3845
[11] Mecklenburg W and Mizrachi L 1984 *Phys. Rev.* D **29** 1709
[12] Kikkawa K and Yamasaki M 1984 *Phys. Lett.* **149B** 357
Nair V P, Shapere A, Strominger A and Wilczek F 1987 *Nucl. Phys.* B **287** 402
[13] Ginsparg P 1989 Applied Conformal Field Theory *Fields, Strings and Critical Phenomena, Les Houches Summer Session XLIX (1988)* ed E Brezin and J Zinn-Justin (Amsterdam: Elsevier)
[14] Giveon A, Rabinovici E and Veneziano G 1989 *Nuc. Phys.* B **322** 167
Shapere A and Wilczek F 1989 *Nucl. Phys.* B **320** 669
[15] Cardy J L 1982 *Nucl. Phys.* B **205** [FS5] 17
[16] Mecklenburg W and Mizrachi L 1983 *Phys. Rev.* D **27** 1922
[17] Mecklenburg W and Mizrachi L 1983 *Phys. Rev.* D **28** 2686
[18] Fock V A 1937 *Phys. Z. Sowjetunion* **12** 404
Schwinger J 1951 *Phys. Rev.* **82** 664
Cronström C 1980 *Phys. Lett.* **90B** 267
Dubovikov M S and Smilga A V 1979 Proc. Int. Seminar on Group Theoretical Methods in Field Theory, Moscow
[19] Kazama Y and Savit R 1980 *Phys. Rev.* D **21** 2916
[20] Kummer W and Weiser J 1986 *Z. Phys.* C **31** 105
[21] Leibbrandt G 1987 *Rev. Mod. Phys.* **59** 1067
[22] Dubovikov M S and Smilga A V 1981 *Nucl. Phys.* B **185** 109
Shifman M A 1980 *Nucl. Phys.* B **173** 13
Nikolaev S N and Radyushkin A V 1982 *Phys. Lett.* B **110** 476; 1983 *Nucl. Phys.* B **213** 285
[23] Mecklenburg W and Mizrachi L 1985 *Phys. Lett.* **150B** 371
[24] 't Hooft G 1980 *Acta Physica Austriaca Suppl.* XXII 531; 1981 *Commun. Math. Phys.* **81** 267
[25] Freitag E 1983 *Siegelsche Modulfunktionen* (Berlin: Springer)
Schoeneberg B 1984 *Elliptic Modular Functions* (Berlin: Springer)
Mumford D 1983, 1984 *Tata Lectures for Theta I, II* (Basel: Birkhäuser)
[26] Verlinde E 1988 *Nucl. Phys.* B **300** [FS22] 360
[27] Moore G and Seiberg N 1988 *Phys. Lett.* **212B** 451
[28] Dolan L 1984 *Phys. Rep.* **109** 1
[29] de Vega H J 1989 *Int. J. Mod. Phys.* A **4** 2371
[30] Dolan L and Grady M 1982 *Phys. Rev.* D **25** 1587
[31] Floratos E G, Iliopoulos J and Tiktopoulos G 1989 *Phys. Lett.* **217B** 285
[32] Hoppe J 1982 *PhD thesis* MIT; *Aachen preprint* PITHA 86/24