Strings, gravity, and the quest for unification/Cordes, gravitation, et la quête d'unification

Super Yang–Mills, matrix models and geometric transitions

Frank Ferrari [a,b]

[a] *Institut de Physique, Université de Neuchâtel, rue A.-L. Bréguet 1, CH-2000 Neuchâtel, Switzerland*
[b] *Service de Physique Théorique et Mathématique, and International Solvay Institutes, Université Libre de Bruxelles, Campus de la Plaine, CP 231, boulevard du Triomphe, B-1050 Bruxelles, Belgium*

Presented by Guy Laval

Abstract

I explain two applications of the relationship between four-dimensional $\mathcal{N} = 1$ supersymmetric gauge theories, zero-dimensional gauged matrix models, and geometric transitions in string theory. The first is related to the spectrum of BPS domain walls or BPS branes. It is shown that one can smoothly interpolate between a D-brane state, whose weak coupling tension scales as $N \sim 1/g_s$, and a closed string solitonic state, whose weak coupling tension scales as $N^2 \sim 1/g_s^2$. This is part of a larger theory of $\mathcal{N} = 1$ quantum parameter spaces. The second is a new purely geometric approach to sum exactly over planar diagrams in zero dimension. It is an example of open/closed string duality. ***To cite this article: F. Ferrari, C. R. Physique 6 (2005).***
© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Résumé

Super Yang–Mills, modèles de matrices et transitions géométriques. J'expose deux applications de la relation entre les théories de jauge supersymétriques $\mathcal{N} = 1$ à quatre dimensions, les modèles de matrices jaugés à zéro dimension, et les transitions géométriques en théorie des cordes. La première traite du spectre des parois ou branes BPS. Il est démontré qu'un état de type D-brane, dont la tension en couplage faible est proportionnelle à $N \sim 1/g_s$, et un état de type soliton de corde fermée, dont la tension en couplage faible est proportionnelle à $N^2 \sim 1/g_s^2$, peuvent être continument déformés l'un en l'autre. Ceci est un aspect de la théorie plus générale des espaces quantiques de paramètres $\mathcal{N} = 1$. La deuxième est une nouvelle méthode, purement géométrique, pour calculer exactement la somme sur les diagrammes planaires en zéro dimension. C'est un exemple de dualité entre cordes ouvertes et cordes fermées. ***Pour citer cet article : F. Ferrari, C. R. Physique 6 (2005).***
© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Keywords: Supersymmetric gauge theories; Matrix models; String theory

Mots-clés : Théories de jauge supersymétriques ; Modèles de matrices ; Théorie des cordes

Version française abrégée

L'étude des théories de jauge supersymétriques en couplage fort a connu deux avancées majeures ces dix dernières années. L'une est fondée sur la notion de dualité S et sur l'utilisation des propriétés d'analyticité de certaines observables pour obtenir des résultats exacts [1–3,10]. L'autre est fondée sur la construction de théories de cordes, mathématiquement équivalentes aux

E-mail address: frank.ferrari@ulb.ac.edu (F. Ferrari).

théories de jauge, mais plus simples à étudier dans le régime de couplage fort [12]. Le point de contact entre ces deux développements [14–16,13] est à l'origine d'une activité intense depuis 2002. Nous souhaitons présenter deux aspects complémentaires de ces recherches, à la suite de travaux publiés dans [17,18].

Parois

Le premier sujet concerne le spectre des parois (ou membranes) BPS dans les théories $\mathcal{N} = 1$. Dans la théorie de jauge pure U(N), ces parois existent en raison de la brisure dynamique de la symétrie chirale $\mathbb{Z}_{2N} \to \mathbb{Z}_2$. Elles ont été étudiées, par exemple, par Witten [19]. Elles se comportent comme des deux-branes de Dirichlet : leur tension est proportionnelle à $N \sim 1/g_s$, et des cordes ouvertes peuvent s'y attacher. Lorsque l'on rajoute à la théorie un multiplet chiral de matière dans la représentation adjointe, avec un superpotentiel W à l'ordre des arbres tel que (8), il apparaît un nouveau type de parois qui interpolent entre les vides $\langle X \rangle = x_+$ et $\langle X \rangle = x_-$. Ces parois peuvent être étudiées en détail dans la théorie classique. Elles se comportent comme des deux-branes solitoniques, avec une tension proportionnelle à $N^2 \sim 1/g_s^2$.

Il est possible de calculer exactement les tensions de ces diverses parois ou membranes, en utilisant soit le formalisme du modèle de matrice [16] en théorie des champs, soit celui de la transition géométrique en théorie des cordes [15]. Le résultat [17], de nature purement non-perturbative, est explicité dans les formules (13), (16), (17). La propriété fondamentale de ces formules [17] est de montrer qu'il est possible de déformer de manière continue les membranes de type Dirichlet en membranes de type soliton et vice versa (voir Fig. 1). De plus, pour certaines valeurs particulières des paramètres, la tension de la membrane peut être proportionnelle à une puissance fractionnaire de N.

Ces propriétés de continuation analytique sont très générales et ont des conséquences drastiques sur la structure de l'espace quantique des paramètres des théories $\mathcal{N} = 1$ [17,21,22]. En particulier, cet espace, qui a de multiples composantes disconnectées au niveau classique, s'avère être connexe dans tous les cas étudiés lorsque les corrections quantiques non-perturbatives sont prises en compte [17,21]. Le résultat dans le cas de la théorie de jauge U(4) est représenté sur la Fig. 2.

L'un de nos résultats les plus important a été de montrer que l'on pouvait déformer continument une D-brane en une brane solitonique et vice versa. Nous avons aussi découvert de nouveaux objets étendus avec une tension proportionnelle à une puissance fractionnaire de la constante de couplage de corde. Un point intéressant est la possibilité de construire des branes solitoniques dans le cadre d'une théorie de jauge classique. Une description explicite de la théorie microscopique sur la membrane peut alors être donnée en principe par la quantification semiclassique. Une telle description microscopique n'a pas pu être obtenue jusqu'à présent dans le cadre de la théorie des cordes.

Diagrammes planaires et géométrie

Nous avons deux techniques à notre disposition pour calculer le potentiel effectif d'une théorie de jauge $\mathcal{N} = 1$. La première consiste à utiliser le modèle de matrice associé (2), l'autre consiste à construire la théorie de jauge sur une D-brane de la théorie des cordes puis à utiliser la dualité entre cordes ouvertes et cordes fermées. L'équivalence des deux procédures implique que les intégrales (2) peuvent être calculées dans la limite planaire par une méthode purement géométrique, que nous nous proposons de décrire [18].

Le point de départ est la variété de Calabi–Yau $\mathcal{M}$, couverte par deux cartes et définie par les fonctions de transition (27). On peut montrer que cette variété non-compacte contient des deux-sphères holomorphes autour desquelles peuvent être enroulées des D5-branes de la théorie de type IIB. La théorie à quatre dimensions qui en résulte est $\mathcal{N} = 1$ avec groupe de jauge U(N), M multiplets chiraux dans la représentation adjointe, et un superpotentiel à l'ordre des arbres donné par (30) [18]. Ce superpotentiel est le plus général jamais construit pour des théories de ce type. Par exemple, pour $M = 2$, on peut obtenir un superpotentiel arbitraire $W(x, y)$ dans le cas $N = 1$ où les deux matrices $X = X_1$ et $Y = X_2$ commutent. Dans le cas général, la formule (30) indique que l'ordre des termes à adopter correspond à celui introduit par Weyl en mécanique quantique (31). La méthode géométrique que nous étudions doit permettre, en principe, de résoudre tous les modèles du type (30), un résultat qui semble impossible à obtenir par les techniques traditionnelles.

En raison de la relation (25) entre couplage de Yang–Mills et volume des deux-sphères, on peut s'attendre à ce que les deux-sphères sur lesquelles s'enroulent les D5-branes aient tendance à rétrécir, puis à disparaître, lorsque l'effet du groupe de renormalisation est pris en compte. Mathématiquement, cette transformation est décrite par une application birationnelle π (32) qui transforme l'espace $\mathcal{M}$ en un espace singulier $\mathcal{M}_0$ (l'application inverse est en général appelée éclatement). Le nouvel espace de Calabi–Yau $\mathcal{M}_0$ est en tout point semblable à l'espace de départ $\mathcal{M}$, si ce n'est que les deux-sphères de $\mathcal{M}$ sont remplacées par des points singuliers de $\mathcal{M}_0$. Le calcul de π, duquel se déduit immédiatement celui de $\mathcal{M}_0$, n'est pas systématique et peut présenter des difficultés. Un point remarquable est que la correspondance avec le modèle de matrice implique que la variété $\mathcal{M}_0$ code complètement la théorie des représentations des équations matricielles (33). Ceci est un résultat puissant et inattendu pour les représentations de dimensions supérieures ou égales à deux. Il montre que la géométrie classique est capable de rendre

compte de la structure non-commutative qui apparaît lorsque l'on considère des D-branes. Cette propriété peut être vérifiée sur des exemples particuliers au prix de calculs détaillés [18].

Pour résoudre le modèle de matrice (2) quand $S \neq 0$, une étape supplémentaire est nécessaire au cours de laquelle l'espace $\mathcal{M}_0$ est déformé en un nouveau Calabi–Yau non-singulier $\hat{\mathcal{M}}$. Les déformations autorisées sont soumises à certaines contraintes de normalisabilité décrites dans [18]. Le nombre $\mathcal{N}_{\mathrm{geo}}$ de paramètres indépendants qui en résultent doit être égal au nombre de paramètres $\mathcal{N}_{\mathrm{alg}}$ dans le modèle de matrices. Ces derniers sont associés au choix d'une solution classique à partir de laquelle les diagrammes de Feynman planaires sont définis, et leur nombre est donc égal au nombre de représentations irréductibles des équations (33). L'identité (38) traduit cette relation profonde entre géométrie classique et algèbre non-commutative. Elle est satisfaite par tous les exemples étudiés [18]. À partir de la connaissance de $\hat{\mathcal{M}}$, on peut aussi calculer explicitement les résolvantes (41) pour les différentes matrices du modèle en utilisant (42), ainsi que la fonction de partition [18].

De nombreuses questions n'ont pas encore été élucidées dans cette correspondance entre modèles de matrices et géométrie. Par exemple, on peut se demander quelles sont les propriétés particulières des modèles (30) qui font qu'ils puissent être reliés à une géométrie de Calabi–Yau (théorie des représentations de (33), équations de boucle, etc...). Le problème ouvert le plus important est de comprendre si l'application π existe toujours pour la géométrie (27). Fournir une preuve d'existence rigoureuse est un problème mathématique difficile [27]. Si l'on pouvait trouver un algorithme permettant de construire π, on pourrait alors résoudre automatiquement tous les modèles de matrices (30). Si, par contre, l'application π n'existait pas en général, alors le cadre des transitions géométriques serait insuffisant pour décrire nos modèles. Ceci aurait des conséquences fondamentales pour notre compréhension des vides $\mathcal{N} = 1$ en théorie des cordes.

1. Introduction

In the last decade, a wealth of exact results in supersymmetric gauge theories have been derived. Beyond the original physical insight based on S-duality [1–3], a great variety of ideas and techniques have been used: singularity theory [4], integrable systems [5], geometric engineering [6], brane constructions [7], etc... The common thread of these developments is that they deal with observables that depend holomorphically on the couplings/moduli/parameters of the problem. This includes the low energy effective actions of $\mathcal{N} = 2$ theories [3,8,9] and the effective superpotentials and U(1) couplings of $\mathcal{N} = 1$ theories [10]. The set of stable BPS states in $\mathcal{N} = 2$ theories can also be determined in this framework [11].

Another major breakthrough has been the explicit construction of gauge theory/string theory dual pairs [12]. This open/closed string duality is in principle exact. However, most of the useful calculations can only be performed approximately, and are valid in a regime where either the closed string theory is tractable (small string coupling or large N, and small curvature or large $g_{\mathrm{YM}}^2 N$), or the open string theory is tractable (small $g_{\mathrm{YM}}^2 N$, which is the usual Feynman perturbation theory).

The point of contact between the two above-mentioned line of developments has generated an intense activity in the last couple of years [13]. The basic observation is that the holomorphic observables in gauge theories correspond to the topological sector of the string theories [14]. The open/closed string duality can be implemented exactly in the context of the topological strings, using the idea of geometric transitions [15]. It turns out that the result is encoded in the sum over planar diagrams of a particular zero-dimensional gauged matrix model [16]. This provides a beautifully simple and unified framework to rederive many of the already known exact results. It also opens up new directions of research. The aim of this article is to discuss two complementary aspects of this subject, based on work first published in [17,18].

We shall consider four-dimensional $\mathcal{N} = 1$ super Yang–Mills theories with gauge group U(N) and M matter chiral multiplets X_i in the adjoint representation. Asymptotic freedom is guaranteed for $M \leqslant 2$. The case $M = 3$ includes the conformally invariant $\mathcal{N} = 4$ theory. The cases $M > 3$ can be defined by embedding the gauge theory in string theory, which provides a natural UV cut-off Λ_0. We consider tree-level superpotentials for the matter fields of the form

$$W = \mathrm{tr}\, V(X_i), \tag{1}$$

where V is a polynomial. Renormalizability is not an issue, because the holomorphic data (F-terms) cannot depend on the UV cut-off Λ_0. Equivalently, we can always introduce auxiliary fields to make W at most cubic. Our super Yang–Mills theories can be geometrically engineered by considering a stack of N D5 branes in type IIB string theory on $\mathbb{R}^4 \times \mathcal{M}$. The space $\mathcal{M}$ is a non-compact Calabi–Yau threefold. The D5 branes span $\mathbb{R}^4$ and wrap two-spheres $\mathbb{P}^1 \subset \mathcal{M}$. We shall explain in Section 3 how to choose the space $\mathcal{M}$ to get an arbitrary number M of adjoint chiral superfields and a large class of superpotentials W of the form (1).

From the gauge theory point of view, the quantum corrections to F-terms are computed, according to Dijkgraaf and Vafa [16], by summing up the planar diagrams of a zero-dimensional gauged matrix model

$$\int_{\mathrm{planar}} \mathrm{e}^{-\frac{n}{S} W} \prod_i \mathrm{d} X_i. \tag{2}$$

The coupling S of the matrix model is identified with the glueball superfield of the gauge theory,

$$S = -\frac{\text{tr}\, W^\alpha W_\alpha}{16\pi^2 N}. \tag{3}$$

From the string theory point of view, the quantum corrections are obtained by using the open/closed string duality. Following [15], the closed string dual is type IIB on a new non-compact Calabi–Yau $\hat{\mathcal{M}}$, with no D-branes, but with three-form flux turned on (consistently with the original D-branes charges). The space $\hat{\mathcal{M}}$ is a deformation of the singular Calabi–Yau obtained by blowing down $\mathcal{M}$. Examples are presented in Section 3.

We thus have two powerful tools to compute the exact effective superpotentials of the gauge theories. An interesting consequence of the results so obtained is described in Section 2. In Section 3 we use the fact that the two approaches, gauge-theoretic and string-theoretic, must be equivalent, to describe a new purely geometric method to sum up planar diagrams in zero dimension.

2. Domain walls

2.1. Pure U(N) super Yang–Mills

The pure U(N) super Yang–Mills theory has a $\mathbb{Z}_{2N}$ chiral symmetry that acts on the gluino field, $\lambda_\alpha \to e^{i\pi k/N}\lambda_\alpha$. The glueball superfield (3) is a good, gauge invariant, order parameter for this symmetry. It turns out that strong quantum effects break $\mathbb{Z}_{2N}$ down to $\mathbb{Z}_2$. This breaking implies an N-fold degeneracy of the space of vacua, which are denoted by $|p\rangle$, $p \in \mathbb{Z}_N$. In the pth vacuum,

$$\langle p|S|p\rangle = \frac{1}{16\pi^2 N}\langle p|\text{tr}\,\lambda^\alpha \lambda_\alpha|p\rangle = \Lambda^3 e^{2i\pi k/N}, \tag{4}$$

where Λ is a conveniently normalized dynamically generated complex mass scale. The existence of a discrete set of degenerate vacua implies the existence of domain walls $(|p\rangle, |q\rangle)$ interpolating between them. These domain walls are discussed for example in [19]. They are BPS two-branes, which implies in particular that their tension can be expressed in terms of the quantum effective superpotential,

$$T_{(|p\rangle,|q\rangle)} = N\big|W_{\text{low}}^{|p\rangle} - W_{\text{low}}^{|q\rangle}\big|, \tag{5}$$

with

$$W_{\text{low}}^{|p\rangle} = N\Lambda^3 e^{2i\pi p/N}. \tag{6}$$

For 'elementary' domain walls, that join the pth and the $(p+1)$th vacua, (5) yields

$$T_{(|p\rangle,|p+1\rangle)} = 2N^2 \Lambda^3 \sin\frac{\pi}{N}. \tag{7}$$

In the large N or small string coupling limit, the tension scales as $N \sim 1/g_s$. This is the behaviour expected for a D-brane, and it was indeed argued in [19] that open strings can end on these domain walls. It is in principle very difficult to study these walls in the framework of gauge theory, because they belong entirely to the strong coupling regime of the theory. On the other hand, we see that in principle a simple description in terms of open strings exists. Of course, the explicit construction of these open strings remains an outstanding unsolved problem.

2.2. U(N) super Yang–Mills with one adjoint X

Let us now add one adjoint chiral multiplet X to the pure theory, with a tree level superpotential W such that

$$dW = g\,\text{tr}\big[(X - x_+)(X - x_-)\,dX\big]. \tag{8}$$

We now have $\frac{1}{6}N(N^2 + 11)$ vacua corresponding to various patterns of gauge and chiral symmetry breaking. For example, for $N = 2$, we have two chirally asymmetric vacua similar to the pure gauge theory case and corresponding to $\langle X\rangle_{\text{cl}} = x_+ I_2$, two other similar vacua corresponding to $\langle X\rangle_{\text{cl}} = x_- I_2$, and a Coulomb phase vacuum corresponding to $\langle X\rangle_{\text{cl}} = \text{diag}(x_+, x_-)$ and pattern of gauge symmetry breaking U(2) $\to$ U(1) $\times$ U(1).

We focus on the $2N$ vacua $|p, \pm\rangle$ corresponding to $\langle X\rangle_{\text{cl}} = x_\pm I_2$ and for which the low energy theory is the pure U(N) theory with scale Λ. We now have to consider two types of domain walls. The first type corresponds to walls interpolating between the vacua $|p, +\rangle$ and $|q, +\rangle$ or $|p, -\rangle$ and $|q, -\rangle$. These are the D-brane states of the pure low energy Yang–Mills

theory discussed before. The second type corresponds to walls interpolating between the vacua $|p, -\rangle$ and $|q, +\rangle$. These walls exist in the classical theory, and in particular they can be quantized semiclassically. It is straightforward to show that they are BPS and to compute their tension in the classical limit. The result shows that the tension scales as $N^2 \sim 1/g_s^2$. These domain walls are thus closed string solitons.

The spectrum of extended objects in the gauge theory with one adjoint is thus very similar to the spectrum in ordinary supersymmetric string theories: we have both D-branes and closed string solitons.

2.3. Interpolation

We now come to our main point. For concreteness, let us specialize to

$$W = \mathrm{tr}\left(\frac{1}{2}mX^2 + \frac{1}{3}gX^3\right), \tag{9}$$

for which

$$x_+ = 0, \qquad x_- = -m/g. \tag{10}$$

If the dynamically generated scale of the full theory is Λ_{u}, then the low energy scales $\Lambda_{\pm}$ in the vacua $|p, \pm\rangle$ are given by

$$\Lambda_{\pm}^3 = \pm m\Lambda_{\mathrm{u}}^2 = \pm\Lambda^3. \tag{11}$$

Let us introduce the dimensionless parameter

$$\lambda = \frac{8g^2\Lambda^3}{m^3}. \tag{12}$$

The effective superpotentials can be computed exactly [20] in the various vacua $|p, \pm\rangle$, by using for example the Dijkgraaf–Vafa matrix model prescription. The result is

$$W_{\mathrm{low}}^{|p,\pm\rangle} = \frac{2N}{3}\Lambda^3\frac{1}{\lambda}\left[1 \mp \left(1 - \lambda e^{2i\pi p/N}\right)^{3/2}\right]. \tag{13}$$

This is a highly non-perturbative formula. The weak coupling expansions, in powers of $\lambda \propto \Lambda^3$, are fractional instanton series. The leading terms yield

$$W_{\mathrm{low}}^{|p,\pm\rangle} = W_{\mathrm{low,cl}}^{|\pm\rangle} + N\Lambda_{\pm}^3 e^{2i\pi p/N} + \mathcal{O}\left(\Lambda^6/m^3\right), \tag{14}$$

consistently with (6). In particular, the tension of the solitonic branes in the classical limit is given by

$$T_{\mathrm{cl}} = N\left|W_{\mathrm{low,cl}}^{|+\rangle} - W_{\mathrm{low,cl}}^{|-\rangle}\right| = N^2\left|\frac{m^3}{6g^2}\right|, \tag{15}$$

which is proportional to N^2 as already advertised. More generally, the exact tensions of the various branes are given in terms of (13) by

$$T_{(|p,\pm\rangle,|q,\pm\rangle)} = N\left|W_{\mathrm{low}}^{|p,\pm\rangle} - W_{\mathrm{low}}^{|q,\pm\rangle}\right|, \tag{16}$$

$$T_{(|p,\pm\rangle,|q,\mp\rangle)} = N\left|W_{\mathrm{low}}^{|p,\pm\rangle} - W_{\mathrm{low}}^{|q,\mp\rangle}\right|. \tag{17}$$

A very important and general feature of the fractional instanton series for W_{low} is that they have a finite radius of convergence. This implies that analyticity is lost at special points and that W_{low} is a multivalued function of the parameters. In our example (13), $W_{\mathrm{low}}^{|p,\pm\rangle}$ is a two-sheeted function. Analyticity is lost at the square root branching point

$$\lambda = \lambda_p = e^{-2i\pi p/N}. \tag{18}$$

By going through the branch cut, we flip the sign of the square root, and thus permute $W_{\mathrm{low}}^{|p,+\rangle}$ and $W_{\mathrm{low}}^{|p,-\rangle}$. For example, by following the path in parameter space depicted in Fig. 1, we join smoothly a weakly coupled region in the vacuum $|p, +\rangle$ to a weakly coupled region in the vacuum $|p, -\rangle$. Exchanging a $|+\rangle$ vacuum with a $|-\rangle$ vacuum by smoothly varying the parameters is not an exploit. This can be achieved classically by permuting the roots x_+ and x_-. The highly non-trivial feature is that we are able to exchange $|p, +\rangle$ and $|p, -\rangle$, *without changing any of the other vacua $|q, \pm\rangle$ for $q \neq p$* [17]. This is a non-perturbative property, made possible by the fact that the branching points (18) and associated branch cuts for the vacua with different values of p are at different locations.
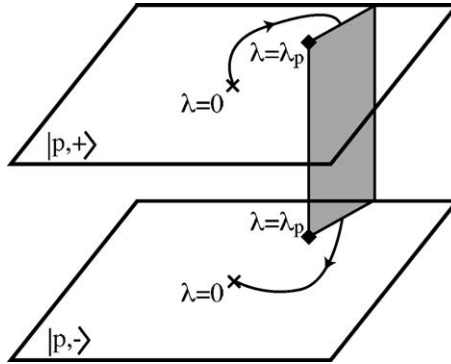
Fig. 1. A path in parameter space interpolating smooothly between the vacua $|p,+\rangle$ and $|p,-\rangle$. The other vacua $|q,\pm\rangle$ for $q \neq p$ are not changed.

The physics underlying the smooth interpolation

$$|p,+\rangle \rightarrow |p,-\rangle \tag{19}$$

depicted in Fig. 1 is best illustrated by looking at the domain walls. Consistently with the multivaluedness of the tension formulas (16), (17), (13), a D-brane is smoothly deformed into a solitonic brane,

$$\big(|p,+\rangle, |p+1,+\rangle\big) \rightarrow \big(|p,-\rangle, |p+1,+\rangle\big), \tag{20}$$

and vice versa. In other words, the multivaluedness implies that there are several possible classical limits, some for which the tension scales as N, and others for which the tension scales as N^2. Obviously, at strong coupling, the distinction between D-branes and solitons does not make sense. This is particularly striking at the point $\lambda = \lambda_p$, for which

$$T_{(|p,+\rangle, |p+1,+\rangle)}(\lambda_p) = T_{(|p,-\rangle, |p+1,+\rangle)}(\lambda_p) \tag{21}$$

$$\underset{N \to \infty}{\propto} \sqrt{N} \sim 1/\sqrt{g_s}. \tag{22}$$

We have discovered a new type of extended object in the theory at $\lambda = \lambda_p$, with a tension proportional to $\sqrt{N} \sim \sqrt{1/g_s}$ [17]. Other fractional powers can be achieved by looking at more complicated branching points [17].

2.4. Quantum parameter space

The physics discussed above is a basic consequence of a very general structure [17,21,22]. Let us consider the space, that we call the classical parameter space $\mathcal{M}_{cl}$, whose points are labeled by the parameters of the gauge theory *and* by a given vacuum. The space $\mathcal{M}_{cl}$ has many disconnected components, or sheets, each labeled by a vacuum. For example, in the theory (8), $\mathcal{M}_{cl}$ has $\frac{1}{6}N(N^2 + 11)$ components parametrized by λ. In the full quantum theory, W_{low} is multivalued and the associated branch cuts glue some of the sheets together. The resulting space is called the quantum parameter space $\mathcal{M}_q$ [17,21]. The effective superpotential W_{low} is always a single-valued function on $\mathcal{M}_q$.

For example, in the U(2) theory, $\mathcal{M}_{cl}$ has five disconnected components corresponding to the four confining vacua $|0,+\rangle$, $|1,+\rangle$, $|0,-\rangle$, $|1,-\rangle$ and the Coulomb vacuum $|C\rangle$. From the above discussion, we know that the sheets labeled by $|0,+\rangle$ and $|0,-\rangle$ on the one hand, and $|1,+\rangle$ and $|1,-\rangle$ on the other hand, are glued together. The Coulomb sheet cannot be smoothly connected to the other sheets, because it is in a different phase. However, there is a connection at points where monopoles become massless. The phase transition is triggered by the condensation of the massless monopoles through a standard Higgs mechanism in magnetic variables. Taking into account both the smooth interpolations through branch cuts and the phase transitions through massless monopole points, we get a fully connected quantum parameter space $\mathcal{M}_q$ [17].

The $\mathcal{N} = 1$ quantum parameter spaces are reminiscent of the $\mathcal{N} = 2$ quantum moduli spaces, but there are important differences. The most notable is that motion on parameter space is not associated with a massless scalar, which makes the $\mathcal{N} = 1$ models much more realistic. Moreover, the quantum corrections in $\mathcal{N} = 1$ have even more drastic effects than in $\mathcal{N} = 2$, since they change the topology of the classical space in addition to producing monopole singularities.

In all cases that have been studied so far, $\mathcal{M}_q$ turns out to be fully connected. It is not known whether this is a general property. For example, the rather intricate case of U(4), with classically eighteen disconnected components, was worked out in [21]. The result is depicted in Fig. 2. There are ten sheets in a confining phase and eight sheets in a Coulomb phase. The
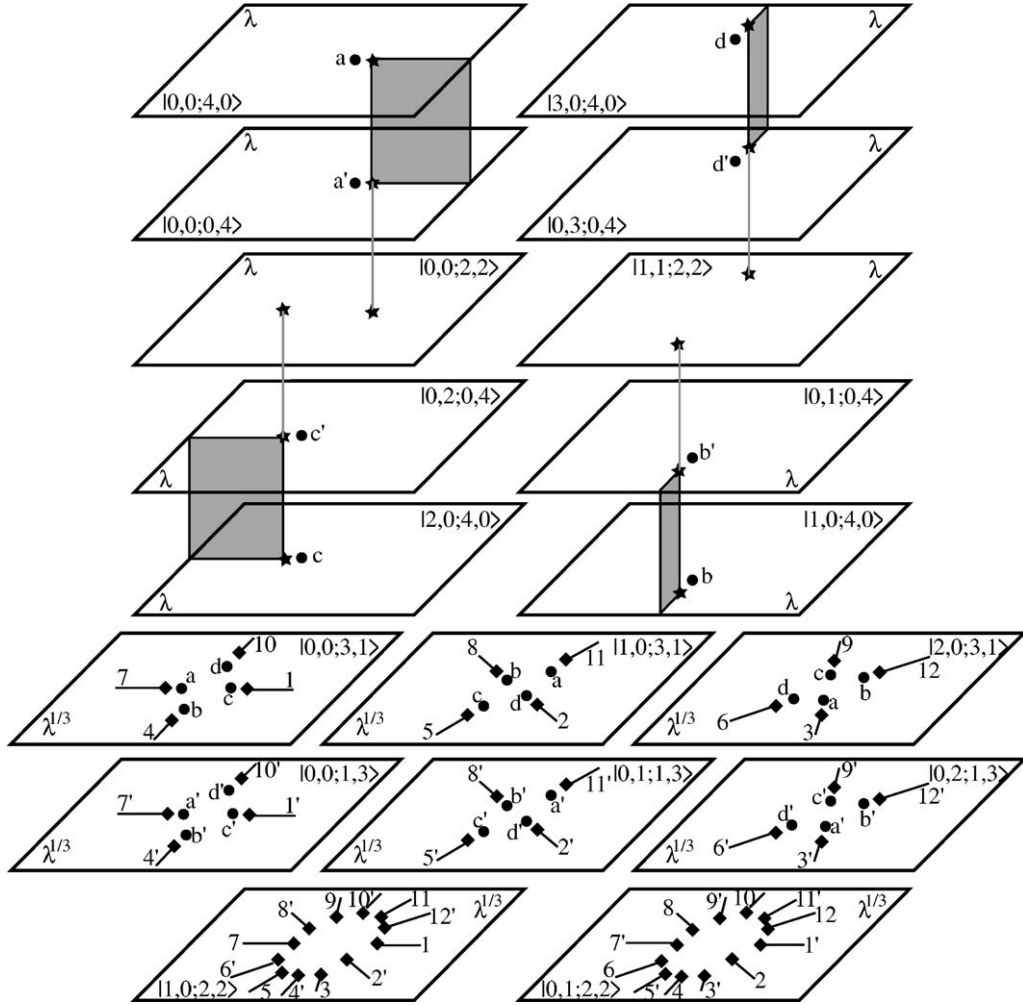
Fig. 2. Sketch of the quantum parameter space $\mathcal{M}_q$ for gauge group U(4). The sheets are parametrized by λ or $\lambda^{1/3}$. The vacua for which X has N_1 eigenvalues at zero and N_2 eigenvalues at $-m/g$ are denoted by $|k_1, k_2; N_1, N_2\rangle$, with $k_j \in \mathbb{Z}_{N_j}$. The dots, squares and stars represent massless monopole points (at $\lambda = (4/5)e^{-ik\pi/2}$), branching points (at $\lambda^{1/3} = (3^{1/4}/2^{1/6})e^{-ik\pi/6}$), or both (at $\lambda = e^{-ik\pi/2}$), respectively. Due to the complexity of the diagram, we have not been able to represent explicitly all the identifications between sheets. It is understood that singularities and branch cuts with the same label are identified.

unbroken gauge group in the Coulomb phase can be either U(2) × U(2) (two vacua) or U(3) × U(1) (six vacua). To show that these eight sheets can be smoothly connected, we can study the analytic structure of W_{low} as already explained. It is often more convenient to work with the derivatives of W_{low}, that yield the expectation values of various chiral operators. Remarkably, these vevs are the solutions of algebraic equations. In our case, it is natural to consider

$$u = -\frac{g}{m}\langle \text{tr } X\rangle. \tag{23}$$

The possible classical values of u in the Coulomb phase are either one (three vacua with unbroken gauge group U(3) × U(1)), two (the two U(2) × U(2) vacua) or three (three U(1) × U(3) vacua). This structure is manifest on the quantum equation satisfied by u [21],

$$(u-1)^3(u-2)^2(u-3)^3 + \frac{\lambda^4}{64} = 0. \tag{24}$$

Varying λ, it is straightforward to show that the eight roots of (24) can be permuted, and thus that the eight Coulomb sheets are glued together by branch cuts. The full $\mathcal{M}_q$ can be found by repeating this kind of argument and studying the possible phase

transitions [21]. In [22], it was emphasized that the smooth interpolations can connect vacua with different patterns of gauge symmetry breaking. In our framework, this is a direct consequence of (24) for example.

2.5. Conclusions

Our main result was to show that D-branes can be smoothly deformed into solitonic branes and vice versa. We have also discovered new extended objects, with tensions scaling as fractional powers of the string coupling.

D-branes are simple in string theory but are hard to study in gauge theory. On the other hand, we have explained that solitonic branes can be quantized semiclassically in gauge theory. This provides a microscopic description of solitonic branes using gauge theories, which is the counterpart of the microscopic description of D-branes using open strings. Since a microscopic description of solitonic branes in string theory has remained elusive, this approach is probably worth pursuing.

3. Planar diagrams and geometry

3.1. Calabi–Yau geometries

A famous non-compact Calabi–Yau geometry is the conifold $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$. By wrapping N D5 branes on the $\mathbb{P}^1$, we engineer the pure $\mathcal{N} = 1$, U(N) gauge theory [23] with a Yang–Mills coupling

$$g_{YM}^2 \sim \frac{1}{\text{vol}(\mathbb{P}^1)}. \tag{25}$$

To add adjoint chiral multiplets, we need to consider a more general local geometry

$$\mathcal{O}(-1 + M) \oplus \mathcal{O}(-1 - M) \to \mathbb{P}^1. \tag{26}$$

This space is Calabi–Yau because the first Chern class 2 of the $\mathbb{P}^1$ and $-1 - M - 1 + M = -2$ of the normal bundle compensate each other. It is not difficult to check, by direct analysis or by using the Riemann–Roch theorem, that the geometry (26) has a M-parameter family of holomorphic two-spheres. These M parameters yield M adjoint chiral multiplets X_i in the world-volume gauge theory of D-branes wrapped on one of these two-spheres.

The existence of an M-dimensional continuous family of $\mathbb{P}^1$s implies that there is no superpotential for the X_i. To turn on a W, we generalize the geometry (26) to a new Calabi–Yau $\mathcal{M}$, covered by two coordinate patches (z, w_1, w_2) and (z', w_1', w_2'), and defined by the transition functions [18]

$$\mathcal{M}: \begin{cases} z' = 1/z, \\ w_1' = z^{1-M} w_1, \\ w_2' = z^{1+M} w_2 + \partial_w E(z, w_1). \end{cases} \tag{27}$$

When $E = 0$, we find (26), but in general the perturbation E is a non-zero function of two complex variables. It can be Laurent expanded in terms of entire functions (often polynomials) E_j in the form

$$E(z, w) = \sum_{j=-\infty}^{+\infty} E_j(w) z^j. \tag{28}$$

From general arguments [24], it is known that perturbing the geometry (26) induces an obstruction to the space of versal deformations of the $\mathbb{P}^1$s. Holomorphic $\mathbb{P}^1$s can only exist at discrete values of the X_i that are solutions of M constraints (interestingly, the fact that the number of constraints is equal to the number of parameters is equivalent to the Calabi–Yau condition). In the case of (27), the constraints can be put in the form $dW = 0$ for a certain holomorphic function $W(X_1, \ldots, X_M)$. Turning on E is the most general known perturbation having this property. Obviously, W is the tree-level superpotential of the world volume gauge theory.

3.2. Superpotential

A formula to compute the superpotential in a variety of cases where branes are wrapped on two-cycles Σ was given by Witten in [19],

$$W(\Sigma) = \int_B \Omega + \text{constant}. \tag{29}$$

The chain B is such that $\partial B = \Sigma - \Sigma_0$ for some fixed two-cycle Σ_0, and Ω is the holomorphic three-form. It is not difficult to check that the equations $\mathrm{d}W = 0$ yield the correct conditions for Σ to be holomorphic. In our case, Σ is a two-sphere in $\mathcal{M}$, parametrized by the X_i, and computing the integral (29) yields $W(X_1, \ldots, X_M)$. This is perfectly good as long as we consider a single wrapped brane. When dealing with two or more branes, the X_i are non-commuting matrices and (29) becomes ambiguous. The correct procedure is then to write directly the conditions for supersymmetry to be preserved. The M equations for non-commuting variables so obtained can still be derived by extremizing a superpotential W [18]

$$W = \frac{1}{2\mathrm{i}\pi} \oint_{C_0} z^{-M-1} \operatorname{tr} E\left(z, \sum_{j=1}^{M} X_j z^{j-1}\right) \mathrm{d}z. \tag{30}$$

The contour integral is over a small loop C_0 encircling the origin. In the commutative limit, (30) and (29) are of course equivalent.

For example, in the case $M = 1$, (30) yields $W(X) = \operatorname{tr} E_1(X)$. In the case $M = 2$, we can consider an arbitrary superpotential $W(x, y) = \sum_{i,j \geqslant 0} a_{ij} x^i y^j$ in the commutative limit. The formula (30) gives the ordering prescription for any given monomial,

$$x^i y^j \mapsto \frac{i!\,j!}{(i+j)!} \oint_{C_0} \frac{\mathrm{d}z}{2\mathrm{i}\pi} z^{-1-j} (X + Yz)^{i+j}. \tag{31}$$

This is the standard Weyl ordering.

The superpotential (30) includes as special cases all the previously engineered superpotentials for gauge theory with adjoints (see for example [25]). The associated multi-matrix integrals (2) are highly non-trivial and far beyond the grasp of ordinary matrix model techniques. Fortunately, we now have an entirely new point of view on the problem: the solution of the model amounts to finding the closed string background $\hat{\mathcal{M}}$ dual to the brane theory on $\mathcal{M}$. We shall now explore this strategy, by describing the two main steps required in the computation of the space $\hat{\mathcal{M}}$.

3.3. Blowing down

Holomorphic two-spheres in the geometry (27) are associated with the critical points of the superpotential (30). For a generic Morse critical point, the adjoint fields are massive and can be integrated out. This means that the normal bundle to the associated $\mathbb{P}^1$ is $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ as in the case of the pure $\mathcal{N} = 1$ theory. More generally, if the corank of the Hessian of W at the critical point is r, we expect the normal bundle to be $\mathcal{O}(-1+r) \oplus \mathcal{O}(-1-r)$, corresponding to a massless theory with r adjoint multiplets. This was checked explicitly in the case $M = 2$ in [18]. When $r \leqslant 2$, the low energy theory is asymptotically free, and we may expect that quantum mechanically the $\mathbb{P}^1$ shrinks to zero size. This heuristic RG argument is in beautiful agreement with a theorem by Laufer [26] that states that shrinkable spheres with normal bundle $\mathcal{O}(-1+r) \oplus \mathcal{O}(-1-r)$ must have $r = 0, 1$ or 2.

The blow down map

$$\pi : \mathcal{M} \to \mathcal{M}_0 \tag{32}$$

is a birational isomorphism except on the $\mathbb{P}^1 \subset \mathcal{M}$ that are mapped onto singular points of $\mathcal{M}_0$. Blowing down is a mathematically precise way to implement the heuristic picture of shrinking spheres. The calculation of the blow down map, for given $\mathcal{M}$, is a non-trivial problem. In the simplest cases, it amounts to finding four globally holomorphic functions $\pi_i(z, w_1, w_2) \in \mathbb{C}$ on $\mathcal{M}$ that maps the $\mathbb{P}^1$ onto points. Global holomorphicity is checked by using the transition functions (27). The four functions π_i depend on three variables and satisfy in general an algebraic equation that defines the singular variety $\mathcal{M}_0$ as a hypersurface in $\mathbb{C}^4$. Several explicit examples can be found in [18].

In the geometric transition picture, the singular space $\mathcal{M}_0$ describes the classical $S \to 0$ matrix model (2), or equivalently the classical equations of motion

$$\mathrm{d}W = \operatorname{tr} \mathrm{d}V(X_i) = 0. \tag{33}$$

When the matrices X_i are commuting (one-dimensional representations), (33) reduces to a set of algebraic equations. By construction of the blow down map, these solutions are automatically associated with singular points in $\mathcal{M}_0$. More generally, the correspondence with the matrix model implies that the higher-dimensional representations of (33) must also be associated with singular points on $\mathcal{M}_0$. This means that classical algebraic geometry must know about the non-commutative structure that emerges when considering D-branes.

Let us illustrate this point on a non-trivial example [18]. We consider the geometry

$$\mathcal{M}: \begin{cases} z' = 1/z, \\ w_1' = z^{-1} w_1, \\ w_2' = z^3 w_2 + w_1^2 - z G(w_1^2/z^2) - z^2 F_1(w_1^2) - z^2 w_1 F_2(w_1^2). \end{cases} \tag{34}$$

The corresponding superpotential is

$$W(X, Y) = \mathrm{tr}\big(XY^2 + A(X) + B(Y)\big), \tag{35}$$

with $A'(X) = -F_1(X^2) - X F_2(X^2)$ and $B'(Y) = -G(Y^2)$. The classical equations of motion are

$$\{X, Y\} = G\big(Y^2\big), \quad Y^2 = F_1\big(X^2\big) + X F_2\big(X^2\big). \tag{36}$$

These equations can be analysed by noting that X^2 and Y^2 are Casimir operators. We find that there are $\max(\deg A' + 2, 2 \deg A' \deg G)$ one-dimensional irreducible representations and $[\frac{1}{2}(\deg A' - 1)]$ two-dimensional irreducible representations (for which X and Y are linear combinations of Pauli matrices). On the other hand, the blow down map can be explicitly constructed [18] and the singular Calabi–Yau is found to be

$$\mathcal{M}_0: x_2\big[(x_2 - F_1(x_1))^2 - x_1 F_2^2(x_1)\big] = x_4^2 - x_1 x_3^2 - G(x_2)\big[(x_2 - F_1(x_1))x_3 + x_4 F_2(x_1)\big]. \tag{37}$$

It is not difficult to check that $\mathcal{M}_0$ has singular points for both one *and* two-dimensional representations.

The fact that the full set of solutions of matrix equations can be encoded in a set of algebraic equations describing the singularities of a Calabi–Yau is a remarkable feature. In all the cases that have been worked out, the solutions to the matrix equations can be characterized by a set of Casimir operators. These Casimirs turn out to be associated with some coordinates on $\mathcal{M}_0$. The special values of the Casimirs labeling the solutions then match the values of the coordinates at the singular points. For example, in the case (37), the coordinate x_1 is associated with the Casimir X^2 and the coordinate x_2 with the Casimir Y^2.

3.4. Deforming

The full $S \neq 0$ matrix model (2) is described by a deformed non-singular geometry $\hat{\mathcal{M}}$ whose equation is obtained from the equation for $\mathcal{M}_0$ by adding certain monomials $\sum c_{abcd} x_1^a x_2^b x_3^c x_4^d$. The allowed monomials must be such that the divergences in the period integrals $\oint \Omega$ of the holomorphic three-form over non-compact cycles are either S-independent or linear in S and logarithmic. This condition is the mathematical counterpart of the renormalization properties of the $\mathcal{N} = 1$ Yang–Mills theories. After implementing this constraint, and taking into account possible coordinate redefinitions, we obtain a certain number $\mathcal{N}_{\mathrm{geo}}$ of independent and freely adjustable coefficients c_{abcd} that parametrize the geometry $\hat{\mathcal{M}}$.

The parameters c_{abcd} should not be confused with the parameters in the superpotential or equivalently in the original geometry (27). They have a very natural interpretation in terms of the matrix model. The matrix integrals (2) are defined by summing over planar Feynman diagrams. The Feynman rules are themselves defined by expanding around a particular solution of the classical equations of motion. The most general classical solution is the direct sum of the irreducible representations of (33), and is thus characterized by parameters, called filling fractions, giving the fraction of eigenvalues in given irreducible representations. The number $\mathcal{N}_{\mathrm{alg}}$ of filling fractions that need to be specified is thus equal to the number of irreducible representations of (33). Consistency requires that there should exist a one-to-one mapping between the filling fractions and the c_{abcd}, and thus that

$$\mathcal{N}_{\mathrm{geo}} = \mathcal{N}_{\mathrm{alg}}. \tag{38}$$

This is a profound equality relating classical geometry and non-commutative algebra: the left-hand side is computed by studying the 'renormalizable' deformation of a certain singular Calabi–Yau space, and the right-hand side is computed by working out the representation theory of a certain matrix algebra. This correspondence follows rather straightforwardly from the physical picture, but remains to be understood at a deeper mathematical level. It was checked explicitly in a variety of non-trivial examples in [18]. For example, in the theory (34), (35), (37), detailed calculations show that

$$\mathcal{N}_{\mathrm{alg}} = \max(\deg A' + 2, 2 \deg A' \deg G) + \big[(\deg A' - 1)/2\big] = \mathcal{N}_{\mathrm{geo}} \tag{39}$$

as expected.

3.5. The solution

In the case of the one-matrix model $W = \mathrm{tr}\, V(X)$, it is well known that the deformed geometry takes the simple form

$$\hat{\mathcal{M}}: x_4^2 = x_3^2 + x_2^2 - V'(x_1)^2 + S\Delta(x_1), \tag{40}$$

where Δ is a polynomial of degree $\deg V' - 1$ describing the deformation of $\mathcal{M}_0$ (in this case it is obvious that $\mathcal{N}_{\text{alg}} = \deg V'$; it is also straightforward to show that $\mathcal{N}_{\text{geo}} = \deg V'$ by computing the periods $\oint \Omega$). It is useful to think about (40) as a fibration over a base parametrized by x_1. The fiber F_{x_1} at x_1 is the simplest ALE space $\mathbb{C}^2/\mathbb{Z}_2$, and contains a single holomorphic two-sphere S^2. The discontinuity of the resolvent

$$g^X(x) = S\left\langle \frac{\text{tr}}{n} \frac{1}{x - X} \right\rangle \tag{41}$$

across its branch cuts is then simply given by

$$g^X(x + \mathrm{i}\varepsilon) - g^X(x - \mathrm{i}\varepsilon) = \int\limits_{S^2 \subset F_x} \mathrm{i}_X \Omega, \tag{42}$$

where i_X denotes the interior product with respect to the vector field $\partial/\partial x_1$.

More general multi-matrix models can always be reduced to one matrix models with a complicated effective potential after integration over all but one matrix. This shows that the structure just described is still relevant, but the fibers F_x are typically much more complicated than $\mathbb{C}^2/\mathbb{Z}_2$, and contains many two-spheres (as a consequence of the multivaluedness of the effective potential). The basic formula (42) is still valid [18] and can be used to find explicitly the resolvents of the various matrices of multi-matrix models. For example, in the case of (35), it is shown with this method that g^X satisfies a degree $\max(3, 2 + 2 \deg G)$ algebraic equation and that g^Y lies at the intersection of d quadrics in a d-dimensional auxiliary space [18]. It appears to be very difficult to obtain these results using standard matrix model technology. Nevertheless, in the special case where G is a constant, the loop equations can be solved [18] and full agreement is found with the geometry.

3.6. Conclusions

We have a rich interplay between algebraic geometry and matrix models, that suggests many non-trivial results. There are many open problems: can we understand the special properties of the matrix models with potentials (30) (irreducible representations of the classical equations of motion, loop equations, etc. . .)?; does the potential (30) describes the most general matrix model that can be engineered in string theory?; can we give a direct mathematical proof of the relation between the singularity structure of $\mathcal{M}_0$ and the representations of (33)?; and so on. The most important question is probably to understand whether the blow down map for the geometry (27) always exists. This is a hard mathematical problem [27]. If it does, and if we can find an algorithm to compute it, then we have solved a huge class of multi-matrix models (30). If it does not, then the picture of geometric transitions cannot deal with our examples in general, with fundamental consequences for our understanding of the structure of $\mathcal{N} = 1$ vacua in string theory.

4. Homework

4.1. Exercises

(a) Prove or disprove that the $\mathcal{N} = 1$ quantum parameter spaces are always connected.
(b) Describe the correct microscopic description of solitonic branes using gauge theory.

4.2. Problem

Consider the geometry (27).

 (i) Find the blow down geometry. Compare with the algebra $\mathrm{d}W = 0$ where W is given by (30).
 (ii) Find the deformed geometry. Check (38).
(iii) Compute the resolvents from the geometry. Compare with the loop equations.

Acknowledgements

References

[1] C. Montonen, D. Olive, Phys. Lett. B 72 (1977) 117.

[2] A. Sen, Phys. Lett. B 329 (1994) 217, hep-th/9402032.

[3] N. Seiberg, E. Witten, Nucl. Phys. B 426 (1994) 19, erratum Nucl. Phys. B 430 (1994) 485, hep-th/9407087.

[4] A. Klemm, W. Lerche, S. Yankielowicz, S. Theisen, Phys. Lett. B 344 (1995) 169, hep-th/9411048.

[5] R.Y. Donagi, Seiberg–Witten integrable systems, alg-geom/9705010.

[6] A. Klemm, W. Lerche, P. Mayr, C. Vafa, N. Warner, Nucl. Phys. B 477 (1996) 746, hep-th/9604034;
    S. Katz, A. Klemm, C. Vafa, Nucl. Phys. B 497 (1997) 173, hep-th/9609239.

[7] A. Giveon, D. Kutasov, Rev. Mod. Phys. 71 (1999) 983, hep-th/9802067.

[8] N. Seiberg, E. Witten, Nucl. Phys. B 431 (1994) 484, hep-th/9408099.

[9] A. Bilal, in: Quantum Fields and Quantum Space–Time, Cargèse, 1996, p. 21, hep-th/9601007;
    L. Alvarez-Gaumé, S.F. Hassan, Fortschr. Phys. 45 (1997) 159, hep-th/9701069.

[10] K. Intriligator, N. Seiberg, Lectures on supersymmetric gauge theories and electric-magnetic duality, Nucl. Phys. Proc. Suppl. 55 (1997) 1, hep-th/9509066.

[11] F. Ferrari, A. Bilal, Nucl. Phys. B 469 (1996) 387, hep-th/9602082;
    F. Ferrari, Nucl. Phys. B 516 (1998) 175, hep-th/9706145.

[12] O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri, Y. Oz, Phys. Rep. 323 (2000) 183, hep-th/9905111.

[13] F. Ferrari, Lectures on supersymmetric gauge theories, matrix models and geometric transitions, NEIP-03-003, LPTENS-04/43, submitted for publication, hep-th/0502nnn.

[14] I. Antoniadis, E. Gava, K.S. Narain, T.R. Taylor, Nucl. Phys. B 413 (1994) 162, hep-th/9307158;
    M. Bershadsky, S. Cecotti, H. Ooguri, C. Vafa, Commun. Math. Phys. 165 (1994) 311, hep-th/9309140.

[15] R. Gopakumar, C. Vafa, Adv. Theor. Math. Phys. 3 (1999) 1415, hep-th/9811131;
    C. Vafa, J. Math. Phys. 42 (2001) 2798, hep-th/0008142;
    F. Cachazo, K. Intriligator, C. Vafa, Nucl. Phys. B 603 (2001) 3, hep-th/0103067.

[16] R. Dijkgraaf, C. Vafa, hep-th/0208048.

[17] F. Ferrari, Phys. Rev. D 67 (2003) 85013, hep-th/0211069.

[18] F. Ferrari, Adv. Theor. Math. Phys. 7 (2003) 619, hep-th/0309151.

[19] E. Witten, Nucl. Phys. B 507 (1997) 658, hep-th/9706109.

[20] F. Ferrari, Nucl. Phys. B 648 (2002) 68, hep-th/0210135.

[21] F. Ferrari, Phys. Lett. B 557 (2003) 290, hep-th/0301157.

[22] F. Cachazo, N. Seiberg, E. Witten, J. High Energy Phys. 0302 (2003) 042, hep-th/0301006.

[23] J.M. Maldacena, C. Nuñez, Phys. Rev. Lett. 86 (2001) 588, hep-th/0008001.

[24] M. Namba, Tôhoku Math. J. 24 (1972) 581.

[25] F. Cachazo, S. Katz, C. Vafa, hep-th/0108120.

[26] H.B. Laufer, Ann. Math. Stud. 100 (1981) 261.

[27] J. Kollár, private communication.