Lectures on nonperturbative aspects of supersymmetric gauge theories*

Nikita A Nekrasov

Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440 Bures-sur-Yvette, France

E-mail: nikita@ihes.fr

Received 9 December 2004
Published 29 March 2005
Online at stacks.iop.org/CQG/22/S77

Abstract
An introduction to (supersymmetric) gauge theories and their applications is followed by a discussion of the recent solution of $\mathcal{N} = 2$ theories by means of direct instanton counting. The gauge/string duality in the context of BPS states and topological strings is reviewed, and the quantum gravitational foam picture emerging out of gauge-theoretic considerations is explained.

Introduction

Gauge theory accurately describes the physics of elementary particles over a wide energy scale. The perturbative expansion is adequate for non-Abelian gauge theories at high energies. For low-energy physics the effective gauge coupling grows strong in the absence of gauge symmetry breaking. As a consequence, non-perturbative effects become important. In particular, low-energy physics is affected by the gauge theory instantons, whose careful treatment is necessary to get non-singular low-energy effective theory. In the realistic gauge theories, such as $SU(3)$ gauge theory, the instanton effects are mixed with those of anti-instantons, and also are dressed by various perturbative corrections. In the supersymmetric gauge theories the situation is simpler, as the perturbation expansion about instanton background is essentially trivial, and there are interesting correlation functions where instantons and anti-instantons do not mix.

In these lectures we shall explain the recent solution of the $\mathcal{N} = 2$ gauge theory by means of direct instanton counting. We shall also discuss the possible applications to the less symmetric $\mathcal{N} = 1$ case.

We shall also discuss gauge theory/string theory duality in the context of the theories above. It turns out that many of the exactly calculable on the gauge theory side quantities have a topological string interpretation.

* Lectures given at the RTN Winter School on Strings, Supergravity and Gauge Fields, which took place in Barcelona, on 12–16 January 2004.

1 On leave of absence from Institute for Theoretical and Experimental Physics, Moscow, 117259, Russia.
We shall conclude with a quantum gravitational foam picture which emerges out of the
gauge theory considerations.

1. Supersymmetric gauge theories

In the first lecture we shall briefly recall the necessity of gauge theories, motivate the
introduction of the supersymmetry (susy), discuss extended susy, and also recall a few advanced
topics, such as twisting.

1.1. Why gauge theory

We shall not give an extensive apology for gauge theories. Suffice it to say that gauge theories
are the most up-to-date accurate descriptions of all the known fundamental interactions except
for gravity. For more details, including renormalizability, asymptotic freedom and other
important aspects, we refer the reader to [27, 31, 65].

1.2. Perturbative expansion versus non-perturbative effects

In quantum field theory with coupling constant $g$, where the Lagrangian can be written as

$$L = L_{\text{free}} + g L_{\text{int}}$$

with quadratic non-degenerate term $L_{\text{free}}$, the correlation functions can be expanded in $g$:

$$\langle O_1(x_1) \cdots O_n(x_n) \rangle = \langle O_1(x_1) \cdots O_n(x_n) \rangle_{\text{free}} + g \int d^d y \langle O_1(x_1) \cdots O_n(x_n) L_{\text{int}}(y) \rangle_{\text{free}} + \frac{g^2}{2!} \int d^d y_1 d^d y_2 \langle O_1(x_1) \cdots O_n(x_n) L_{\text{int}}(y_1) L_{\text{int}}(y_2) \rangle_{\text{free}} + \cdots.$$  

(2)

However, this expansion is formal, and in fact does not make sense, as generally the free
theory correlation functions have divergences as $y \to x_i, i = 1, \ldots, n$, and also as $y_1 \to y_2$,
etc. These divergences are taken care of by allowing the operators $O_i$ and the interaction
Lagrangian $L_{\text{int}}$ to depend on $g$ in a non-trivial way. The dependence on $g$ involves the
ultraviolet cut-off, which also enters the proper definition of the integrals in equation (2). The
procedure is known as regularizat ion and renormalization.

If the theory is regularized and renormalized the terms of fixed order in the coupling
constant, e.g. $\approx g^n$, are finite. They depend, however, on the details of the renormalization,
i.e. on the normalization point $\mu$.

In addition to these perturbative contributions to the correlation function, one has various
non-perturbative effects. They come from the fact that the action $S = L_{\text{free}} + g L_{\text{int}}$ may
have several critical points, and in general all of these contribute to the path integral. The
contribution of a non-trivial critical point $\varphi_*$ goes like

$$Z_{\varphi_*} = e^{-\frac{1}{2} S(\varphi_*)} \frac{\text{Pfaff}(\delta^2 S_F)}{\sqrt{\text{Det} \delta^2 S_B}} (1 + A_2(\varphi_*) g^2 + \cdots)$$

(3)

where $\delta^2 S_F$ is the fermionic quadratic part of the action expanded about $\varphi_*$, and $\delta^2 S_B$ is the
bosonic one.

In the case of pure gauge theory, with the action

$$S = \frac{1}{4 \pi^2 g^2} \int \text{Tr} \ F \wedge \ast F$$

(4)
by rescaling the gauge field $A \rightarrow g \cdot A$ one brings the Lagrangian to the form
\[ L = \|dA\|^2 + g \text{Tr} dA \wedge \star[A, A] + g^2 \text{Tr}[A, A] \wedge \star[A, A] \] (5)
which can be used (after gauge fixing) to write the perturbative amplitudes. However, the action (4) possesses non-trivial solutions of equations of motion with finite action, and these solutions (instantons and anti-instantons) lead to extra terms in the low-energy effective action. To the standard Yang–Mills action (4) one can add a topological term, which does not affect equations of motion, nor the perturbation theory, which measures the instanton charge:
\[ \frac{i}{8\pi^2} \int \text{Tr} F \wedge F. \] (6)
This term depends on the extra parameter, $\Theta$, also known as theta, or the instanton angle. The physics is periodic with respect to the shifts $\Theta \rightarrow \Theta + 2\pi$, as (6) enters gauge theory amplitudes only in the exponentiated form, while $\frac{i}{8\pi^2} \int \text{Tr} F \wedge F$ is quantized².

As was shown by Callan, Dashen and Gross [13], and also by ’t Hooft [32], the effects of the instantons can be systematically taken into account, at least for the single instanton. The correction to the effective action, which comes from the single instanton, has the form of an integral over the moduli space of centred instantons. The remaining integral over the position of the instanton is the integral which converts the Lagrangian into the action.

So, on general grounds, one can write³
\[ L_{\text{eff}} = L_{\text{cl}} + g^2 L_{\text{1-loop}} + g^4 L_{\text{2-loop}} + \cdots + e^{-\frac{1}{8\pi^2} \Theta} L_{\text{1-inst}}(g). \] (7)

1.3. Wilsonian effective action

There are several ways of packaging the information in quantum field theory. One is the so-called effective action, the generating function $\Gamma(\phi, \mu)$ of all 1PI diagrams, which is the Legendre transform of the free energy $W[J, \mu]$ as a functional of sources $J$. Both depend on the normalization point $\mu$ which is used to define the renormalized vertices. The momenta in the loop integrals defining $W[J, \mu]$ are integrated over from zero to the ultraviolet cut-off $\Lambda_{\text{UV}}$, which is then taken to infinity, while at the same time adjusting the bare coupling constants.

This object is almost impossible to calculate, and it contains too much information. Another object is the Wilsonian effective action. It is defined, roughly, by integrating all the fields of momenta of magnitude higher than a given value (and less the ultraviolet cut-off). This value, $\mu$, is the parameter of the Wilsonian action. The Wilsonian action $S_{\text{eff}}[\phi^{\text{slow}}, \mu]$ is the functional of the remaining, unintegrated fields. By construction, they are slowly varying. For the theories with massive particles only the Wilsonian action and the 1PI effective actions are quite similar, for $\mu$ much less than the mass scale. However, in the presence of massless particles the Wilsonian action is much better defined; it does not suffer from the infrared divergences present in $\Gamma[\phi, \mu]$. These divergences lead to various unwanted effects.

The Wilsonian action can be expanded in $\frac{1}{\mu^2} \partial^2$, i.e. in the derivatives of the fields.

1.4. Supersymmetry

Conventional gauge theory is defined on flat Minkowski space. Lorentz invariance allows us to define it on a general curved manifold. However, one can define it on more exotic manifolds.

---

² We use the standard notation for simple Lie groups: $\text{Tr} = \frac{1}{h} \text{Tr}_{\text{adj}}$.
³ $\cos \Theta L_{1-\text{inst}}$ should really read $\frac{1}{2} (e^{i\Theta} L_{1-\text{inst}} + e^{-i\Theta} L_{\text{anti-inst}})$. 
namely on noncommutative spaces. The simplest non-trivial example of such a manifold is the superspace.

Gauge theory on the superspace is the supersymmetric (susy) gauge theory. The fields of the theory are functions of the ordinary, bosonic, four spacetime coordinates $x^\mu$, and the odd, fermionic, coordinates $\vartheta^\alpha$, $\alpha = 1, 2$ (for minimal susy), and their conjugates $\bar{\vartheta}^\alpha$. The minimal supersymmetry in four dimensions is generated by two complex supercharges $Q_\alpha, \bar{Q}_\alpha$, which form the usual Poincaré superalgebra:

\[
\{ Q_\alpha, \bar{Q}_\beta \} = -2i \sigma^\mu_{\alpha \beta} \partial_\mu, \quad \{ Q_1, Q_2 \} = 0.
\]  

(8)

The superspace meaning of the supercharges (8) is the superanalogue of the ordinary translations $\partial_\mu = \frac{\partial}{\partial x^\mu}$:

\[
Q_\alpha = \frac{\partial}{\partial \vartheta^\alpha} - i \sigma^\mu_{\alpha \beta} \vartheta^\beta \frac{\partial}{\partial x^\mu}, \quad \bar{Q}_\alpha = \frac{\partial}{\partial \bar{\vartheta}^\alpha} - i \sigma^\mu_{\alpha \beta} \bar{\vartheta}^\beta \frac{\partial}{\partial x^\mu}.
\]

(9)

To write down supersymmetric Lagrangians one makes use of the superspace covariant derivatives:

\[
D_\alpha = \frac{\partial}{\partial \vartheta^\alpha} + i \sigma^\mu_{\alpha \beta} \vartheta^\beta \frac{\partial}{\partial x^\mu}, \quad \bar{D}_\alpha = \frac{\partial}{\partial \bar{\vartheta}^\alpha} + i \sigma^\mu_{\alpha \beta} \bar{\vartheta}^\beta \frac{\partial}{\partial x^\mu}.
\]

(10)

The remarkable property of the supersymmetric theories is the cancellation of various quantum corrections thanks to the competition between the bosonic and fermionic loops. In particular, vacuum energy (if the susy is not unbroken) is exactly zero.

The more supersymmetry has the theory, the softer are its ultraviolet properties. The maximal rigid supersymmetry in four dimensions, $N = 4$, is so constrained that the theory has no parameters except for the choice of the gauge group $G$, and the (complexified) gauge coupling for each simple factor of $G$. Moreover, these couplings do not get renormalized.

Less constrained is the $N = 2$ supersymmetry. In this case one has two sets of $N = 1$ supercharges. The superalgebra is generated by $Q^I_\alpha, I = 1, 2$, and $\bar{Q}^I_{\bar{\alpha}}, I = 1, 2$. The indices $i, j$ label the two components of the two-dimensional representation of the global symmetry group $SU(2)_I$ which rotates two $N = 1$ subalgebras of $N = 2$ superalgebra among themselves:

\[
\{ Q^I_\alpha, \bar{Q}^J_{\bar{\alpha}} \} = -2i \delta^I_J \epsilon^\mu_{\alpha \beta} \partial_\mu, \quad \{ Q^I_\alpha, Q^J_\beta \} = \epsilon^{IJ} \delta_{\alpha \beta} Z
\]

where $Z$ is the central element.

1.4.1. Representations of the supersymmetry algebra. It is straightforward to work out the low-spin representations of the algebras (11), (8). In the minimal susy case the representations crucially depend on whether the particle it describes is massive or massless.

In the massive case we can choose the Lorentz frame where $P_0 = M, P_i = P_2 = 0$. Writing $Q_\alpha = \sqrt{2M} b_\alpha, \bar{Q}_\alpha = \sqrt{2M} \bar{b}_\alpha$, we get $[b_\alpha, \bar{b}^I_{\bar{\alpha}}] = \delta_{\alpha \bar{\beta}}, \{ b_1, \bar{b}_2 \} = 0$. In other words, we have two fermionic oscillators. The minimal representation has four helicity states: the vacuum $\{ \downarrow \}$, which has spacetime helicity $-\frac{1}{2}$, two zero helicity states $\bar{b}_I^1 | \downarrow \rangle$ and $\bar{b}_2^1 | \downarrow \rangle$, and one $+\frac{1}{2}$ helicity state $b_1^1 | \downarrow \rangle$.

In the massless case we can choose the Lorentz frame with $P_0 = P_3 = E, P_i = P_2 = 0$. In this case we should introduce the oscillators as follows: $Q_1 = \sqrt{2E} b, \bar{Q}_1 = \sqrt{2E} \bar{b}^I, \{ b, \bar{b}^I \} = 1$ and $Q_2, \bar{Q}_2$ anticommuting with everything. The minimal representation would be two dimensional, with the states $\{ \downarrow \}$ and $b_1^1 | \downarrow \rangle$. The CPT invariant multiplet will have two such conjugate representations in it. Depending on the lowest helicity of the state $\{ \downarrow \}$ in each of them we would get the chiral multiplet, with helicities $-\frac{1}{2}, 0, 0, +\frac{1}{2}$, or the vector multiplet, with helicities $-1, -\frac{1}{2}, +\frac{1}{2}, +1$. 


1.4.2. Superfields. The chiral superfield is a compact way of packaging the chiral multiplet representation. Introducing the chiral coordinates \( y^\mu = x^\mu + i \theta \sigma^\mu \bar{\theta} \), we write

\[
\Phi(y, \theta) = \phi(y) + \sqrt{2} \theta \psi(y) - \theta \theta f.
\]

(12)

In the physical space coordinates the components of the chiral superfield are

\[
\Phi_1(y, \theta) = \phi(x) + \sqrt{2} \theta \psi(x) - \theta \theta f.
\]

(13)

The vector superfield is longer. In the so-called Wess–Zumino gauge it can be expanded as

\[
V = \theta \sigma^\mu \bar{\theta} A_\mu(x) + i \theta^2 \bar{\theta} \lambda(x) - i \bar{\theta}^2 \theta \lambda(x) + \frac{1}{2} \theta^2 \bar{\theta}^2 D(x).
\]

(14)

The vector superfield is real: \( V^\dagger = V \). This field transforms non-trivially under gauge transformations:

\[
V \rightarrow V + \phi + \phi^\dagger,
\]

for real \( \alpha \). Its non-Abelian generalization has all components in the Lie algebra \( g \) of the gauge group \( G \). The gauge transformations look as follows:

\[
e^{2iV} \rightarrow e^{i\phi^\dagger} e^{2iV} e^{-i\phi},
\]

(15)

To write the gauge-invariant Lagrangian one uses the chiral-looking gauge-covariant superfield:

\[
W_\alpha = -\frac{1}{4} \bar{D}^2 (e^{2iV} D_\alpha e^{2iV}), \quad \bar{W}_\dot{\alpha} = \frac{1}{4} D^2 (e^{2iV} \bar{D}_{\dot{\alpha}} e^{-2iV}),
\]

(16)

which expands as follows:

\[
W_\alpha = -2i \lambda_\alpha + 2 \partial_\alpha D(y) + 2i (\sigma^{\mu \nu} \theta)_{\alpha} F_{\mu \nu} + \partial^2 (\sigma^\mu \bar{D}_\mu \lambda(y))_\alpha.
\]

(17)

1.4.3. Susy Lagrangians. The simplest supersymmetric theory is the \( \mathcal{N} = 1 \) susy sigma model. It describes maps of \( R^{4|4} \) into some Kahler manifold \( X \). The Lagrangian is written in terms of the following geometric data: the Kahler metric on \( X \), which locally can be described using Kahler potential \( K(z, \bar{z}) \), where \( z^I \) are local holomorphic coordinates on \( X \), and the holomorphic function \( W(z) \) on \( \mathcal{X} \), called the superpotential. Here is the Lagrangian:

\[
L_{\mathcal{N} = 1\text{-model}} = \int d^2 \theta d^2 \bar{\theta} K(\Phi, \Phi^\dagger) + \int d^2 \theta W(\Phi) + \int d^2 \bar{\theta} \bar{W}(\Phi^\dagger).
\]

(18)

The simplest supersymmetric gauge theory Lagrangian has the following structure:

\[
L_{\mathcal{N} = 1\text{-YM}} = \operatorname{Im} \left( \int d^2 \theta \tau \operatorname{Tr} W_a W^a \right)
\]

(19)

where

\[
\tau = \frac{\Theta}{2\pi} + \frac{4\pi i}{g^2}
\]

is the complexified gauge coupling \( g \) being the ordinary Yang–Mills gauge coupling and \( \Theta \) the instanton angle.

One can also add some charged matter, with coupling to the Yang–Mills multiplet. The setup is the following: the Lie group \( G \) acts on \( X \) by isometries. Its complexification \( G_C \) acts by holomorphic transformations:

\[
z \mapsto g \cdot z.
\]

(20)

The compact subgroup \( G \) preserves the Kahler metric and, in particular, the Kahler form \( \omega = \partial \bar{\partial} K \). In this context one has the so-called moment map

\[
\mu : X \rightarrow g^*_C
\]

(21)

which plays an important role in symplectic geometry.
The gauge-invariant $\sigma$-model Lagrangian:

$$L_{\text{matter}} = \int d^2\vartheta \, d^2\bar{\vartheta} \, K(\Phi^1, (e^{2V} \cdot \Phi)) + \left( \int d^2\vartheta \, W(\Phi) + c.c. \right)$$

(22)

where the Kahler potential is analytically continued to the holomorphic function on $X \times \bar{X}$ (actually, to the formal neighbourhood of the diagonal). The superpotential $W$ should be $G_C$-invariant.

Renormalizable field theories are obtained for $X$ a vector space, which is some representation of $G$. This representation has to be real, $X^* \approx X$, to avoid gauge anomaly.

Effective field theories for supersymmetric theories are described by supersymmetric Lagrangians. The sigma model part of such a Lagrangian may very well correspond to nonlinear $X$, which would nevertheless have to be Kahler (for $N = 1$), hyper-Kahler or special Kahler (for $N = 2$) or flat manifold (or orbifold) (for $N = 4$).

1.4.4. Extended susy. Extended supersymmetry multiplets behave in an interesting way.

The representations of $N = 2$ superalgebra are obtained by combining the representations of $N = 1$ algebra.

In the absence of central extension ($Z = 0$) they contain 16 helicity states. If the central charge of the representation is non-zero, then one can get shorter multiplets for massive particles. This happens precisely when the so-called BPS bound is saturated:

$$M = \frac{1}{2} |Z|$$

1.4.6. Topological gauge theory. Twisted theory can be formulated on any four-manifold $X$, while preserving at least one fermionic symmetry, namely $\bar{Q}$. In this way the observables annihilated by $\bar{Q}$ become distinguished. For one thing, their correlation functions are independent of the metric on $X$ and provide the smooth structure invariants of $X$. For $G = SO(3)$, and the observables constructed out of $\text{Tr} \, \Phi^2$ these are the celebrated Donaldson
Lectures on nonperturbative aspects of supersymmetric gauge theories

invariants. For \( G = U(1) \) with a single hypermultiplet the corresponding correlators are the so-called Seiberg–Witten monopole invariants. If \( X \) has extra structure the twisted gauge theory might have more fermionic symmetries. If \( X \) is Kähler, then one gets one more supercharge: \( \bar{Q}_\omega = \omega_{\mu\nu} \bar{Q}^{\mu\nu} \). If \( X \) has isometries, say generated by the vector fields \( V_a = V^\mu_a \partial_\mu \), then one gets extra supercharges: \( Q_a = V^\mu_a \bar{Q}_\mu \).

2. Prepotential of the \( \mathcal{N} = 2 \) gauge theory

2.1. \( \mathcal{N} = 2 \) rigid special geometry\n
The four-dimensional \( \mathcal{N} = 2 \) supersymmetric action can be written as an integral over the superspace:

\[
\mathcal{L}_{\mathcal{N}=2} = \int d^4 \theta \mathcal{F}(\Psi) + \int d^4 \bar{\theta} \bar{\mathcal{F}}(\bar{\Psi}).
\]  

The classical action has the form (25) with \( \mathcal{F}(\Psi) = \frac{2\pi i}{4} \text{Tr} \Psi^2 \), where \( \Psi \) is the superfield taking values in the adjoint representation of the Lie algebra \( g \). The effective Wilsonian action is an infinite expansion in derivatives. However, the leading term, with two derivatives and four fermions, at most has the form (25) with another \( \mathcal{F} \), and with \( \Psi = (A^i), i = 1, \ldots, r \), in the Cartan subalgebra \( t \subset g \), \( r = \dim t = \text{rk} g \). The superfield \( A \) can be expanded, in this case, as

\[
A^i = a^i + \theta \psi^i + \partial \partial F^{-i} + \ldots
\]  

and the Lagrangian becomes

\[
\mathcal{L}_{\text{eff}}^{\mathcal{N}=2} = \tau_{ij}(a) F^{-i} \wedge F^{-j} + \bar{\tau}_{ij}(a) F^{+i} \wedge F^{+j} + \text{Im} \tau_{ij} da^i \wedge \ast da^j
\]  

The effective Abelian theory has a very special target space \( \mathcal{M} \), which is the moduli space of vacua of the effective theory. In fact, its geometry is special Kahler. Geometrically, \( \mathcal{M} = (t \otimes \mathbb{C}) / W \), where \( W \) is the Weyl group of \( G \). The metric on \( \mathcal{M} \) can be read off the Lagrangian (27):

\[
g_{ij} = \text{Im} \tau_{ij}.
\]

The special feature of \( \mathcal{M} \) is the integral affine structure it carries. What it means is that \( a^i \) are special coordinates, defined up to some discrete transformations. On an ordinary manifold, with no special structure on it, the local coordinates are defined up to arbitrary smooth changes of coordinates. On the complex manifold we only allow holomorphic changes of coordinates. Our friend \( \mathcal{M} \) has coordinates defined up to \( \text{Sp}(2r, \mathbb{Z}) \) transformations. It is convenient to add to \( a^i \) extra \( r \) ‘dual’ coordinates \( a^D,i \), which are related to \( a^i \) via

\[
a_{D,i} = \frac{\partial \mathcal{F}}{\partial a^i}.
\]  

Then \( \text{Sp}(2r, \mathbb{Z}) \) acts linearly on the vector

\[
(a^i, a^D,i)^t.
\]
2.2. Higher dimensional point of view

Theories with extended supersymmetry can be viewed as reductions of higher dimensional minimally supersymmetric field theories. This point of view can be useful, especially if one considers compactification instead of reduction. One should take care of the regularization of the higher dimensional theory. As we shall discuss later in these lectures all the constructions we employ have string theory realization, which makes them well defined in the ultraviolet.

2.2.1. Lift to five dimensions. The $\mathcal{N} = 2$ gauge theory in four dimensions is a dimensional reduction of the $\mathcal{N} = 1$ five-dimensional theory. The latter theory needs an ultraviolet completion to be well defined. However, some features of its low-energy behaviour are robust [67].

In particular, the effective gauge coupling runs because of the vacuum polarization by the charged particles. These particles are W-bosons (for non-Abelian theory), four-dimensional instantons, viewed as solitons in five-dimensional theory, and the bound states thereof.

We see from (27) that the effective gauge couplings are determined by the prepotential. We conclude, therefore, that the prepotential can be viewed as some sort of generating function of the beta function contributions of the charged particles in the theory spectrum. Moreover, one can show that only BPS particles make a nonvanishing contribution to the prepotential, the long multiplets giving total contribution zero [53].

To calculate the effective couplings we need to know the multiplicities, the masses, the charges and the spins of the BPS particles present in the spectrum of the theory [24, 53]. This can be done, in principle, by careful quantization of the moduli space of collective coordinates of the soliton solutions (which are four-dimensional gauge instantons). Now suppose the theory is compactified on a circle. Then the one-loop effect of a given particle consists of a bulk term, present in the five-dimensional theory, and a new finite-size effect, having to do with the loops wrapping the circle in spacetime [53].

If in addition the noncompact part of the spacetime in going around the circle is rotated, as we shall arrange below, then the loops wrapping the circle would have to be localized near the origin in the spacetime. Physically, the multiplicities of the BPS states are accounted for by the supersymmetric character-valued index [24]:

$$\sum_{\text{solitons}} \text{Tr}(-)^F e^{-rP_5} e^{\Omega M} e^{A I}$$

where $P_5$ is the momentum in the fifth direction, $M$ is the generator of the Lorentz rotations, $I$ is the generator of the $R$-symmetry rotations and $r$ is the circumference of the fifth circle. Under certain conditions on $\Omega$ and $A$ this trace has some supersymmetry which allows us to evaluate it. In the process one gets some integrals over the instanton collective coordinates, as in [11, 49, 69, 70]. As in [49], these integrals are exactly calculable, thanks to the equivariant localization.

2.2.2. Lift to six dimensions and twisted compactification. Consider lifting the $\mathcal{N} = 2$ four-dimensional theory to the $\mathcal{N} = (1, 0)$ six-dimensional theory. This is done in a unique fashion. Vector multiplets lift to vector multiplets, and hypers lift to hypers.

Now compactify the $\mathcal{N} = (1, 0)$ theory on a 2-torus with the twisted boundary conditions (along both $A$ and $B$ cycles). As we go around a non-contractible loop $\ell \sim nA + mB$, the spacetime and the fields of the gauge theory charged under the $R$-symmetry group $SU(2)_I$ are rotated by the element $(e^{i(na_1+m_1)\sigma_3}, e^{i(na_2+m_2)\sigma_3}, e^{i(na_2+m_2)\sigma_3}) \in SU(2)_L \times SU(2)_R \times SU(2)_I = Spin(4) \times SU(2)_I$. In other words, we compactify the six-dimensional $\mathcal{N} = 1$
susy gauge theory on the manifold with the topology \( T^2 \times \mathbb{R}^4 \) with the metric and the \( R \)-symmetry gauge field Wilson line:

\[
\begin{align*}
\text{d}s^2 &= r^2 \text{d}z \text{d} \bar{z} + (\text{d}x^\mu + \Omega^{\mu\nu} x^\nu \text{d}z + \bar{\Omega}^{\mu\nu} x^\nu \text{d} \bar{z})^2, \\
A^a &= (\Omega^{\mu\nu} \text{d}z + \bar{\Omega}^{\mu\nu} \text{d} \bar{z}) \eta^a_{\mu\nu}, \quad \mu = 1, 2, 3, 4, \quad a = 1, 2, 3
\end{align*}
\]

(30)

where \( \eta \) is the anti-self-dual 't Hooft symbol. It is convenient to combine \( a_{1,2} \) and \( b_{1,2} \) into two complex parameters \( \varepsilon_{1,2} \):

\[
\varepsilon_1 - \varepsilon_2 = 2(a_1 + i b_1), \quad \varepsilon_1 + \varepsilon_2 = 2(a_2 + i b_2).
\]

(31)

The antisymmetric matrices \( \Omega, \bar{\Omega} \) are given by

\[
\Omega = \frac{\varepsilon_1 + \varepsilon_2}{2} J^L + \frac{\varepsilon_1 - \varepsilon_2}{2} J^R, \quad \bar{\Omega} = \frac{\bar{\varepsilon}_1 + \bar{\varepsilon}_2}{2} J^L + \frac{\bar{\varepsilon}_1 - \bar{\varepsilon}_2}{2} J^R
\]

(32)

where we have used the decomposition \( SU(2)_L \times SU(2)_R \) of the rotation group \( Spin(4) \).

Clearly, \( [\Omega, \bar{\Omega}] = 0 \). In the limit \( r \to 0 \) we get four-dimensional gauge theory. We could also take the limit to the five-dimensional theory by considering the degenerate torus \( T^2 \). We note in passing that the complex structure of the torus \( T^2 \) could be kept finite. The resulting four-dimensional theory (for gauge group \( SU(2) \)) is related to the theory of the so-called E-strings [22, 46]. The instanton contributions to the correlation functions of the chiral operators in this theory are related to the elliptic genera of the instanton moduli space [7] and could be summed up, giving rise to the Seiberg–Witten curves for these theories. We discuss neither elliptic nor trigonometric limits, even though they lead to interesting integrable systems [44].

2.2.3. \( \Omega \)-background. The prepotential of the low-energy effective action can be made observable. To this end one should consider the theory in the so-called \( \Omega \)-background. It is obtained in the \( r \to 0 \) limit of the compactification on the torus with the metric (30).

The action of the four-dimensional theory in the limit \( r \to 0 \) is not that of the pure supersymmetric Yang–Mills theory on \( \mathbb{R}^4 \). Rather, it is a deformation of the latter by the \( \Omega, \bar{\Omega} \)-dependent terms. We shall write down here only the terms with bosonic fields (for simplicity, we have set the bare theta angle to zero):

\[
S(\Omega)_{\text{bos}} = -\frac{1}{2g_0^2} \text{Tr} \left( \frac{1}{2} F_{\mu\nu}^2 + (D_\mu \Phi - \Omega^\mu_x x^x F_{\mu
u}) (D_\nu \Phi - \bar{\Omega}^\nu_x \bar{x}^x F_{\mu\nu}) + \mathcal{D}^2 \right)
\]

(33)

\[
\mathcal{D} = [\Phi, \bar{\Phi}] + \Omega^\mu_x x^x D_\mu \Phi - \bar{\Omega}^\nu_x \bar{x}^x D_\nu \bar{\Phi}.
\]

We call the theory (33) an \( \mathcal{N} = 2 \) theory in the \( \Omega \)-background. By construction it is the deformation of the ordinary \( \mathcal{N} = 2 \) super-Yang–Mills theory. This deformation violates Poincaré invariance. It also violates supersymmetry. However, one can construct certain fermionic symmetries of the theory, which will commute to the spacetime rotation, instead of translation. In this way the \( \Omega \)-background supersymmetry is similar to the anti-de Sitter supersymmetry.

It will prove useful that \( \Omega \)-deformation can be described as a superspace-dependent bare coupling \( \tau_0 \):

\[
\tau_0(x, \theta; \Lambda_{\text{UV}}) = \tau_0 (\Lambda_{\text{UV}}) + \Omega_{ab} \theta_1^a \theta_2^b + \Omega_{\mu\nu} \bar{\Omega}_a^{\nu} x^a x^\nu.
\]

(34)

It is worth mentioning at this point that the (super)spacetime-dependent bare coupling is similar, in spirit, to the spacetime-dependent regulator mass, which was found by 't Hooft to be quite useful in analysing the one-loop corrections to the effective action in the instanton background [32].
2.3. Chiral observables and higher times

We are going to study the correlation functions of chiral observables. These observables are gauge-invariant holomorphic functions of the superfield $\Psi$. Viewed as a function on the superspace, every such observable $O$ can be decomposed:

$$O[\Psi(x, \theta)] = O^{(0)} + O^{(1)} \theta + \cdots + O^{(4)} \theta^2_2.$$  \hspace{1cm} (35)

The component $O^{(4)}$ can be used to deform the action of the theory; this deformation is equivalent to the addition of $O$ to the bare prepotential. Thus ultimately the prepotential of the effective theory will depend on the infinite set of couplings $\tilde{t}$, called higher times, which are in one-to-one correspondence with the gauge-invariant polynomials on the Lie algebra of the gauge group:

$$\mathcal{F}_0(\tilde{a}, \tilde{t}, \Lambda).$$

We shall discuss them more thoroughly in the next lecture.

2.4. The prepotential comes out of the closet

The low-energy effective action depends on the bare gauge coupling, via dimensional transmutation. Moreover, the axial anomaly makes it also bare theta angle dependent. $\mathcal{N} = 2$ susy relates these two dependences, by making the prepotential a holomorphic function of the complex combination $\tau_0$ of the bare gauge coupling and the theta angle.

So the low-energy effective action has the form ($\beta_1$ is the one-loop beta function coefficient)

$$S_{\text{low-energy}} = \int d^4x d^4\theta \mathcal{F}_0(\tilde{A}, \Lambda) + \text{c.c.}$$

$$\Lambda = \Lambda_{\text{UV}} \exp\left(\frac{2\pi i \tau_0(\Lambda_{\text{UV}})}{\beta_1}\right)$$ \hspace{1cm} (36)

where the difference with (25) is that we made the $\Lambda$ dependence explicit. The significance of this is that expression (36) also makes sense when $\Lambda$ is promoted to the superfield. In fact, from the string theory perspective $\log \Lambda$ should be viewed as the scalar component of yet another vector multiplet.

For our immediate purposes it suffices to make $\Lambda$ superspace dependent, by substituting (34) into (36).

In this background one can completely integrate the vector multiplets remaining in the Wilsonian action, the only parameter left being the value of the Higgs vev $\tilde{a}$. Indeed, for constant $\Lambda$ the Wilsonian action contains the vector multiplets because they are massless, and integrating them produces the usual infrared divergences. But in the $\Omega$-background (34) there are no infrared divergences left, and the non-trivial finite value of the integral

$$S_{\text{eff}}(\tilde{a}) = \int d^4x d^4\theta \mathcal{F}_0(\tilde{a}, \Lambda(x, \theta))$$

is a good indication that (in the ordinary, $\Lambda = \text{const}$, background this integral would give $\infty \times 0$):

$$S_{\text{eff}}(\tilde{a}, \Lambda) = \mathcal{F}_0(\tilde{a}, \Lambda) \frac{\epsilon_1 \epsilon_2}{\epsilon_1 \epsilon_2} + \cdots$$ \hspace{1cm} (37)

where $\Lambda = \Lambda(0, 0)$, and $\cdots$ stands for less singular, as $\epsilon \to 0$, terms, which come from the higher derivative couplings in the effective action. Similarly, the partition function of the
theory in the $\Omega$-background would behave as

$$Z(\vec{a}, \Lambda, \epsilon_1, \epsilon_2) = \exp\left(\frac{1}{\epsilon_1\epsilon_2}F_0(\vec{a}, \Lambda) + \frac{\epsilon_1 + \epsilon_2}{\epsilon_1\epsilon_2}F_1(\vec{a}, \Lambda) + \frac{\epsilon_1 + \epsilon_2^2}{\epsilon_1\epsilon_2}G_1(\vec{a}, \Lambda) + \cdots\right).$$  \hfill (38)

2.5. Instanton moduli space integration

The nice property of the chiral observables is the independence of their correlation functions of the anti-chiral deformations of the theory, in particular of $\bar{\tau}$. We can, therefore, consider the limit $\bar{\tau} \to \infty$. In this limit the term $\bar{\tau} \parallel F \parallel^2$ in the action localizes the path integral onto the instanton configurations. In the fixed instanton sector the index (29) can be, therefore, represented by the finite-dimensional integral over the moduli space of instantons.

2.5.1. Localization. The vev of the Higgs field shrinks these instantons to pointlike ones. This phenomenon is of course well known [32].

In addition, the $\Omega$-background further localizes the measure on the instantons, invariant under rotations. This is a relatively new idea. As a result, all integrations are eliminated, reducing them to the single sum over the pointlike-invariant instantons. One then has to desingularize and partially compactify the instanton moduli spaces to make these invariant configurations non-degenerate. This can be done explicitly for $G = U(N)$ and implicitly for all classical gauge groups, using ADHM construction [6], and for all Lie groups using Atiyah’s construction.

2.5.2. Factorization. Finally, the instanton moduli space has the following factorization property. The instantons $A_1$ and $A_2$ of charges $k_1$ and $k_2$ respectively, which are widely separated, can be superposed in a sense that by a small perturbation $A_1 + A_2$ can be made to solve the instanton equations, thus giving rise to the gauge field of the instanton charge $k_1 + k_2$. In this sense the moduli spaces of instantons allow some sort of multiplication (which gives a semi-group structure at the level of homotopy) [73]:

$$\mathcal{M}_{k_1} \times \mathcal{M}_{k_2} \hookrightarrow \mathcal{M}_{k_1 + \cdots + k_p}. \hfill (39)$$

In the sense of this multiplication the universal moduli space of instantons is an exponent

$$\mathcal{M} = \bigcup_{k=0}^{\infty} \mathcal{M}_k = \exp \left[ R^4 \times \bigcup_{k=0}^{\infty} \mathcal{M}_k^{\text{centre}} \right] \hfill (40)$$

where $\mathcal{M}_k^{\text{centre}}$ is the quotient of the ordinary instanton moduli space by the action of the group of translations. This exponentiation is nothing but the physically well-known phenomenon of the cluster decomposition. When applied to the partition function $Z(\vec{a}, \epsilon_1, \epsilon_2, \Lambda)$ the cluster property (40) implies (38).

To recapitulate, by appropriately deforming the theory (in a controllable way) we achieve that the path integral has isolated saddle points and, thanks to the supersymmetry, is exactly given by the WKB approximation. The final answer is then the sum over these critical points of the ratio of bosonic and fermionic determinants. This sum is shown to be equal to the partition function of an auxiliary statistical model, describing the random growth of the Young diagrams.

\footnote{However, beware of the holomorphic anomaly.}
diagrams (for the gauge groups which are the products of unitary groups\(^5\)). We describe this model in detail in the next lecture.

### 3. Random partitions and limit shapes

The partition function of the four-dimensional \(\mathcal{N} = 2\) gauge theory in the \(\Omega\)-background can be reduced to the sum of equivariant integrals over instanton moduli spaces:

\[
Z(\vec{a}, \Lambda, \varepsilon_1, \varepsilon_2) = \sum_{k=0}^{\infty} \Lambda^{2kh} Z_k(\vec{a}, \varepsilon_1, \varepsilon_2). \tag{41}
\]

For \(G = SU(N)\) these integrals can be evaluated explicitly [57]. In the limit \(\varepsilon_1, \varepsilon_2 \to 0\) where the \(\Omega\)-background approaches flat space the partition function \(Z\) has the essential singularity:

\[
Z(\vec{a}, \Lambda) = \exp \frac{1}{\varepsilon_1 \varepsilon_2} F_0(\vec{a}, \Lambda) + \text{less singular terms} \tag{42}
\]

where \(F_0\) is the sought-for prepotential of the effective theory.

#### 3.1. Instanton integrals

The trace (29) is given by the integral over the moduli space of instantons \(\mathcal{M}_k\), which calculates the trace of the element of the symmetry group of \(\mathcal{M}_k\) on the space of holomorphic functions on \(\mathcal{M}_k\). The moduli space \(\mathcal{M}_k\) is acted upon by the group \(G\) of constant gauge transformations, and by the group of isometries of \(\mathbb{R}^4\). Actually the group \(SO(4,1)\) of conformal transformations of \(S^4\) acts there. However, the trace is only non-trivial when restricted to the compact subgroup of the global symmetry group. There, it is sufficient to consider only the elements of the maximal torus, which is the product of the maximal torus \(T\) of \(G\) and \(U(1) \times U(1)\) which comes from the group \(SO(4)\) of rotations:

\[
Z^{5d}_k(\vec{a}, \varepsilon_1, \varepsilon_2; \beta) = \text{Tr}_{\mathcal{M}_k} ((e^{\beta \vec{a}} \times (e^{\beta \varepsilon_1}, e^{\beta \varepsilon_2})) \tag{43}
\]

where \(\vec{a} \in \text{Lie} T\). To get the four-dimensional theory we should take the limit \(\beta \to 0\). In this limit the trace (43) scales as

\[
Z^{5d}_k(\vec{a}, \varepsilon_1, \varepsilon_2; \beta) \approx \frac{1}{\beta^{2kh}} Z^{4d}_k(\vec{a}, \varepsilon_1, \varepsilon_2). \tag{44}
\]

For the classical gauge groups the moduli space \(\mathcal{M}_k\) is the quotient of an affine algebraic variety \(\mathcal{N}_k\) by an algebraic group \(G_k\), where \(G^D\) stands for the Corrigan–Goddard dual group \((U(k)\) for \(G = SU(N), SO\) for \(Sp\) and \(Sp\) for \(SO\)) [14]. This allows us to represent the trace in (43) as the integral over the maximal compact subgroup of the trace on the space of holomorphic functions on the affine algebraic variety, see [59] for details.

#### 3.2. Statistical mechanics of the instanton gas: random partitions

In the case of \(G = SU(N)\) the integrals can be evaluated using residues. The poles are in one-to-one correspondence with \(N\)-tuples of partitions, \(\lambda = (\lambda_1, \ldots, \lambda_N)\). The instanton charge \(k\) is equal to the sum of the sizes of all partitions:

\[
k = |\lambda| \equiv \sum_{l=1}^{N} |\lambda_l|. \tag{45}
\]

\(^5\) It seems that one gets the sums over Young diagrams for all classical gauge groups, but at the time of writing this has not been shown to complete satisfaction, see [48, 59] for the current state of affairs.
The partition $\lambda_l$ can be identified with the set of pairs of positive integers $(i, j)$, $1 \leq j \leq \lambda_l,i$, where the non-negative integers $\lambda_{l,1} \geq \lambda_{l,2} \geq \lambda_{l,3} \cdots$ represent the partition $\lambda_l$ of the number $|\lambda_l| = \sum_i \lambda_{l,i}$.

To describe the residue corresponding to $\vec{\lambda}$ we need further notation. Let

$$W = \sum_{l=1}^{N} e^{a_{l}}, \quad V_{\vec{\lambda}} = \sum_{l=1}^{N} e^{a_{l}} \sum_{(i,j) \in \lambda_{l}} e^{\epsilon_1 (i-1) + \epsilon_2 (j-1)},$$

$$E_{\vec{\lambda}} = W - V_{\vec{\lambda}}(1 - e^{\epsilon_1})(1 - e^{\epsilon_2}), \quad M_{\vec{\lambda}} = \frac{WW^* - E_{\vec{\lambda}}E_{\vec{\lambda}}^*}{(1 - e^{-\epsilon_1})(1 - e^{-\epsilon_2})}$$

where we use the notation for $X = \sum_{\alpha} e^{x_{\alpha}}, X^* = \sum_{\alpha} e^{-x_{\alpha}}$. Then

$$Z_{\text{inst}}(a, \epsilon_1, \epsilon_2) = \sum_{\vec{\lambda}, |\vec{\lambda}| = k} \mu_{\vec{\lambda}}(\vec{a}, \epsilon_1, \epsilon_2)$$

where

$$M_{\vec{\lambda}} = \sum_{a} e^{m_{\vec{\lambda},a}}.$$  

The full partition function is obtained by summing (47) and then multiplying it by

$$Z(\text{pert})(a, \epsilon_1, \epsilon_2; \Lambda) = \exp \left( -\sum_{l,n} \gamma (a_l - a_n; \epsilon_1, \epsilon_2) \right)$$

which can be viewed as the regularized contribution of the $WW^*/(1 - e^{-\epsilon_1})(1 - e^{-\epsilon_2})$ part in (46):

$$Z(a, \epsilon_1, \epsilon_2; \Lambda) = Z(\text{pert})(a, \epsilon_1, \epsilon_2; \Lambda) \times \sum_{k=0}^{\infty} \Lambda^{2kN} Z_{\text{inst}}^{k}(a, \epsilon_1, \epsilon_2).$$

### 3.3. Limit shape and Seiberg–Witten curves

For small $\epsilon_1, \epsilon_2$ the sum over $\vec{\lambda}$ in equation (50) has a saddle point [55]. The partitions which contribute most to the partition function are of large size, $|\lambda_l| \sim \frac{1}{\epsilon_1 \epsilon_2}$, for which the profiles of their Young diagrams can be approximated by smooth curves. These curves can be glued together to a single real analytic curve. Here are the formulae:

$$f(x) = \sum_{l} |x - a_l| + \sum_{l,i} |x - a_l - \epsilon_1 (i - 1) - \epsilon_2 \lambda_{l,i}| - |x - a_l - \epsilon_1 (i - 1)|$$

$$- |x - a_l - \epsilon_1 i - \epsilon_2 \lambda_{l,i}| + |x - a_l - \epsilon_1 i|.$$  

The maximizing configuration $f^*(x)$ obeys a certain equation [55]:

$$\int dy (y - x) f^*_n(y) \log \left( \frac{x - y}{\epsilon \Lambda} \right) = \sigma^*(f^*_n(x)).$$
where $e = 2.71828\ldots$, and $\sigma(y)$ is a concave, piecewise-linear function on $[-N, N]$ such that

$$\sigma'(y) = \xi_l, \quad y \in [-N + 2(l - 1), -N + 2l],$$

and

$$\sigma(-N) = -\sigma(N) = -\sum_l \xi_l.$$

Finally, the numbers $\xi_1 > \cdots > \xi_N$ are the Lagrange multipliers, which should be arranged so that the profile $f_\nu(x)$ corresponds to the eigenvalues $\alpha_l$ of the Higgs field.

Equation (52) should be satisfied for any point $x$, for which $f'_\nu(x)$ is a point of continuity of $\sigma'$. When

$$f'_\nu(x) \in \{-N + 2l | l = 1, \ldots, N - 1\}$$

then considering the left and right derivatives separately one obtains the inequalities

$$Xf_\nu(x) \in (\sigma'(f'_\nu(x) - 0), \sigma'(f'_\nu(x) + 0)), \quad (53)$$

where, by definition,

$$[Xf](x) = \int_{y \neq x} dy \left( \log \frac{|y - x|}{\Lambda} - 1 \right) f''(y). \quad (54)$$

The transform $Xf'$ is closely related to the standard Hilbert transform

$$[Hg](x) = \frac{1}{\pi} \int_{y \neq x} \frac{g(y)}{y - x}. \quad (55)$$

Indeed,

$$[Xf]' = \pi H(f'').$$

3.3.1. Turning on the higher times. As we discussed before, the $\mathcal{N} = 2$ gauge theories have a rich set of exactly calculable correlation functions. Namely, the correlators of the observables (35) are saturated by instantons, and are given exactly by the one-loop approximation around instanton solutions.

In what follows $\nu, \lambda_l$ denote partitions, $l = 1, \ldots, \tilde{N}$, $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_N)$. With each partition $\nu$ one associates the gauge theory operator $O_\nu$, which is simply the Schur function $s_\nu$ evaluated on the Higgs field $\phi$ eigenvalues,

$$Z(\vec{\alpha}; \vec{\epsilon}_1, \vec{\epsilon}_2) = \sum_{\vec{\lambda}} \mu^{\vec{\lambda}}(\vec{\alpha}, \vec{\epsilon}_1, \vec{\epsilon}_2) \exp O[\phi; \vec{\tau}] \quad (56)$$

where $O[X; \vec{\tau}]$ is a symmetric function of the eigenvalues of the matrix $X$ of indefinite size, and $\vec{\tau}$ are some coordinates on the space of all symmetric functions. For example, one could take $\vec{\tau} = (t_\nu)$,

$$O[X; \vec{\tau}] = \sum_\nu t_\nu s_\nu(X)$$

with $s_\nu$ being Schur functions. Or, large $\tilde{N}$ gauge theory suggests expansion in multi-trace operators: $\vec{\tau} = (t_\nu; t_{kl}; t_{klm}; \ldots)$,

$$O[X; \vec{\tau}] = \sum_k t_k \text{Tr} X^k + \sum_{k,l} t_{kl} \text{Tr} X^k \text{Tr} X_l^l + \sum_{k,l,m} t_{klm} \text{Tr} X^k \text{Tr} X_l^l \text{Tr} X_m^m.$$
Finally, the matrix $\phi_\lambda$ is characterized by the Chern character:
\[
\text{Tr} e^{\beta \phi_\lambda} = \sum_l e^{\beta a_l} + \sum_{l,i} e^{\beta (a_l + \epsilon_1 (i-1))} (e^{\beta \epsilon_2} - 1) (1 - e^{\beta \epsilon_2})
\]
(57)
\[
Z(\vec{a}; \vec{t}, \epsilon_1, \epsilon_2) = \exp \left( \frac{1}{\epsilon_1 \epsilon_2} \mathcal{F}_0(\vec{a}, \vec{t}) + \cdots \right).
\]
(58)

The limit shape changes as a function of times $\vec{t}$. In fact, the limit shape should be viewed not as a curve, but rather as a (non-compact) calibrated Calabi–Yau manifold, i.e. a pair (curve, meromorphic differential). Krichever has associated the so-called Whitham integrable hierarchy with these data, and the natural expectation is that $\vec{t}$ prove to be the times of this hierarchy [19, 26, 42, 47, 55]. However, this has not been established so far.

3.4. Extended symmetry point

It turns out that the partition function $Z(\vec{a}; \vec{t}, \epsilon_1, \epsilon_2)$ has some magnificent properties when $\epsilon_1 = -\epsilon_2 = \bar{\hbar}$. We therefore introduce a special notation:
\[
Z(\vec{a}; \vec{t}, \bar{\hbar}) = Z(\vec{a}; \vec{t}, \bar{\hbar}, -\bar{\hbar}).
\]
(59)

3.4.1. Dual partition function. The first special property is the relation to free fermions and integrable hierarchies of Toda type: define
\[
Z_D^p(\vec{\xi}; \vec{t}, \bar{\hbar}) = \sum_{\vec{M} \in \mathbb{Z}^N, \sum M = p \mod N} e^{\frac{\bar{\hbar}}{2} \vec{M} \cdot \vec{\xi}} Z(h(\vec{M} + \vec{\rho}); \vec{t}, \bar{\hbar}),
\]
(60)
which is the ‘quantum’ electro-magnetic dual partition function. The dual partition function $Z_D$ can be rewritten in terms of the $N$-tuples of partitions, as the original $Z$. Statistical sum could. However, and this is quite remarkable, one can also repackage these $N$-tuples of partitions into a single one, and then use the fermion representation of the Plancherel measure on the partitions to write $Z_D$ as a free fermion correlator [55]:
\[
Z_D^p(\vec{\xi}; \vec{t}, \bar{\hbar}) = \langle p | e^{\frac{\bar{\hbar}}{2} \vec{J}_1} e^{\frac{\bar{\hbar}}{2} \vec{J}_2} e^{\frac{\bar{\hbar}}{2} \vec{J}_3} e^{-\frac{\bar{\hbar}}{2} \vec{J}_4} | p \rangle
\]
(61)
where $\vec{J}_0$ stands for the zero modes of the $\hat{u}_1(N)$ current algebra, $\vec{J}_- \cdot \vec{J}_3$ represents the $\hat{u}_N(1)$ current modes and $\vec{W}$ stands for the $W$ generators, corresponding to the symmetric functions.

3.4.2. Relation to topological strings. The second miracle that happens at the point $\epsilon_1 + \epsilon_2 = 0$ is the stringy nature of the $\epsilon$-expansion. We shall of course discuss it in more detail in the next lectures, but here we simply mention that in the expansion
\[
Z(\vec{a}, \vec{t}, \bar{\hbar}) = \exp \left( \sum_{g=0}^\infty \hbar^{2g-2} \mathcal{F}_g(\vec{a}, \vec{t}) \right)
\]
(62)
the terms $\mathcal{F}_g$ turn out to be the genus $g$ topological string amplitudes. Moreover, there are several topological string theories which are behind (62). There are topological strings on local Calabi–Yau threefolds [37], which engineer the $\mathcal{N} = 2$ gauge theory on $\mathbb{R}^4$. These show up most naturally in the considerations of $Z^{5d}$—the five-dimensional analogue of $Z$. The role of higher times $\vec{t}$ is not understood in this case.

There are also topological strings on non-Calabi–Yau manifolds, such as $\mathbb{P}^1$, where the higher times are mapped to the gravitational descendents [41], thanks to the free fermion representation found in [62].
3.5. More general \( \mathcal{N} = 2 \) theories: hypermultiplets and higher dimensions

3.5.1. Theories with matter. The \( \mathcal{N} = 2 \) theory with hypermultiplets in some representation \( R \) of the gauge group can be treated similarly to the pure theory case. Consider the decomposition of \( R \) into the sum of irreducible representations:

\[
R = \bigoplus_f M_f
\]

where \( M_f \) is the multiplicity space. In the representation ring of \( U(N) \) the representation \( R \) can be expressed as a polynomial in the fundamental representation \( N \) and its conjugate \( \bar{N} \):

\[
R = \bigoplus_{f} m_f N^f \otimes \bar{N}^f
\]

where \( m_f \) are the multiplicity spaces (they can be related to \( M_f \), of course). The partition function of the \( \mathcal{N} = 2 U(N) \) gauge theory with the matter hypermultiplet in the representation (64) is given by

\[
Z(\vec{a}; \, \mathcal{A}, \, h) = \sum_{\vec{\lambda}} \Lambda(\vec{\lambda}) \mu(\vec{\lambda}, h) v(\vec{\lambda}, \mathcal{A}, h)
\]

\[
v(\vec{\lambda}, \mathcal{A}, h) = \prod_{\ell} \langle \ell(\vec{\lambda}, \mathcal{A}, \mathcal{A}, h) \rangle
\]

\[
\sum_{f, l} \text{Tr}_{m_f} e^{h l} \frac{E^*_f E^*_l}{(1 - e^{-h})(1 - e^{h})} = \sum_{f, l} e^{h(\vec{a}, \mathcal{A}, h)}
\]

and everything is understood with the regularization as in the case of pure gauge theory (i.e. with the functions \( \gamma \) in place of the infinite products and so on).

Formula (65) is a straightforward consequence of the fact that the matter contribution to the instanton integral is the equivariant Euler class of the bundle over the moduli space of instantons, whose sections are the solutions of the Dirac equation in the instanton background, where the spinor transforms in the matter representation.

3.5.2. Quiver theories. One can similarly treat quiver theories. We consider the simplest case: the \( U(N) \times U(N) \) theory with bi-fundamental matter: \( (\mathbb{N}, \bar{\mathbb{N}}) \). The generalization to other cases is straightforward. The partition function is given by

\[
Z(\vec{a}_1, \vec{a}_2; \, \Lambda_1, \Lambda_2; \, h) = \sum_{\vec{\lambda}_1, \vec{\lambda}_2} \Lambda^*_{\vec{\lambda}_1} \Lambda_{\vec{\lambda}_2} \mu^*_{\vec{\lambda}_1} (\vec{a}_1, h) \mu_{\vec{\lambda}_2} (\vec{a}_2, h) v_{\vec{\lambda}_1, \vec{\lambda}_2} (\vec{a}_1, \vec{a}_2, h)
\]

\[
v_{\vec{\lambda}_1, \vec{\lambda}_2} (\vec{a}_1, \vec{a}_2, h) = \prod_{\ell} \langle \ell(\vec{\lambda}_1, \vec{\lambda}_2; \vec{a}_1, \vec{a}_2, h) \rangle
\]

\[
\frac{E^*_{\ell_1} E^*_{\ell_2}}{(1 - e^{-h})(1 - e^{h})} = \sum_{\ell} e^{h(\vec{a}_1, \vec{a}_2, h)}
\]

Note that the \( \mathbb{Z}_2 \) ‘orbifold’ of the \( \mathcal{N} = 4 \) theory [36, 40], which corresponds to the theory on the regular D3 branes at the \( \mathbb{C}^2/\mathbb{Z}_2 \) singularity, has the gauge group \( U(N) \times U(N) \) and the matter in the representations \( (\mathbb{N}, \bar{\mathbb{N}}) \times (\mathbb{N}, \bar{\mathbb{N}}) \). For that theory in (66) one should use \( E^*_{\ell_1} E^*_{\ell_2} \) instead of \( E^*_{\ell_1} E^*_{\ell_2} \). The perturbative part of the partition function of this theory is independent of the regulator mass if \( \vec{a}_1 = \vec{a}_2 \). This is in agreement with the well-known fact that the perturbative beta function in this theory vanishes for the coupling which is the descendant of the \( \mathcal{N} = 4 \) one. But we see that even for \( \vec{a}_1 = \vec{a}_2 \) instantons produce non-trivial non-perturbative corrections to the effective couplings.

\[\text{Note that from the } \mathcal{N} = 1 \text{ susy point of view the matter chiral multiplets transform in the representation } R \otimes R^*.\]
3.5.3. Higher dimensional theories. The five-dimensional \( \mathcal{N} = 1 \) supersymmetric theory compactified on a circle, and the six-dimensional \( \mathcal{N} = 1 \) theory compactified on a torus can be regularized by embedding in the string theory, as we describe below. The low-energy physics is four dimensional, but the prepotential will depend on the radii of the compactification. See [55] for the five-dimensional partition function, [53] for the original integrable system description of the prepotentials and [35] for the six-dimensional theory description.

3.6. Breaking susy down to \( \mathcal{N} = 1 \): Dijkgraaf–Vafa proposal

\( \mathcal{N} = 2 \) supersymmetry can be softly broken down to \( \mathcal{N} = 1 \) by adding a superpotential \( W_0(\Phi) \) to the action

\[
\mathcal{L}_{\mathcal{N}=1} = \mathcal{L}_{\mathcal{N}=2} + \int d^2 \theta \ W_0(\Phi) + \text{c.c.}
\]

The choice of two out of four chiral thetas corresponds to the choice of \( \mathcal{N} = 1 \) subalgebra within \( \mathcal{N} = 2 \) superalgebra.

In their series of papers, originated in paper [17], Dijkgraaf and Vafa made a proposal for the superpotential of the effective low-energy \( \mathcal{N} = 1 \) theory. It basically consists of the following: take the original superpotential \( W_0 \) and consider the matrix model with the single \( \hat{N} \times \hat{N} \) matrix \( \hat{\Phi} \) with the action

\[
I_{\hat{N}} = \frac{1}{\text{Vol}(\hat{N})} \int D\hat{\Phi} \ e^{-\frac{1}{g_s} \text{Tr} W_0(\hat{\Phi})}.
\]

The effective superpotential is a function of glueball superfields \( S_1, \ldots, S_p \), which correspond to the gauge symmetry breaking pattern: \( U(N) \to U(N_1) \times \cdots \times U(N_p) \), according to the minima of the tree level superpotential \( W_0 \).

4. String theory

In this lecture we discuss the following string-theory-related issues: realizations of \( \mathcal{N} = 2 \) four-dimensional theories as low-energy sectors of string theory compactifications or D-brane arrangements; \( \Omega \)-background realizations; topological strings and their relation to the gauge theory F-terms. In the next lecture we shall discuss the topological string calculations on \( \mathbb{R}^6 \) in more detail, relate them to the quantum spacetime foam picture and also present its six-dimensional gauge theory realization.

4.1. String theory realizations of \( \mathcal{N} = 2 \) theories

The \( \mathcal{N} = 2 \) theory can arise as a low-energy limit of the theory on a stack of D branes in the type II string theory. A stack of \( N \) parallel D3 branes in IIB theory in flat \( \mathbb{R}^{1,9} \) carries the \( \mathcal{N} = 4 \) supersymmetric Yang–Mills theory [78]. A stack of parallel D4 branes in IIA theory in flat \( \mathbb{R}^{1,9} \) carries \( \mathcal{N} = 2 \) supersymmetric Yang–Mills theory in five dimensions. Upon compactification on a circle the latter theory reduces to the former in the limit of zero radius.

Now consider the stack of \( N \) D4 branes in the geometry \( S^1 \times \mathbb{R}^{1,8} \) with the metric

\[
ds^2 = dx^\mu \ dx^\mu + r^2 \ d\phi^2 + dv^2 + |dZ_1 + mrZ_1 \ d\phi|^2 + |dZ_2 - mrZ_2 \ d\phi|^2.
\]

Here \( x^\mu \) denote the coordinates on the Minkowski space \( \mathbb{R}^{1,3} \), \( \phi \) is the periodic coordinate on the circle of circumference \( r \), \( v \) is a real transverse direction, and \( Z_1 \) and \( Z_2 \) are the holomorphic coordinates on the remaining \( \mathbb{C}^2 \). The worldvolume of the branes is \( S^1 \times \mathbb{R}^{1,3} \), which is located at \( Z_1 = Z_2 = 0 \), and \( v = v_i, i = 1, \ldots, N \). Together with the Wilson loop
eigenvalues \( e^{i\sigma_1}, \ldots, e^{i\sigma_N} \) around \( S^1 \), \( v_i \) form \( N \) complex moduli \( w_1, \ldots, w_N \), parametrizing the moduli space of vacua. In the limit \( r \to 0 \) the \( N \) complex moduli lose periodicity.

It is easy to check that the worldvolume theory has \( \mathcal{N} = 2 \) susy, with the massive hypermultiplet in the adjoint representation (of mass \( m \)). This realization is T-dual to the standard realization with the NS5 branes, as in [79]. Note that the background (30) is similar to (69). However, the D branes are differently located, a fact which leads to very interesting geometries upon T-dualities and lifts to M-theory, providing (hopefully) another useful insight.

However, in our story we want to analyse the pure \( \mathcal{N} = 2 \) supersymmetric Yang–Mills theory. This can be achieved by taking the \( m \to \infty \) limit, and at the same time taking the weak string coupling limit. The resulting brane configuration can be described using two parallel NS5 branes and \( N \) D4 branes suspended between them, as in [79], or, alternatively, as a stack of \( N \) D3 (fractional) branes stuck at the \( \mathbb{C}^2/\mathbb{Z}_2 \) singularity, as in [15]. In fact the precise form of the singularity is irrelevant, as long as it corresponds to a discrete subgroup of \( SU(2) \), and all the fractional branes are of the same type. The relation between these two pictures is through the T-duality of the resolved \( \mathbb{C}^2/\mathbb{Z}_2 \) singularity. The fractional D3 branes blow up into D5 branes wrapping a non-contractible 2-sphere. The resolved space \( T^* \mathbb{CP}^1 \) has a \( U(1) \) isometry, with two fixed points (the North and South poles of the non-contractible 2-sphere). Upon T-duality these turn into two NS5 branes. The D5 branes dualize to D4 branes suspended between NS5s.

The instanton effects in this theory are due to the fractional D(\( -1 \)) instantons, which bind to the fractional D3 branes, in the IIB description. The ‘worldvolume’ theory on these D(\( -1 \)) instantons is the supersymmetric matrix integral, which we describe with the help of the ADHM construction below. In the IIA picture the instanton effects are due to Euclidean D0 branes, which ‘propagate’ between two NS5 branes.

The IIB picture with the fractional branes corresponds to the metric (before \( \Omega \) is turned on)

\[
\text{ds}^2 = \text{dx}^\mu \text{dx}^\mu + \text{dw} \text{d}\bar{w} + \text{ds}^2_{\mathbb{C}^2/\mathbb{Z}_2}. \tag{70}
\]

The singularity \( \mathbb{C}^2/\mathbb{Z}_2 \) has five moduli in the IIB string theory: three parameters of the geometric resolution of the singularity, and the fluxes of the NSNS and RR 2-forms through the 2-cycle, which becomes visible upon the blowup; otherwise all of these show up as the twisted sector massless field. The NSNS and RR fluxes are responsible for the gauge couplings on the fractional D3 branes, as in [40]:

\[
\tau_0 = \int_{S^2} B_{\text{RR}} + \tau_{\text{IIB}} \int_{S^2} B_{\text{NSNS}}. \tag{71}
\]

Our conjecture is that turning on the higher Casimirs (and gravitational descendants on the dual closed string side) corresponds to a ‘holomorphic wave’, where \( \tau_0 \) holomorphically depend on \( w \). From [38] this is known to be a solution of IIB supergravity.

We shall return to the fractional brane picture later on. Right now let us mention another stringy effect. By turning on the constant NSNS B-field along the worldvolume of the D3 branes we deform the super-Yang–Mills on \( \mathbb{R}^4 \) to the super-Yang–Mills on the noncommutative \( \mathbb{R}^4_\zeta \), see [12, 13, 68]. On the worldvolume of the D(\( -1 \)) instantons the noncommutativity acts as a Fayet–Illiopoulos term, deforming the ADHM equations, as in [1, 2, 58], and resolving the singularities of the moduli space, as in [51]. This deformation is a technical tool, so we shall not describe it in much detail. The necessary references can be found in [68].

At this point we remark that even for \( N = 1 \) the instantons are present in the D-brane picture. They become visible in the gauge theory when noncommutativity is turned on. Remarkably, the actual value of the noncommutativity parameter \( \zeta \) does not affect the expectation values of the chiral observables, thus simplifying our life enormously.
So far we presented the D-brane realization of the $N = 2$ theory. There exists another useful realization, via local Calabi–Yau manifolds, introduced in [37]. This realization is useful in relating the prepotential to the topological string amplitudes. If the theory is embedded in the IIA string on local Calabi–Yau, then the interesting physics comes from the worldsheet instantons, wrapping some 2-cycles in the Calabi–Yau. In the mirror IIB description one gets a string without worldsheet instantons contributing to the prepotential, and effectively reducing to some field theory. This field theory is known in the case of global Calabi–Yau. But it is not known explicitly in the case of local Calabi–Yau. As we shall show, it can sometimes be identified with the free fermion theory on an auxiliary Riemann surface (cf [16, 55, 57]).

The relation to the geometrical engineering of [37] is also useful in making contact between our $\Omega$-deformation and the sugra backgrounds with graviphoton field strength. Indeed, our construction involved a lift to five or six dimensions. The first case embeds easily to the IIA string theory where this corresponds to the lift to M-theory. To see the whole six-dimensional picture (30) one should use IIB language and the lift to F-theory (one has to set $\bar{\Omega}_1 = 0$, though).

Let us consider the five-dimensional lift. We have M-theory on the 11-fold with the metric:

$$ds^2 = (dx^\mu + \Omega_2^a x^a d\psi)^2 + r^2 d\psi^2 + ds^2_{\text{CY}}.$$  \hfill (72)

Here we assume, for simplicity, that $\epsilon_1 = - \epsilon_2$, so that $\Omega = \Omega^-$ generates an $SU(2)$ rotation, thus preserving half of suSY. Now let us reduce on the circle $S^1$ and interpret the background (72) in the type IIA string. Using [77] we arrive at the following IIA background:

$$g_s = (r^2 + \|\Omega \cdot x\|)^2 \frac{1}{4}$$ \hfill (73)

$$A^{\text{grav}} = \frac{1}{r^2 + \|\Omega \cdot x\|^2} \Omega_{\mu\nu} x^\mu dx^\nu$$ \hfill (74)

$$ds_{10}^2 = \frac{1}{\sqrt{r^2 + \|\Omega \cdot x\|^2}} \left( r^2 dx^2 + \Omega_2^a \Omega_2^b (x^2 \delta^{ab} \delta_{ij} - x^i x^j dx^i dx^j) \right)$$

$$+ \sqrt{r^2 + \|\Omega \cdot x\|^2} ds^2_{\text{CY}}$$

where the graviphoton $U(1)$ field is turned on. The IIA string coupling becomes strong at $x \to \infty$. However, the effective coupling in the calculations of $F_g$ is

$$\hbar \sim g_s \sqrt{\|dA^{\text{grav}}\|^2} \sim (r^2 + \|\Omega \cdot x\|)^{-\frac{1}{2}} \to 0, \quad x \to \infty. \hfill (75)$$

4.2. Topological strings

In this section we recall the appearance of the topological strings in the physical string calculations. We consider a type II superstring compactified on a Calabi–Yau threefold. The effective four-dimensional supergravity action contains the terms which are calculable by the topological string.

4.2.1. Calabi–Yau compactifications. A critical superstring lives in ten dimensions. We live in four macroscopic dimensions, so the remaining six dimensions should be unobservably small. One option is to have them compactified on a manifold $M$, which is severely constrained by various symmetry requirements. For example, if we want to have four-dimensional supersymmetry, the manifold $M$ has to be Calabi–Yau (if we want $\mathcal{N} = 1$ susy we must also turn on some fluxes). In what follows we assume it is Calabi–Yau. It means that $M$ is a
complex manifold, of complex dimension 3, which can be endowed with a Ricci-flat Kahler metric. Yau’s theorem states that for a given complex structure, and the cohomology class of the Kahler form $\kappa$, the Ricci-flat metric which has the corresponding Kahler form, exists and is unique. Existence of the Ricci-flat metric implies that $c_1(M) = 0$, which, in turn, implies the existence of nowhere vanishing holomorphic 3-form $\Omega$, which is unique up to a multiplication by a complex number.

4.2.2. Effective supergravity in four dimensions. In these circumstances the low-energy physics in four dimensions is described by the $\mathcal{N} = 2$ supergravity (sugra) theory. Such a theory contains a gravity supermultiplet, whose bosonic fields are the metric $g_{mn}$, $m, n = 1, 2, 3, 4$, and the vector field $b_{mn}$, called a graviphoton. In addition, there are matter multiplets, vectors and hypers. Vector multiplets contain complex scalars $\phi^a, a = 1, \ldots, r$, and $U(1)$ vector fields $A^a_m$. The hypermultiplets contain two complex scalars $Q^i, \tilde{Q}^i, i = 1, \ldots, s$. In the IIA string $r = h_{1,1}(M)$, and $s = h_{1,2}(M)$. Mirror symmetry would exchange IIA with IIB and $M$ with $\tilde{M}$—the mirror Calabi–Yau.

4.2.3. Vector multiplet moduli space. Let us work, for definiteness, with the IIA string. Then the vector fields in four dimensions come by reduction of the $U(1)$ RR gauge field $C_1$ in ten dimensions and by reduction of the RR 3-form $C_3$ along 2-cycles in CY $M$. The vector field $C_1$ becomes $b$ and falls into the gravity multiplet. The graviphoton field strength $T = db$ plays a special role in what follows. The scalars of the vector multiplets come from the complexified Kahler moduli of Calabi–Yau $M$—the real part parametrizes the Kahler class $\kappa$, which fixes some of the metric moduli, while the imaginary part comes from the periods of the NSNS $B$-field: $\phi^\omega = \int_\beta \omega, \omega = \kappa + iB$.

The geometry of the vector multiplet moduli space $\mathcal{M}_v$ is governed by the holomorphic function, the prepotential $F_0(\phi^a)$. The effective Lagrangian for the vector multiplets looks as follows:

$$\tau_{ab} = \frac{\partial^2 F_0}{\partial \phi^a \partial \phi^b}.$$  \hfill (76)

The geometry in (76) is projective special Kahler. It differs from the special Kahler geometry (27) in the following respect: the special coordinates $\phi^a$ in supergravity are defined up to rescaling, and the prepotential is the homogeneous degree 2 function of these homogeneous coordinates. Globally the special coordinates are defined up to the $Sp(2r + 2, \mathbb{Z}) \times \mathbb{C}^*$ transformations. In the rigid supersymmetry case the special coordinates are defined up to $Sp(2r, \mathbb{Z})$ transformations.

In addition, there are couplings between the vector multiplets and the gravity multiplet:

$$\sum_{\mu=1}^\infty F_\mu(\phi^a)T^{2\mu-2} R \wedge R$$  \hfill (77)

where $T^2 = T_{mn}T^{mn}$. Note that the string dilaton does not show up among the vector multiplets.

4.2.4. Hypermultiplet moduli space. The moduli space $\mathcal{M}_h$ of hypermultiplets has quaternionic-Kahler geometry. The hypermultiplet scalars come from dilaton, axion (dual to the $B_{mn}$ in four dimensions), and the Calabi–Yau $C_3$ period, corresponding to the $(3,0)$-form $\Omega$: $\int_M C_3 \wedge \Omega$, and $\int_M C_3 \wedge \tilde{\Omega}$ (this gives rise to the so-called universal hypermultiplet); and then $h_{1,2}(M)$ multiplets where $Q^i$ corresponds to the complex structure deformations of $M$, and $\tilde{Q}_1$ to the remaining periods of $C_3$.  

4.2.5. Topological strings and vector multiplets. The vector multiplet couplings $F_g(\phi)$, $g \geq 0$, can be calculated using a simplified version of string theory. In the IIA superstring context it would be the A type topological string, and in the IIB context it would be, not surprisingly, the B type topological string. The topological strings were introduced by Witten [80], and their importance for the superstring effective theory calculations was understood in [3, 8]. The type A string calculates $F_g$ by summing over the holomorphic maps of the genus $g$ Riemann surfaces into $M$:

$$F_g = \sum_{\beta \in H^2(M, \mathbb{Z})} e^{-\int_{\mathcal{M}(M, \beta)} 1}$$

(78)

where $\mathcal{M}(M, \beta)$ is the moduli space of stable maps, introduced by Kontsevich [39]. The integral of 1 counts these maps, if the moduli space of stable maps consists of isolated points. Its expected dimension is zero, so this is indeed the answer in the generic situation. However, it could happen that the actual dimension is positive. Then one has to integrate the Euler class of the obstruction sheaf over the real moduli space. In this way, say, every isolated rational curve in the primitive homology class $\beta$ would contribute to $F_0$ not just a single exponential $e^{-\int_{\beta} \omega}$ but the whole series of multicover contributions:

$$F_0 = \sum_{\beta \in H^2_{\text{primitives}}(M, \mathbb{Z})} n_{\beta} \log \left( e^{-\int_{\beta} \omega} \right).$$

(79)

The B model definition of the $F_g$ amplitudes is rather tricky, except for the genus zero part. Suffice it to say that it can be computed by purely classical means, by calculating the periods of the $(3, 0)$-form $\Omega$.

The partition function $Z(\phi; \hbar)$ is expected to obey various remarkable equations [8], and also to have some modular properties.

4.2.6. Kodaira–Spencer theory of gravity. The great simplification of the topological string compared to the physical string is the reduction, for the B type topological string, of the integrals over the moduli space of Riemann surfaces to the locus of maximally degenerate surfaces [8]. The combinatorics of such surfaces is that of trivalent graphs. Thus the generating function of the B type topological string amplitudes can be written as the sum over Feynman diagrams of some field theory. This theory seems to be the so-called Kodaira–Spencer theory of gravity, introduced in [8]:

$$L_{KS} = \frac{1}{2\mathbb{R}^2} \left[ A^\vee \frac{1}{\delta} \delta A^\vee + \frac{1}{3} (A \vee A) \vee \wedge A^\vee \right]$$

(81)

where

$$A \in \Omega^{-1,1}(M), \quad A^\vee = t_A \Omega, \quad (A \wedge A)^\vee = t_{A \wedge A} \Omega \in \Omega^{1,2}(M).$$
4.2.7. **Kahler gravity.** The mirror version of the theory (81) is much less understood. The problem is that the worldsheet instantons contribute to the effective action. However, in the large radius limit the theory looks very much like (81) modulo the standard mirror replacements: $A^i \leftrightarrow \omega, \bar{A} \leftrightarrow d, \bar{\partial} \leftrightarrow dc$ [9]:

$$L_{KG} = \frac{1}{2\kappa^2} \left[ \omega \frac{1}{d^6} d\omega + \frac{1}{3} \omega \wedge \omega \wedge \omega \right].$$  (82)

4.3. **Topological string amplitudes and $\mathcal{N} = 2$ theories in $\Omega$-background**

Now let us return to the discussion of $\mathcal{N} = 2$ gauge theories and their F-terms, i.e. the terms in the effective action, which are obtained by the integration over half of the superspace. The F-terms have two faces. On the one hand, they are calculable within gauge theory, by doing the instanton moduli space integration, the way we discussed in the first two sections. On the other hand, in the string theory realization of the gauge theory they are given by the topological string amplitudes.

There is one non-trivial point in this identification. As we argued before, using the lift to M-theory, it is the partition function in the $\Omega$-background that captures the spin-charge particle content of the gauge theory, and contains all the information about the effective couplings. On the other hand, the simplest way to relate the physical and topological strings proceeds via consideration of the constant graviphoton field strength background [3, 8].

Note that the general $\Omega$-background depends on two parameters $\epsilon_1, \epsilon_2$, since generic $SO(4)$ rotation is characterized by two angles. As we discussed in (30), one has to turn on the R-symmetry Wilson lines in order to preserve fermionic symmetry, if $\epsilon_1 + \epsilon_2 \neq 0$. For generic type II string theory compactifications R-symmetry does not exist as a global symmetry. So we may hope to relate simply the [3, 8] calculations in string theory to our calculations in gauge theory only for the special $\Omega$-backgrounds, which do not touch R-symmetry. To understand the general case, with two parameters $\epsilon_1, \epsilon_2$, one has to work harder. Now, for the special case of the $\Omega$-background, $\epsilon_1 = -\epsilon_2 = \hbar$, using its M-theory realization and the resulting type II sugra background (74), we can observe that one has the non-trivial graviphoton field strength. Even though it is not constant, one may hope that since the effective string coupling goes to zero at distances much larger than $\frac{\hbar}{|\epsilon|}$, the higher genus string does not probe these regions of $\mathbb{R}^4$; it is effectively confined near the origin, where the graviphoton field strength is almost constant, and the twisting arguments of [8] apply.

4.3.1. **Localization.** The most powerful method of exact calculations in the supersymmetric theories is the localization on the $Q$-fixed points, where $Q$ is the conserved global supercharge. The localization can be applied to the calculations of $Z(\phi; \hbar)$ for toric varieties $X$, using the worldsheet definition [39].

Imagine calculating the Kahler gravity partition function in some background. Assuming the background has isometries, one can set up an equivariant supercharge whose fixed points would correspond to the geometries with the same asymptotics as the background, and in addition invariant under the isometry group action. This is in complete parallel with the $\mathcal{N} = 2$ supersymmetric gauge theory partition function calculation [55, 57].

Going back to our gravity problem, we consider toric Calabi–Yau $X_0$ as our background. The toric CY has $\mathbb{T}^3$ as a group of isometries. The Kahler gravity partition function $Z_{KG}(X_0)$, evaluated by localization with respect to $\mathbb{T}^1$, would be expressed as a sum over all toric $X$, with the same asymptotics as $X_0$.

Vafa points out that such parameters are present, if at all, only for the compactifications on rigid CYs.
As we shall explain in the next lecture, this method of evaluating the partition function leads to the quantum spacetime foam picture in the topological string theory.

5. Quantum gravitational foam

5.1. Spacetime geometry in string theory

Since the remarkable realization by Scherk and Schwarz [66] that a string spectrum contains gravitons one is looking for the stringy approach to quantum gravity. One feature of quantum gravity—strong fluctuations of spacetime topology at Planck scales, advocated by Hawking [29]—has been very elusive. The reason is that the conventional approach to string quantization produces a $S$-matrix in the form of the expansion in the string coupling constant. The $S$-matrix describes, at best, the scattering of the gravitational waves in some stringy background, most commonly the Minkowski background (however, recent advances in AdS/CFT duality allow us to hope for other homogeneous spaces as well). At finite order of string perturbation theory one has a finite number of gravitons, so it is unlikely that the strong fluctuation of the spacetime topology would be seen. The only hope is to resume string perturbation theory, perhaps analytically continuing in the string coupling:

$$\exp \sum_{g=0}^{\infty} g_s^{2g-2} F_g(a) + O(e^{-1/g_s}) = \int_{\text{bndy cndts}} Dg_{\mu\nu} D(...) \exp -S_{\text{eff}}(g_{\mu\nu}, ...)$$

where $g_s$ denotes string coupling, and $\ldots$ stand for superpartners and massive string modes. But to do that one needs an example of all-loop string calculation. For conventional graviton scattering on $\mathbb{R}^{10}$ no-one has been able to go beyond two loops so far [30]. However, recently, using the advances in the topological string theory, where all-loop exact string amplitudes are not uncommon, the quantum foam picture has indeed emerged. It is the purpose of this lecture to explain how it came about.

5.2. Type A string on $\mathbb{R}^6$

The simplest, but already inspirational example comes from the Gromov–Witten theory on $\mathbb{C}^3$—the simplest Calabi–Yau manifold. This theory needs regularization because of the noncompactness of $\mathbb{C}^3$, but this can be cured by using the torus $T^3$ action. Then the explicit calculation [20] gives

$$F_g = \int_{\text{M}^3_{g-1}} c_{g-1}(H) = \frac{B_{2g}B_{2g-2}}{2^g(2g-2)(2g-2)!}.$$  

The all-genus partition function is then

$$Z_{\text{top}}(g_s) = \exp \sum_{g=0}^{\infty} F_g g_s^{2g-2} \sim M^2(q)$$

$$M(q) = \prod_{n=1}^{\infty} \frac{1}{(1 - q^n)^n}$$

$$q = -e^{-i\pi}.$$

To pass from $Z_{\text{top}}(g_s)$ to $M(q)$ requires adding some unstable, e.g. $F_0 = \zeta(3)$, as well as non-perturbative terms, which are fixed in [24, 64].
5.3. Crystal interpretation

5.3.1. MacMahon function. The function $M(q)$ which enters (85) is well known in combinatorics [43, 63]. It counts three-dimensional partitions:

$$M(q) = \sum_{\pi \text{ 3d partitions}} q^{\abs{\pi}}$$

(86)

5.3.2. Donaldson–Thomas (DT) theory. The reason three-dimensional partitions pop up in our calculation is explained in [45, 56]—they are $T^3$-invariant ideal sheaves: the partition $\pi$ corresponds to the ideal $I_\pi$ in $\mathbb{C}[x, y, z]$: $(i, j, k) \in \mathbb{Z}^3 \setminus \pi \iff x^{i-1}y^{j-1}z^{k-1} \in I_\pi$. The summation over the partitions is actually the fixed-point formula for the integral of 1 over the moduli space of ideal sheaves, weighted by $q^{\text{ch}_1(I_\pi)}$.

For more general toric $X_0$ the DT theory produces the partition function which is

$$Z_{\text{DT}}(X_0) = \int_{\mathcal{M}_{\text{ideal}}(X_0)} q^{\text{ch}_1(I)} e^{-\int_{X_0} \omega_0 \wedge \text{ch}^2(I)}$$

(87)

where $\mathcal{M}_{\text{ideal}}(X_0)$ is the moduli space of ideal sheaves $I$, $\text{ch}_1(I) = 0$, on $X_0$.

5.3.3. From two- to three-dimensional partitions: gauge theory in higher dimensions. Ideal sheaves have showed up in physics in some other problems already. In some sense they are singular limits of $U(1)$ instantons. In this same sense the ch$_p$ of a sheaf corresponds to $rac{1}{p!} (F^p)^{\wedge p}$, where $F$ is the curvature of the gauge field.

In the previous sections we saw that the instanton calculus in four-dimensional gauge theories with $N = 2$ supersymmetry reduces (for gauge group $SU(N)$) to the statistical model on $N$-tuples of partitions. In two lines the idea of the correspondence is the following: instantons correspond to holomorphic bundles, and moduli of bundles are compactified by torsion-free sheaves; the latter differ from bundles by ideals of zero-dimensional schemes. When working equivariantly with respect to the torus action on the moduli space the integrals of interest reduce to the summation over the torus-invariant schemes. These are labelled by two-dimensional partitions.

For three complex dimensional manifolds the analogous considerations suggest that three-dimensional partitions should come into play as labelling the fixed points of the torus action on the compactified moduli space of holomorphic bundles.

The corresponding field theory problem looks as follows. One studies connections $A$ on a bundle $E$ over Kahler three complex dimensional manifold $X$. In addition, one takes $\varphi \in \Omega^{0,1} \otimes \text{ad} E$, and its complex conjugate. The BPS equations are

$$F^{0,2} + \bar{\theta}^1 \varphi = 0, \quad F^{2,0} + \bar{\theta}^0 \varphi = 0, \quad F^{1,1} \wedge k + [\varphi, \bar{\varphi}] = 0$$

(88)

where $k$ is a Kahler form.

When we work on $\mathbb{R}^6$ we can again turn on the $\Omega$-background, which this time will depend on three parameters $\varepsilon_1, \varepsilon_2, \varepsilon_3$. Instead of the summation over two-dimensional partitions, as in (47), we would get the sum over three-dimensional ones (for $U(1)$ theory):

$$Z(q, \varepsilon_1, \varepsilon_2, \varepsilon_3) = \sum_{\pi} q^{\abs{\pi}} \mu_\pi(\varepsilon_1, \varepsilon_2, \varepsilon_3)$$

(89)

where the three-dimensional measure on partitions, the so-called equivariant vertex [54] measure, is defined as follows.
For the partition $\pi$ define the following functions of three variables $(q_1, q_2, q_3)$:

\[
V_\pi(q_1, q_2, q_3) = \sum_{(i,j,k) \in \pi} q_i^{1-1} q_j^{1-1} q_k^{1-1} \\
H_\pi(q_1, q_2, q_3) = \frac{1}{(1 - q_1)(1 - q_2)(1 - q_3)} - V_\pi \\
M_\pi(q_1, q_2, q_3) = -(1 - q_1)(1 - q_2)(1 - q_3)H_\pi(q_1, q_2, q_3)H_\pi(q_1^{-1}, q_2^{-1}, q_3^{-1})
\]

which describe the content of the partition, its complement and the deformation complex of the corresponding ideal sheaf. The equivariant vertex measure is given by

\[
\mu_\pi(\epsilon_1, \epsilon_2, \epsilon_3) = \prod_{\ell_+} \ell_+^{\epsilon_1, \epsilon_2, \epsilon_3} \prod_{\ell_-} \ell_-^{\epsilon_1, \epsilon_2, \epsilon_3}
\]

where the linear functions of three variables $\ell_\pm$ are found from

\[
M_\pi(q_1, q_2, q_3) = \frac{q_1 q_2 q_3}{(1 - q_1)(1 - q_2)(1 - q_3)} - \sum_{\ell_+} \ell_+^{\epsilon_1, \epsilon_2, \epsilon_3} + \sum_{\ell_-} \ell_-^{\epsilon_1, \epsilon_2, \epsilon_3}
\]

\[
q_1 = e^{-\epsilon_1}, \quad q_2 = e^{-\epsilon_2}, \quad q_3 = e^{-\epsilon_3}.
\]

The partition function (89) is shown in [45] to be equal to

\[
\mathcal{Z}(q, \epsilon_1, \epsilon_2, \epsilon_3) = M(q) \frac{\prod_{\ell_+} \ell_+^{\epsilon_1, \epsilon_2, \epsilon_3} \prod_{\ell_-} \ell_-^{\epsilon_1, \epsilon_2, \epsilon_3}}{\mathcal{Z}(q_1, q_2, q_3)}.
\]

In particular, if the $\Omega$-background is special, in the sense that the $SO(6)$ rotation generated by $(\epsilon_1, \epsilon_2, \epsilon_3)$ belongs to $SU(3)$, i.e. preserves the Calabi–Yau structure on $\mathbb{R}^6$, then the partition function reduces to the MacMahon function (86) which, as we saw already, counts the type A topological string amplitudes on $\mathbb{R}^6$ (up to an important issue of squaring the partition function).8

The fully equivariant partition function (93), not very surprisingly, agrees with the equivariant Gromov–Witten (GW) partition function (the calculation on the GW side is almost identical to (84), thanks to Mumford’s relations).

5.4. Kahler gravity interpretation

The three-dimensional partition point of view allows us to relate the topological string to the would-be Kahler gravity calculation. The idea is to interpret $\omega_0 - i g_s F$ as the Kahler form $\omega_X$ on the manifold $X$, which is obtained from $\mathbb{C}^3$ by the blowup along the ideal $I_\pi$. The weight

\[
e^{-\int_X \omega_0 \wedge \omega_0 / 6 g_s} q_1^{\epsilon_1} q_2^{\epsilon_2} q_3^{\epsilon_3}
\]

corresponds to

\[
\exp \left( -\frac{1}{6 g_s^2} \int_X \omega \wedge \omega \wedge \omega \right)
\]

which is the value of the Kahler gravity action (82).

5.5. Summary of the last two lectures

We have shown that the all-genus topological string partition function can be interpreted as a Kahler gravity partition function [33], which contains the sum over various space-time topologies, in accordance with the quantum foam expectations [29]. It has been shown in [56] that the non-perturbative completion of the type A string on Calabi–Yau manifold $M$ must include the D-brane contributions which in turn can be expressed using the topological

8 To get it fair: square.
The existence of such duality suggests a more fundamental theory, perhaps higher dimensional, which would explain this duality geometrically. Indeed, in the physical superstring context the S-duality is a geometric symmetry of M-theory.

The full topological string partition function would therefore depend on both the Kahler and complex moduli of Calabi–Yau, in other terms, on all the metric moduli. The natural way to combine these moduli is in the $G_2$ moduli of the seven-dimensional manifold $B = S^1 \times M$. Recently, Hitchin has proposed a functional on the space of closed 3-forms on $B$ whose critical points are the metrics of $G_2$ holonomy. Its quantization is currently being investigated, see [18, 60], and also [23].

6. Summary

In these lectures we have covered in some detail the exact non-perturbative calculations in four-dimensional $\mathcal{N} = 2$ systems, both from the gauge theory and the string theory points of view.

We gave an introduction to the gauge theories, their perturbation theory, effective actions, non-perturbative effects, notably those coming from instantons. We then proceeded with supersymmetry, studied the minimal and extended susy, their representations, twisted supersymmetry, discussed topological gauge theories and invariants of four-manifolds. After that we introduced the $\Omega$-backgrounds and studied the partition functions of the gauge theories in these backgrounds, including the chiral observables. We mostly talked about $\mathcal{N} = 2$ theories, but also mentioned briefly the superpotential calculations of Dijkgraaf and Vafa in the $\mathcal{N} = 1$ context.

We then discussed the string theory realizations of $\mathcal{N} = 2$ systems and the string theory point of view on the same partition functions. In particular, we saw that the instanton calculations in the gauge theory are mapped to the topological string calculations on the string theory side.

This is one of the manifestations of the existence of $Z$-theory, the framework which encompasses all the F-term calculations in string and gauge theories [60].

We then discussed an interesting corner of $Z$-theory, where the Gromov–Witten theory (the type $A$ topological closed string) meets the Donaldson–Thomas theory (the type $B$ topological open string). The Donaldson–Thomas computation was interpreted using the Kahler theory of gravity as realizing the summation over the space–time topologies, which one long suspected takes place in any quantum gravity theory. Thus, one of the outcomes of $Z$-theory is the string theory realization of the quantum space–time foam.
References

[34] D’Hoker E and Phong D 2001 *Preprint* hep-th/0110247
[38] Hooft G ’t 1998 *Preprint* hep-th/9812203
[40] Iqbal A, Kashani-Poor A-K and Vafa C 2003 Quantum foam and topological strings *Preprint* hep-th/0312022
[38] Klebanov I and Nekrasov N 1999 Preprint hep-th/9911096
Polchinski J 2000 Preprint hep-th/0011193
Kontsevich M and Manin Yu I 1994 Preprint hep-th/9402147
[40] Lawrence A, Nekrasov N and Vafa C 1998 Preprint hep-th/9803015
Losev A, Nekrasov N and Shatashvili S 1997 Testing Seiberg–Witten solution Strings, Branes and Dualities (NATO ASI Cargese) pp 359–72
Braden H and Marshakov A 2000 Preprint hep-th/0009060
Hollowood T 2003 Preprint hep-th/0302165
Maulik D, Nekrasov N, Okounkov A and Pandharipande R 2004 Preprint math.AG/0406092
Eguchi T and Sakai K 2002 Preprint hep-th/0203025
Eguchi T and Sakai K 2002 Preprint hep-th/0211213
Nakajima H and Yoshioka K 2003 Preprint math.AG/0306198
Nakajima H and Yoshioka K 2003 Preprint math.AG/0311058
Lawrence A and Nekrasov N 1997 Preprint hep-th/9706025
Nekrasov N and Okounkov A Equivariant vertex in preparation
Nekrasov N 2003 Preprint hep-th/0301184
Okounkov A and Pandharipande R 2002 Preprint math.AG/0204305
Okounkov A and Pandharipande R 2002 Preprint math.AG/0207233
Okounkov A and Pandharipande R 2003 Preprint math.AG/0308097
Okounkov A and Pandharipande R 2001 Preprint math.AG/0101147
Okounkov A and Pandharipande R 2002 Preprint math.AG/0204305
Okounkov A and Pandharipande R 2002 Preprint math.AG/0207233
Okounkov A and Pandharipande R 2003 Preprint math.AG/0308097
[53] Okounkov A and Reshetikhin N 2001 Correlation function of Schur process with application to local geometry of a random 3-dimensional Young diagram Preprint math.CO/0107056
Lectures on nonperturbative aspects of supersymmetric gauge theories

Okun L Introduction to gauge theories ITEP-84-043
[66] Okun L 1982 Leptons and Quarks (Amsterdam: North-Holland)


[70] Smilga A 1986 Yad. Fiz. 43 215–8
[72] Strominger A 2004 Lecture at Strings’04
[74] Vafa C 2000 Preprint hep-th/0008142