On localization and some of its applications in supersymmetric gauge theories

Frank FERRARI

Université Libre de Bruxelles
International Solvay Institutes

INTERNATIONAL SOLVAY INSTITUTES
Brussels

Lisbon, June 30th, 2009

Our goal is to describe methods that allow rigorous, first-principle derivations of non-perturbative results in four dimensional super Yang-Mills theories. Our main interest will be in the minimally supersymmetric N=1 case. This is ambitious but remarkably we shall have some success in this case. This is based on research published over the last two years, but also on work in progress.

We shall also discuss the case of N=2, for pedagogical reasons and also because the N=2 technology is an extension of the N=1 technology.

To achieve our goal, we have to answer several obvious questions:

1) If we claim that we do derivations from first principles of non-perturbative results, then we must have a non-perturbative definition of the theory to start with...

2) We also need a computational tool that allows to make calculation that are usually (naïvely?) believed to be intractable...

What allows us to answer both of these questions is a tool called LOCALIZATION.

The most important mathematical result for physics may well be the formula for the gaussian integral,

$$\int \mathrm{d}X \exp\left(-\frac{1}{2}{}^t X A X + {}^t J X\right) = \frac{(2\pi)^{n/2}}{\sqrt{\det A}} \exp\left(\frac{1}{2}{}^t J A^{-1} J\right)$$

This formula summarizes all what we know (at least in principle) from perturbative quantum field theory...

To go beyond perturbation theory, we need to learn how to compute integrals that are more complicated (at least superficially) than simple gaussians.

Usually, we "compute" more complicated integrals by doing a saddle-point approximation, which reduces the problem to gaussian integrals.

Remarkable property: it is known that in some cases, the saddle-point approximation yields the exact result.

Examples:

- The gaussian integral
- The Harish-Chandra Itzykson-Zuber integral

$$\int \mathrm{d}U \exp \mathrm{Tr}\, AUBU^{-1} = \frac{\det\left(e^{a_i b_j}\right)}{\prod_{1\leq i<j\leq n}(a_i - a_j)(b_i - b_j)}$$

- The canonical partition function of a classical integrable system.

This can be explained as follows.

Duistermaat and Heckman 1982, 1983
Berline and Vergne 1982, 1983

Let $\mathscr{M}$ be a 2p-dimensional manifold with a U(1) action generated by V, and differential forms $\omega(a)$ that are polynomials in a variable a (that can be viewed as an element of the Lie algebra of U(1) is one wish). Assign degree 2 to a and define the equivariant differential to be $\delta = \mathrm{d} + \iota_{aV}$. Taking into account the degree of a, this operator increases the total degree of a form by one. Note also that $\delta^2 = L_{aV}$ and that δ acts as a graded differential.

An equivariant differential is a form invariant under the U(1) action, $L_V\omega = 0$. On these forms, $\delta^2 = 0$ and we can define the equivariant cohomology at some degree.

This is a useful notion because even on a topologically trivial space (like $\mathbb{R}^{2p}$) the equivariant cohomology can be non-trivial: we can find closed equivariant forms that are not exact.

This is nice because the integrals of equivariantly-closed forms

$$I(a) = \int_{\mathscr{M}} \omega(a)$$

can be computed exactly using the following trick.

Deform the integral to $I(a,\epsilon) = \int_{\mathscr{M}} \omega(a) \wedge \exp(-\frac{\delta\eta}{\epsilon})$

η must be an odd-degree equivariant form such that $\delta\eta$ has a positive definite zero-form part $(\delta\eta)_0$.

Then $I(a,\epsilon)$ does not depend on ϵ.

$$\frac{\partial I}{\partial \epsilon} = -\int_{\mathscr{M}} \omega(a) \wedge \delta\eta \wedge \exp\Big(-\frac{\delta\eta}{\epsilon}\Big)$$

$\delta\eta$ is an even form

$$= -\int_{\mathscr{M}} \delta\Big(\omega(a) \wedge \eta \wedge \exp\Big(-\frac{\delta\eta}{\epsilon}\Big)\Big)$$

ω and $\delta\eta$ are equivariantly closed

$$= -\int_{\mathscr{M}} \mathrm{d}\Big(\omega(a) \wedge \eta \wedge \exp\Big(-\frac{\delta\eta}{\epsilon}\Big)\Big) = 0\,.$$

$\delta = d$ since we pick the top form

with good boundary conditions

Since $I(a,\epsilon)$ does not depend on ϵ, we can compute it in the $\epsilon \to 0$ limit, for which a saddle-point approximation becomes exact.

euclidean metric invariant under U(1)

$$\eta = a g_{\mu\nu} V^\nu \mathrm{d}x^\mu$$

$$\delta\eta = a^2 |V|^2 + a \nabla_\nu V_\mu \mathrm{d}x^\nu \wedge \mathrm{d}x^\mu$$

The integral then localizes on the fixed point of the U(1) action (corresponding to V=0) and we obtain the Berline-Vergne formula,

$$\int_{\mathscr{M}} \omega(a) = (-2\pi)^p \sum_{x^*} \omega_0(a)|_{x^*} \frac{\sqrt{\det g(x^*)}}{\mathrm{Pf}\, \partial_\mu V_{a\nu(x^*)}}$$

zero-form part

Let's rephrase the discussion in a suggesting language.

$$\omega = \frac{1}{q!}\omega_{\mu_1\cdots\mu_q}\theta^{\mu_1}\cdot\wedge\theta^{\mu_q}\wedge \mathrm{d}x^{\mu_1}\cdot\wedge\mathrm{d}x^{\mu_q}$$

anticommuting Grassman variables

odd form $\longleftrightarrow$ fermionic

even form $\longleftrightarrow$ bosonic

$$\mathrm{d} = \theta^\mu \frac{\partial}{\partial x^\mu}$$

$$\iota_V = V^\mu \frac{\partial}{\partial \theta^\mu}$$

$$\delta = V^\mu \frac{\partial}{\partial \theta^\mu} + \theta^\mu \frac{\partial}{\partial x^\mu} = \mathcal{Q}$$

supercharge

Field theory set-up: $\int \mathscr{D}\mu\, \mathcal{O}\, e^{-S}$

Assume that we have a fermionic operator $\mathscr{Q}$ (typically a suitable linear combination of supercharges). Its square is a bosonic transformation that vanishes on a suitable class of operators (e.g. if it vanishes on gauge invariant operators we are dealing with the equivariant cohomology of the gauge group).

$\mathscr{Q} \cdot S = 0$ (SUSY invariance)

$\mathscr{Q} \cdot \mathcal{O} = 0$ (restriction on the type of operators that we can consider: "chiral sector", "SUSY-preserving sector", "topological observables")

$$S = S_0 + \mathcal{Q} \cdot \Pi\,, \quad \mathcal{Q} \cdot S_0 = 0$$

As far as the integral converges, Π can be chosen at will, and if the purely bosonic part (i.e. independent of fermions) of Π is definite positive, the path integral localizes on the set of field configurations satisfying

$$\mathcal{Q} \cdot \Pi = 0$$

This yields typically a set of partial differential equations and/or algebraic constraints, whose most general solution is parametrized by some parameters ("moduli") μ . The original path integral is then replaced by an integral over this moduli space (which can be infinite-dimensional).

N=2 super Yang-Mills (in the euclidean)

$$\mathrm{SU}(2)_1 \times \mathrm{SU}(2)_2 \times \mathrm{SU}(2)_\mathrm{R} \times \mathrm{U}(1)_\mathrm{A}$$

μ-indices α-indices $\dot{\alpha}$-indices $\underline{\alpha}$-indices

$$\mathrm{SU}(2)_1 \times \mathrm{SU}(2)_\mathrm{diag}$$

4D rotation group
(euclidean Lorentz group)

$$A_\mu\,,\ \phi\,,\ \phi^\dagger \qquad\qquad \lambda_{\underline{\alpha}\alpha}\,,\ \bar{\lambda}_{\underline{\alpha}\dot{\alpha}}$$

$$\eta = \epsilon^{\underline{\alpha}\dot{\alpha}}\bar{\lambda}_{\underline{\alpha}\dot{\alpha}}\,,\ \chi_{\mu\nu} = (\bar{\sigma}_{\mu\nu})^{\dot{\alpha}}_{\ \underline{\alpha}}\bar{\lambda}^{\underline{\alpha}}_{\dot{\alpha}}\,,\ \psi_\mu = (\sigma_\mu)_{\alpha\underline{\alpha}}\lambda^{\underline{\alpha}\alpha}$$

$$\bar{\lambda}^{\dot{\alpha}}_{\underline{\alpha}} = \frac{1}{2}\delta^{\dot{\alpha}}_{\underline{\alpha}}\eta - \frac{1}{2}(\bar{\sigma}_{\mu\nu})^{\dot{\alpha}}_{\ \underline{\alpha}}\chi_{\mu\nu}\,,\ \lambda^{\underline{\alpha}\alpha} = -\frac{1}{2}\bar{\sigma}^{\underline{\alpha}\alpha}_{\mu}\psi_\mu$$

$\mathscr{Q} = \epsilon^{\underline{\alpha}\dot{\alpha}} \bar{Q}_{\underline{\alpha}\dot{\alpha}}$ $\qquad$ $\mathscr{Q}^2 =$ gauge transformation

$$\mathscr{Q} \cdot A_\mu = \psi_\mu \,, \;\; \mathscr{Q} \cdot \psi_\mu = 2\sqrt{2} \nabla_\mu \phi \,, \mathscr{Q} \cdot \phi = 0$$

$$\mathscr{Q} \cdot \phi^\dagger = -\sqrt{2}\eta \,, \;\; \mathscr{Q} \cdot \eta = 2i[\phi, \phi^\dagger]$$

$$\mathscr{Q} \cdot \chi_{\mu\nu} = -2F^-_{\mu\nu} - 2D_{\mu\nu}$$

$$\mathscr{Q} \cdot D_{\mu\nu} = -\left(\nabla_\mu \psi_\nu - \nabla_\nu \psi_\mu\right)^- + i\sqrt{2}[\phi, \chi_{\mu\nu}]$$

$$\mathscr{L} = \frac{1}{g^2} \operatorname{tr}\left(\frac{1}{2}F^2 + 2|\nabla\phi|^2 + [\phi, \phi^\dagger]^2 - D^2 + \text{fermions}\right)$$

$$- \frac{i\theta}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \operatorname{tr} F_{\mu\nu} F_{\rho\sigma}$$

$$= \frac{1}{g^2} \mathscr{Q} \cdot \Pi - \frac{i\tau}{16\pi} \epsilon_{\mu\nu\rho\sigma} \operatorname{tr} F_{\mu\nu} F_{\rho\sigma} \qquad \tau = \frac{\theta}{2\pi} + i\frac{4\pi}{g^2}$$

$$\Pi = \Pi_1 + \Pi_2 + \Pi_3$$

$$\Pi_1 = \frac{1}{\sqrt{2}} \operatorname{tr} \psi_\mu \nabla_\mu \phi^\dagger \,,$$

$$\Pi_2 = -\frac{i}{2} \operatorname{tr} \eta [\phi, \phi^\dagger] \,,$$

$$\Pi_3 = \frac{1}{2} \operatorname{tr} \chi_{\mu\nu} (D_{\mu\nu} - F_{\mu\nu})$$

$$\mathscr{Q} \cdot \Pi_1 = 2 \operatorname{tr} |\nabla \phi|^2 + \text{fermions} \,,$$

$$\mathscr{Q} \cdot \Pi_2 = \operatorname{tr} [\phi, \phi^\dagger]^2 + \text{fermions} \,,$$

$$\mathscr{Q} \cdot \Pi_3 = \operatorname{tr} (|F^-|^2 + |D|^2) + \text{fermions}$$

On constrained instantons

The usual approach to N=2 (Seiberg-Witten) is to perform a semi-classical approximation at weak coupling (instantons) and resum the instanton series, which is supposed to yield the exact results.

People have tried to use this method from 1994-1995 to check SW from direct instanton calculations.

The problem is: strictly speaking, there is no instanton solutions (no smooth solution to the classical e.o.m.) for non-zero vev of the scalar field, whereas the semi-classical approximation should be valid at large vev:

$$\langle \phi \rangle = \mathrm{diag}(a_1, \ldots, a_N) \,, \;\; |a_i - a_j| \gg \Lambda$$

So people have devised some sort of modified semi-classical expansion: instead of using exact solutions to the e.o.m., which do not exist, they use solutions of modified equations, then argue that in the semi-classical limit the neglected terms are subleading. This is known as the "constrained instanton" method.

N. Dorey et al., hep-th/0206063

A basic feature of this method is that, since the field configurations one integrate over are not exact solutions, there is a potential on the would-be "moduli" space.

Somewhat mysteriously, this kind of awkward approximation scheme is supposed to yield the exact result...

Let us see that this is actually a consequence of localization.

F. F and V. Wens, unpublished

$$\mathscr{L} = \frac{1}{g^2}\mathscr{Q}\cdot\Pi - \frac{i\tau}{16\pi}\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,F_{\mu\nu}F_{\rho\sigma}$$

The standard semiclassical approximation corresponds to localization on the solutions to

$$\mathscr{Q}\cdot\Pi = 0$$

which is the empty set (or more precisely which corresponds only to singular field configurations). But the localization theorem clearly does not work a priori if the fixed points are singular points of the manifold over which we integrate.

However, one has $\Pi = \Pi_1 + \Pi_2 + \Pi_3$. One could try to use any of the three terms to localize.

$$\mathcal{Q} \cdot \Pi_1 = 2\,\mathrm{tr}\,|\nabla\phi|^2 + \text{fermions}\,,$$
$$\mathcal{Q} \cdot \Pi_2 = \mathrm{tr}[\phi, \phi^\dagger]^2 + \text{fermions}\,,$$
$$\mathcal{Q} \cdot \Pi_3 = \mathrm{tr}(|F^-|^2 + |D|^2) + \text{fermions}$$

Let us use the third term only. Then the fixed point equations are simply

$$F^- = 0\,,\ D = 0$$

The fixed points that contribute must have finite action and this implies that the boundary conditions must be such that the gauge field is pure gauge at infinity. With these b.c., the most general solution is given by the ADHM construction.

So the integral over the gauge fields in reduced to an integral over the ADHM moduli space.

It is not too difficult to deal with the fermionic path integrals in a general instanton background.

One then find precisely the prescription of the constrained instanton formalism, now derived in a clear way from a localization theorem.

The potential "on the moduli space" now comes simply from the fact that the solutions to $\mathscr{Q} \cdot \Pi_3 = 0$ yield non-zero $\mathscr{Q} \cdot \Pi_1$ and $\mathscr{Q} \cdot \Pi_2$.

It is also clear that the result is exact (there can be no other non-perturbative contribution).

Summary

$$\int \mathscr{D}\varphi\, \mathcal{O}\, e^{-S_0 - \mathscr{Q}\cdot\int \mathrm{d}^4x\, \Pi_3}$$

$\downarrow$ localization on $\mathscr{Q}\cdot\Pi_3|_{\text{bosonic}} = 0$

$$\int_{\mathscr{M}} \mathscr{D}\mu\, \mathcal{O}(\mu)\, e^{-S_0(\mu)}$$

$\downarrow$ $\mathscr{M} = \cup_{k\geq 0}\mathscr{M}_k$, k=topological charge

$$\sum_{k\geq 0}\int_{\mathscr{M}_k} \mathrm{d}\mu_k\, \mathcal{O}(\mu)\, e^{-S_0(\mu)}$$

finite dimensional integral

So we see an example where the idea of localization yields a definition of what we mean by the non-perturbative quantum field theory (in the particular sector of Q-closed operators).

Note that consistency implies that the instanton series must converge! Indeed, we don't have any other possible contribution to the path integral and thus if we had only an asymptotic series then it would mean that the theory do not have a non-perturbative completion.

The series usually has a finite radius of convergence, which means that we need to analytically continue the series. This is a mathematically unambiguous procedure that is associated with a lot of physics. We shall have more to say about this later.

The spaces $\mathscr{M}_k$ have singularities corresponding to zero-size instantons. These singularities actually correspond to the singular fixed points of the "big" localization prescription $\mathscr{Q} \cdot \Pi = 0$.

One can show that all the physical information of the theory is contained in operators for which the singularities are actually integrable.

On the other hand, the singularities play a role for operators that are ambiguous in the quantum theory. Since this is an important point of principle that we shall encounter later, let us give more details on this point.

$u_k = \mathrm{tr}\,\phi^k$

Because the gauge group is U(N) (or SU(N)), we are dealing with NXN matrices and thus there are relations of the form $u_{N+p} = P_p(u_1, \ldots, u_N)$ that imply that all the physical information is contained in u_k for $1 \leq k \leq N$.

It turns out that the operator u_k is integrable for $k<2N$. However, for $k \geq 2N$, there is an ambiguity of the form zeroXinfinity that shows up at one instanton. This ambiguity is physical: the operator can indeed be redefined

$$u_{2N} \to u_{2N} + aq$$

consistently with all the U(1) charges of the theory. In particular,

$$u_{2N} = P_N(u_1, \ldots, u_N) + aq = \mathscr{P}_N(u_1, \ldots, u_N)$$

More generally,one can define

$$u_{N+p} = \mathscr{P}_p(u_1, \ldots, u_N)$$

It should be obvious that $\mathscr{P}_p$, corresponding to a particular arbitrary choice of what is meant by the operator u_{N+p}, can be chosen arbitrarily (as long as it has the correct classical limit and is consistent with the U(1) symmetries of the theory). In particular, if doesn't make any sense to claim that one computes $\mathscr{P}_p$!

One can choose a regularization of the instanton moduli space (for example: the non-commutative deformation), and this will fix unambiguously the definition of all the operators.

This is similar to the definition of composite operators in ordinary perturbation theory, by choosing a particular regularization and renormalization scheme.

To remember:

Singularities on the moduli space are associated with arbitrary choices in the field theory (corresponding to field redefinitions) but are not associated with non-trivial physics.

One must be careful in interpreting the result of calculations, by making the distinction between non-trivial physical content and arbitrary choices!

Historical remark

Constrained instantons are relevant because we formulate the theory on $\mathbb{R}^4$ and we must specify the b.c. at infinity for the scalar.

One may ask whether the full structure of the action as S = topological term + Q( ) is useful.
This structure was actually used in 1988 by Witten to formulate the theory on an arbitrary compact four-manifold, in which the coupling to gravity is made by using the "twisted" Lorentz that we have singled-out in our notation. On a compact manifold, there is no problem of b.c.

In this sense the theory on $\mathbb{R}^4$ is more subtle than on a compact manifold - and was indeed solved using localization much later.

$$\sum_{k\geq 0}\int_{\mathscr{M}_k} \mathrm{d}\mu_k\, \mathcal{O}(\mu)\, e^{-S_0(\mu)} \qquad q = e^{2i\pi\tau}$$

These integrals are still far too complicated to compute.

$$e^{-S_0(\mu)} = q^k e^{-\mathscr{Q}\cdot(\pi_1+\pi_2)|_\mu}$$

One can use this structure to localize the integrals over a subspace of $\mathscr{M}_k$ (Hollowood 2002) but the resulting integrals are still to hard to evaluate.

In 2002, Nekrasov found a remarkable way to enhance the power of the localization technique. The idea is to introduce a deformation of the theory, that preserves the nice features of the underformed theory (existence of Q etc) while breaking the rotation group (going to the equivariant cohomology of the rotation group).

The deformation we need is called the Omega background. The best way to understand it is by doing a Scherk-Schwarz dimensional reduction from the six dimensional N=1 gauge theory.

$$\mathrm{SU}(2)_1 \times \mathrm{SU}(2)_2 \times \mathrm{SU}(2)_\mathrm{R} \times \mathrm{U}(1)_\mathrm{A}$$

$$\mathrm{SU}(2)_1 \quad \times \quad \mathrm{SU}(2)_\mathrm{diag}$$

5th direction -> use Ω_1 in $\mathrm{SU}(2)_1 \times \mathrm{SU}(2)_\mathrm{diag}$

6th direction -> use Ω_2 in $\mathrm{SU}(2)_1 \times \mathrm{SU}(2)_\mathrm{diag}$

$$[\Omega_1, \Omega_2] = 0$$

$$\Omega = \Omega_1 + i\Omega_2 = \Omega^+ + \Omega^-$$

self-dual anti self-dual

This point of view on the Omega background allows to find easily the supercharges that are preserved: they correspond to the supercharges of the 6D theory that are invariant under the Omega-rotations.

$\Omega^+ = 0$: four supercharges

In general: a deformation of Q is preserved and was used by Nekrasov.

Actually there is a second supercharge, Q', that is preserved! It will play a crucial rôle in the N=1 story.

$\mathcal{Q}^2 =$ gauge transformation + rotation $\qquad W_\mu = \Omega_{\mu\nu} x^\nu$

$$\mathcal{Q} \cdot A_\mu = \psi_\mu \,, \;\; \mathcal{Q} \cdot \chi_{\mu\nu} = -2F^-_{\mu\nu} - 2D_{\mu\nu}$$

$$\mathcal{Q} \cdot \phi = -W_\mu \psi_\mu \,, \;\; \mathcal{Q} \cdot \psi_\mu = 2\sqrt{2}\big(\nabla_\mu \phi + W_\nu F_{\mu\nu}\big)$$

$$\mathscr{L} = \frac{1}{g^2} \operatorname{tr}\Big(\frac{1}{2}F^2 + 2|\nabla\phi - \iota_W F|^2 + \big([\phi,\phi^\dagger]$$

$$+ i(\nabla_{W^*}\phi - \nabla_W \phi^\dagger) + i W^*_\mu W_\nu F_{\mu\nu}\big)^2 - D^2 + \text{fermions}$$

$$\mathscr{L} = \frac{1}{g^2} \operatorname{tr}\Big(\frac{1}{2}F^2 + 2|\nabla\phi - \iota_W F|^2 + \big([\phi,\phi^\dagger] + i(\nabla_{W^*}\phi - \nabla_W \phi^\dagger)$$

$$+ i W^*_\mu W_\nu F_{\mu\nu}\big)^2 - D^2 + \text{fermions}\Big) - \frac{i\theta}{32\pi^2}\epsilon_{\mu\nu\rho\sigma} \operatorname{tr} F_{\mu\nu} F_{\rho\sigma}$$

$$= \frac{1}{g^2}\mathcal{Q} \cdot (\Pi_1 + \Pi_2 + \Pi_3) - \frac{i\tau}{16\pi}\epsilon_{\mu\nu\rho\sigma} \operatorname{tr} F_{\mu\nu} F_{\rho\sigma}$$

Localization using

$$\Pi_3 = \frac{1}{2}\,\mathrm{tr}\,\chi_{\mu\nu}(D_{\mu\nu} - F_{\mu\nu})$$

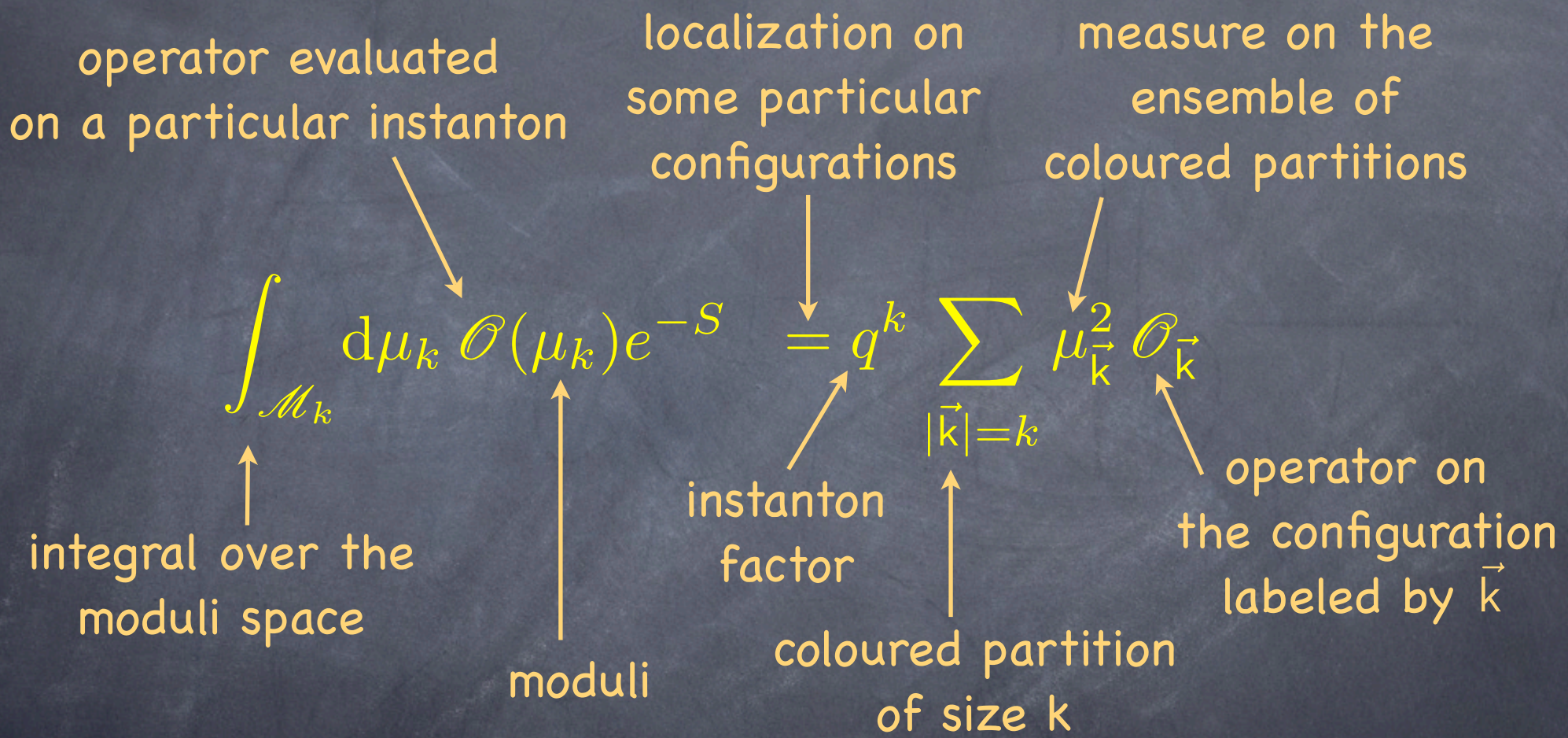
yields constrained instantons in the Omega background.

$$\sum_{k\geq 0}\int_{\mathscr{M}_k} \mathrm{d}\mu_k\,\mathcal{O}(\mu)\,e^{-S_0(\mu)}$$

$$e^{-S_0(\mu)} = q^k\,e^{-\mathscr{Q}\cdot(\pi_1+\pi_2)|_\mu}$$

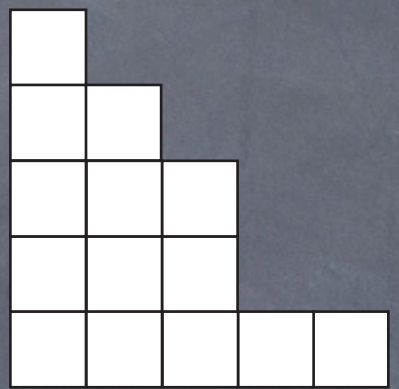
Using this structure, the integrals over $\mathscr{M}_k$ for fixed k localize on a finite set of fixed points!

The end result has a remarkable form.


operator evaluated on a particular instanton
localization on some particular configurations
measure on the ensemble of coloured partitions
$\int_{\mathscr{M}_k} \mathrm{d}\mu_k \, \mathscr{O}(\mu_k) e^{-S} = q^k \sum_{|\vec{\mathrm{k}}|=k} \mu_{\vec{\mathrm{k}}}^2 \, \mathscr{O}_{\vec{\mathrm{k}}}$
integral over the moduli space
moduli
instanton factor
coloured partition of size k
operator on the configuration labeled by $\vec{\mathrm{k}}$

The ensemble of (coloured) partitions

Ordinary partition k



14 = 5 + 3 + 3 + 2 + 1

Number of boxes = |k| instanton charge

$$\mu_{\mathrm{k}} = \frac{1}{|\mathrm{k}|!\epsilon^{|\mathrm{k}|}} \dim R_{\mathrm{k}}$$

measure

deformation parameter

dimension of the associated irrep. of the symmetric group

A coloured partition is a collection of N ordinary partition, $\vec{k} = (k_1, \ldots, k_N)$, with a measure factor that generalizes the case of the ordinary partitions.

When the deformation parameter ϵ (i.e. the Omega background) goes to zero, the sum over coloured partitions can be evaluated by using a suitable saddle-point approximation (the sum is dominated by a single very large coloured partition, and the shape of the associated generalized Young tableau can be computed explicitly).

This formalism was used by Nekrasov et al. in 2002-2003 to solve for the N=2 gauge theories (Seiberg-Witten).

It was a truly remarkable achievement, crowning many years of developments in instanton calculus.

N. Nekrasov, Adv. Theor. Math. Phys. 7 (2004) 831, hep-th/0206161
hep-th/0306211
N. Nekrasov and A. Okounkov, hep-th/0306238
N. Nekrasov and S.Shadchin, Comm. Math. Phys. 252 (2004) 359,
hep-th/0404225
A. Marshakov and N. Nekrasov, JHEP 01 (2007) 104, hep-th/0612019

Dorey, Hollowood, Khoze, Mattis, Slater
Bellisai, Fucito, Morales, Poghossian, Tanzini, Travaglini

Application of localization: computation of Wilson loops in N=4 and N=2 gauge theories

Remember: $\mathcal{Q} \cdot A_\mu = \psi_\mu \,, \;\; \mathcal{Q} \cdot \phi = -W_\mu \psi_\mu$

Consider a circle C in the 1-2 plane, of radius R, centered at the origin. The tangent vector is given by

$$t_\mu = -\Omega_{\mu\nu} x_\nu \,, \;\; \Omega_{12} = -\Omega_{21} = \frac{1}{R}$$

Then $$\mathcal{Q} \cdot \oint_C \left(A + \phi \mathrm{d}s \right) = 0$$

We can compute the corresponding Wilson loop using localization,

$$W(C) = P \, \mathrm{tr} \exp\left(i \oint_C \left(A + \phi \mathrm{d}s \right) \right)$$

If you put the theory on a sphere S^4

$$\langle W(C)\rangle = \frac{\int \mathrm{d}a\, |Z_{\rm pert}(a,R)|^2\, e^{-4\pi^2 R^2 \,\mathrm{tr}\, a^2/g^2}\, \mathrm{tr}\, e^{2i\pi Ra}}{\int \mathrm{d}a\, |Z(a,R)|^2\, e^{-4\pi^2 R^2 \,\mathrm{tr}\, a^2/g^2}}$$

asymptotics of ϕ at infinity

includes determinant of fluctuations, includes instantons

Amazing formula (Pestun, 2007): includes both an infinite set of perturbative corrections and an infinite set of non-perturbative corrections (this is the only formula of this type that I know of in 4D gauge theories).

$$\langle W(C)\rangle_{\mathcal{N}=4} = \frac{\int \mathrm{d}a\, e^{-4\pi^2 R^2 \,\mathrm{tr}\, a^2/g^2}\, \mathrm{tr}\, e^{2i\pi Ra}}{\int \mathrm{d}a\, e^{-4\pi^2 R^2 \,\mathrm{tr}\, a^2/g^2}}$$

Erickson, Semenoff, Zarembo 2000

Going to N=1???

F.F., V. Wens et al., 2007-now

One needs an extra structure to go to N=1, because an N=2-breaking superpotential term is not Q-closed in general.

It turns out that the Omega-deformed theory preserves a second supercharge,

$$\mathscr{Q}' = \Omega^{-}_{\mu\nu}(\bar{\sigma}_{\mu\nu})^{\dot{\alpha}}{}_{\underline{\alpha}}\bar{Q}^{\underline{\alpha}}_{\dot{\alpha}}$$

and we can use a combination of Q and Q' to study N=1.

This is actually reminiscent of the N=1 twist by Witten (1994) on Kahler manifold. He was using the fact that the holonomy on a Kahler manifold is

$$\mathrm{U}(1) \times \mathrm{SU}(2)_2 \subset \mathrm{SU}(2)_1 \times \mathrm{SU}(2)_2$$

to construct a Q'.

In our case, we are using the fact that the Omega-rotation is in the same $\mathrm{U}(1) \times \mathrm{SU}(2)_2$ subgroup of the Lorentz group.

There is still a fundamental problem in trying to apply localization in N=1 theories, which is actually probably the main reason why it was not attempted before very recently.

In N=1 theories, there is always a corner of the moduli space where the theory is arbitrarily weakly coupled and thus the instanton series converge.

This is NOT the case in general for theories with N=1 supersymmetry. N=1 theories generically do not have a moduli space, but a discrete set of vacua. Some of these vacua are intrinsically strongly coupled. Observables can be given by fractional instanton series in these vacua.

Morally speaking, this means that viewing the infinite dimensional instanton moduli space as a union of finite dimensional pieces,

$$\mathscr{M} = \cup_{k \geq 0} \mathscr{M}_k$$

is probably not the best way of approaching the problem.

Yet, it is the only point of view we know how to work with. How can we reconcile this with our goal which is to solve N=1 theories?

Fundamental idea: the variables a, which play such a prominent rôle in N=2, are actually also playing an important rôle in N=1!

Instead of trying to compute directly the correlators, we use a two-step procedure. We start by considering off-shell correlators obtained by computing path integrals with fixed boundary conditions at infinity.

F. F., JHEP 10 (2007) 065, arXiv:0707.3885

$$\langle \boldsymbol{a}|\mathscr{O}|\boldsymbol{a}\rangle = \frac{\int_{X_\infty=\mathrm{diag}\,\boldsymbol{a}} \mathrm{d}\mu\, \mathscr{O} e^{-\mathcal{S}_\mathrm{E}}}{\int_{X_\infty=\mathrm{diag}\,\boldsymbol{a}} \mathrm{d}\mu\, e^{-\mathcal{S}_\mathrm{E}}}$$

$$\boldsymbol{a} = (a_1, \ldots, a_N)$$

By choosing appropriately a ($|a_i - a_j| \gg \Lambda$) the path integrals can be forced to be weakly coupled, and the result is then given exactly by a convergent instanton series. For arbitrary a, the answer is obtained by analytic continuation.

This formalism, based on using the variables a in the N=1 theories, I will call the "microscopic formalism."

The variables a play a role in the microscopic formalism which is in some sense analogous to the filling fractions in the Dijkgraaf-Vafa matrix model formalism.

Of course, to compute the physical correlators, we need to fix the variables a to physical on-shell values. In other words, there is a quantum superpotential for a, that we call the microscopic superpotential $W_{\rm mic}(a)$.

This microscopic superpotential has two fundamental properties that distinguish it from all the previously used quantum superpotentials.

1) it can always be computed exactly in the instanton approximation for some values of a and then obtained for arbitrary values of a by analytic continuation.

2) its critical points are in one-to-one correspondence with the full set of vacua of the theory (for example, the solutions corresponding to any number of cuts in the matrix model are obtained as critical points of a single superpotential)

x = solution a* to $\mathrm{d}W_{\mathrm{mic}} = 0$

a-plane

Note: expanding around the strongly coupled solutions a* automatically produce fractional instanton series for the correlators.

To make things very concrete, let us explain again the procedure, in the N=1 theory obtained by deforming N=2 by adding the tree-level superpotential,

$$W = \frac{1}{2}\,\mathrm{tr}\,\phi^2$$

focusing on the gluino condensate $\langle\lambda\lambda\rangle$.

1) compute the off-shell correlator $\langle a|\lambda\lambda|a\rangle$ as a function of a.

2) compute the microscopic superpotential. In the present case it is simply given by

$$W_{\mathrm{mic}} = \frac{1}{2} m \langle a|\,\mathrm{tr}\,\phi^2|a\rangle$$

3) extremize this superpotential. We find N distinct solutions a_k, $0 \le k \le N-1$.

4) plug the values a_k into $\langle a|\lambda\lambda|a\rangle$ and find the physical correlators,

$$\langle a_k|\lambda\lambda|a_k\rangle = mq^{1/N}e^{2i\pi k/N}$$

We could of course take the $m \to \infty$ limit and obtain the gluino condensate (and thus a derivation of chiral symmetry breaking) in the pure N=1 gauge theory.

We now want to be more ambitious and actually compute any chiral correlator. As we shall see, the resulting structure is very interesting.

We limit the discussion for simplicity to the theory with no matter hypermultiplets, but an arbitrary polynomial tree-level superpotential $W(\phi)$.

The most general chiral operator in the theory is a polynomial in the basic operators

$$u_k = \mathrm{tr}\,\phi^k\,,\;\; v_k = -\frac{1}{16\pi^2}\,\mathrm{tr}\,\lambda\lambda\phi^k$$

for $1\leq k\leq N$ and $0\leq k\leq N-1$ respectively.

It is convenient to introduce generating functions:

Off-shell:

$$R_{\rm mic}(z;a) = \sum_{k\geq 0}\frac{\langle a|u_k|a\rangle}{z^{k+1}}\,,\;\; S_{\rm mic}(z;a) = \sum_{k\geq 0}\frac{\langle a|v_k|a\rangle}{z^{k+1}}$$

solve $\mathrm{d}W_{\rm mic} = 0$

On-shell:

$$R^{|0\rangle}(z;a) = \sum_{k\geq 0}\frac{\langle 0|u_k|0\rangle}{z^{k+1}}\,,\;\; S^{|0\rangle}(z;a) = \sum_{k\geq 0}\frac{\langle 0|v_k|0\rangle}{z^{k+1}}$$

Let us remember that in principle, there is another (conjectured) way to compute these chiral operators, by using the matrix model (which is itself a direct consequence of the closed string description of the gauge theory).

In this formalism, one uses generating functions

$$R_{\rm MM}(z;s)=\sum_{k\geq 0}\frac{\langle s|u_k|s\rangle}{z^{k+1}}\ ,\ S_{\rm MM}(z;s)=\sum_{k\geq 0}\frac{\langle s|v_k|s\rangle}{z^{k+1}}$$

that are not to be confused with the generating functions $R_{mic}(z;a)$ and $S_{mic}(z;a)$ of the microscopic approach (open string approach) that we have introduced previously.

In the MM formalism, the link with the physical correlators is made by extremizing effective glueball superpotentials that depend on the filling fractions. These superpotentials (one for each given number of cuts) are conjectured (they follow directly from the closed string description where they coincide with flux superpotentials).

The glueball superpotentials should not be confused with our microscopic superpotential (they do not even depend on the same variables).

We shall come back to the relation between the MM formalism and our approach later.

Application: generalized Konishi anomaly equations

This is a set of algebraic constraints on the operators in the theory.

$$\mathrm{Tr}\,(W'(X)\mathscr{O}) = 0$$

A complete set of equations is obtained in the classical theory by using $\mathscr{O} = X^{n+1}$, $n \geq -1$.

arbitrary operator

This amounts to considering the variations $\delta X \sim X^{n+1}$ which are generated by the Virasoro operators

$$L_n = -X^{n+1}\frac{\delta}{\delta X}\,, \quad [L_n, L_m] = (n-m)L_{n+m}$$

Using $\delta X^k = -X^{n+k+1}$, or $L_n \cdot u_k = -k u_{n+k}$ we get the equations

$$L_n \cdot \mathrm{Tr}\, W = L_n \cdot \sum_{k\geq 0} \frac{g_k}{k+1} u_{k+1} = -\sum_{k\geq 0} g_k u_{n+k+1} = 0$$

if $W'(X) = \sum_{k\geq 0} g_k X^k$.

$$W_n'(z) = R(V) = \text{polynomial} + N \sum_{k \geq 0} g_k \omega_{n+k+1}(z)$$

$$R(z) = \sum_I \frac{c_I}{z - a_I}$$

This is not the expected solution, because the c_I are arbitrary complex numbers, not positive integers. The closed string description is missing something. Intuitively, open strings (matrices) can be built from the closed strings (gauge invariant operators) only when some quantization conditions are satisfied.

The fundamental idea is that the u_k are not independent: there must exist relations of the form

$$u_{N+p} = P_p(u_1, \ldots, u_N)$$

for all $p \geq 1$, for some polynomials P_p that are easy to compute.

It is not difficult to prove (simple algebraic lemma) that the above relations are consistent with

$$R(z) = \sum_I \frac{c_I}{z - a_I}$$

if and only if the c_I are positive integers. As we shall see, this idea has deep consequences for the closed string formulation of the quantum theory.

Let us now discuss the quantum perturbative theory.

$$L_n = -X^{n+1}\frac{\delta}{\delta X}\,, \quad [L_n, L_m] = (n-m)L_{n+m}$$

$$\underbrace{-N\sum_{k\geq 0} g_k u_{n+k+1}}_{\text{classical term}} + \underbrace{2\sum_{k_1+k_2=n} u_{k_1}v_{k_2}}_{\substack{\text{one-loop anomaly} \\ \text{(generalized Konishi anomaly)}}} = 0$$

$$J_n = \frac{W^2}{16\pi^2}\frac{\delta}{\delta X}\,, \quad [L_n, J_m] = (n-m)J_{n+m}\,, \quad [J_n, J_m] = 0$$

$$-N\sum_{k\geq 0} g_k v_{n+k+1} + \sum_{k_1+k_2=n} v_{k_1}v_{k_2} = 0$$

This is the loop equation of the planar matrix model: the v_k are identified with matrix averages.

$$-N\sum_{k\geq 0} g_k u_{n+k+1} + 2\sum_{k_1+k_2=n} u_{k_1} v_{k_2} = 0$$

$$-N\sum_{k\geq 0} g_k v_{n+k+1} + \sum_{k_1+k_2=n} v_{k_1} v_{k_2} = 0$$

But we also have the constraints

$$u_{N+p} = P_p(u_1, \ldots, u_N)$$

which are automatically valid to all orders of perturbation theory. This is a fundamental ingredient that distinguishes the planar $n \to \infty$ matrix model and the finite N gauge theory. The solution of the above three sets of equations is just the classical theory,

$$R(z) = \sum_I \frac{N_I}{z - a_I} \qquad S(z) = 0$$

So we see that the anomaly equations DO get non-perturbative corrections (otherwise we would not have quantum corrections to the correlators). For example,

$$-N\sum_{k\geq 0} g_k u_{n+k+1} + 2\sum_{k_1+k_2=n} u_{k_1} v_{k_2} +$$
$$-N\sum_{k\geq 0} g_k u_{n+k+1} + 2\sum_{k_1+k_2=n} u_{k_1} v_{k_2} = 0$$
$$\sum_{r\geq 1} q^r \left(\sum_{k\geq 0} g_k A^{(r)}_{n,k}(u_p) + \sum_{t\geq 0} C^{(r)}_{n,t}(u_p)\, v_t \right) = 0$$

In particular, the generators L_n and J_n, and the algebra they generate, do get strong quantum corrections that we shall compute explicitly later.

So what is the strongest result that we can expect at the non-perturbative level?

The non-perturbative anomaly theorem:

F. F., arXiv:0709.0472

It is possible to absorb the non-perturbative quantum corrections in the anomaly equations by a suitable redefinition of the variables that enter the equations.

$$-N\sum_{k\geq 0} g_k u_{n+k+1} + 2\sum_{k_1+k_2=n} u_{k_1} v_{k_2} +$$

$$-N\sum_{k\geq 0} g_k u_{n+k+1} + 2\sum_{k_1+k_2=n} u_{k_1} v_{k_2} = 0$$

$$\sum_{r\geq 1} q^r \left(\sum_{k\geq 0} g_k A^{(r)}_{n,k}(u_p) + \sum_{t\geq 0} C^{(r)}_{n,t}(u_p)\, v_t\right) = 0$$

$$u_{N+p} = \mathscr{P}_p(u_1, \ldots, u_N, q)$$

$$u_{N+p} \to u_{N+p} + \sum_{r\geq 1} q^r c^{(r)}_p(u_k)$$

$$v_{N+p} \to v_{N+p} + \sum_{r\geq 1} q^r d^{(r)}_p$$

A possible route to address the problem was suggested in the original CDSW paper.

1) Make an ansatz for the possible non-perturbative corrections. For example,

$$L_n \cdot u_m = -m u_{n+m} + \sum_{k\geq 1} q^k r^{(k)}_{n,m}, \ \ldots$$

2) Use Wess-Zumino consistency conditions

But this does not work. $L_n \cdot u_m$ turns out to be multivalued (not in the chiral ring!), and thus no simple ansatz can be guessed to perform the analysis (it can be shown that single-valuedness + WZ would imply no quantum corrections to the correlators!).

We need to do better (of course the result will satisfy WZ consistency conditions, but for a strongly quantum corrected algebra that in particular do not close on the original super-Virasoro operators).

Before we do that, we need to tackle another problem. The most general solution to the anomaly equations depends on a finite number of arbitrary parameters: the constants c_I that we had at the classical level, and also the matrix model filling fractions (gluino condensates) in the quantum theory.

In the usual approach, the c_I are postulated to be positive integers, and the gluino condensates are found by extremizing a suitable glueball superpotential that includes Veneziano-Yankielowicz terms.

These results follow directly from the microscopic approach. However, it is very interesting to understand that they are actually consequences of consistency conditions, if we assume the non-perturbative anomaly theorem.

Chiral ring consistency theorem

F. F., hep-th/0701220

F. F. and V. Wens, arXiv:0710.2978

A general solution to the anomaly equations is not consistent with the existence of relations of the form

$$u_{N+p} = \mathscr{P}_p(u_1, \ldots, u_N, q)$$

between the variables.

Consistency is achieved only for special values of the filling fractions.

These special values are precisely the one derived from the glueball superpotentials with the correct (uniquely specified) Veneziano-Yankielowicz terms. They are equivalent to the quantization of the periods of the meromorphic one-form Rdz.

Microscopic point of view on the anomaly equations

Let W_{mic} be the microscopic effective quantum superpotential. The anomaly polynomials

$$\mathscr{A}_n = -N \sum_{k\geq 0} g_k u_{n+k+1} + 2 \sum_{k_1+k_2=n} u_{k_1} v_{k_2}$$

$$\mathscr{B}_n = -N \sum_{k\geq 0} g_k v_{n+k+1} + \sum_{k_1+k_2=n} v_{k_1} v_{k_2}$$

should be expressed as suitable variations of W_{mic} (textbook version of anomalies).

$$\mathscr{L}_n \cdot W_{\rm mic} = \mathscr{A}_n \,, \qquad \mathscr{J}_n \cdot W_{\rm mic} = \mathscr{B}_n$$

microscopic quantum version of the classical (perturbative) variation operators L_n and J_n discussed previously

$$[\mathscr{L}_n, \mathscr{L}_m] = (n-m)\mathscr{L}_{n+m} + \text{non-perturbative corrections}$$

etc...

Note: in the planar limit, we could use the matrix model free energy and the usual matrix model Virasoro operators. The fundamental point is that the anomaly equations must be exact at FINITE N in the gauge theory! This cannot work with uncorrected operators.

The non-perturbative anomaly theorem

$$\mathscr{A}_n = -N \sum_{k\geq 0} g_k u_{n+k+1} + 2 \sum_{k_1+k_2=n} u_{k_1} v_{k_2}$$

$$\mathscr{B}_n = -N \sum_{k\geq 0} g_k v_{n+k+1} + \sum_{k_1+k_2=n} v_{k_1} v_{k_2}$$

$$\mathscr{L}_n \cdot W_{\rm mic} = \mathscr{A}_n\,, \qquad \mathscr{J}_n \cdot W_{\rm mic} = \mathscr{B}_n$$

$$\mathscr{L}_n = \sum_i \ell_i(\boldsymbol{a}) \frac{\partial}{\partial a_i}\,, \quad \mathscr{J}_n = \sum_i j_i(\boldsymbol{a}) \frac{\partial}{\partial a_i}$$

first order partial differential operators

$$\ell_i(\boldsymbol{a}) \sim \oint_{\alpha_i} z^{n+1} R_{\rm mic} \mathrm{d}z\,, \quad j_i(\boldsymbol{a}) \sim \oint_{\alpha_i} z^{n+1} S_{\rm mic} \mathrm{d}z$$

The corrections to the super-Virasoro algebra can be computed straightforwardly (the quantum algebra does not close on the operators $\mathscr{L}_n$ and $\mathscr{J}_n$).

Final remark: the quantum corrections to the perturbative operators L_n and J_n are very strong in the sense that $\mathscr{L}_n$ and $\mathscr{J}_n$ acting on a chiral operator like u_k or v_k do not yield a chiral operator in the ordinary sense.

$\mathscr{L}_n \cdot u_k$ is a good function of a but a multi-valued function of the us and vs (this phenomenon is similar to the multi-valuedness on the SW moduli space due to the non-trivial monodromies: operators are good function of a but not the other way around).

This explains why the simple polynomial ansatz for the quantum corrected operators cannot work.

So we have a new formalism to solve the N=1 theories, based on sums over coloured partitions and the microscopic superpotential. This is the open string formalism.

On the other hand, the closed string formalism is based on summing over hermitian matrices and the Dijkgraaf-Vafa glueball superpotential.

The two formalisms are clearly totally different, but there are some formal similarities when one exchange scalar and glueball operators, coloured partitions and matrices, identities and equations of motion etc.

When both formalisms are on-shell, they yield the same correlators.

F. F., JHEP 11 (2007) 001, arXiv:0709.0472

The matrix model recipe

$$v_k(\boldsymbol{s}) \sim \langle\langle \operatorname{Tr} X^k \rangle\rangle$$

average over hermitian matrices

$$u_k(\boldsymbol{s}) \sim \langle\langle \operatorname{Tr} X^k \operatorname{Tr} t''(X) \rangle\rangle - \langle\langle \operatorname{Tr} X^k \rangle\rangle \langle\langle \operatorname{Tr} t''(X) \rangle\rangle$$

$$S(z;\boldsymbol{s})\,,\quad R(z;\boldsymbol{s})$$

$$\mathrm{d}W_{\mathrm{DV}}(\boldsymbol{s}=\boldsymbol{s}^*)=0$$

$$\langle v_k \rangle = v_k(\boldsymbol{s}^*)$$

$$\langle u_k \rangle = u_k(\boldsymbol{s}^*)$$

The microscopic approach

$$u_k(\boldsymbol{a}) \sim \lfloor \operatorname{Tr} X^k \rfloor$$

average over coloured partitions

$$v_k(\boldsymbol{a}) \sim \lfloor \operatorname{Tr} X^{k+2} \operatorname{Tr} W(X) \rfloor - \lfloor \operatorname{Tr} X^{k+2} \rfloor \lfloor \operatorname{Tr} W(X) \rfloor$$

$$S(z;\boldsymbol{a})\,,\quad R(z;\boldsymbol{a})$$

$$\mathrm{d}W_{\mathrm{mic}}(\boldsymbol{a}=\boldsymbol{a}^*)=0$$

$$\langle v_k \rangle = v_k(\boldsymbol{a}^*)$$

$$\langle u_k \rangle = u_k(\boldsymbol{a}^*)$$

totally different functions of z
(different analytic structures, etc...)

MUST BE THE SAME

How is the equivalence supposed to work?

It is useful to remember the anomaly consistency theorem.

The "philosophy" of this theorem is that the open strings can be built from the closed strings (u_k=Tr X^k) only when some quantization conditions, that here follow from the glueball/flux superpotential, are satisfied in the closed string theory.

The formula $u_k = \mathrm{Tr}\, X^k$ are of course automatically implemented in the open string/microscopic description. This means that non-trivial dynamics in the closed string framework is exchanged with trivial identitied in the open string formulation.

On the other hand, off-shell indentities in the closed string/matrix model formulation, like the generalized anomaly equations, should correspond to non-trivial dynamical identities valid only on-shell in the open string/microscopic description.

This is precisely what happens when one study mathematically how the equivalence between the formalism works.

Microscopic

$$R_{\rm mic}(z,\boldsymbol{a})\,,\;\; S_{\rm mic}(z,\boldsymbol{a})$$

$$\oint R_{\rm mic}\mathrm{d}z \in 2i\pi\mathbb{Z}$$

identity in the open string formalism
(simply follows from the definition of the variables induced by the NC regularization of the instanton moduli space)

$$\mathrm{d}W_{\rm mic} \sim \oint S'_{\rm mic}\,\mathrm{d}z$$

anomaly $\sim \delta W_{\rm mic}$

Macroscopic

$$R_{\rm mac}(z,\boldsymbol{s})\,,\;\; S_{\rm mac}(z,\boldsymbol{s})$$

$$\mathrm{d}W_{\rm mac} \sim \oint R_{\rm mac}\mathrm{d}z \;\mathrm{mod}\; 2i\pi\mathbb{Z}$$

ensures consistency with the open string formulation of the (chiral ring consistency theorem)

$$\oint S_{\rm mac}\,\mathrm{d}z = 0$$

direct consequence of the anomaly equations

$$\boxed{R_{\rm mic}(z,\boldsymbol{a}^*) = R_{\rm mac}(z,\boldsymbol{s}^*)\,,\;\; S_{\rm mic}(z,\boldsymbol{a}^*) = S_{\rm mac}(z,\boldsymbol{s}^*)}$$

Some conclusions

1) we have now a completely rigorous, microscopic, formalism to solve the chiral sector of both N=2 and N=1 gauge theories.

2) this formalism provides interesting insights, we have discussed in some details the case of the generalized Konishi anomaly equations.

3) there are many interesting open problems, for example can we compute N=1 Wilson loops using the formalism?.

Some perspectives

In doing instantons, we always start by using the decomposition of the infinite dimensional moduli space as a union of finite dimensional spaces, according to the topological charge,

$$\mathscr{M} = \cup_k \mathscr{M}_k$$

However, the most interesting results correspond to "collective" instanton behaviour. We have seen that this can yield an answer which is a fractional power of the usual instanton factor.

It would be extremely useful to try to understand the space $\mathscr{M}$ as a whole, independently of the decomposition in terms of the topological charge.

I'm not sure what this could be, some kind of field theory of instantons.

In such a set-up, one may be able to get important insights into S-duality.

In $N=1^*$ for example, we can apply the microscopic formalism and find correlators that are modular forms of the coupling.

We can then use two instanton factors:

$$q = e^{2i\pi\tau} \ , \ q' = e^{-2i\pi/\tau}$$

Expanding in terms of q or q′ corresponds to two distinct decompositions of $\mathscr{M}$,

$$\mathscr{M} = \bigcup_k \mathscr{M}_k = \bigcup_{k'} \mathscr{M}'_{k'}$$

Any good description of $\mathscr{M}$ should make the existence of these inequivalent decompositions manifest. This is an outstanding problem for the future.

This problem is related to the question of understanding the "quantization" of systems that have distinct, inequivalent, classical limits.

Thank you for this wonderful conference!!!