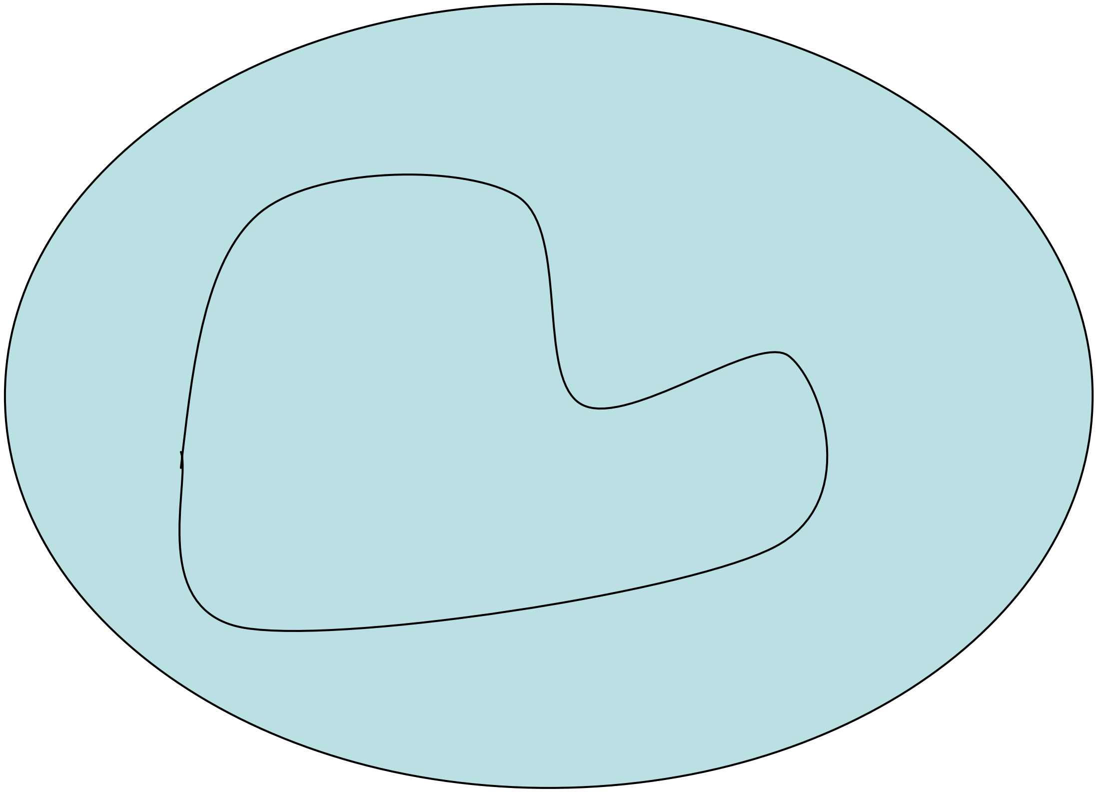
SURFACE OPERATORS IN GAUGE THEORY

Edward Witten

Naples

Important operators in four-dimensional gauge theory are the ‘t Hooft and Wilson line operators – which are supported on curves in spacetime:



They are important, for example, as order parameters for probing confinement and the Higgs mechanism.

I've been working with Sergei Gukov on analogous operators in gauge theory that are supported in codimension two, that is (in four dimensions) on a two-dimensional surface.

One primary motivation, which is the basis for our forthcoming paper, has to do with extending the approach to the "geometric Langlands program" via gauge theory – which was developed recently

Electric-Magnetic Duality And The Geometric Langlands Program.
Anton Kapustin (Caltech) , Edward Witten (Princeton, Inst. Advanced Study) . Apr 2006

– to the so-called "ramified case"

There is another possible motivation involving trying to understand in gauge theory some new knot invariants discovered by mathematicians

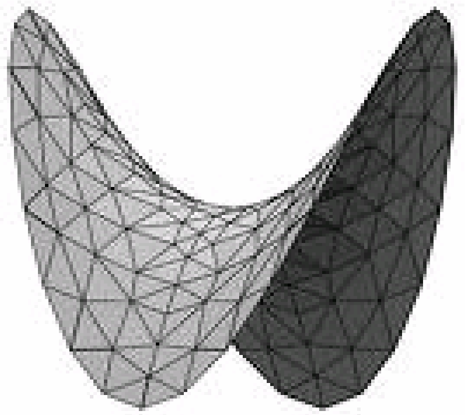
(Khovanov, ...; Gukov, Schwarz, Vafa)

These new invariants somehow generalize the knot invariants related to three-dimensional Chern-Simons gauge theory. We feel that they should have a gauge theory description, which may involve surface operators – but we have not understood it yet. (Gukov may explain something about this on Friday.)

A third motivation (thanks here to N. Seiberg for helpful remarks) is that, like ‘t Hooft and Wilson operators, the surface operators might conceivably have applications as probes in understanding the dynamics of four-dimensional gauge theory.

However, not much has happened in this direction yet, though I will say a few words later.

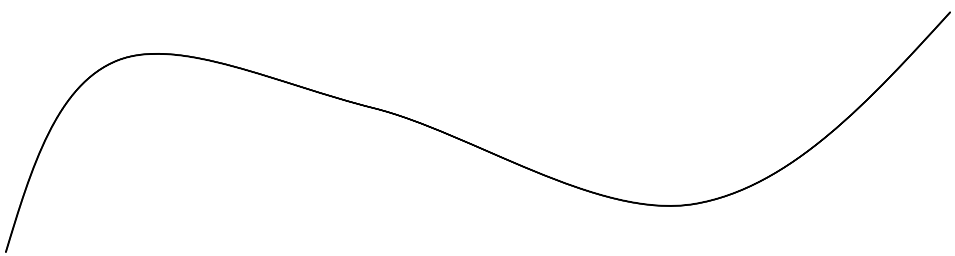
Surface operators could appear physically if cosmic strings are discovered. Depending on the microscopic details, the coupling to QCD of the world-volume of a cosmic string might give us an interesting surface operator probing QCD.



For example, there might be colored fields living only on the surface.

That is one prototype for what a surface operator might mean physically.

We get another prototype if we imagine that the string is a sort of magnetic flux line with color magnetic flux as well as ordinary magnetic flux. Then a quark circling the string



experiences a color-dependent Aharonov-Bohm effect.

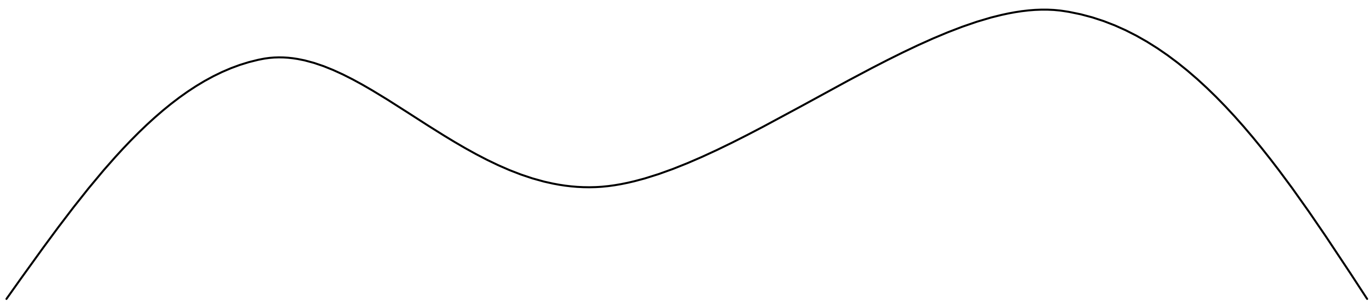
In either of these cases, we get a surface operator coupled to QCD if the string is so heavy that its motion can be considered fixed and known. The problem is to determine the behavior of QCD in the presence of this kind of impurity.

Sometimes, the same cosmic string can be described (in its influence on QCD) in either of these two approaches.

In our work, we found the “Aharonov-Bohm” approach to be more useful. This means that we characterize our surface operator by the singularity it produces in the gauge fields around it. The simplest picture is that this singularity is simply a monodromy.

So we pick a conjugacy class in G, represented by a group element U, and we say that the monodromy of the surface operator around the string must be conjugate to U.

In this way, we define a surface operator that depends on the choice of a conjugacy class in G.



It can also be conveniently described in formulas. Locally, we pick coordinates so that the string is located at

$$x_1 = x_2 = 0$$

And then, setting

$$x_1 + ix_2 = re^{i\theta}$$

we ask that the gauge field, near $r = 0$ should look like

$$A = \alpha \, d\theta + \dots$$

(where the dots refer to less singular terms) and α is some chosen element of the Lie algebra.

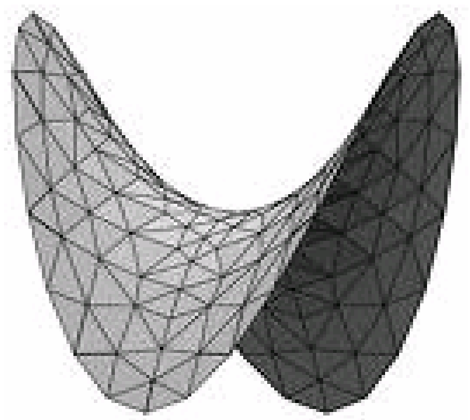
We choose α so that the resulting monodromy, which is

$$U = \exp(-2\pi\alpha)$$

is in the desired conjugacy class. So we can specify our surface operator by talking about the monodromy, or by specifying the singular part of the gauge field A.

Up to a gauge transformation, we can assume that U lives in the maximal torus T of G … or more or less equivalently that α takes values in the abelian Lie algebra of the torus. Let us assume moreover that U is chosen generically, so that the subgroup of G that commutes with U is precisely equal to T. It is also interesting to consider what happens for non-generic choices of U, but for today we will omit this.

What happens now is that we can define more parameters that the surface operator depends on.   The reason is that along the surface



the gauge group G is explicitly broken to the abelian subgroup T by the choice of U or α

So we are in a situation that is a little like abelian gauge theory in two dimensions. In two-dimensional abelian gauge theory, we can introduce a "theta angle", but I am going to call it η to avoid confusion with the "bulk" four-dimensional theta angle.

If the gauge group is just U(1), we would include in the integrand of the path integral a factor

$$\exp\left(i\eta\int_D F\right)$$

Here D is a surface with gauge field strength F. I normalized η to be defined mod 1.

How do we generalize this to any G? On D, we have a T-bundle, which is classified topologically by picking a U(1) bundle over D of “degree 1”, and then mapping it to T by a homomorphism

$$\rho : U(1) \to T$$

So the bundles are classified by

$$M = \mathrm{Hom}(U(1), T)$$

And consequently we can usefully think of the theta-like angles η as an element of

$$W = \mathsf{Hom}(M, U(1))$$

W is a torus, but to understand better which torus, we need to review electric-magnetic duality.

Let us recall the basic idea of such duality.

For every gauge group G, there is a "dual group" $\tilde{G}$, which physicists call the GNO dual and mathematicians call the Langlands dual group. The basic relation among them is that electric charge of G is magnetic charge of $\tilde{G}$, and vice-versa.

We write T for the maximal torus of G
and $\tilde{T}$ for the maximal torus of $\tilde{G}$

Electric charge of G is a representation of T
(i.e. every representation of G has a highest weight, which is a representation of T).

So electric charge lives in

$$E = \mathrm{Hom}(T, U(1))$$

Where does magnetic charge live?

Magnetic charge can be measured by the magnetic flux through a large two-sphere at infinity, and I already told you that in general the magnetic flux through a two-manifold D is classified by

$$M = \mathrm{Hom}(U(1), T)$$

To put the two side by side, therefore, electric and magnetic charge take values respectively in

$$E = \mathrm{Hom}(T, U(1))$$

and

$$M = \mathrm{Hom}(U(1), T)$$

Electric-magnetic duality is supposed to exchange G with $\tilde{G}$ while also exchanging E and M. So looking back to the definitions of E and M, we find the fundamental relation between T and $\tilde{T}$:

$$\mathrm{Hom}(T, U(1)) = \mathrm{Hom}(U(1), \tilde{T})$$

This means that T and $\tilde{T}$ are "dual tori"

Going back to the theta angles of the surface operator or “cosmic string”, we found that they take values in

$$W = \mathrm{Hom}(M, U(1))$$

where we now say

$$M = \mathrm{Hom}(\tilde{T}, U(1))$$

which is dual to the previous definition

This probably sounds a little abstract, but it actually just means that the mysterious torus W where the two-dimensional theta angles live is none other than $\tilde{T}$

Here I am using “Pontryagin duality,” which says that if

$$A = \mathrm{Hom}(B, U(1))$$

then

$$B = \mathrm{Hom}(A, U(1))$$

In other words, since

$$M = \mathsf{Hom}(\tilde{T}, U(1))$$

it follows that

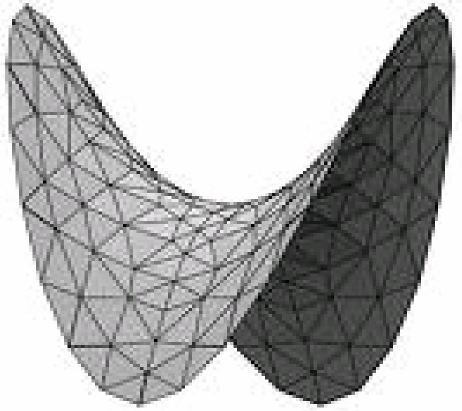
$$W = \mathsf{Hom}(M, U(1))$$

is none other than

$$\tilde{T}$$

Now, hoping not to have lost you in the technicalities, what we have learned is that

the surface operator



is characterized by two types of variables… α takes values in T and η takes values in $\tilde{T}$.

All this makes sense without any assumption of electric-magnetic duality. The dual torus was introduced just to give a convenient description of the mysterious torus W where the theta-like angles live.

However, it becomes more interesting if we do assume electric-magnetic duality – the prime case being N=4 super Yang-Mills theory in four dimensions.

In this case, we have the electric-magnetic duality transformation

$$S : \tau \rightarrow -1/\tau$$

that exchanges the group G with the dual group $\tilde{G}$ and likewise the torus T with the dual torus $\tilde{T}$

In this theory, our construction gives a surface operator that is "1/2 BPS," i.e. it preserves half the supersymmetry

It depends on the variables α and η

(and on some other variables that specify the singular behavior of the scalar fields near the surface … these other variables will not concern us today, though they are important in the geometric Langlands program).

Under S-duality it must somehow map to another surface operator.

Since the variables are

$$(\alpha, \eta) \in T \times \tilde{T}$$

where S exchanges the two factors on the right, the obvious hypothesis is that S acts by

$$(\alpha, \eta) \to (\eta, -\alpha)$$

(The minus sign is to agree with known action of S^2, which is charge conjugation.)

This has been the basic hypothesis in the work that Gukov and I have been doing on the ramified case of the geometric Langlands correspondence, and it seems to work very nicely. In that context, one studies N=4 super Yang-Mills on a four-manifold $M = \Sigma \times C$ where C is the Riemann surface on which one want to study geometric Langlands. At low energies, with Σ large compared to C, one gets a two-dimensional (4,4) sigma model on Σ with target a hyper-Kahler manifold that parameterizes gauge fields and scalar fields on C.

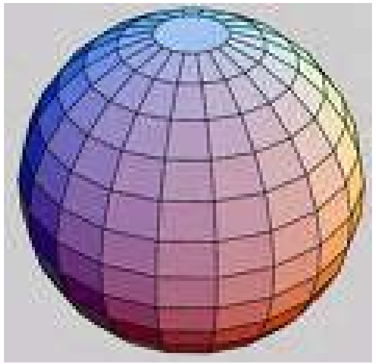
As Kapustin and I showed, the geometric Langlands program is all about what S-duality does in this sigma model, and how it acts on Wilson and 't Hooft operators.

To get to the "ramified case," Gukov and I have merely considered the same construction …

but now with the addition of a “surface operator” supported on $\Sigma_p = \Sigma \times p$ where p is a point in Σ

When we apply standard physical understanding of S-duality (enriched by the proposal I explained for how it acts on the parameters of the surface operator) we get the ramified extension of the geometric Langlands program, as understood by mathematicians.

How would we try to use these operators as order parameters for probing gauge theories? I don't know if we can get anything useful, but the obvious idea is to pick some values of α and η and then consider a surface operator with those parameters that is supported on a closed two-manifold, for example a two-sphere,



with area A and enclosing a volume V.

Then, imitating what is usually done for Wilson and ‘t Hooft operators, we ask if the expectation value of such an operator behaves for large A and V as

exp(- f A) or as exp(-f V)

for some “constant” f – which one would expect to depend on α, η The range of α, η for which one gets an area or volume law might distinguish different phases of gauge theory.

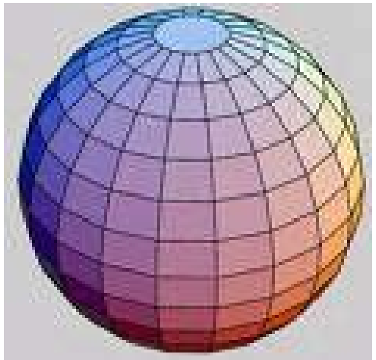
For most purposes, such new order parameters do not really seem to be needed (since Wilson and ‘t Hooft operators are adequate), but there are some known situations in which four dimensional gauge theories appear to have distinct phases that are hard to distinguish using the standard order parameters. (For example, N=1 Super Yang-Mills with adjoint matter and a general single-trace superpotential.)

The only cases in which I can offer any computation to show if a surface operator has an area or volume law are rather trivial.

Conformal invariance makes a volume law impossible, so we will not get a volume law in N=4 super YM in the vacuum in which the scalar fields vanish. Even if we give VeV's to scalars, the low energy theory is just a conformally invariant abelian N=4 theory – no volume law.

To find a volume law, it suffices to consider the Higgs phase of for example U(1) gauge theory. In the Higgs phase, to minimize the action, the Higgs field ϕ must be covariantly constant, $D_\mu\phi = 0$

If α is nonzero, then this is only possible if $\phi = 0$ along a "domain wall" that ends on the "string worldvolume." So there is a volume law for $\alpha \neq 0$



And I believe there is an area rather than volume law for $\alpha = 0$

By duality, we seem to learn that in the confining phase, there is a volume law unless $\eta = 0$

I don’t think I have any intuition about this.

Now let us look at the duality conjecture more closely. As I described it, we had the transformation $S : \tau \to -1/\tau$ acting by $(\alpha, \eta) \to (\eta, -\alpha)$

However, actually, S is part of an infinite discrete symmetry group that roughly speaking is $SL(2, \mathbf{Z})$ To be precise, the duality group is exactly $SL(2, \mathbf{Z})$ essentially only for the case of E_8

For the case of E_8 , the action of the duality group on (α, η) is (we believe)

$$(\alpha, \eta) \longrightarrow (\alpha, \eta) \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

for any element

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbf{Z})$$

For $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, this formula gives

the transformation that I suggested before,
and for the other generator $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

it gives

$$(\alpha, \eta) \longrightarrow (\alpha, \eta + \alpha)$$

This last one can be demonstrated explicitly,
similarly to computing the electric charge

of a magnetic monopole.

The idea is that the transformation

$$T : \tau \rightarrow \tau + 1$$

is the same as $$\theta \rightarrow \theta + 2\pi$$

where θ is the four-dimensional theta angle, appearing in

$$L_\theta = \frac{\theta}{8\pi^2}\int_M \mathrm{Tr}\, F \wedge F$$

Usually $\theta \to \theta + 2\pi$ is a symmetry because

$$N = \frac{1}{8\pi^2} \int_M \mathrm{Tr}\, F \wedge F$$

is integer-valued, so under $\theta \to \theta + 2\pi$ the action shifts by an integer multiple of 2π

But once we introduce the singularity with α it is no longer true that N is an integer. It differs from one by

$$\frac{1}{2\pi} \int_D \mathrm{Tr}\, \alpha F$$

The result of this is that by itself $\theta \rightarrow \theta + 2\pi$ is not a symmetry. It changes the action by

$$I \rightarrow I - \frac{1}{2\pi}\int_D \mathrm{Tr}\,\alpha F \mod 2\pi$$

To get a symmetry, one must also shift the two-dimensional theta angle η, which appears in a term in the action $\int_D \mathrm{Tr}\,\eta F$

Indeed, $\theta \rightarrow \theta + 2\pi$ must be accompanied by $\eta \rightarrow \eta + \alpha$

So this explains the action of T.

Once the action of T is pinned down, the proposal for the action of S is (we think) the only reasonable ansatz for the action of SL(2,Z) on (α, η)

A further bit of evidence in its favor is that for G=U(1), one can explicitly compute that the parameters transform as claimed.

I think I've said enough about surface operators in gauge theory, so to the extent that there is any time left, I am going to use it to say a few words on "what is the geometric Langlands program," for whatever cultural value the answer to this question may have.

The *Langlands program* is an attempt to unify many old and new results in number theory

– ranging from quadratic reciprocity, proved around 1800, to the modern proof of Fermat's Last Theorem.

Number theory is difficult because calculus is powerful ….

And so mathematicians have sought to find geometric analogs of problems in number theory.

Number field --- Riemann surface C

Prime number --- point on C

The Langlands program, too, has a geometric analog, which has been intensively developed.

Even in its geometric form, the Langlands program involves statements that at first sight are likely to look unrecognizable to physicists.

But, if one probes a little more deeply, the Langlands program has many obvious and less obvious analogies with quantum field theory.

For one thing, Langlands introduced (ca. 1970) a "duality" between a simple Lie group G and a "dual" group often called ^{L}G. However, this relationship, which pairs SU(N) with $SU(N)/Z_N$, E_8 with itself, SO(2n+1) with Sp(n), etc., also plays a very distinguished role in four-dimensional gauge theory…

In fact, the Langlands dual group ${}^{L}G$ is precisely the magnetic group introduced in 1976 by Goddard, Nuyts, and Olive to classify magnetic monopoles:

Their idea was that if electric charges are classified by representations of a group G, then the corresponding magnetic charges are classified by a representation of the dual group, which in fact coincides with the dual ${}^{L}G$ of Langlands.

The GNO work was of course the jumping off point for Montonen-Olive duality, an extremely fruitful idea in which the two groups G and $^L G$ enter in a completely symmetric fashion.

By contrast, in the Langlands program the roles of G and $^L G$ are at first sight bafflingly unlike. An object of one type associated with $^L G$ is classified by a completely different type of object associated with G.

That is a bit disconcerting at first sight, but on the positive side, the objects of study on the two sides are both objects that are familiar in QFT – or at least they are once we make the translation from number theory to geometry.

On the left hand side of the Langlands correspondence, we have a flat connection, on a Riemann surface C, with gauge group ${}^LG$. In gauge theory, flat connections are those with least energy, most supersymmetry, etc.

The right hand side of the Langlands correspondence involves an "automorphic representation" of G, a notion which when suitably translated to geometry is very closely related to the "conformal blocks" of current algebra, with symmetry group G, on a Riemann surface C.

Moreover, in recent years mathematicians (Beilinson, Drinfeld, Frenkel, Gaitsgory…) working on geometric Langlands have relied heavily on two-dimensional current algebra – albeit with a focus on aspects (such as the exceptional behavior at level k=-h) that appear extremely exotic from a physical point of view.

As I already mentioned, A. Kapustin and I recently described geometric Langlands in terms of gauge theory (and S. Gukov and I have been extending this to the "ramified case")

To simplify a bit, our analysis is based on six ideas:

(1) A certain twisting of N=4 super Yang-Mills theory gives a family of TQFT's parameterized by CP^1. S-duality acts on this CP^1

(2) If this theory is compactified to two dimensions on a Riemann surface C, one gets a sigma model whose target is "Hitchin's moduli space"; S-duality turns into, roughly speaking, mirror symmetry of the sigma model.

(3) Wilson and ‘t Hooft operators of the four-dimensional gauge theory can be interpreted as operators on branes – that is operators that act a brane to give a new brane (as opposed to conventional operators acting on quantum states

As a result, we can define a new notion of “electric eigenbrane” (eigenbrane of Wilson operators) or “magnetic eigenbrane” (eigenbrane of ‘t Hooft operators)

(4) The electric eigenbranes correspond to representations of the fundamental group in ${}^LG$. This is the left hand side of the geometric Langlands correspondence.

(5) S-duality will transform electric eigenbranes into magnetic eigenbranes ….and it turns out that the magnetic eigenbranes are what are called "Hecke eigensheaves" in the jargon

(6) Finally, another trick involving branes enables us to interpret the "magnetic eigenbranes" as the "D-modules" that appear on the right hand side of the geometric Langlands program.

(1) A certain twisting of N=4 super Yang-Mills theory gives a family of TQFT's parameterized by CP^1. S-duality acts on this CP^1

I think I will not explain this today as the basic idea of "twisting" a supersymmetric field theory to get a TQFT is fairly well known. We just apply this procedure to N=4. There are a few unusual features but we needn't explain them today.

(2) *If this theory is compactified to two dimensions on a Riemann surface C, one gets a sigma model whose target is "Hitchin's moduli space"; S-duality turns into, roughly speaking, mirror symmetry of the sigma model.*

This step largely follows BJSV and HMS (1995). We take the four-manifold M_4 to be a product of Riemann surfaces Σ and C. In the limit that Σ is very large compared to C, we get an effective two dimensional σ-model on Σ.

Our four-dimensional topological field theory reduces in two dimensions essentially to a familiar type of A- or B-model of the sigma model (or a hybrid made possible by the fact that in this case the target is hyper-Kahler).

In one of the complex structures, the S-duality reduces essentially to a mirror symmetry.

Our four-dimensional topological field theory reduces in two dimensions essentially to a familiar type of A- or B-model of the sigma model (or a hybrid made possible by the fact that in this case the target is hyper-Kahler).

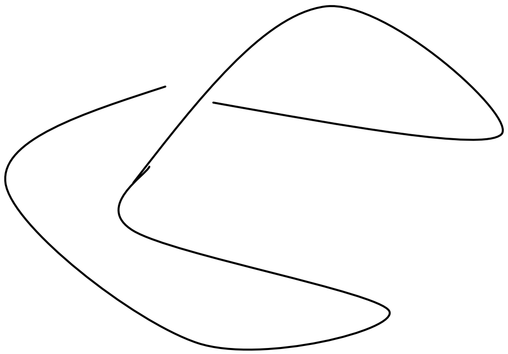
In one of the complex structures, the S-duality reduces essentially to a mirror symmetry.

To summarize the remaining points (3)-(6):

In the twisted N=2 super Yang-Mills (which is related to Donaldson theory of four-manifolds), the important operators are local operators and their “descendants,” but in the present problem…

The important operators are not local operators, but rather Wilson and ‘t Hooft “line operators”

S



associated with an open or closed loop S in spacetime.

The important objects that the line operators act on are “branes.”

The branes are defined in the four-dimensional gauge theory by local boundary conditions at the boundary of the four-manifold M.

In two dimensions, they reduce to the ordinary branes of the sigma model, including the usual A-branes and B-branes of the appropriate topological field theories.

Now the key fact is that line operators behave as operators acting on branes… A brane (the solid line) with a line operator (the dashed line) gives us a new "composite" brane.

This gives an action of line operators on "theories," more abstract than the usual action of operators on states.

If L is a line operator and B is a brane, we say that B is an "eigenbrane" of L if L acting on B gives us back several copies of B.

In formulas

$$LB = B \otimes V$$

where V is a vector space.

We call a brane that is an eigenbrane for the Wilson operators an electric eigenbrane, and a brane that is an eigenbrane for the ‘t Hooft operators a magnetic eigenbrane.

Obvious, S-duality will map electric eigenbranes to magnetic eigenbranes.

With a little effort, one shows that electric eigenbranes are classified by representations of the fundamental group of Σ into ${}^L G$

And then one goes on to show that their S-duals, which are the “magnetic eigenbranes”, do indeed correspond to the

“D-modules” that are claimed by the geometers.

To get the "ramified case," we just include a surface operator….