

GAUGE THEORY AND LANGLANDS DUALITY

by **Edward FRENKEL**

INTRODUCTION

In the late 1960s Robert Langlands launched what has become known as the Langlands Program with the ambitious goal of relating deep questions in Number Theory to Harmonic Analysis [L]. In particular, Langlands conjectured that Galois representations and motives can be described in terms of the more tangible data of automorphic representations. A striking application of this general principle is the celebrated Shimura–Taniyama–Weil conjecture (which implies Fermat’s Last Theorem), proved by A. Wiles and others, which says that information about Galois representations associated to elliptic curves over \mathbb{Q} is encoded in the Fourier expansion of certain modular forms on the upper-half plane.

One of the most fascinating and mysterious aspects of the Langlands Program is the appearance of the *Langlands dual group*. Given a reductive algebraic group G , one constructs its Langlands dual ${}^L G$ by applying an involution to its root data. Under the Langlands correspondence, automorphic representations of the group G correspond to Galois representations with values in ${}^L G$.

Surprisingly, the Langlands dual group also appears in Quantum Physics in what looks like an entirely different context; namely, the *electro-magnetic duality*. Looking at the Maxwell equations describing the classical electromagnetism, one quickly notices that they are invariant under the exchange of the electric and magnetic fields. It is natural to ask whether this duality exists at the quantum level. In quantum theory there is an important parameter, the electric charge e . Physicists have speculated that there is an electro-magnetic duality in the quantum theory under which $e \longleftrightarrow 1/e$. Under this duality the electrically charged particle should be exchanged with a magnetically charged particle, called magnetic monopole, first theorized by P. Dirac (so far, it has not been discovered experimentally).

In modern terms, Maxwell theory is an example of 4D *gauge theory* (or Yang–Mills theory) which is defined, classically, on the space of connections on various G_c -bundles

on a four-manifold M , where G_c is a compact Lie group.⁽¹⁾ Electromagnetism corresponds to the simplest, abelian, compact Lie group $U(1)$. It is natural to ask whether there is a non-abelian analogue of the electro-magnetic duality for gauge theories with non-abelian gauge groups.

The answer was proposed in the late 1970s, by Montonen and Olive [MO], following Goddard, Nuyts and Olive [GNO] (see also [EW, O]). A gauge theory has a coupling constant g , which plays the role of the electric charge e . The conjectural non-abelian electro-magnetic duality, which has later become known as *S-duality*, has the form

$$(0.1) \quad (G_c, g) \longleftrightarrow ({}^L G_c, 1/g).$$

In other words, the duality states that the gauge theory with gauge group G_c (more precisely, its “ $N = 4$ supersymmetric” version) and coupling constant g should be equivalent to the gauge theory with the Langlands dual gauge group ${}^L G_c$ and coupling constant $1/g$ (note that if $G_c = U(1)$, then ${}^L G_c$ is also $U(1)$). If true, this duality would have tremendous consequences for quantum gauge theory, because it would relate a theory at small values of the coupling constant (weak coupling) to a theory with large values of the coupling constant (strong coupling). Quantum gauge theory is usually defined as a power series expansion in g , which can only converge for small values of g . It is a very hard problem to show that these series make sense beyond perturbation theory. *S-duality* indicates that the theory does exist non-perturbatively and gives us a tool for understanding it at strong coupling. That is why it has become a holy grail of modern Quantum Field Theory.

Looking at (0.1), we see that the Langlands dual group shows up again. Could it be that the Langlands duality in Mathematics is somehow related to *S-duality* in Physics?

This question has remained a mystery until about five years ago. In March of 2004, DARPA sponsored a meeting of a small group of physicists and mathematicians at the Institute for Advanced Study in Princeton (which I co-organized) to tackle this question. At the end of this meeting Edward Witten gave a broad outline of a relation between the two topics. This was explained in more detail in his subsequent joint work [KW] with Anton Kapustin. This paper, and the work that followed it, opened new bridges between areas of great interest for both physicists and mathematicians, leading to new ideas, insights and directions of research.

The goal of these notes is to describe briefly some elements of the emerging picture. In Sections 1 and 2, we will discuss the Langlands Program and its three flavors, putting it in the context of André Weil’s “big picture”. This will eventually lead us to a formulation of the geometric Langlands correspondence as an equivalence of certain categories of sheaves in Section 3. In Section 4 we will turn to the *S-duality* in topological twisted $N = 4$ super–Yang–Mills theory. Its dimensional reduction gives rise to the Mirror Symmetry of two-dimensional sigma models associated to the Hitchin

⁽¹⁾We will use the notation G for a complex Lie group and G_c for its compact form. Note that physicists usually denote by G a compact Lie group and by $G_{\mathbb{C}}$ its complexification.

moduli spaces of Higgs bundles. In Section 5 we will describe a connection between the geometric Langlands correspondence and this Mirror Symmetry, following [KW], as well as its ramified analogue [GW1]. In Section 6 we will discuss subsequent work and open questions.

Acknowledgments. I thank Sergei Gukov, Vincent Lafforgue, Robert Langlands, and Edward Witten for inspiring discussions. I also thank S. Gukov and V. Lafforgue for their comments on a draft of this paper.

I am grateful to DARPA (and especially Benjamin Mann) for generous support which has been instrumental not only for my research, but for the development of this whole area. I also thank Fondation Sciences Mathématiques de Paris for its support during my stay in Paris.

1. LANGLANDS PROGRAM

In 1940 André Weil was put in jail for his refusal to serve in the army. There, he wrote a letter to his sister Simone Weil (a noted philosopher) in response to her question as to what really interested him in his work [We]. This is a remarkable document, in which Weil tries to explain, in fairly elementary terms (presumably, accessible even to a philosopher), the “big picture” of mathematics, the way he saw it. I think this sets a great example to follow for all of us.

Weil writes about the role of *analogy* in mathematics, and he illustrates it by the analogy that interested him the most: between Number Theory and Geometry.

On one side we look at the field \mathbb{Q} of rational numbers and its algebraic closure $\overline{\mathbb{Q}}$, obtained by adjoining all roots of all polynomial equations in one variable with rational coefficients (like $x^2 + 1 = 0$). The group of field automorphisms of $\overline{\mathbb{Q}}$ is the *Galois group* $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. We are interested in the structure of this group and its finite-dimensional representations. We may also take a more general *number field* – that is, a finite extension F of \mathbb{Q} (such as $\mathbb{Q}(i)$) – and study its Galois group and its representations.

On the other side we have Riemann surfaces: smooth compact orientable surfaces equipped with a complex structure, and various geometric objects associated to them: vector bundles, their endomorphisms, connections, etc.

At first glance, the two subjects are far apart. However, it turns out that there are many analogies between them. The key point is that there is another class of objects which are in-between the two. A Riemann surface may be viewed as the set of points of a projective algebraic curve over \mathbb{C} . In other words, Riemann surfaces may be described by algebraic equations, such as the equation

$$(1.1) \quad y^2 = x^3 + ax + b,$$

where $a, b \in \mathbb{C}$. The set of complex solutions of this equation (for generic a, b for which the polynomial on the right hand side has no multiple roots), compactified by a point

at infinity, is a Riemann surface of genus 1. However, we may look at the equation (1.1) not only over \mathbb{C} , but also over other fields – for instance, over finite fields.

Recall that there is a unique, up to an isomorphism, finite field \mathbb{F}_q of q elements for all q of the form p^n , where p is a prime. In particular, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \simeq \{0, 1, \dots, p-1\}$, with the usual arithmetic modulo p . Let a, b be elements of \mathbb{F}_q . Then the equation (1.1) defines a curve over \mathbb{F}_q . These objects are clearly analogous to algebraic curves over \mathbb{C} (that is, Riemann surfaces). But there is also a deep analogy with number fields!

Indeed, let X be a curve over \mathbb{F}_q (such as an elliptic curve defined by (1.1)) and F the field of rational functions on X . This *function field* is very similar to a number field. For instance, if X is the projective line over \mathbb{F}_q , then F consists of all fractions $P(t)/Q(t)$, where P and Q are two relatively prime polynomials in one variable with coefficients in \mathbb{F}_q . The ring $\mathbb{F}_q[t]$ of polynomials in one variable over \mathbb{F}_q is similar to the ring of integers and so the fractions $P(t)/Q(t)$ are similar to the fractions p/q , where $p, q \in \mathbb{Z}$.

Thus, we find a *bridge*, or a “turntable” – as Weil calls it – between Number Theory and Geometry, and that is the theory of algebraic curves over finite fields.

In other words, we can talk about three parallel tracks

Number Theory

Curves over \mathbb{F}_q

Riemann Surfaces

Weil’s idea is to exploit it in the following way: take a statement in one of the three columns and translate it into statements in the other columns [We]: “my work consists in deciphering a trilingual text; of each of the three columns I have only disparate fragments; I have some ideas about each of the three languages: but I know as well there are great differences in meaning from one column to another, for which nothing has prepared me in advance. In the several years I have worked at it, I have found little pieces of the dictionary.” Weil went on to find one of the most spectacular applications of this “Rosetta stone”: what we now call the *Weil conjectures* describing analogues of the Riemann Hypothesis in Number Theory in the context of algebraic curves over finite fields.

It is instructive to look at the Langlands Program through the prism of Weil’s big picture. Langlands’ original formulation [L] concerned the two columns on the left. Part of the Langlands Program may be framed as the question of describing n -dimensional representations of the Galois group $\text{Gal}(\overline{F}/F)$, where F is either a number field (\mathbb{Q} or its finite extension) or the function field of a curve over \mathbb{F}_q .⁽²⁾ Langlands proposed that such representations may be described in terms of *automorphic representations* of the group $GL_n(\mathbb{A}_F)$, where \mathbb{A}_F is the ring of adèles of F . I will not attempt to explain this here referring the reader to the surveys [Ge, F1, F2].

However, it is important for us to emphasize how the Langlands dual group appears in this story. Let us replace $GL_n(\mathbb{A}_F)$ by $G(\mathbb{A}_F)$, where G is a general reductive algebraic

⁽²⁾Langlands’ more general “functoriality principle” is beyond the scope of the present article.

group (such as orthogonal or symplectic, or E_8). In the case when $G = GL_n$ its automorphic representations are related to the n -dimensional representations of $\text{Gal}(\overline{F}/F)$, that is, homomorphisms $\text{Gal}(\overline{F}/F) \rightarrow GL_n$. The general Langlands conjectures predict that automorphic representations of $G(\mathbb{A}_F)$ are related, in a similar way, to homomorphisms $\text{Gal}(\overline{F}/F) \rightarrow {}^L G$, where ${}^L G$ is the Langlands dual group to G .⁽³⁾

It is easiest to define ${}^L G$ in the case when G , defined over a field k , is split over k , that is, contains a maximal split torus T (which is the product of copies of the multiplicative group GL_1 over k). We associate to T two lattices: the weight lattice $X^*(T)$ of homomorphisms $T \rightarrow GL_1$ and the coweight lattice $X_*(T)$ of homomorphisms $GL_1 \rightarrow T$. They contain the sets of roots $\Delta \subset X^*(T)$ and coroots $\Delta^\vee \subset X_*(T)$ of G , respectively. The quadruple $(X^*(T), X_*(T), \Delta, \Delta^\vee)$ is called the root data for G over k . The root data determines the split group G up to an isomorphism.

Let us now exchange the lattices of weights and coweights and the sets of simple roots and coroots. Then we obtain the root data

$$(X_*(T), X^*(T), \Delta^\vee, \Delta)$$

of another reductive algebraic group over \mathbb{C} (or $\overline{\mathbb{Q}}_\ell$), which is denoted by ${}^L G$.⁽⁴⁾ Here are some examples:

| | |
|-------------|------------------------|
| G | ${}^L G$ |
| GL_n | GL_n |
| SL_n | PGL_n |
| Sp_{2n} | SO_{2n+1} |
| $Spin_{2n}$ | SO_{2n}/\mathbb{Z}_2 |
| E_8 | E_8 |

In the function field case we expect to have a correspondence between homomorphisms⁽⁵⁾ $\text{Gal}(\overline{F}/F) \rightarrow {}^L G$ and automorphic representations of $G(\mathbb{A}_F)$, where \mathbb{A}_F is the ring of adèles of F ,

$$\mathbb{A}_F = \prod'_{x \in X} F_x,$$

$F_x \simeq \mathbb{F}_{q_x}((t_x))$ being the completion of the field of functions at a closed point x of X , and the prime means that we take the restricted product, in the sense that for all but finitely many x the element of F_x belongs to its ring of integers $\mathcal{O}_x \simeq \mathbb{F}_x[[t_x]]$. We have a natural diagonal inclusion $F \subset \mathbb{A}_F$ and hence $G(F) \subset G(\mathbb{A}_F)$. Roughly

⁽³⁾More precisely, ${}^L G$ should be defined over $\overline{\mathbb{Q}}_\ell$, where ℓ is relatively prime to q , and we should consider homomorphisms $\text{Gal}(\overline{F}/F) \rightarrow {}^L G(\overline{\mathbb{Q}}_\ell)$ which are continuous with respect to natural topology (see, e.g., Section 2.2 of [F2]).

⁽⁴⁾In Langlands' definition [L], ${}^L G$ also includes the Galois group of a finite extension of F . This is needed for non-split groups, but since we focus here on the split case, this is not necessary.

⁽⁵⁾More precisely, the Galois group should be replaced by its subgroup called the Weil group.

speaking, an irreducible representation of $G(\mathbb{A}_F)$ is called automorphic if it occurs in the decomposition of $L^2(G(F)\backslash G(\mathbb{A}_F))$ (with respect to the right action of $G(\mathbb{A}_F)$).

For $G = GL_n$, in the function field case, the Langlands correspondence is a bijection between equivalence classes of irreducible n -dimensional (ℓ -adic) representations of $\text{Gal}(\overline{F}/F)$ (more precisely, the Weil group) and cuspidal automorphic representations of $GL_n(\mathbb{A}_F)$. It has been proved by V. Drinfeld [Dr1, Dr2] for $n = 2$ and by L. Lafforgue [LafL] for $n > 2$. A lot of progress has also been made recently in proving the Langlands correspondence for GL_n in the number field case.

For other groups the correspondence is expected to be much more subtle; for instance, it is not one-to-one. Homomorphisms from the Weil group of F to ${}^L G$ (and more general parameters introduced by J. Arthur, see Section 6.2) should parametrize certain collections of automorphic representations called “ L -packets.” This has only been proved in a few cases so far.

2. GEOMETRIC LANGLANDS CORRESPONDENCE

The above discussion corresponds to the middle column in the Weil big picture. What should be its analogue in the right column – that is, for complex curves?

In order to explain this, we need a geometric reformulation of the Langlands correspondence which would make sense for curves defined both over a finite field and over \mathbb{C} . Thus, we need to find geometric analogues of the notions of Galois representations and automorphic representations.

The former is fairly easy. Let X be a curve over a field k and $F = k(X)$ the field of rational functions on X . If $Y \rightarrow X$ is a covering of X , then the field $k(Y)$ of rational functions on Y is an extension of the field $F = k(X)$ of rational functions on X , and the Galois group $\text{Gal}(k(Y)/k(X))$ may be viewed as the group of “deck transformations” of the cover. If our cover is *unramified*, then this group is a quotient of the (arithmetic) fundamental group of X . For a cover ramified at points x_1, \dots, x_n , it is a quotient of the (arithmetic) fundamental group of $X \setminus \{x_1, \dots, x_n\}$. From now on (with the exception of Section 5.4) we will focus on the unramified case. This means that we replace $\text{Gal}(\overline{F}/F)$ by its maximal unramified quotient, which is nothing but the (arithmetic) fundamental group of X . Its geometric analogue, when X is defined over \mathbb{C} , is $\pi_1(X)$.

Thus, the geometric counterpart of a (unramified) homomorphism $\text{Gal}(\overline{F}/F) \rightarrow {}^L G$ is a homomorphism $\pi_1(X) \rightarrow {}^L G$.

From now on, let X be a smooth projective connected algebraic curve defined over \mathbb{C} . Let G be a complex reductive algebraic group and ${}^L G$ its Langlands dual group. Then homomorphisms $\pi_1(X) \rightarrow {}^L G$ may be described in differential geometric terms as bundles with a flat connection (the monodromy of the flat connection gives rise to a homomorphism $\pi_1(X) \rightarrow {}^L G$). Let E be a smooth principal ${}^L G$ -bundle on X . A flat connection on E has two components. The $(0, 1)$ component, with respect to the complex structure on X , defines holomorphic structure on E , and the $(1, 0)$ component defines

a holomorphic connection ∇ . Thus, an ${}^L G$ -bundle with a flat connection on X is the same as a pair (E, ∇) , where E is a holomorphic (equivalently, algebraic) ${}^L G$ -bundle on X and ∇ is a holomorphic (equivalently, algebraic) connection on E .

Thus, for complex curves the objects on one side of the Langlands correspondence are equivalence classes of flat (holomorphic or algebraic) ${}^L G$ -bundles (E, ∇) .

What about the other side? Here the answer is not quite as obvious. I will sketch it briefly referring the reader to Section 3 of [F2] for more details.

Recall that automorphic representations of $G(\mathbb{A}_F)$ (where F is a function field of a curve X defined over \mathbb{F}_q) are realized in functions on the quotient $G(F)\backslash G(\mathbb{A}_F)$. An unramified automorphic representation (which corresponds to an unramified homomorphism $\text{Gal}(\overline{F}/F) \rightarrow {}^L G$) gives rise to a function on the double quotient $G(F)\backslash G(\mathbb{A}_F)/G(\mathcal{O}_F)$, where $\mathcal{O}_F = \prod_{x \in X} \mathcal{O}_x$. A key observation (which is due to Weil) is that this double quotient is precisely the set of isomorphism classes of principal G -bundles on our curve X .⁽⁶⁾ This statement is also true if the curve X is defined over \mathbb{C} . Thus, geometric analogues of unramified automorphic representations should be some geometric objects which “live” on a moduli space of G -bundles.

Unfortunately, for a non-abelian group G there is no algebraic variety whose set of \mathbb{C} -points is the set of isomorphism classes of G -bundles on X (for $G = GL_1$ we can take the Picard variety). However, there is an *algebraic moduli stack* denoted by Bun_G . It is not an algebraic variety, but it looks locally like the quotient of an algebraic variety by the action of an algebraic group (these actions are not free, and therefore the quotient is no longer an algebraic variety). It turns out that this is good enough for our purposes.

So which geometric objects on Bun_G will replace unramified automorphic representations? Here we need to recall that the function on the double quotient $G(F)\backslash G(\mathbb{A}_F)/G(\mathcal{O}_F)$ attached to an unramified automorphic representation has a special property: it is an eigenfunction of the so-called Hecke operators. Those are cousins of the classical Hecke operators one studies in the theory of modular forms (which is in the left column of Weil’s big picture). The geometric objects we are looking for will be certain sheaves on Bun_G satisfying an analogue of the Hecke property. We will call them *Hecke eigensheaves*.

More precisely, these sheaves are \mathcal{D} -modules on Bun_G . Recall (see, e.g., [KS, GM]) that a \mathcal{D} -module on a smooth algebraic variety Z is a sheaf of modules over the sheaf \mathcal{D}_Z of differential operators on Z . An example of a \mathcal{D} -module is the sheaf of sections of a flat vector bundle on Z . The sheaf of functions on Z acts on sections by multiplication, so it is an \mathcal{O}_Z -module. But the flat connection also allows us to act on sections by vector fields on Z . This gives rise to an action of the sheaf \mathcal{D}_Z , because it is generated by vector fields and functions. Thus, we obtain the structure of a \mathcal{D} -module.

In our case, Bun_G is not a variety, but an algebraic stack, but the (derived) category of \mathcal{D} -modules on it has been defined in [BD]. On this category act the so-called *Hecke functors*. These are labeled by pairs (x, V) , where $x \in X$ and V is a finite-dimensional

⁽⁶⁾From now on we will only consider algebraic bundles.

representation of the dual group ${}^L G$, and are defined using certain modifications of G -bundles.

Instead of giving a general definition (which may be found in [BD] or [F2]) we will consider two examples. First, consider the abelian case when $G = GL_1$ (thus, we have $G(\mathbb{C}) = \mathbb{C}^\times$). In this case Bun_G may be replaced by the Picard variety Pic which parametrizes line bundles on X . Given a point $x \in X$, consider the map $h_x : \text{Pic} \rightarrow \text{Pic}$ sending a line bundle \mathcal{L} to $\mathcal{L}(x)$ (the line bundle whose sections are sections of \mathcal{L} which are allowed to have a pole of order 1 at x). By definition, the Hecke functor $H_{1,x}$ corresponding to x and $1 \in \mathbb{Z}$ (which we identify with the set of one-dimensional representations of ${}^L G = GL_1$), is given by the formula

$$H_{1,x}(\mathcal{F}) = h_x^*(\mathcal{F}).$$

Next, consider the case of $G = GL_n$ and $V = V_{\check{\omega}_1}$, the defining n -dimensional representation of ${}^L G = GL_n$. In this case Bun_{GL_n} is the moduli stack Bun_n of rank n bundles on X . There is an obvious analogue of the map h_x , sending a rank n bundle \mathcal{M} to $\mathcal{M}(x)$. But then the degree of the bundle jumps by n . It is possible to increase it by 1, but we need to choose a line ℓ in the fiber of \mathcal{M} at x . We then define a new rank n bundle \mathcal{M}' by saying that its sections are the sections of \mathcal{M} having a pole of order 1 at x , but the polar part has to belong to ℓ . Then $\deg \mathcal{M}' = \deg \mathcal{M} + 1$. However, we now have a \mathbb{P}^{n-1} worth of modifications of \mathcal{M} corresponding to different choices of the line ℓ . The Hecke functor $H_{V_{\check{\omega}_1},x}$ is obtained by “integrating” over all of them.

More precisely, let $\mathcal{H}ecke_{\check{\omega}_1,x}$ be the moduli stack of pairs $(\mathcal{M}, \mathcal{M}')$ as above. It defines a correspondence over $\text{Bun}_n \times \text{Bun}_n$:

$$(2.1) \quad \begin{array}{ccc} & \mathcal{H}ecke_{\check{\omega}_1,x} & \\ h_x^{\leftarrow} \swarrow & & \searrow h_x^{\rightarrow} \\ \text{Bun}_n & & \text{Bun}_n \end{array}$$

By definition,

$$(2.2) \quad H_{\check{\omega}_1,x}(\mathcal{F}) = h_{x*}^{\rightarrow} h_x^{\leftarrow*}(\mathcal{F}).$$

For irreducible representations $V_{\check{\lambda}}$ of ${}^L G$ with general dominant integral highest weights $\check{\lambda}$ there is an analogous correspondence in which the role of the projective space \mathbb{P}^{n-1} is played by the Schubert variety in the affine Grassmannian of G corresponding to $\check{\lambda}$ (see [BD, MV], and [F2] for a brief outline).

Allowing the point x to vary, we obtain a correspondence between Bun_G and $X \times \text{Bun}_G$ and Hecke functors acting from the category of \mathcal{D} -modules on Bun_G to the (derived) category of \mathcal{D} -modules on $X \times \text{Bun}_G$, which we denote by $H_V, V \in \text{Rep } {}^L G$.

Now let $\mathcal{E} = (E, \nabla)$ be a flat ${}^L G$ -bundle on X . A \mathcal{D} -module \mathcal{F} on Bun_G is called a *Hecke eigensheaf* with respect to \mathcal{E} (or with “eigenvalue” \mathcal{E}) if we have a collection of isomorphisms

$$(2.3) \quad H_V(\mathcal{F}) \simeq V_{\mathcal{E}} \boxtimes \mathcal{F},$$

compatible with the tensor product structures. Here

$$V_{\mathcal{E}} = \mathcal{E} \times_{\mathcal{L}G} V$$

is the flat vector bundle on X associated to \mathcal{E} and V , viewed as a \mathcal{D} -module. Thus, in particular, we have a collection of isomorphisms

$$H_{V,x}(\mathcal{F}) \simeq V \otimes \mathcal{F}, \quad x \in X.$$

When we vary the point x , the “eigenvalues”, which are all isomorphic to the vector space underlying V , combine into the flat vector bundle $V_{\mathcal{E}}$ on X .

The *geometric Langlands conjecture* may be stated as follows: for any flat $\mathcal{L}G$ -bundle \mathcal{E} there exists a non-zero \mathcal{D} -module $\mathcal{F}_{\mathcal{E}}$ on Bun_G with eigenvalue \mathcal{E} .

Moreover, if \mathcal{E} is irreducible, this \mathcal{D} -module is supposed to be irreducible (when restricted to each connected component of Bun_G) and unique up to an isomorphism (it should also be holonomic and have regular singularities). But if \mathcal{E} is not irreducible, we might have a non-trivial (derived) category of Hecke eigensheaves, and the situation becomes more subtle.

Thus, at least for irreducible \mathcal{E} , we expect the following picture:

$$(2.4) \quad \boxed{\begin{array}{c} \text{flat} \\ \mathcal{L}G\text{-bundles on } X \end{array}} \longrightarrow \boxed{\begin{array}{c} \text{Hecke eigensheaves} \\ \text{on } \text{Bun}_G \end{array}}$$

$$\mathcal{E} \longrightarrow \mathcal{F}_{\mathcal{E}}.$$

The geometric Langlands correspondence has been constructed in many cases. For $G = GL_n$ the Hecke eigensheaves corresponding to irreducible \mathcal{E} have been constructed in [FGV, Ga], building on the work of P. Deligne for $n = 1$ (explained in [Lau1] and [F2]), V. Drinfeld [Dr1] for $n = 2$, and G. Laumon [Lau1] (this construction works for curves defined both over \mathbb{F}_q or \mathbb{C}).

For all simple algebraic groups G the Hecke eigensheaves have been constructed in a different way (for curves over \mathbb{C}) by A. Beilinson and V. Drinfeld [BD] in the case when \mathcal{E} has an additional structure of an *oper* (this means that \mathcal{E} belongs to a certain half-dimensional locus in $\text{Loc}_{\mathcal{L}G}$). It is interesting that this construction is also closely related to quantum field theory, but in a seemingly different way. Namely, it uses methods of 2D Conformal Field Theory and representation theory of affine Kac–Moody algebras of critical level. For more on this, see Part III of [F2].

3. CATEGORICAL VERSION

Looking at the correspondence (2.4), we notice that there is an essential asymmetry between the two sides. On the left we have flat $\mathcal{L}G$ -bundles, which are points of a moduli stack $\text{Loc}_{\mathcal{L}G}$ of flat $\mathcal{L}G$ -bundles (or local systems) on X . But on the right we have Hecke

eigensheaves, which are objects of a category; namely, the category of \mathcal{D} -modules on Bun_G . Beilinson and Drinfeld have suggested a natural way to formulate it in a more symmetrical way.

The idea is to replace a point $\mathcal{E} \in \text{Loc}_{LG}$ by an object of another category; namely, the skyscraper sheaf $\mathcal{O}_{\mathcal{E}}$ at \mathcal{E} viewed as an object of the category of coherent \mathcal{O} -modules on Loc_{LG} . A much stronger, categorical, version of the geometric Langlands correspondence is then a conjectural equivalence of derived categories⁽⁷⁾

$$(3.1) \quad \boxed{\text{derived category of } \mathcal{O}\text{-modules on } \text{Loc}_{LG}} \longleftrightarrow \boxed{\text{derived category of } \mathcal{D}\text{-modules on } \text{Bun}_G}$$

This equivalence should send the skyscraper sheaf $\mathcal{O}_{\mathcal{E}}$ on Loc_{LG} supported at \mathcal{E} to the Hecke eigensheaf \mathcal{F}_E . If this were true, it would mean that Hecke eigensheaves provide a good “basis” in the category of \mathcal{D} -modules on Bun_G , so we would obtain a kind of spectral decomposition of the derived category of \mathcal{D} -modules on Bun_G , like in the Fourier transform. (Recall that under the Fourier transform on the real line the delta-functions δ_x , analogues of $\mathcal{O}_{\mathcal{E}}$, go to the exponential functions e^{itx} , analogues of $\mathcal{F}_{\mathcal{E}}$.)

This equivalence has been proved by G. Laumon [Lau2] and M. Rothstein [R] in the abelian case, when $G = GL_1$ (or a more general torus). They showed that in this case this is nothing but a version of the Fourier–Mukai transform. Thus, the categorical Langlands correspondence may be viewed as a kind of non-abelian Fourier–Mukai transform (see [F2], Section 4.4).

Unfortunately, a precise formulation of such a correspondence, even as a conjecture, is not so clear because of various subtleties involved. One difficulty is the structure of Loc_{LG} . Unlike the case of ${}^L G = GL_1$, when all flat bundles have the same groups of automorphisms (namely, GL_1) and Loc_{GL_1} is smooth, for a general group ${}^L G$ the groups of automorphisms are different for different flat bundles, and so Loc_{LG} is a complicated stack. For example, if ${}^L G$ is a simple Lie group of adjoint type, then a generic flat ${}^L G$ -bundle has no automorphisms, while the group of automorphisms of the trivial flat bundle is isomorphic to ${}^L G$. In addition, unlike Bun_G , the stack Loc_{LG} has singularities. All of this has to be reflected on the other side of the correspondence, in ways that have not yet been fully understood.

Nevertheless, the diagram (3.1) gives us a valuable guiding principle to the geometric Langlands correspondence. In particular, it gives us a natural explanation as to why the skyscraper sheaves on Loc_{LG} should correspond to Hecke eigensheaves.

The point is that on the category of \mathcal{O} -modules on Loc_{LG} we also have a collection of functors W_V , parametrized by the same data as the Hecke functors H_V . Following

⁽⁷⁾ It is expected (see [FW], Sect. 10) that there is in fact a \mathbb{Z}_2 -gerbe of such equivalences. This gerbe is trivial, but not canonically trivialized. One gets a particular trivialization of this gerbe, and hence a particular equivalence, for each choice of the square root of the canonical line bundle K_X on X .

physics terminology, we will call them *Wilson functors*. These functors act from the category of \mathcal{O} -modules on Loc_{L_G} to the category of sheaves on $X \times \mathrm{Loc}_{L_G}$, which are \mathcal{D} -modules along X and \mathcal{O} -modules along Loc_{L_G} .

To define them, observe that we have a tautological L_G -bundle \mathcal{T} on $X \times \mathrm{Loc}_{L_G}$, whose restriction to $X \times \mathcal{E}$, where $\mathcal{E} = (E, \nabla)$, is E . Moreover, ∇ gives us a partial connection on \mathcal{T} along X . For a representation V of L_G , let $V_{\mathcal{T}}$ be the associated vector bundle on $X \times \mathrm{Loc}_{L_G}$, with a connection along X .

Let $p : X \times \mathrm{Loc}_{L_G} \rightarrow \mathrm{Loc}_{L_G}$ be the projection onto the second factor. By definition,

$$(3.2) \quad W_V(\mathcal{F}) = V_{\mathcal{T}} \otimes p^*(\mathcal{F})$$

(note that by construction $V_{\mathcal{T}}$ carries a connection along X and so the right hand side really is a \mathcal{D} -module along X).

Now, the conjectural equivalence (3.1) should be compatible with the Wilson/Hecke functors in the sense that

$$(3.3) \quad C(W_V(\mathcal{F})) \simeq H_V(C(\mathcal{F})), \quad V \in \mathrm{Rep} L_G,$$

where C denotes this equivalence (from left to right).

In particular, observe that the skyscraper sheaf $\mathcal{O}_{\mathcal{E}}$ at $\mathcal{E} \in \mathrm{Loc}_{L_G}$ is obviously an eigensheaf of the Wilson functors:

$$W_V(\mathcal{O}_{\mathcal{E}}) = V_{\mathcal{E}} \boxtimes \mathcal{O}_{\mathcal{E}}.$$

Indeed, tensoring a skyscraper sheaf with a vector bundle is the same as tensoring it with the fiber of this vector bundle at the point of support of this skyscraper sheaf. Therefore (3.3) implies that $\mathcal{F}_{\mathcal{E}} = C(\mathcal{O}_{\mathcal{E}})$ must satisfy the Hecke property (2.3). In other words, $\mathcal{F}_{\mathcal{E}}$ should be a Hecke eigensheaf on Bun_G with eigenvalue \mathcal{E} . Thus, we obtain a natural explanation of the Hecke property of $\mathcal{F}_{\mathcal{E}}$: it follows from the compatibility of the categorical Langlands correspondence (3.1) with the Wilson/Hecke functors.

Let us summarize: the conjectural equivalence (3.1) gives us a natural and convenient framework for the geometric Langlands correspondence. It is this equivalence that Kapustin and Witten have related to the S -duality of 4D super–Yang–Mills.

4. ENTER PHYSICS

We will now add a fourth column to Weil’s big picture, which we will call “Quantum Physics”:

| | | | |
|----------------------|----------------------------------------------|-------------------------|------------------------|
| <i>Number Theory</i> | <i>Curves over \mathbb{F}_q</i> | <i>Riemann Surfaces</i> | <i>Quantum Physics</i> |
|----------------------|----------------------------------------------|-------------------------|------------------------|

In the context of the Langlands Program, the last column means S -duality and Mirror Symmetry of certain 4D and 2D quantum field theories, which we will now briefly describe following [KW].

We start with the pure 4D Yang–Mills (or gauge) theory on a Riemannian four-manifold M_4 . Let G_c be a compact connected simple Lie group. The classical (Euclidean) action is a functional on the space of connections on arbitrary principal G_c -bundles \mathcal{P} on M_4 given by the formula

$$I = \frac{1}{4g^2} \int_{M_4} \text{Tr} F_A \wedge \star F_A + \frac{i\theta}{8\pi^2} \int_{M_4} \text{Tr} F_A \wedge F_A.$$

Here F_A is the curvature of the connection A (a \mathfrak{g} -valued two-form on M_4), \star is the Hodge star operator, and Tr is the invariant bilinear form on the Lie algebra \mathfrak{g} normalized in such a way that the second term is equal to $i\theta k$, where k could be an arbitrary integer, if G_c is simply-connected. The second term is equal to $i\theta$ times the second Chern class $c_2(\mathcal{P})$ of the bundle \mathcal{P} and hence is topological. Correlation functions are given by path integrals of the form $\int e^{-I}$ over the space of connections modulo gauge transformations. Hence they may be written as Fourier series in $e^{i\theta}$ (or its root if G_c is not simply-connected) such that the coefficient in front of $e^{i\theta n}$ is the sum of contributions from bundles \mathcal{P} with $c_2(\mathcal{P}) = -n$.

It is customary to combine the two parameters, g and θ , into one complex coupling constant

$$\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}.$$

Next, we consider $N = 4$ supersymmetric extension of this model. This means that we add fermionic and bosonic fields in such a way that the action of the Lorentz group (we will work in Euclidean signature, where this group becomes $SO(4)$) is extended to an action of an appropriate supergroup (see [KW] or the books [D] for a background on supersymmetric quantum field theory).

The S -duality of this theory is the statement that the theory with gauge group G_c and complex coupling constant τ is equivalent to the theory with the Langlands dual gauge group ${}^L G_c$ and coupling constant ${}^L \tau = -1/n_{\mathfrak{g}}\tau$:

$$(4.1) \quad (G_c, \tau) \longleftrightarrow ({}^L G_c, -1/n_{\mathfrak{g}}\tau),$$

where $n_{\mathfrak{g}}$ is the lacing number of the Lie algebra \mathfrak{g} (equal to 1 for simply-laced Lie algebras, 2 for B_n, C_n and F_4 , and 3 for G_2). This is an extension of the duality (0.1) of [MO] discussed in the introduction to non-zero values of θ (with g normalized in a slightly different way). In addition, for simply-laced G_c the path integral is a Fourier series in $e^{i\theta}$, so θ may be shifted by an integer multiple of 2π without changing the path integral. Thus, we also have the equivalence

$$(G_c, \tau) \longleftrightarrow (G_c, \tau + 1).$$

For general non-simply connected groups we have instead a symmetry $\tau \mapsto \tau + n\mathbb{Z}$, where n is a certain integer. Thus, we obtain an action of a subgroup of $SL_2(\mathbb{Z})$ on

the super–Yang–Mills theories with gauge groups G_c and ${}^L G_c$.⁽⁸⁾ As we discussed in the Introduction, this is a striking statement because it relates a theory at strong coupling to a theory at weak coupling.

We want to focus next on a “topological sector” of this theory. This means that we pick an element Q in the Lie superalgebra \mathfrak{s} of the super-Lorentz group (the supergroup extension of $SO(4)$) such that $Q^2 = 0$, and such that the stress tensor (which is a field responsible for variation of the metric on M_4) is equal to the commutator of Q and another field. Let us restrict ourselves to those objects (fields, boundary conditions, etc.) in the theory which commute with this Q . This is a particular (and relatively small) sector of the full quantum field theory, in which all quantities (such as correlation functions) are topological, that is, metric-independent. This sector is what is usually referred to as Topological Field Theory (TFT).

There is a problem, however. For this $Q \in \mathfrak{s}$ to be well-defined on an arbitrary manifold M_4 , it has to be invariant under the action of the Lorentz group $SO(4)$ – more precisely, its double cover $Spin(4)$. Unfortunately, there are no such elements in our Lie superalgebra \mathfrak{s} . In order to obtain such an element, one uses a trick, called *twisting* (see, e.g., [Wi6]). Our theory has an additional group of automorphisms commuting with the action of $Spin(4)$, called R -symmetry; namely, the group $Spin(6)$. We can use it to modify the action of $Spin(4)$ on the fields of the theory and on the Lie superalgebra \mathfrak{s} as follows: define a new action of $Spin(4)$ equal to the old action together with the action coming from a homomorphism $Spin(4) \rightarrow Spin(6)$ and the action of $Spin(6)$ by R -symmetry. One might then be able to find a differential $Q \in \mathfrak{s}$ invariant under this new action of $Spin(4)$.

There are essentially three different choices for doing this, as explained in [VW]. The first two are similar to the twists used in Witten’s construction of a topological field theory that yields Donaldson invariants of four-manifolds (which is a topological twist of an $N = 2$ supersymmetric Yang–Mills theory) [Wi1]. It is the third twist, studied in detail in [KW], that is relevant to the geometric Langlands. For this twist there are actually two linearly independent (and anti-commuting with each other) operators, Q_l and Q_r , which square to 0. We can therefore use any linear combination

$$Q = uQ_l + vQ_r$$

as the differential defining the topological field theory (for each of them the stress tensor will be a Q -commutator, so we will indeed obtain a topological field theory). We obtain the same theory if we rescale both u and v by the same number. Hence we obtain a family of topological field theories parametrized by \mathbb{P}^1 .

Let $t = v/u$ be a coordinate on this \mathbb{P}^1 . We will refer to it as the “twisting parameter.” The S -duality (4.1) should be accompanied by the change of the twisting parameter

⁽⁸⁾In general, it is a proper subgroup of $SL_2(\mathbb{Z})$ for two reasons: first, we have the coefficient n_g in formula (4.1) for non-simply laced G_c , and second, the dual of a simply-connected Lie group is not simply-connected in general, in which case the transformation $\tau \rightarrow \tau + 1$ is not a symmetry.

according to the rule

$$(4.2) \quad t \mapsto \frac{\tau}{|\tau|} t.$$

How can we test S -duality of these topological theories? Vafa and Witten have earlier tested the S -duality of a different (Donaldson type) topological twisting of $N = 4$ super–Yang–Mills by showing that the partition functions of these theories (depending on τ) are modular forms [VW]. This proves the invariance of the partition functions under the action of a subgroup of $SL_2(\mathbb{Z})$ on τ . What turns out to be relevant to the geometric Langlands Program is the study of boundary conditions in these topological field theories.

Kapustin and Witten assume that the four-manifold M_4 has the form

$$M_4 = \Sigma \times X,$$

where X is a closed Riemann surface (this will be the algebraic curve of the geometric Langlands) and Σ is a Riemann surface with a boundary (which we may simply take to be a half-plane). They study the limit of the topological gauge theory on this manifold when X becomes very small (this is called “compactification of the theory on X ”). In this limit the theory is described by an effective *two-dimensional* topological field theory on Σ . In earlier works [BJSV, HMS] the latter theory was identified with the (twisted) topological sigma model on Σ with the target manifold $\mathcal{M}_H(G)$, the *Hitchin moduli space* of Higgs G -bundles on X . Moreover, the S -duality of the supersymmetric gauge theories on $\Sigma \times X$ (for particular values of τ and t) becomes Mirror Symmetry between the topological sigma models with the targets $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^L G)$.

Next, we look at the boundary conditions in the gauge theories, which give rise to *branes* in these sigma models. S -duality yields an equivalence of the categories of branes for $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^L G)$ (also known, after M. Kontsevich, as Homological Mirror Symmetry). Kapustin and Witten have related this equivalence to the categorical geometric Langlands correspondence (3.1). Thus, they establish a link between S -duality and geometric Langlands duality. We describe this in more detail in the next section.

5. MIRROR SYMMETRY OF HITCHIN MODULI SPACES

In [Hi1] N. Hitchin introduced a remarkable hyper-Kähler manifold $\mathcal{M}_H(G)$ for each smooth projective complex algebraic curve X and reductive Lie group G . It is easiest to describe it in its complex structure I , in which it is the moduli space of semi-stable *Higgs bundles* on X . Recall that a Higgs G -bundle on X is a pair (E, ϕ) , where E is a (algebraic) G -bundle on X and ϕ is a Higgs field on it, that is,

$$\phi \in H^0(X, \mathfrak{g}_E \otimes K_X),$$

where $\mathfrak{g}_E = E \times_G \mathfrak{g}$ is the adjoint vector bundle.

In the complex structure J , however, $\mathcal{M}_H(G)$ is described as the moduli space of semi-stable *flat bundles*, that is, pairs (E, ∇) , where E is again (algebraic) G -bundle on X and ∇ is (algebraic) connection on E . To distinguish between it as a complex algebraic variety from the moduli space of Higgs bundles we will denote it by $\mathcal{Y}(G)$. The two are isomorphic as real manifolds (this is the statement of non-abelian Hodge theory [Hi1, C, S1]), but not as complex (or algebraic) manifolds.

There are two types of twisted supersymmetric two-dimensional sigma models with Kähler target manifolds: A -model and B -model (see [Wi2]). The former depends on the symplectic structure on the target manifold and the latter depends on the complex structure.

Kapustin and Witten start with two topological twisted super–Yang–Mills theories on $\Sigma \times X$. One has gauge group G_c , twisting parameter $t = 1$, and $\theta = 0$. The other, S -dual theory, has gauge group ${}^L G_c$, the twisting parameter ${}^L t = i$, and ${}^L \theta = 0$ (neither of these topological theories depends on g [KW]).⁽⁹⁾ They show that after compactification on X the first theory becomes the A -model with the target manifold $\mathcal{M}_H(G)$ and the symplectic structure ω_K , which is the Kähler form for the complex structure K on $\mathcal{M}_H(G)$. This symplectic structure has a nice geometric description. Note that the Higgs field ϕ is an element of $H^0(X, \mathfrak{g}_E \otimes K_X)$, which is isomorphic to the cotangent space to E , viewed as a point of Bun_G , the moduli stack of G -bundles on X . Thus, $\mathcal{M}_H(G)$ is almost the cotangent bundle to Bun_G ; “almost” because we impose the semi-stability condition on the Higgs bundle. The symplectic form ω_K comes from the standard symplectic form on the cotangent bundle (which is the imaginary part of the holomorphic symplectic form).

The second gauge theory becomes, after compactification on X , the B -model with the target manifold $\mathcal{Y}({}^L G)$; that is, $\mathcal{M}_H({}^L G)$ with respect to the complex structure J .

After dimensional reduction from 4D to 2D, the S -duality of super–Yang–Mills theories becomes Mirror Symmetry between the A -model with the target manifold $\mathcal{M}_H(G)$ (and symplectic structure ω_K) and the B -model with the target manifold $\mathcal{Y}({}^L G)$ (and complex structure J).

Remark. As explained in [KW], there is also Mirror Symmetry between A - and B -models with respect to other symplectic and complex structures. For instance, there is Mirror Symmetry, studied in [DP, Hi3, Ari], between the B -models on $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^L G)$ with respect to the complex structures I on both of them. In what follows we will not discuss these additional dualities.

5.1. Dual Hitchin fibrations

In order to understand better this Mirror Symmetry, we recall the construction of the *Hitchin map*. For any Higgs bundle (E, ϕ) and an invariant polynomial P of degree d on the Lie algebra \mathfrak{g} , we can evaluate P on ϕ and obtain a well-defined section $P(\phi)$

⁽⁹⁾As explained in [KW, Ka], for some other values of parameters one obtains the so-called quantum geometric Langlands correspondence (see [F2], Section 6.3).

of $K_X^{\otimes d}$. The algebra $(\text{Fun}(\mathfrak{g}))^G$ of invariant polynomial functions on \mathfrak{g} is a graded free polynomial algebra with $\ell = \text{rank}(\mathfrak{g})$ generators of degrees $d_i, i = 1, \dots, \ell$, where the d_i are the exponents of \mathfrak{g} plus 1. Let us choose a set of generators $P_i, i = 1, \dots, \ell$. Then we construct the Hitchin map [Hi1, Hi2]

$$p : \mathcal{M}_H(G) \rightarrow \mathbf{B} = \bigoplus_{i=1}^{\ell} H^0(X, K_X^{\otimes d_i}),$$

$$(E, \phi) \mapsto (P_1(\phi), \dots, P_\ell(\phi)).$$

This is slightly non-canonical, because there is no canonical choice of generators P_i in general. More canonically, we have a map to

$$\mathbf{B} := H^0(X, (\mathfrak{g}/G)_{K_X}), \quad (\mathfrak{g}/G)_{K_X} := K_X^{\times} \times_{\mathbb{C}^{\times}} \mathfrak{g}/G,$$

where K_X^{\times} denotes the \mathbb{C}^{\times} -bundle associated to K_X , and $\mathfrak{g}/G := \text{Spec}((\text{Fun}(\mathfrak{g}))^G)$ is the graded vector space on which the P_i are coordinate functions (the \mathbb{C}^{\times} -action on it comes from the grading).

By Chevalley's theorem, $\mathfrak{g}/G = \mathfrak{h}/W := \text{Spec}((\text{Fun}(\mathfrak{h}))^W)$, where \mathfrak{h} is a Cartan subalgebra. By definition, ${}^L\mathfrak{h} = \mathfrak{h}^*$. Hence ${}^L\mathfrak{g}/{}^L G = {}^L\mathfrak{h}/W = \mathfrak{h}^*/W$. Choosing a non-degenerate invariant bilinear form κ_0 on \mathfrak{g} , we identify \mathfrak{h} and \mathfrak{h}^* , and hence \mathbf{B} and ${}^L\mathbf{B}$. Any other invariant bilinear form is proportional to κ_0 . Hence replacing κ_0 by another non-zero bilinear form would correspond to a \mathbb{C}^{\times} -action on the base. But this action can be lifted to a \mathbb{C}^{\times} -action on the total space $\mathcal{M}_H(G)$ (rescaling the Higgs field ϕ). Hence the ambiguity in the choice of κ_0 is not essential; it may be absorbed into an automorphism of one of the two Hitchin moduli spaces.

As the result, we obtain two fibrations over the same base \mathbf{B} :

$$(5.1) \quad \begin{array}{ccc} \mathcal{M}_H({}^L G) & & \mathcal{M}_H(G) \\ & \searrow & \swarrow \\ & \mathbf{B} & \end{array}$$

For generic $b \in \mathbf{B}$ (the connected components of) the fibers ${}^L\mathbf{F}_b$ and \mathbf{F}_b of these Hitchin fibrations are smooth tori, which are in fact isomorphic to abelian varieties in the complex structure I . For instance, for $G = SL_n$, \mathbf{F}_b is the generalized Prym variety of the spectral curve associated to b , which is a smooth degree n cover of X if b is generic.

Moreover, the tori ${}^L\mathbf{F}_b$ and \mathbf{F}_b (again, for generic b) are *dual* to each other. This can be expressed in the following way which will be convenient for our purposes: there is a bijection between points of ${}^L\mathbf{F}_b$ and flat unitary line bundles on \mathbf{F}_b .⁽¹⁰⁾ The duality

⁽¹⁰⁾ This bijection depends on the choice of base points in the fibers. Using the Hitchin section, one obtains such base points, but for this one may have to choose a square root of the canonical line bundle K_X . This is closely related to the \mathbb{Z}_2 -gerbe ambiguity in the equivalence (3.1) discussed in the footnote on page 10.

of tori is the simplest (abelian) example of Mirror Symmetry, also known as *T-duality*. Thus, we expect that the Mirror Symmetry of the Hitchin moduli spaces $\mathcal{M}_H({}^L G)$ and $\mathcal{M}_H(G)$ is realized via fiberwise *T-duality* (for generic fibers of the dual Hitchin fibrations). This is an example of a general proposal of Strominger, Yau and Zaslow [SYZ] that Mirror Symmetry of two Calabi–Yau manifolds X and Y should be realized as *T-duality* of generic fibers in special Lagrangian fibrations of X and Y (what happens for the singular fibers is *a priori* far less clear; but see Section 6.1). It is in this sense that “*S-duality* reduces to *T-duality*” [HMS].

We are interested in the study of the *B-model* with the target $\mathcal{M}_H({}^L G)$, with respect to the complex structure J (that is, the moduli space $\mathcal{Y}({}^L G)$ of flat bundles), and the *A-model* with the target $\mathcal{M}_H(G)$, with respect to the symplectic structure ω_K . These two topological field theories are expected to be equivalent to each other. Therefore anything we can say about one of them should have a counterpart in the other. For instance, their cohomologies may be interpreted as the spaces of vacua in these field theories, and hence they should be isomorphic. This has indeed been verified by Hausel and Thaddeus [HT] in the case when $G = SL_n$, ${}^L G = PGL_n$ (since the Hitchin moduli spaces are non-compact, special care has to be taken to properly define these cohomologies, see [HT]).

To make contact with the geometric Langlands correspondence, Kapustin and Witten study in [KW] the *categories of branes* in these two topological field theories.

5.2. Categories of branes

Branes in two-dimensional sigma models are certain generalizations of boundary conditions. When writing path integral for maps $\Phi : \Sigma \rightarrow M$, where Σ has a boundary, we need to specify boundary conditions for Φ on $\partial\Sigma$. We may also “couple” the sigma model to another quantum field theory on $\partial\Sigma$ (that is, modify the action by a boundary term) which may be interpreted as a decoration of the boundary condition. In topological field theory these conditions should preserve the supersymmetry, which leads to natural restrictions.

A typical example of a boundary condition is specifying that $\Phi(\partial\Sigma)$ belongs to a submanifold $M' \subset M$. In the *B-model* the target manifold M is a complex manifold, and in order to preserve the supersymmetry M' has to be a complex submanifold. In the *A-model*, M is a symplectic manifold and M' should be Lagrangian. Coupling to field theories on $\partial\Sigma$ allows us to introduce into the picture a holomorphic vector bundle on M' in the case of *B-model*, and a flat unitary vector bundle on M' in the case of *A-model*.

More generally, the category of branes in the *B-model* with a complex target manifold M (called *B-branes*) is the (derived) category of coherent sheaves on M , something that is fairly well understood mathematically. The category of branes in the *A-model* with a symplectic target manifold M (called *A-branes*) is less understood. It is believed to contain what mathematicians call the Fukaya category, typical objects of which are

pairs (L, ∇) , where $L \subset M$ is a Lagrangian submanifold and ∇ is a flat unitary vector bundle on L . However (and this turns out to be crucial for applications to the geometric Langlands), it also contains more general objects, such as coisotropic submanifolds of M equipped with vector bundles with unitary connection.

Under the Mirror Symmetry between the sigma models with the target manifolds $\mathcal{Y}({}^L G)$ and $\mathcal{M}_H(G)$ we therefore expect to have the following equivalence of (derived) categories of branes (often referred to, after Kontsevich, as Homological Mirror Symmetry):

$$(5.2) \quad \boxed{B\text{-branes on } \mathcal{Y}({}^L G)} \longleftrightarrow \boxed{A\text{-branes on } \mathcal{M}_H(G)}$$

It is this equivalence that Kapustin and Witten have related to the categorical Langlands correspondence (3.1). The category on the left in (5.2) is the (derived) category of coherent sheaves on $\mathcal{Y}({}^L G)$, which is the moduli space of semi-stable flat ${}^L G$ -bundles on X . It is closely related to the (derived) category of coherent sheaves (or, equivalently, \mathcal{O} -modules) on $\text{Loc}_{{}^L G}$, which appears on the left of (3.1). The difference is that, first of all, $\text{Loc}_{{}^L G}$ is the moduli *stack* of flat ${}^L G$ -bundles on X , whereas $\mathcal{Y}({}^L G)$ is the moduli *space* of semi-stable ones. Second, from the physics perspective it is more natural to consider coherent sheaves on $\mathcal{Y}({}^L G)$ with respect to its complex analytic rather than algebraic structure, whereas in (3.1) we consider algebraic \mathcal{O} -modules on $\text{Loc}_{{}^L G}$. These differences aside, these two categories are very similar to each other. They certainly share many objects, such as skyscraper sheaves supported at points corresponding to stable flat ${}^L G$ -bundles which we will discuss momentarily.

5.3. Triangle of equivalences

The categories on the right in (3.1) and (5.2) appear at first glance to be quite different. But Kapustin and Witten have suggested that they should be equivalent to each other as well. Thus, we obtain the following triangle of derived categories:

$$(5.3) \quad \begin{array}{ccc} & \boxed{A\text{-branes on } \mathcal{M}_H(G)} & \\ & \nearrow & \downarrow \\ \boxed{B\text{-branes on } \mathcal{Y}({}^L G)} & & \boxed{\mathcal{D}\text{-modules on } \text{Bun}_G} \\ & \searrow & \end{array}$$

The upper arrow represents Homological Mirror Symmetry (5.2) whereas the lower arrow represents the categorical Langlands correspondence (3.1).

According to [KW], Section 11, the vertical arrow is another equivalence that has nothing to do with either Mirror Symmetry or geometric Langlands. It should be a general statement linking the (derived) category of \mathcal{D} -modules on a variety M and the (derived) category of A -branes on its cotangent bundle T^*M (recall that $\mathcal{M}_H(G)$ is almost equal to $T^*\text{Bun}_G$). Kapustin and Witten have proposed the following functor from the category of A -branes on T^*M with respect to the symplectic structure $\text{Im } \Omega$ (where Ω is the holomorphic symplectic form on T^*M) to the category of \mathcal{D} -modules on M :

$$(5.4) \quad \mathcal{A} \mapsto \text{Hom}(\mathcal{A}_{cc}, \mathcal{A}),$$

where \mathcal{A}_{cc} is a “canonical coisotropic brane” on T^*M . This is T^*M itself (viewed as a coisotropic submanifold) equipped with a line bundle with connection satisfying special properties. They argued on physical grounds that the right hand side of (5.4) may be “sheafified” along M , and moreover that the corresponding sheaf of rings $\text{Hom}(\mathcal{A}_{cc}, \mathcal{A}_{cc})$ is nothing but the sheaf of differential operators on $\mathcal{M}_H(G)$.⁽¹¹⁾ Hence the (sheafified) right hand side of (5.4) should be a \mathcal{D} -module. While this argument has not yet been made mathematically rigorous, it allows one to describe important characteristics of the \mathcal{D} -module associated to an A -brane, such as its reducibility, the open subset of M where it is represented by a local system, the rank of this local system, and even its monodromy (see Section 4 of [FW]).

An alternative (and mathematically rigorous) approach to establishing an equivalence between the categories of A -branes and \mathcal{D} -modules has also been proposed by D. Nadler and E. Zaslow [NZ, Nad].

Though a lot of work still needs to be done to distill this connection and reconcile different approaches, this is clearly a very important and beautiful idea on its own right.

Thus, according to Kapustin and Witten, the (categorical) geometric Langlands correspondence (3.1) may be obtained in two steps. The first step is the Homological Mirror Symmetry (5.2) of the Hitchin moduli spaces for two dual groups, and the second step is the above link between the A -branes and \mathcal{D} -modules.

There is actually more structure in the triangle (5.3). On each of these three categories we have an action of certain functors, and all equivalences between them are supposed to commute with these functors. We have already described the functors on two of these categories in Section 3: these are the Wilson and Hecke functors. The functors acting on the categories of A -branes, were introduced in [KW] as the two-dimensional shadows of the ’t Hooft loop operators in 4D super–Yang–Mills theory. Like the Hecke functors, they are defined using modifications of G -bundles, but only those modifications which preserve the Higgs field. The S -duality of super–Yang–Mills theories is supposed to exchange the ’t Hooft operators and the Wilson operators (whose

⁽¹¹⁾More precisely, it is the sheaf of differential operators acting on a square root of the canonical line bundle on Bun_G , but we will ignore this subtlety here.

two-dimensional shadows are the functors described in Section 3), and this is the reason why we expect the equivalence (5.2) to commute with the action of these functors.

As explained above, the central objects in the geometric Langlands correspondence are Hecke eigensheaves attached to flat ${}^L G$ -bundles. Recall that the Hecke eigensheaf $\mathcal{F}_{\mathcal{E}}$ is the \mathcal{D} -module attached to the skyscraper sheaf $\mathcal{O}_{\mathcal{E}}$ supported at a point \mathcal{E} of $\text{Loc}_{{}^L G}$ under the conjectural equivalence (3.1). These \mathcal{D} -modules have very complicated structure. What can we learn about them from the point of view of Mirror Symmetry?

Let us assume first that \mathcal{E} has no automorphisms other than those coming from the center of ${}^L G$. Then it is a smooth point of $\mathcal{Y}({}^L G)$. These skyscraper sheaves $\mathcal{O}_{\mathcal{E}}$ are the simplest examples of B -branes on $\mathcal{Y}({}^L G)$ (called 0-branes). What is the corresponding A -brane on $\mathcal{M}_H(G)$?

The answer is surprisingly simple. Let $b \in \mathbf{B}$ be the projection of \mathcal{E} to the base of the Hitchin fibration. For \mathcal{E} satisfying the above conditions the Hitchin fiber ${}^L \mathbf{F}_b$ is a smooth torus (it is actually an abelian variety in the complex structure I , but now we look at it from the point of view of complex structure J , so it is just a smooth torus). It is identified (possibly, up to a choice of the square root of the canonical line bundle K_X , see the footnote on page 16) with the moduli space of flat unitary line bundles on the dual Hitchin fiber \mathbf{F}_b , which happens to be a Lagrangian submanifold of $\mathcal{M}_H(G)$. The Mirror Symmetry sends the B -brane $\mathcal{O}_{\mathcal{E}}$ to the A -brane which is the pair $(\mathbf{F}_b, \nabla_{\mathcal{E}})$, the Lagrangian submanifold \mathbf{F}_b of $\mathcal{M}_H(G)$, together with the flat unitary line bundle on it corresponding to \mathcal{E} :

$$\mathcal{O}_{\mathcal{E}} \mapsto (\mathbf{F}_b, \nabla_{\mathcal{E}}).$$

Since $\mathcal{O}_{\mathcal{E}}$ is obviously an eigenbrane of the Wilson functors (as we discussed in Section 3), the A -brane $(\mathbf{F}_b, \nabla_{\mathcal{E}})$ should be an eigenbrane of the 't Hooft functors. This may in fact be made into a precise mathematical conjecture, and Kapustin and Witten have verified it explicitly in some cases.

Thus, the A -branes associated to the simplest B -branes turn out to be very nice and simple. This is in sharp contrast with the structure of the corresponding \mathcal{D} -modules, which is notoriously complicated in the non-abelian case. Therefore the formalism of A -branes developed in [KW] has clear advantages. It replaces \mathcal{D} -modules with A -branes that are much easier to “observe experimentally” and to analyze explicitly. One can hope to use this new language in order to gain insights into the structure of the geometric Langlands correspondence. It has already been used in [FW] for understanding what happens in the endoscopic case as explained in the next section.

5.4. Ramification

Up to now we have considered the *unramified* case of the geometric Langlands correspondence, in which the objects on the Galois side of the correspondence are holomorphic ${}^L G$ -bundles on our curve X with a holomorphic connection. These flat bundles give rise to homomorphisms $\pi_1(X) \rightarrow {}^L G$. In the classical Langlands correspondence one looks at more general homomorphisms $\pi_1(X \setminus \{x_1, \dots, x_n\}) \rightarrow {}^L G$. Thus, we look

at holomorphic ${}^L G$ -bundles on X with *meromorphic* connections which have poles at finitely many points of X . The connections with poles of order one (regular singularities) correspond to *tame* ramification in the classical Langlands Program. Those with poles of orders higher than one (irregular singularities) correspond to *wild* ramification.

Mathematically, the ramified geometric Langlands correspondence has been studied in [FGa] and follow-up papers (see [F3] for an exposition), using the affine Kac–Moody algebras of critical level and generalizing the Beilinson–Drinfeld approach [BD] to allow ramification.

S. Gukov and E. Witten [GW1] have explained how to include tame ramification in the S -duality picture. Physicists have a general way of including into a quantum field theory on a manifold M objects supported on submanifolds of M . An example of this is the *surface operators* in 4D super–Yang–Mills theory, supported on two-dimensional submanifolds of the four-manifold M_4 . If we include such an operator, we obtain a certain modification of the theory. Let us again take $M_4 = \Sigma \times X$ and take this submanifold to be of the form $\Sigma \times x, x \in X$. Gukov and Witten show that for a particular class of surface operators the dimensional reduction of the resulting theory is the sigma model on Σ with a different target manifold $\mathcal{M}_H(G, x)$, the moduli space of semi-stable Higgs bundles with regular singularity at $x \in X$. It is again hyper-Kähler, and in the complex structure J it has a different incarnation as the moduli space of semi-stable bundles with a connection having regular singularity (see [S]).

The moduli space $\mathcal{M}_H(G, x)$ has parameters (α, β, γ) , which lie in the (compact) Cartan subalgebra of \mathfrak{g} (see [S]). For generic parameters, this moduli space parametrizes semi-stable triples (E, ϕ, \mathcal{L}) , where E is a (holomorphic) G -bundle, ϕ is a Higgs field which has a pole at x of order one whose residue belongs to the regular semi-simple conjugacy class of $\frac{1}{2}(\beta + i\gamma)$, and \mathcal{L} is a flag in the fiber of E at x which is preserved by this residue (the remaining parameter α determines the flag). Various degenerations of parameters give rise to similar moduli spaces in which the residues of the Higgs fields could take arbitrary values. The moduli spaces $\mathcal{M}_H(G, x)$ and $\mathcal{M}_H({}^L G, x)$ (with matching parameters) are equipped with a pair of mirror dual Hitchin fibrations, and the Mirror Symmetry between them is again realized as fiberwise T -duality (for generic fibers which are again smooth dual tori).

The S -duality of the super–Yang–Mills theories, associated to the dual groups G_c and ${}^L G_c$, with surface operators gives rise to an equivalence of categories of A - and B -branes on $\mathcal{M}_H(G, x)$ and $\mathcal{M}_H({}^L G, x)$. The mirror dual for a generic 0-brane on $\mathcal{M}_H({}^L G, x)$ is the A -brane consisting of a Hitchin fiber and a flat unitary line bundle on it, as in the unramified case.

The analysis of [GW1] leads to many of the same conclusions as those obtained in [FGa] by using representations of affine Kac–Moody algebras and two-dimensional conformal field theory.

Gukov and Witten also considered [GW2] more general surface operators associated to coadjoint orbits in \mathfrak{g} and ${}^L \mathfrak{g}$. The S -duality between these surface operators leads

to some non-trivial and unexpected relations between these orbits. Gukov and Witten present many interesting examples of this in [GW2] drawing connections with earlier work done by mathematicians.

In [Wi3], Witten has generalized the analysis of [GW1, GW2] to the case of wild ramification.

6. MORE GENERAL BRANES

In the previous section we discussed applications of Mirror Symmetry of the dual Hitchin fibrations to the geometric Langlands correspondence. We saw that the A -branes associated to the B -branes supported at the generic flat ${}^L G$ -bundles have very simple description: these are the Hitchin fibers equipped with flat unitary line bundles. But what about the B -branes supported at more general flat ${}^L G$ -bundles? Can we describe explicitly the A -branes dual to them?

This question goes to the heart of the subtle interplay between physics and mathematics of Langlands duality. Trying to answer this question, we will see the limitations of the above analysis and a way for its generalization incorporating more general branes. This will lead us to surprising physical interpretation of deep mathematical concepts such as endoscopy and Arthur’s SL_2 .

The generic flat ${}^L G$ -bundles, for which Mirror Symmetry works so nicely, are the ones that have no automorphisms (apart from those coming from the center of ${}^L G$). They correspond to smooth points of $\mathcal{Y}({}^L G)$ such that the corresponding Hitchin fiber is also smooth. We should consider next the singularities of $\mathcal{Y}({}^L G)$. The simplest of those are the orbifold singularities. We will discuss them, and their connection to endoscopy, in the next subsection, following [FW]. We will then talk about more general singularities corresponding to flat ${}^L G$ -bundles with continuous groups of automorphisms, and what we can learn about the corresponding categories from physics.

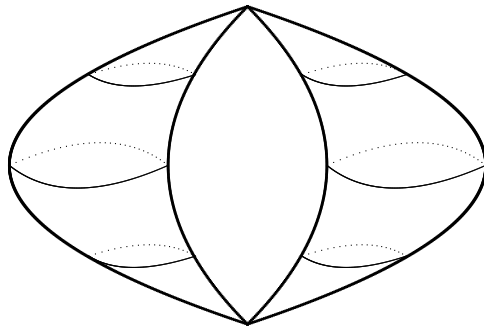
6.1. Geometric Endoscopy

We start with the mildest possible singularities in $\mathcal{Y}({}^L G)$; namely, the orbifold singular points. The corresponding flat ${}^L G$ -bundles are those having finite groups of automorphisms (modulo the center). In the classical Langlands correspondence the analogous Galois representations are called endoscopic. They, and the corresponding automorphic representations, play an important role in the stabilization of the trace formula.

The simplest example, analyzed in [FW] and dubbed “geometric endoscopy”, arises when ${}^L G = PGL_2$, which contains $O_2 = \mathbb{Z}_2 \times \mathbb{C}^\times$ as a subgroup. Suppose that a flat PGL_2 -bundle \mathcal{E} on our curve X is reduced to this subgroup. Then generically it will have the group of automorphisms $\mathbb{Z}_2 = \{1, -1\} \subset \mathbb{C}^\times$, which is the center of O_2 (note that the center of PGL_2 itself is trivial). Therefore the corresponding points of $\mathcal{Y}({}^L G)$ are \mathbb{Z}_2 -orbifold points. This means that the category of B -branes supported at such a

point is equivalent to the category $\text{Rep}(\mathbb{Z}_2)$ of representations of \mathbb{Z}_2 .⁽¹²⁾ Thus, it has two irreducible objects. Therefore we expect that the dual category of A -branes should also have two irreducible objects. In fact, it was shown in [FW] that the dual Hitchin fiber has two irreducible components in this case, and the sought-after A -branes are *fractional branes* supported on these two components.

This was analyzed very explicitly in [FW] in the case when X is an elliptic curve. Here we allow a single point of tame ramification (along the lines of Section 5.4) – this turns out to be better for our purposes than the unramified case. The corresponding Hitchin moduli spaces are two-dimensional. They fiber over the same one-dimensional vector space, and the fibers over all but three points in the base are smooth elliptic curves. The three pairs of dual singular fibers look as follows:



Singular Hitchin fiber in
the A -model, $G = SL_2$.



Singular Hitchin fiber in
the B -model, ${}^L G = SO_3$.

The fiber on the B -model side is a projective line with a double point, corresponding to a flat PGL_2 -bundle that is reduced to the subgroup O_2 . It is a \mathbb{Z}_2 -orbifold point of the moduli space. The dual fiber on the A -model side is the union of two projective lines connected at two points. These two singular points of the fiber are actually smooth points of the ambient moduli space.

There are two irreducible B -branes supported at each of the \mathbb{Z}_2 -orbifold points, corresponding to two irreducible representations of \mathbb{Z}_2 . Let us denote them by \mathcal{B}_+ and \mathcal{B}_- . The corresponding fiber of the Hitchin fibration for SL_2 is the union of two components \mathbf{F}_1 and \mathbf{F}_2 , and accordingly in the dual A -model there are two irreducible A -branes, \mathcal{A}_1 and \mathcal{A}_2 supported on these components (each component is a copy of \mathbb{P}^1 , and therefore the only flat unitary line bundle on it is the trivial one). These A -branes are dual to the B -branes \mathcal{B}_+ and \mathcal{B}_- . Unlike \mathcal{B}_+ and \mathcal{B}_- , they are indistinguishable. An apparent contradiction is explained by the fact that in the equivalence (5.2) of the categories of A -branes and B -branes there is a twist by a \mathbb{Z}_2 -gerbe which is not canonically trivialized (see Section 9 of [FW] and the footnotes on pages 10 and 16). In order to set up an equivalence, we need to pick a trivialization of this gerbe, and this breaks the

⁽¹²⁾The corresponding derived category has a more complicated structure, but we will not discuss it here.

symmetry between \mathcal{A}_1 and \mathcal{A}_2 . We also have a similar picture when X has higher genus (see [FW]).

What happens when we act on \mathcal{B}_+ or \mathcal{B}_- by the Wilson operator $W_x, x \in X$, corresponding to the three-dimensional adjoint representation of ${}^L G = PGL_2$? Since \mathcal{B}_+ and \mathcal{B}_- both have skyscraper support at the same point \mathcal{E} of $\mathcal{Y}({}^L G)$, W_x acts on either of them by tensor product with the three-dimensional vector space \mathcal{E}_x , the fiber of \mathcal{E} at x (in the adjoint representation). However, we should be more precise to keep track of the \mathbb{Z}_2 -action. Recall that the structure group of our flat PGL_2 -bundle \mathcal{E} is reduced to the subgroup $O_2 = \mathbb{Z}_2 \times \mathbb{C}^\times$. Denote by U the defining two-dimensional representation of O_2 . Then $\det U$ is the one-dimensional sign representation induced by the homomorphism $O_2 \rightarrow \mathbb{Z}_2$. The adjoint representation of PGL_2 decomposes into the direct sum

$$(\det U \otimes I) \oplus (U \otimes S)$$

as a representation of $O_2 \times \mathbb{Z}_2$, where \mathbb{Z}_2 is the centralizer of O_2 in PGL_2 (the center of O_2), S is the sign representation of \mathbb{Z}_2 , and I is the trivial representation of \mathbb{Z}_2 . Therefore we have the following decomposition of the corresponding flat vector bundle:

$$(6.1) \quad (\det U_{\mathcal{E}} \otimes I) \oplus (U_{\mathcal{E}} \otimes S),$$

and there is a decomposition $U_{\mathcal{E}}|_x \oplus \det U_{\mathcal{E}}|_x$, where the non-trivial element of \mathbb{Z}_2 acts as -1 on the first summand and as $+1$ on the second summand. So we have

$$(6.2) \quad W_x \cdot \mathcal{B}_{\pm} = (\mathcal{B}_{\mp} \otimes U_{\mathcal{E}}|_x) \oplus (\mathcal{B}_{\pm} \otimes \det U_{\mathcal{E}}|_x).$$

Thus, individual branes \mathcal{B}_+ and \mathcal{B}_- are not eigenbranes of the Wilson operators; only their sum $\mathcal{B}_+ \oplus \mathcal{B}_-$ (corresponding to the regular representation of \mathbb{Z}_2) is.

The mirror dual statement, shown in [FW], is that we have a similar formula for the action of the corresponding 't Hooft operators on the \mathcal{A} -branes \mathcal{A}_1 and \mathcal{A}_2 . Again, only their sum (or union), which gives the entire Hitchin fiber, is an eigenbrane of the 't Hooft operators.

Since an eigenbrane \mathcal{A} decomposes into two irreducible branes \mathcal{A}_1 and \mathcal{A}_2 , the corresponding Hecke eigensheaf \mathcal{F} on Bun_G should also decompose as a direct sum of two \mathcal{D} -modules, \mathcal{F}_1 and \mathcal{F}_2 , corresponding to \mathcal{A}_1 and \mathcal{A}_2 , respectively. Furthermore, these two \mathcal{D} -modules should then separately satisfy an analogue of formula (6.2), which is a natural modification of the standard Hecke property. We called it in [FW] the *fractional Hecke property*, and the \mathcal{D} -modules \mathcal{F}_1 and \mathcal{F}_2 *fractional Hecke eigensheaves*. We have also generalized this notion to other groups in [FW].

Thus, the Mirror Symmetry picture leads us to predict the existence of fractional Hecke eigensheaves for endoscopic flat ${}^L G$ -bundles.⁽¹³⁾ This has non-trivial consequences even for curves over \mathbb{F}_q , some of which have been verified in [FW]. In addition, we obtain a relation between the group $\pi_0(\mathbf{P}_b)$ of components of the generalized Prym variety \mathbf{P}_b ,

⁽¹³⁾In the case of SL_2 (as well as GSp_4) the existence of these \mathcal{D} -modules follows from the work of Lysenko [Ly], but for other groups this is still a conjecture.

which is an open dense part of the singular Hitchin fiber \mathbf{F}_b arising in the A -model, and the group of automorphisms of the endoscopic flat ${}^L G$ -bundles which are the singular points in the dual Hitchin fiber ${}^L \mathbf{F}_b$ of the B -model (this relation was independently observed by B.C. Ngô). Roughly speaking, elements of $\pi_0(\mathbf{P}_b)$ label the components of \mathbf{F}_b , and hence fractional A -branes. The B -branes dual to them correspond to characters of the group of automorphisms of an endoscopic flat ${}^L G$ -bundle, viewed as a point in ${}^L \mathbf{F}_b$. Hence $\pi_0(\mathbf{P}_b)$ should be dual (as an abelian group) to this group of automorphisms.

The upshot of all this is that by analyzing the categories of A -branes supported on the singular Hitchin fibers, we learn many things about the geometric Langlands correspondence (and even the Langlands correspondence for curves over finite fields) which would have been very difficult to see directly using the conventional formalism of \mathcal{D} -modules. This is a good illustration of the power of this new method.

There is a link between our analysis and the classical theory of endoscopy, due to the fact that the geometry we use is similar to that exploited by B.C. Ngô in his recent proof of the fundamental lemma [Ngô]. Ngô has discovered a striking connection between the orbital integrals appearing on the geometric side of the trace formula and the (ℓ -adic) cohomology of the Hitchin fibers in the moduli space $\mathcal{M}_H(G)$, but for curves over \mathbb{F}_q ; more specifically, its decomposition under the action of the group $\pi_0(\mathbf{P}_b)$.⁽¹⁴⁾

However, there are important differences. First of all, we work over \mathbb{C} , whereas Ngô works over \mathbb{F}_q . In the latter setting there is no obvious analogue of the Homological Mirror Symmetry between $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^L G)$. Second, and more importantly, the objects we assign to the connected components of the singular Hitchin fiber \mathbf{F}_b – the A -branes – are objects of automorphic nature; we hope to relate them to Hecke eigen-sheaves and ultimately to the automorphic functions in the classical theory. Thus, these objects should live on the *spectral side* of the trace formula. On the other hand, in Ngô’s work Hitchin fibers appear on the *geometric side* of the trace formula (more precisely, its Lie algebra version), through orbital integrals.

This raises the following question: could there be a more direct link between individual Hitchin fibers in the moduli space $\mathcal{M}_H(G)$ over \mathbb{F}_q and automorphic representations? In other words, could it be that the passage from A -branes to Hecke eigensheaves discussed above has an analogue in the classical theory as a passage from orbital integrals to Hecke eigenfunctions? If so, then the Mirror Symmetry picture would give us valuable insights into the Langlands correspondence.

6.2. S -duality of more general boundary conditions

More general flat ${}^L G$ -bundles have continuous groups of automorphisms. For instance, generic flat bundles reduced to a Cartan subalgebra ${}^L H$ have the group of automorphisms ${}^L H$. Or consider the trivial flat ${}^L G$ -bundle, whose group of automorphisms is ${}^L G$ itself. What are the A -branes corresponding to these flat ${}^L G$ -bundles?

⁽¹⁴⁾More precisely, Ngô considers a generalization of $\mathcal{M}_H(G)$ parametrizing meromorphic Higgs fields with a sufficiently large divisor of poles.

The picture of two dual Hitchin fibrations discussed above is too naive to answer this question. The reason is that even if the flat bundles with continuous groups of automorphisms are semi-stable (which is not necessarily the case), they correspond to points of $\mathcal{Y}({}^L G)$ with singularity so severe that the category of B -branes corresponding to it cannot be described solely in terms of the moduli space $\mathcal{Y}({}^L G)$. In fact, the definition of the sigma model itself is problematic for singular target manifolds.

As an illustration, consider the quotient $\mathbb{C}^n/\mathbb{C}^\times$. The origin has the group of automorphisms \mathbb{C}^\times . What is the category of B -branes associated to this point? Because it is a singular point, there is no obvious answer (unlike the case of smooth points or orbifold points, discussed above). However, we can resolve the singularity by blowing it up. On general ground one can argue that this resolution will not change the category of B -branes. The category of B -branes after the resolution of singularities is the category of coherent sheaves on \mathbb{P}^{n-1} . So the singular point in $\mathbb{C}^n/\mathbb{C}^\times$ has “swallowed” an entire projective space! Likewise, singular points in $\mathcal{Y}({}^L G)$ also have complicated “inner structure” which needs to be uncovered to do justice to the corresponding categories of B -branes.

In order to understand better what is going on we should go back to the four-dimensional gauge theory and look more closely at the S -duality of boundary conditions there. From the physics perspective, this is the “master duality” and everything should follow from it. The Mirror Symmetry of the Hitchin fibrations is but the first approximation to the S -duality when we compactify the theory to two dimensions.

It is instructive to recall how one obtains the Hitchin moduli spaces in the first place: Each of the S -dual gauge theories has a differential Q such that $Q^2 = 0$, and we study the corresponding topological field theories. In the topological theory the path integral localizes on the moduli space of solutions to the “BPS equations”, which read $Q \cdot \Psi = 0$, for all fermionic fields Ψ of our theory (since Q is fermionic and we want the equations on the bosonic degrees of freedom). After that we make dimensional reduction of these equations. This means that we assume that our four-manifold has the form $\Sigma \times X$ and the fields on Σ vary “slowly” along Σ . The corresponding equations have been written in [KW]:

$$(6.3) \quad \begin{aligned} F_A - \phi \wedge \phi &= 0, \\ d_A \phi &= d_A \star \phi = 0, \end{aligned}$$

where d_A is the exterior derivative corresponding to the connection A , and \star is the Hodge star operator. These are precisely the Hitchin equations [Hi1] describing the moduli spaces $\mathcal{M}_H(G)$ or $\mathcal{M}_H({}^L G)$ (depending on which side of S -duality we are on). For example, points of $\mathcal{M}_H({}^L G)$ in the complex structure J are semi-stable flat ${}^L G$ -bundles on X . The flat connection on this bundle is given by the formula $\nabla = A + i\phi$ (the flatness of ∇ is a corollary of (6.3)). This is how the (B -twisted) sigma model on Σ with values in $\mathcal{M}_H({}^L G)$ appears in this story. One obtains the (A -twisted) sigma model with target $\mathcal{M}_H(G)$ in the complex structure I in a similar way.

However, as Kapustin and Witten explain in [KW], this derivation breaks down when we encounter singularities of the Hitchin moduli spaces. Thus, the sigma models with the targets $\mathcal{M}_H({}^L G)$ and $\mathcal{M}_H(G)$ are only approximations to the true physical theory. To understand what happens at the singularities we have to go back to the four-dimensional theory and analyze it more carefully (for more on this, see [Wi4]).

There is also another problem: in the above derivation we have not taken into account all the fields of the super–Yang–Mills theory. In fact, there are additional scalar fields, denoted by σ and $\bar{\sigma}$ in [KW], which we have ignored so far. The field σ is a section of the adjoint bundle \mathfrak{g}_E on X , and $\bar{\sigma}$ is its complex conjugate. On the B -model side, which we have so far approximated by the sigma model with the target $\mathcal{M}_H({}^L G)$, we obtain from the BPS equations that σ is annihilated by the flat connection $\nabla = A + i\phi$, that is, $\nabla \cdot \sigma = 0$. In other words, σ belongs to the Lie algebra of infinitesimal automorphisms of the flat bundle (E, ∇) .

Up to now we have considered generic flat ${}^L G$ -bundles which have no non-trivial infinitesimal automorphisms. For such flat bundles we therefore have $\sigma \equiv 0$, and so we could safely ignore it. But for flat bundles with continuous automorphisms this field starts playing an important role.

The upshot of this discussion is that when we consider most general flat bundles there are new degrees of freedom that have to be taken into account. In order to find a physical interpretation of the geometric Langlands correspondence for such flat bundles we need to consider the S -duality of boundary conditions in the four-dimensional gauge theory with these degrees of freedom included.

A detailed study of S -duality of these boundary conditions has been undertaken by Gaiotto and Witten [GaW1, GaW2]. We will only mention two important aspects of this work.

First of all, Gaiotto and Witten show that in the non-abelian gauge theory the S -dual of the Neumann boundary condition is not the usual Dirichlet boundary condition as one might naively hope, but a more complicated boundary condition in which the field σ has a pole at the boundary. This boundary condition corresponds to a solution of the Nahm equations, which is in turn determined by an embedding of the Lie algebra \mathfrak{sl}_2 into ${}^L \mathfrak{g}$ (see [Wi6]).

This is parallel to the appearance of Arthur’s SL_2 in the classical Langlands correspondence. Arthur has conjectured that the true parameters for (L -packets of) unitary automorphic representations of $G(\mathbb{A})$ are not homomorphisms $\text{Gal}(\bar{F}/F) \rightarrow {}^L G$, but rather $\text{Gal}(\bar{F}/F) \times SL_2 \rightarrow {}^L G$. The homomorphisms whose restriction to the SL_2 factor is trivial should correspond to the so-called tempered representations. (In the case of GL_n all cuspidal unitary representations are tempered, and that is why Arthur’s SL_2 does not appear in the theorem of Drinfeld and Lafforgue quoted in Section 1.) An example of non-tempered unitary representation is the trivial representation of $G(\mathbb{A})$. According to [Art], the corresponding parameter is the homomorphism

$\mathrm{Gal}(\overline{F}/F) \times SL_2 \rightarrow {}^L G$, which is trivial on the first factor and is the principal embedding of the SL_2 factor.

In the geometric Langlands correspondence, Arthur’s SL_2 may be observed in the following way. The analogue of the trivial representation is the constant sheaf \mathbf{C} on Bun_G . It is a Hecke eigensheaf, but the “eigenvalues” are complexes of vector spaces with cohomological grading coming from the Cartan subalgebra of the principal SL_2 in ${}^L G$. For example, consider the case of $G = GL_n$ and let us apply the Hecke operators $H_{\tilde{\omega}_1, x}$ defined by formula (2.2) to the constant sheaf. It follows from the definition that

$$H_{\tilde{\omega}_1, x}(\mathbf{C}) \simeq H^\bullet(\mathbb{P}^{n-1}, \mathbf{C}) \otimes \mathbf{C}.$$

Thus, the eigenvalue is a graded n -dimensional vector space with one-dimensional pieces in cohomological degrees $0, 2, 4, \dots, 2(n-1)$. In the standard normalization of the Hecke operator the cohomological grading is shifted to $-(n-1), -(n-3), \dots, (n-1) -$ as in the grading on the n -dimensional representation of GL_n coming from the principal SL_2 .

The A -brane corresponding to the constant sheaf on Bun_G is the Lagrangian submanifold of $\mathcal{M}_H(G)$ defined by the equation $\phi = 0$ (this is the zero section of the cotangent bundle to the moduli space of semi-stable G -bundles inside $\mathcal{M}_H(G)$). This A -brane corresponds to a Neumann boundary condition in the 4D gauge theory with gauge group G_c . According to Gaiotto and Witten, the dual boundary condition (in the theory with gauge group ${}^L G_c$) is a generalization of the Dirichlet boundary condition, in which the field σ has a pole at the boundary solving the Nahm equations corresponding to the principal SL_2 embedding into ${}^L G$. Thus, we obtain a beautiful interpretation of Arthur’s SL_2 from the point of view of S -duality of boundary conditions in gauge theory. For more on this, see [Wi6, FGU].

Another important feature discovered in [GaW1, GaW2] is that the S -duals of the general Dirichlet boundary conditions involve coupling the 4D super–Yang–Mills to 3D superconformal QFTs at the boundary. This means that there are some additional degrees of freedom that we have to include to describe the geometric Langlands correspondence.

What we learn from all this is that the true moduli spaces arising in the S -duality picture are not the Hitchin moduli spaces $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^L G)$, but some enhanced versions $\widetilde{\mathcal{M}}_H(G)$ and $\widetilde{\mathcal{M}}_H({}^L G)$ thereof, including, in addition to the Higgs bundle (E, ϕ) , an element σ in the Lie algebra of its infinitesimal automorphisms as well as other data. (This will be discussed in more detail in the forthcoming paper [FGU].) Physically, the field σ has non-zero “ghost number” 2. Mathematically, this means that these additional degrees of freedom have cohomological grading 2, and so $\widetilde{\mathcal{M}}_H(G)$ and $\widetilde{\mathcal{M}}_H({}^L G)$ are actually differential graded (DG) stacks. Similar DG stacks have been recently studied in the context of the categorical Langlands correspondence by V. Lafforgue [LafV].

Thus, S -duality of super–Yang–Mills theory offers new insights into the Langlands correspondence and surprising new connections to geometry. Ultimately, we have to

tackle the biggest question of all: what is the underlying reason for the Langlands duality? On the physics side the corresponding question is: why S -duality? In fact, physicists have the following elegant explanation (see [Wi5]): there is a mysterious six-dimensional quantum field theory which gives rise to the four-dimensional super–Yang–Mills upon compactification on an elliptic curve. Roughly speaking (the argument should be modified slightly for non-simply laced groups), this elliptic curve is $E = \mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$, where τ is the coupling constant of this Yang–Mills theory. Since the action of $SL_2(\mathbb{Z})$ on τ does not change the elliptic curve E , we obtain that there are equivalences between the super–Yang–Mills theories whose coupling constants are related by the action of $SL_2(\mathbb{Z})$. This should explain the S -duality, which corresponds to the transformation $\tau \mapsto -1/\tau$, and hence the geometric Langlands correspondence. But that is a topic for a future Séminaire Bourbaki.

REFERENCES

- [Ari] D. ARINKIN – *Moduli of connections with a small parameter on a curve*, Preprint arXiv:math.AG/0409373.
- [Art] J. ARTHUR – *Unipotent automorphic representations: conjectures*, Astérisque **171-172** (1989), 13–71.
- [BD] A. BEILINSON and V. DRINFELD – *Quantization of Hitchin’s integrable system and Hecke eigensheaves*, available at <http://www.math.uchicago.edu/~mitya/langlands/hitchin/BD-hitchin.pdf>
- [BJSV] M. BERSHADSKY, A. JOHANSEN, V. SADOV, and C. VAFA, *Topological reduction of 4D SYM to 2D sigma models*, Nucl. Phys. **B448** (1995), 166–186.
- [C] K. CORLETTE – *Flat G -bundles with canonical metrics*, J. Diff. Geom. **28** (1988), 361–382.
- [D] P. DELIGNE, e.a. – *Quantum fields and strings: A course for mathematicians*, Volumes 1 and 2, AMS, 1999.
- [DP] R. DONAGI and T. PANTEV – *Langlands duality for Hitchin systems*, Preprint arXiv:math.AG/0604617.
- [Dr1] V.G. DRINFELD – *Two-dimensional ℓ -adic representations of the fundamental group of a curve over a finite field and automorphic forms on $GL(2)$* , Amer. J. Math. **105** (1983), 85–114.
- [Dr2] V.G. DRINFELD – *Langlands conjecture for $GL(2)$ over function field*, Proc. of Int. Congress of Math. (Helsinki, 1978), pp. 565–574; *Moduli varieties of F -sheaves*, Funct. Anal. Appl. **21** (1987), 107–122; *The proof of Petersson’s conjecture for $GL(2)$ over a global field of characteristic p* , Funct. Anal. Appl. **22** (1988), 28–43.
- [EW] F. ENGLERT and P. WINDEY – *Quantization condition for ’t Hooft monopoles in compact simple Lie groups*, Phys. Rev. **D14** (1976), 2728–2731.

- [F1] E. FRENKEL – *Recent advances in the Langlands program*, Bull. Amer. Math. Soc. **41** (2004), 151–184.
- [F2] E. FRENKEL – *Lectures on the Langlands program and conformal field theory*, in *Frontiers in number Theory, Physics and Geometry II*, eds. P. Cartier, e.a., pp. 387–536, Springer, 2007, arXiv:hep-th/0512172.
- [F3] E. FRENKEL – *Langlands correspondence for loop groups*, Cambridge Studies in Advanced Mathematics **103**, Cambridge University Press, 2007.
- [FGa] E. FRENKEL and D. GAITSGORY – *Local geometric Langlands correspondence and affine Kac–Moody algebras*, in *Algebraic Geometry and Number Theory*, Progress in Math. **253**, pp. 69–260, Birkhäuser Boston, 2006.
- [FGV] E. FRENKEL, D. GAITSGORY, and K. VILONEN, *On the geometric Langlands conjecture*, Journal of AMS **15** (2001), 367–417.
- [FGu] E. FRENKEL and S. GUKOV, *S-duality of branes and geometric Langlands correspondence*, to appear.
- [FW] E. FRENKEL and E. WITTEN – *Geometric endoscopy and mirror symmetry*, Communications in Number Theory and Physics, **2** (2008), 113–283, arXiv:0710.5939.
- [GaW1] D. GAIOTTO and E. WITTEN – *Supersymmetric boundary conditions in $N = 4$ super Yang-Mills theory*, Preprint arXiv:0804.2902.
- [GaW2] D. GAIOTTO and E. WITTEN – *S-duality of boundary conditions in $N = 4$ super Yang-Mills theory*, Preprint arXiv:0807.3720.
- [Ga] D. GAITSGORY – *On a vanishing conjecture appearing in the geometric Langlands correspondence*, Ann. Math. **160** (2004) 617–682.
- [Ge] S. GELBART – *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. **10** (1984) 177–219.
- [GM] S.I. GELFAND and Yu.I. MANIN – *Homological algebra*, Encyclopedia of Mathematical Sciences **38**, Springer, 1994.
- [GNO] P. GODDARD, J. NUYTS and D. OLIVE – *Gauge theories and magnetic charge*, Nuclear Phys. **B125** (1977) 1–28.
- [GW1] S. GUKOV and E. WITTEN – *Gauge theory, ramification, and the geometric Langlands program*, in *Current developments in mathematics*, 2006, pp. 35–180, Int. Press, 2008, arXiv:hep-th/0612073.
- [GW2] S. GUKOV and E. WITTEN – *Rigid surface operators*, Preprint arXiv:0804.1561.
- [HMS] J.A. HARVEY, G.W. MOORE, and A. STROMINGER – *Reducing S-duality to T-duality*, Phys. Rev. **D52** (1995), 7161–7167.
- [HT] T. HAUSEL and M. THADDEUS – *Mirror symmetry, Langlands duality, and the Hitchin system*, Invent. Math. **153** (2003), 197–229.
- [Hi1] N. HITCHIN – *The self-duality equations on a Riemann surface*, Proc. London Math. Soc. **55** (1987), 59–126.

- [Hi2] N. HITCHIN – *Stable bundles and integrable systems*, Duke Math. J. **54** (1987), 91–114.
- [Hi3] N. HITCHIN – *Langlands duality and G_2 spectral curves*, Quat. J. Math. **58** (2007), 319–344.
- [Ka] A. KAPUSTIN – *A note on quantum geometric Langlands duality, gauge theory, and quantization of the moduli space of flat connections*, Preprint arXiv:0811.3264.
- [KW] A. KAPUSTIN and E. WITTEN – *Electric-magnetic duality and the geometric Langlands program*, Communications in Number Theory and Physics **1** (2007), 1-236, arXiv:hep-th/0604151.
- [KS] M. KASHIWARA and P. SCHAPIRA – *Sheaves on manifolds*, Springer, 1990.
- [LafL] L. LAFFORGUE – *Chtoucas de Drinfeld et correspondance de Langlands*, Invent. Math. **147** (2002), 1–241.
- [LafV] V. LAFFORGUE – *Quelques calculs reliés à la correspondance de Langlands géométrique pour \mathbb{P}^1* , available at <http://people.math.jussieu.fr/~vlafforg/geom.pdf>
- [Lau1] G. LAUMON – *Correspondance de Langlands géométrique pour les corps de fonctions*, Duke Math. J. **54** (1987) 309-359.
- [Lau2] G. LAUMON – *Transformation de Fourier généralisée*, Preprint arXiv:alg-geom/9603004.
- [L] R.P. LANGLANDS – *Problems in the theory of automorphic forms*, in Lect. Notes in Math. **170**, pp. 18–61, Springer Verlag, 1970.
- [Ly] S. LYSENKO – *Geometric Waldspurger periods*, Compos. Math. **144** (2008) 377–438; *Geometric theta-lifting for the dual pair SO_{2m}, Sp_{2n}* , Preprint arXiv:math/0701170; *Geometric theta-lifting for the dual pair GSp_{2n}, GO_{2m}* , Preprint arXiv:0802.0457.
- [MV] I. MIRKOVIĆ and K. VILONEN – *Geometric Langlands duality and representations of algebraic groups over commutative rings*, Ann. of Math. **166** (2007) 95–143.
- [MO] C. MONTONEN and D.I. OLIVE – *Magnetic monopoles as gauge particles?* Phys. Lett. **B 72** (1977), 117–120.
- [Nad] D. NADLER – *Microlocal branes are constructible sheaves*, Preprint arXiv:math/0612399.
- [NZ] D. NADLER and E. ZASLOW – *Constructible sheaves and the Fukaya category*, Preprint arXiv:math/0604379.
- [Ngô] B.C. NGÔ, *Le Lemme Fondamental pour les algèbres de Lie*, Preprint arXiv:0801.0446.
- [O] H. OSBORN, *Topological charges for $N = 4$ supersymmetric gauge theories and monopoles of Spin 1*, Phys. Lett. **B 83** (1979), 321–326.
- [R] M. ROTHSTEIN – *Connections on the total Picard sheaf and the KP hierarchy*, Acta Applicandae Mathematicae **42** (1996), 297–308.

- [S1] C. SIMPSON – *Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization*, Journal of AMS **1** (1988), 867–918.
- [S] C. SIMPSON – *Harmonic bundles on noncompact curves*, Journal of AMS **3** (1990) 713–770.
- [SYZ] A. STROMINGER, S.-T. YAU and E. ZASLOW – *Mirror symmetry is T-duality*, Nucl. Phys. **B479** (1996), 243–259.
- [VW] C. VAFA and E. WITTEN – *A strong coupling test of S-duality*, Nucl. Phys. **B431** (1994), 3–77.
- [We] A. WEIL – *A 1940 letter of André Weil on analogy in mathematics*, Translated by Martin H. Krieger, Notices of AMS **52** (2005), 334–341.
- [Wi1] E. WITTEN – *Topological quantum field theory*, Comm. Math. Phys. **117** (1988), 353–386.
- [Wi2] E. WITTEN – *Mirror manifolds and topological field theory*, in Essays on Mirror manifolds, Ed. S.-T. Yau, pp. 120–158, International Press, 1992.
- [Wi3] E. WITTEN – *Gauge theory and wild ramification*, Anal. Appl. (Singap.) **6** (2008), 429–501.
- [Wi4] E. WITTEN – *Mirror symmetry, Hitchin’s equations, and Langlands duality*, Preprint arXiv:0802.0999.
- [Wi5] E. WITTEN – *Geometric Langlands from six dimensions*, Preprint arXiv:0905.2720.
- [Wi6] E. WITTEN – *Geometric Langlands and the equations of Nahm and Bogomolny*, to appear.

Edward FRENKEL

Department of Mathematics
 University of California
 Berkeley, CA 94720-3840 – U.S.A.
E-mail : frenkel@math.berkeley.edu