FUNCTORIALITY AND RECIPROCITY

Two lectures at the Institute for Advanced Study, March, 2011[1,2]

Robert P. Langlands

These are, for me, the two major issues in the theory of automorphic forms and related topics. Both appear in three guises, each of which has two forms, global and local. Thus there are, in total, six theories. The global forms arise over global fields F, of which there are three types: (1) fields of algebraic numbers of finite degree over $\mathbb{Q}$; (2) function fields of nonsingular curves over finite fields; (3) function fields of nonsingular curves over the complex field. The corresponding local fields F_v are well known: (1) the real field, the complex fields, and p-adic fields (2) fields of formal Laurent series over finite fields; (3) the field of formal Laurent series over $\mathbb{C}$.

My main concern in these lectures is a description of proposed strategies — in so far as they are available — for establishing the generally accepted conjectures. For functoriality the strategy has some promise but a number of central issues are still elusive. For reciprocity — also referred to as the correspondence — as described here, the difficulties are considerably more daunting and, so far as I know, no convincing strategy has been proposed. Indeed, it is seldom, if ever, explained what is wanted. Functoriality was introduced in my 1967 letter to Weil, although without the name, and matured rather rapidly in the following few years. The correspondence was proposed somewhat later, very diffidently in [Pr] but more boldly after I began in 1970 to study Shimura varieties, although never with any serious attempt at precision. The precision ultimately desired would still be premature, but I do want to draw attention here to some of the issues. They are often overlooked.

An additional topic is the possibility of realizing the geometric correspondence over $\mathbb{C}$, thus the third guise of the correspondence, in the context of mirror symmetry and gauge theory. This is, for me, largely, but not entirely, a matter of idle curiosity, but I do want to discuss it briefly, principally for my own sake, but also because it has become the most popular, or at least most widely known, form of the correspondence. An important question is whether whatever insights the field-theory context offers are pertinent for the basic mathematical problem: the formulation of the correct statements of functoriality and reciprocity in all six guises and the construction of their proofs.

[1] These notes are unfinished, extremely and deliberately informal, and may contain errors and premature surmises. I see no good reason to be coy about the possibilities envisaged or to conceal my shallow knowledge of various fundamental contemporary mathematical theories. I am too old for that kind of vanity. Since the purpose of these lectures and these notes is to orient myself — and anyone else prepared to take me at all seriously — with respect to current developments and prospects, any comments and criticisms would be examined with attention. The fate of the notes will likely be to change constantly and never to assume a final form.

[2] March 8, 2011

1. Functoriality.

We can describe the objects in these six possible cases readily, although it is best not to try for too much precision at first.

Global

1.a) Automorphic representations over F often restricted in some way or another by their spectral properties in the sense of the theory of Eisenstein series. For the present purposes we always take automorphic representations to be irreducible.

2.a) The description is the same as for (1.a) but the field F is of a different type.

3.a) The basic objects are Hecke eigensheaves on smooth curves over $\mathbb{C}$. Since we want to consider a complete theory in which singularities, regular and irregular, in the associated holonomic systems are allowed, details are postponed.

Local

1.b) Irreducible complex representations of $G(F_v)$ perhaps restricted in some way, again by their spectral properties. It can be demanded that they are tempered. For other purposes, the class introduced by Arthur, to which I refer here as the Arthur class, is more appropriate.

2.b) The same objects as in (1.b) but with F_v given by a formal power-series field.

3.b) The local form of (3.a) but that too will be described below.

The correspondence is between these six "theories" taken on one side, say the right side, and another six "objects" or theories on the left. The objects on the right are all associated to a reductive group G over F or over F_v. Functoriality in the L-group, which for reasons to be described below it is best to regard as a possibility that with sufficient effort and imagination can be realized, although certainly not from one day to the next, for all six theories, allows the introduction of categories which, for lack of a better term, I will refer to as mock Tannaka categories. Informally — and these lectures are certainly intended to be informal — this is a category of vector spaces over a field of characteristic 0 that can be formulated as the category of representations of some "group", the word being placed in quotation marks because this group can become, if one is not careful, a very large inverse limit of reductive algebraic groups. In fact, this large inverse limit can for all practical purposes be replaced by one — chosen according to the problems under consideration — of the collection of finite-dimensional groups over which the limit is taken.

Automorphic represenions are representations of $G(\mathbb{A}_F)$, where G is a reductive group over the global field F or the local field F_v. In all six cases, there is an L-group ${}^LG$ attached to G. As I implicitly suggested in the paper [ST], in each of the six cases functoriality would allow us to isolate the elements of hadronic type, a notion introduced there and in the lecture [SP], where they are referred to as thick. The explanations of these two papers are not meant to be more than intimations of a possibility, but they allow us to isolate the generating elements of the relevant mock Tannakian categories, stable classes of pairs $({}^LH, \pi_H^{\mathrm{st}})$, where π_H^{st} is what has come to be called an L-packet, a notion introduced many years ago and required by the distinction between *conjugacy class* and *stable conjugacy class*. The generating elements are those for which π_H^{st} is hadronic. All others, say $({}^LG, \pi_G^{\mathrm{st}})$

are obtained one of hadronic type by the functorial transfer associated to a homomorphism $\phi : {}^LH \to {}^LG$. Even after functoriality has been established, there will be delicate — and interesting — issues remaining. There is no reason to believe that ϕ is uniquely determined, although there is reason to hope that H is, although in conformity with the principles of the stable trace formula only if it is taken to be quasi-split. Moreover there are fastidious questions connected with the center of H that it would be premature to attempt to resolve. There are also questions of multiplicity. See the papers [LP] and [WS]. So long as we are concerned with only a finite collection of automorphic representations or representations of local groups $G(F_v)$, we can account for all possibilities with a fixed H and a hadronic π_H^{st}, provided we allow all homomorphisms of ${}^LH$ to other L-groups. Questions of fusions along the centers aside, the group, sometimes denoted $\mathcal{G}_L$, defining the mock Tannaka category will be in each case the inverse limit over finite sets S of hadronic pairs $({}^LH, \pi_H^{\mathrm{st}})$ of $\mathcal{G}_L^S = \prod_S {}^LH$.

It may seem premature to insist on this description, but it is not unimportant when reflecting on what the correspondence is to be. It is important for this to recall that categories of mock Tannakian type are associated to a field. For the six theories on the right this is the field of complex numbers.

2. Reciprocity

The objects on the left will have, in principle, a natural functorial structure. For those of type (1.b) and (2.b) it is the category of finite-dimensional complex representations of an appropriately thickened Weil group. This too has a natural Tannakian structure over $\mathbb{C}$. Over $F_v = \mathbb{R}, \mathbb{C}$, the thickened group is the Weil group W_{F_v} itself. Over a nonarchimedean the thickened group is $W_{F_v} \times SL(2, \mathbb{C})$. I would be quite happy to refer to this group as the Weil-Deligne group, but unfortunately that name was given to an object that does not suit the purpose. For the class of representations introduced by Arthur one takes the thickened Weil group multiplied by a second $SL(2, \mathbb{C})$. Call it the Weil-Arthur group.

So there is no objection in principle to having a motivic correspondence or reciprocity between the theories on the right and those on the left in cases (1.b) and (2.b). At least the field of the motives are the same on both sides. There is also of course considerable evidence for the existence of the correspondence, complete for $\mathbb{R}$ and $\mathbb{C}$, and partial, but strong, for local nonarchimedean fields, whether of characteristic 0 or of positive characteristic.

On the other hand, for cases (1.a) and (2.a), thus in the context of global fields, the first possibility for the left side that presents itself is, or so it seems to me, the category of motives over the field $\mathbb{Q}$, thus motives whose associated field in the sense of Grothendieck is $\mathbb{Q}$ but which are deduced from algebraic varieties over the field F, which I stress is either an algebraic number field finite over $\mathbb{Q}$ or the function field of a nonsingular complete curve over a finite field. To have the correspondence envisaged with the theories on the right, the theory on the left has to be tensored with $\mathbb{C}$. This is perfectly reasonable, but it conflicts with some preconceived ideas. Presumably the ℓ-adic representations are a realization of these two motives but only after they are tensored with $\mathbb{Q}_\ell$. So we have a diagram, not extremely precise,

$$(1) \qquad \ell\text{-adic representations} \overbrace{\longleftarrow}^{\otimes \mathbb{Q}_\ell} \text{motives}/F \overbrace{\longrightarrow}^{\otimes \mathbb{C}} \text{automorphic representations},$$

in which there is no suggestion of a direct relation between the extreme elements.

Although I have observed it before, but perhaps only in personal correspondence, I repeat here that this suggests that the theory of motives for varieties over number fields or over finite fields will be developed in tandem with the correspondence. It suggests in addition that the theory of motives for varieties over $\mathbb{C}$ will be a consequence of the theory for varieties over number fields.

More to the point at the moment, the information available at the moment raises the question whether the category of motives over F is the appropriate category. The diagram of formula (1) could be perfectly correct but inadequate as it stands since it may be valid for a larger category than that of motives. This is suggested by the theory of Shimura varieties.

There is a consequence of diagram 1 that is usually emphasized more than the diagram itself. Although the two extreme categories in diagram 1 are in principle incomparable, the L-functions defined by an object on the left might be equal to the L-function defined by an object on the right. This would mean that the monic polynomials in q_v^{-s} that appear as the denominators of the local factors would have to have coefficients from the same field, or from two different fields, one of which can be imbedded in another, but only with an imbedding that is fundamentally arbitrary. In the context of Shimura varieties, we can arrange that the field on the left is reduced to $\mathbb{Q}$ because it has been established that conjugates of Shimura varieties are again Shimura varieties. So we can presumably take sums of the L-functions or "motives" over conjugates of the field of definition of the Shimura variety in order to arrange that we are dealing with a mock Tannakian category associated to the field $\mathbb{Q}$. On the other hand, the L-functions considered in the theory of Shimura varieties are associated to a representation of the fundamental group over a number field. So they are defined by sheaves of vector bundles with a flat connection. These vector bundles with connection are normally defined over a number field and, thanks to the device indicated, could even be defined over $\mathbb{Q}$. So what we have are perverse sheaves of geometric origin — a notion apparently introduced in [BBD], but which I have still to learn. I suppose, or I would hope, that it is possible, at first formally but ultimately precisely, to attach to this larger group of objects — perverse sheaves of geometric origin over varieties, everything being defined over a given algebraic number field — a theory of motives with fiber functors and associated Tannkakian categories. They are what should appear on the left of (1) for the correspondence in cases (1.a) and (2.a), although it would alse be legitimate to restrict attention to motives defined by algebraic varieties over F. It would be useful to have a clearer notion of what the theory, for motives in the narrower sense or for motives in the larger sense, entails, for example, some form of the Hodge conjecture or, for the correspondence over function fields of curves over finite fields, a form of the Tate conjecture.

Although the local reciprocity for types (1) and (2) is first formulated principally for unitary representations on both sides, the nature of algebro-geometric zeta-functions is such that global reciprocity will occur in a wider context. I also underline here before continuing with functoriality that neither the global correspondence nor the local correspondence is simply a question of comparing L-functions.

3. Establishing functoriality.

This is a serious problem in all of the six cases. The local problem is of a different nature than the global problem and might be considered less difficult were it not that local abelian class field theory had not been first established simultaneously with the global theory or as a consequence of it. Over the real or complex numbers, local functoriality can be established as a consequence of Harish-Chandra's analytic theory, but that is neither easy nor generally familiar to contemporary mathematicians. It may however be considered as a purely analytic problem in the spectral theory of differential equations.

The local theory over $\mathbb{R}$ or over $\mathbb{C}$ is a spectral theory and reposes on our understanding of the characters of irreducible representations as conjugation-invariant functions that satisfy — as distributions — a differential equation whose formal solutions we know. Such an understanding is missing over nonarchimedean fields. It has to be found. Leave this central problem aside for the moment and consider the proposals, described in [BE] and [FLN], for attacking global functoriality with the help of the stable trace formula.

The strategy in the papers [BE,FLN,ST] can be briefly recalled. The L-functions $L(s,\pi^{\mathrm{st}},\rho)=\prod_v L_v(s,\pi^{\mathrm{st}},\rho)$ are attached to a stable classe π^{st} of automorphic representations of the group G and an algebraic representation ρ of the complex algebraic group ${}^LG$. If S is a finite set of places, usually but not necessarily taken to include all archimedean places, we can also form the incomplete L-functions $L_S(s,\pi^{\mathrm{st}},\rho)=\prod_{v\notin S} L_v(s,\pi^{\mathrm{st}},\rho)$. Choose S and for v in S fix a smooth compactly supported function f_v^G on $G(F_v)$. In the three papers studied, the analysis of the behavior of

$$\sum_{\pi_G^{\mathrm{st}}} m_{\pi_G^{\mathrm{st}}} L_S(s,\pi_G^{\mathrm{st}},\rho)\prod_{v\in S}\operatorname{tr}\pi_{G_v}^{\mathrm{st}}(f^{G_v}) \tag{2}$$

was begun. The stable multiplicity $m_{\pi_G^{\mathrm{st}}}$ is a notion that is not yet available in general. It belongs to the theory of endoscopy and first appeared in [LL] for the group $SL(2)$. What is pertinent here is the behavior of (2) on an interval $[1,1+\epsilon)$, $\epsilon>0$. As discussed in the papers this is only possible when the representations not of Ramanujan type are eliminated from the sum. This is not an easy matter and, as anyone who has examined the papers can attest, has hardly been broached. As indicated in [ST], there is, at best, a very great deal of analysis yet to be undertaken, not only to eliminate the representations not of Ramanujan type from (2), but also to get a handle on the asymptotic behavior as $s\searrow 1$. Ali Altuğ has made some progress with the simplest case, that of $SL(2)$. It is by no means easy and a complete treatment of this group would be a major accomplishment.

The techniques introduced, but not sufficiently developed, in the three papers mentioned are intended to lead to an expression of the following form for the asymptotic behavior of the expression (2) at $s = 1$.

$$\text{(3)} \qquad \sum_{\phi\colon {}^LH\to{}^LG} \zeta_S(s,F)^{m_1(\rho\cdot\phi)}\{{\sum_{\pi_H^{\mathrm{st}}}}' m_{\pi_H^{\mathrm{st}}} m_2(s,\rho\cdot\phi)\prod_{v\in S}\operatorname{tr}\pi_{H_v}^{\mathrm{st}}(f^{H_v})\}.$$

The function $\zeta_S(s,F)$ is the incomplete zeta-function with respect to S of the field F. The factor $m_1(\rho\cdot\phi)$ is the multiplicity of the trivial representation in $\rho\cdot\phi$, a representation of the group ${}^LH$. If σ is the quotient of $\rho\cdot\phi$ by the trivial m_1-dimensionsal subrepresentation, then

$$m_2(s,\rho\cdot\phi) = L_S(s,\pi_H^{\mathrm{st}},\sigma),$$

a function that will have to be shown inductively to be holomorphic and nonzero at $s = 1$. Although it is more than a little risky, I have removed the factors $m_{\pi_G^{\mathrm{st}}}$ that appear in (2) in the expectation, to some extent justified by the results of [WS] that they will reappear as a result of the sum over ϕ in (3). The sum over ϕ is, roughly speaking, a sum over inequivalent imbeddings of L-groups ${}^LH$ in the given L-group G, but there are subtleties with centers that it would be premature to discuss here. There is also a related question of telescoping, with iterated imbeddings

$${}^LH_1 \to {}^LH_2 \to {}^LG,$$

that it would, once again, be premature to discuss. It is to some extent removed by the prime on the internal sum in (3) which expresses the condition that the sum is over hadronic π_H^{st}. It is the information in the poles of (3) that is useful for our purposes. The representations π_H^{st} hadronic for ${}^LH = {}^LG$ contribute — presumably — nothing to the principal part of the pole of (2) at $s = 1$.

In a nutshell, the plan is to use (3) to attach, in part thanks to the freedom in the choice of S and the functions f_v^G, $v \in S$, to each stable class of automorphic representations of $G(\mathbb{A}_F)$ dimension data in the sense of [LP]. From this dimension data we can, thanks to their theorems, construct the group ${}^LH_{\pi^{\mathrm{st}}}$

4. Stable transfer and the local theory.

Even if it is possible over the course of the years to carry out this analysis, it will most likely not be adequate for establishing functoriality. There is one additional difficulty in the analysis that is, I believe, very serious. It is the transfer $f^G \to f^H$, a local issue over the fields F_v, so that we should write $f^G = f_v^G$, and the central difficulty that we left aside at the beginning of the previous section. The notion of transfer was introduced in [TF] for $G = SL(2)$ and for H a torus of dimension one. It was not discussed for $H = \{1\}$, when ϕ can be a nontrivial imbedding of $\mathrm{Gal}(K/F)$, $[K : F] < \infty$, in ${}^LG = PGL(2)$. Then f^H is a number, thus a function on $\{1\}$. The number is equal to $\operatorname{tr}\pi_G^{\mathrm{st}}(f^G)$ if π_G^{st} is the L-packet associated to ϕ. So it is given by the character, and characters over nonarchimedean local fields are notoriously elusive.

Although this transfer is very similar to endoscopic transfer it is something quite different. As with endoscopic transfer $f_v^H = f^H$ will not be determined uniquely by $f_v^G = f^G$. Only its stable orbital integrals are determined by the relation. In contrast to endoscopy, they are determined by the stable orbital integrals of f^G, not by its unstable orbital integrals. As explained in [ST] for the simple example of $SL(2)$, they are determined by the stable characters. These can be regarded as functions on the Steinberg-Hitchin base of [FLN]. As observed in the previous section the characters are eigenfunctions of differential operators. In other words they are determined by a $\mathcal{D}$-module. Of course they are a little more and for several reasons. There are boundary conditions and jump conditions to take into account. Moreover, the underlying field may be real or complex. It is perhaps not too gross an oversimplification to state that Harish-Chandra's theory of characters is a part of the theory of $\mathcal{D}$-modules.

I do not believe that I am the first to think that the local theory of characters and transfer might be developed in a similar framework, but as $\mathcal{D}$-modules are not available, it would have to be in the context of perverse sheaves. The two notions would seem to be equivalent over $\mathbb{C}$. At the moment, I am certainly in no position to reflect on such matters in a serious way. What the papers [Wa, N] suggest to me is that if p is the residual characteristic of the pertinent local field then the pertinent perverse sheaves will be sheaves on algebraic varieties over finite fields of characteristic p. Although it is more than a little frivolous to suggest anything at this point, the next step for me being rather to acquire some serious understanding of the basic notions of perverse sheaves and a much better grasp of Waldspurger's reduction of the fundamental lemma to the fundamental lemma for Lie algebras over fields $\kappa((t))$ of Laurent series, not to speak of greater familiarity with the current understanding of the asymptotic behavior of orbital orbital integrals and of characters over nonarchimedean local fields, there is an obvious point to make. As I recall, Waldspurger reduction functions in layers $t^{n-1}\mathfrak{g}[[t]]/t^n\mathfrak{g}[[t]]$, $n = 1, 2, 3, \ldots$. So at each stage the weight of a point is q^{-n}. On the other hand, perverse sheaves are sums of sheaves, each associated to a variety of some dimension.

The characters and the orbital integrals are functions on the Steinberg-Hitchin base $\mathfrak{A}$, which for $SL(2)$ is the line over the local field F_v. After multiplication by an appropriate normalizing factor they become, in accordance with Shalika's theorem, sums $a + b|\Delta|^{1/2}$, where a and b are constants. A constant is pretty much taken care of as the direct image of the constant sheaf under $G \mapsto \mathfrak{A}$. What about $|\Delta|^{1/2}$. However, if I am not mistaken, the perverse sheaves function at the levels $b + \mathcal{O}_v\varpi^n$, $n = 0, 1, 2, 3, \ldots$, $a \in \mathfrak{A}$. If $b = 2 + x$ and the residual characteristic is not 2, then $b/2 \pm \sqrt{b^2 - 4}/2 = 1 \pm \sqrt{x} + o(\sqrt{x}$ and $|\Delta(b)|^{1/2} = |x|^{1/2}$. Somehow the second term in the Shalika expansion at b seems to be picked up at the level n, where $|\sqrt{x}| = |\varpi|^{n/2}$, and from a sheaf concentrated at a point. It is a problem on which to reflect.

For any group transfer from $H = \{1\}$ to an arbitrary G has a special significance. I hope and, to some extent, expect that the variant on the study of Kummer extensions, local and global, described in the last section of [ST] will be pertinent here. In some more general form, it may be necessary for higher dimensional H as well.

Certainly for the purpose of constructing a general arithmetic theory of automorphic representations the most urgent matter at present is, in my view, local transfer. For those who have persuaded themselves that the use of the Steinberg-Hitchin base and the stable transfer to pass from the trace formula to the Poisson represents, no matter how artificial or unlikely it may appear, the first genuinely promising possibility — outside of endoscopy, whose limits are severe and clear to those who understand it — for establishing functoriality, the next most urgent, although perhaps more accessible, problem is to find methods for studying the asymptotic behavior of (2).

5. The geometric theory for curves over the complex numbers.

This theory has a different appeal than the arithmetic theories of types (1) and (2). It can be presented in a form that is strongly categorical and cohomological, but it is the connection with the theory of ordinary differential equations with irregular singular points as developed by Poincaré and G. D. Birkhoff and presented in [CL] that strikes my fancy. There are still few concrete examples, but those available are very appealing[G,FG]. The first question is whether the methods introduced in the past few years for tackling the issue of functoriality for the types (1) and (2) has any chance of succeeding for type (3).

For type (3) the left side of the correspondence, whether local or global, has a straightforward analytical description. On the left side the global object is an ordinary differential equation on a complete nonsingular curve X over $\mathbb{C}$. The equation may have a finite number of singular points, regular or singular, but it describes movement in a ${}^LG$-bundle on X. Locally there is no ambiguity about where the singularity — when present — is located. Globally the location is arbitrary. There are ambiguities that I have not resolved, neither for type (3) nor even for types (1) and (2), and neither locally nor globally. For types (1) and (2), however, many of them are resolved by the stable trace formula and endoscopy, from which the possibility of referring all groups to quasi-split groups is deduced. There seems here to be either some uncertainty about the boundaries between the form of the group G to be taken — how much twisting is permitted — and the twisting in the ${}^LG$-bundle. To be blunt, to straighten matters out for myself, I shall have to review the relevant cohomological results over all three forms of the global fields in question. For the moment we can take G to be a Chevalley form of the group over the function field F of the curve or, at worse, a quasi-split form defined over a finite extension.

A more important matter is whether there is a trace formula available in the context of type (3) and whether it can in any sense lead to the passage to a formula of Poisson type as in [FLN,ST], and if it does whether this can be used to establish functoriality or reciprocity. For all three types, (1), (2) and (3), both local and global, the question of recognizing the trivial — or almost trivial, especially the cohomology in dimension $2n$ of an n-dimensional variety — representation in the mock Tannakian category on the left arises. For types (1.b) and (2.b), this is just a question of recognizing the trivial representation of the Weil group, so that it demands no explicit attention. For types (1.a) and (2.a) the Hodge conjecture, or even some extension of the Hodge conjecture, raises its forbidding head. For (3.a) the question does not seem so formidable, except perhaps for questions of direct sum decompositions. The trivial object is the trivial flat ${}^LG$ bundle over X. The local question, thus the question for (3.b), is just as easy to answer. We are again dealing with the trivial bundle. On the other hand, on the right side, the corresponding object is in cases (1.b) and (2.b) easy to construct as a well-known representation of the principal series. In the same way, the objects on the right for (1.a) and (2.a) are autmorphic representations whose existence is guaranteed by the theory of Eisenstein series. For (3.a) and (3.b), especially (3.a) the corresponding object seems to be a cause of considerable puzzlement, cf. §6.1 of [CFT].

For types (1.a) and (2.a) the triviality on the left but more importantly for analytic purposes on the right is recognized by the zeta-functions, thus by the order of the poles of $L(s, \pi^{\mathrm{st}}, \rho)$ as in formula (3) of §3. It is not clear what is to be substituted for these L-functions for type (3.a). The absence of a way to identify — to define if one prefers — even if only in a hypothetical way, the multiplicity of the trivial object in whatever functoriality attaches as image to the pair formed by an object for (3.a) and a representation ρ of ${}^LG$ is a major defect of our reflections so far. Even for the local version (3.b) it is not so clear what to do.

There is a good deal written about the right side of the correspondence of type (3), some useful, indeed indispensable, for our present purposes, some which is certainly relevant but not always central. The basic information can be gleaned from [CFT,CLG]. It is clear enough when there is no ramification — thus no singularities anywhere. The difficulty is to keep track of the various spaces, the structure on none of which is very clear to me: on the left side, the moduli space of ${}^LG$ bundles on X with a flat connection (and thus, in some sense with a perverse sheaf); on the right side (perverse) sheaves and therefore $\mathcal{D}$-modules on the the moduli space of G-bundles on X. It is all just a little too much for an old-fashioned mathematician. Certainly functoriality in ${}^LG$ on the left is as clear for type (3), both local and global, as it was for the first two types. It is the functoriality on the right that is doubtful.

If we are persuaded by the theory suggested above for types (1) and (2) there are several questions to address when treating the right side for type (3). Some have been formulated above, some not.

(i) What are the local objects? Are they representations of some kind of some group?

(ii) What are the global objects?

(iii) How do we recognize when a local or, more seriously, a global object contains the trivial object and how often?

(iv) What is the analogue of formula (3)? In particular, what takes the place of the trace formula?

(v) If answering (iv) requires some form of the trace formula, what is it? Is there a pertinent local representation theory?

It is clear from these questions that I am not prepared to accept as the ultimate statements much that is currently offered, as in [CLG,CFT]. On the other hand, these are still among my basic sources for the geometric theory and they contain a great many interesting suggestions. The global objects are certainly the Hecke eigensheaves of §6.1 of [CFT], although we have to be prepared to accept singularities at finitely many points of X.

For the group $GL(1)$, Serre's book [S] is available as a reference, to which I shall return. It is also possible to consider a general local theory. Such a theory may even exist, but, so far as I can see, there is no allusion to it in [CFT,CLG]. There is, nevertheless, a possibility that suggests itself it. Take F to be the field of formal Laurent series over $\mathbb{C}$

$$\sum_{n\geq M} a_n t^n,$$

where M is any integer, positive or negative. Then $\mathcal{O}$ can be the ring of formal Taylor series over $\mathbb{C}$. Let $\mathcal{P}$ be its maximal ideal. If $n \geq 0$ let $G_n(\mathcal{O})$ be the group of elements $g \in G(\mathcal{O})$ such that $g \equiv I \pmod{\mathcal{P}^n}$, a notation that implies G is defined over $\mathcal{O}$, a notion that we can accept without demanding any great precision. Thus $G(\mathcal{O}) = G_0(\mathcal{O})$. We can also expect that their will be a theory of elementary divisors or a Hecke theory so that $G(\mathcal{O})\backslash G(F)/G(\mathcal{O})$ is finite. We can hope in addition that the space $G(\mathcal{O})/\mathcal{G}_n(\mathcal{O}) = G_0(\mathcal{O})/G_n(\mathcal{O})$ — I hope this notation does not conflict with any standard notations — is the disjoint union of algebraic variety of finite dimension, presumably even smooth, and that it carries an essentially unique invariant Haar measure. In any case this measure $d\mu$ and any smooth compactly supported $G_n(\mathcal{O})$ bi-invariant function f on $\mathcal{H}_n$ defines an object $F_f = f d\mu$. As n grows the set of these objects grows and we can define their sums and products, forming a ring $H = \cup H_n$. We would like to introduce complex representations π of these rings, often and, indeed, almost always infinite-dimensional but with the property that $\pi(H_f)$, $f \in H_n$ was always of trace class, even with finite-dimensional range. More precisely, π would be a sequence of compatible representations π_n on spaces $\mathfrak{H}_1 \subset \mathfrak{H}_2 \subset \mathfrak{H}_3 \ldots$, π_n being a representations of H_n.

The first question is whether one can find an adequate number of π for which $\pi(F_f)$ is always of trace class. The second question is then whether the linear form given by the trace is a distribution on the smooth compactly supported functions. I suppose we would have to take a smooth compactly supported g on $\mathcal{G}_n$, integrate it over $G_n(\mathcal{O})$ to obtain f and F_f and then treat the trace of $\pi_n(F_f)$ as a linear form in the variable g. The question is whether this linear form is a distribution. It seems to me that there will be an adequate number of admissible π, but I haven't thought much about it.

For $G = GL(1)$, the group $G(F)/G(\mathcal{O})$ is just $\mathbb{Z}$ and the group $G_0(\mathcal{O})/G_n(\mathcal{O})$ a finite dimensional abelian Lie group. It is very likely that a theory along the above lines is available in the literature. It seems very likely that it can be constructed for general G. Up until now, I have avoided, for clear reasons, all but reductive Lie groups.

The global question for type (3) entails the same kind of consideration as for sum (3). For types (1) and (2), the Eisenstein series depend on a continuous parameter, but the motives over F do not. For this and for other reasons, namely the existence of Maaß forms for type (1), for these two types of theories the image of the right-hand arrow in (1) is relatively small. None the less, a more careful explanation of the trace formula would have allowed for the appearance in (3) of Eisenstein contributions. For type (3), the global objects on the left are ${}^LG$-bundles with flat connections — and singularities. I do not know what the moduli theory of such bundles is like or where to find it, but I suppose one has to expect continuous parameters, even when the singularities are given. How are we to prove the existence of these continuous parameters on the right. According to §3.8 of [CFT], it is possible to construct geometric Eisenstein series, but should one expect that they exhaust the continuous spectrum on the left? In the reciprocity for type (c), the continuous parameter on the left will come from two sources: (i) flat bundles that are not stable, even very unstable; (ii) continuous families of stable flat bundles if they exist. The continuous parameter on the right will come, as in the theory of Eisenstein series, from the unstable bundles, those lying near the cusps, but is there another sort?

As an example, not to illustrate the real possibilities but just to illustrate the nature of the constraints built into the definitions, consider the unramified case over $\mathbb{P}^1$. Local systems are defined in §3.1 of [CFT]. On $\mathbb{P}^1$ a local system for ${}^LGL(1) = GL(1)$ is trivial. On the other hand, the variety Bun_1 of $GL(1)$-bundles on $\mathbb{P}^1$ is the set $\mathbb{Z}$. There is one bundle of each degree. What the Hecke operator does is shift to the left or the right, according to the conventions. Thus $z \mapsto z+1$. The one Hecke eigensheaf in the sense of §3.8 and equation (3.9) is the constant sheaf. It certainly does not seem to exhaust the possibilities. If we pass to functions on $\mathbb{Z}$, then any function $z \mapsto e^{\alpha z}$, $\alpha \in \mathbb{C}$ serves as an eigenfunction.

There is a lesson to be learned from this. We see once again, as we saw for type (1) — this may or may not be so for type (2) — that the left side is smaller than the right. How much smaller is not clear to me. It is pretty clear that a good many unramified Eisenstein series will be missing on the left, and this must be so also for type (2). The flatness condition on a local system excludes them no matter what the genus of X is. As observed in §4.4 of [CFT] local systems for $GL(1)$ are line bundles of degree 0. The reason that, as asserted in the same section, every line bundle of degree 0 admits a holomorphic connection is not immediately clear to me, although that the fibers, if not empty, are affine spaces over $H^0(X,\Omega)$ is of course clear. A question for type (3), and to some extent also for type (2), that does not lack interest is whether there are objects on the right excluded on the left for other reasons and that are not in any sense Eisenstein series.

As recalled in §3.2 of [CFT] bundles for $GL(n)$ on X can as a set be parametrized by $GL(n,F)\backslash GL(n,\mathbb{A}_F)/GL(n,\mathcal{O}_F^\infty)$, where $\infty \in X$ is any point with coefficients in $\mathbb{C}$, $\mathcal{O}_F^\infty$ is the set of rational functions on X with no poles outside of $\{\infty\}$, and where

$$GL(n,\mathbb{A}_F) = \lim_{\infty\in S} \prod_{v\in S} GL(n,F_v)/GL(n,\mathcal{O}_v) \tag{4}$$

is a limit over finite sets S containing ∞. This is presumably true for very general G. We could consider Hecke operators on functions on this space, functions whose properties (continuity, support) would have to be appropriate to the topological structure underlying (4). We could also adapt this formulation to allow for ramification at a finite set of places. It may be a good idea to introduce if possible some analytic structure into the situation, replacing the sheaves of, for example, §3.8 of [CFT] by $\mathcal{D}$-modules, thus by differential equations and to show that the dimension of the spectrum of the resulting problem is equal to the dimension of the moduli space for ${}^LG$ local systems.

Analytically the spectra typically treated are the spectra of a commuting family of differential operators. If the underlying space is compact, they are discrete. If the underlying space has no compact element and is of dimension equal to the number of algebraically independent elements in the family, the dimension of the spectrum is — typically — equal to this dimension. Although general theories assure us of that there is an equivalence between perverse sheaves and $\mathcal{D}$-modules, the presence of a large group of symmetries entails, not alone for type (3) but also for types (1) and (2) a good deal of redundancy. Like it or not this redundancy has to be removed, clumsily by restricting to classical automorphic forms or Maaß forms — thus, implicitly, spherical functions — but, in my view, more elegantly by exploiting concepts discovered by Dedekind and Frobenius, thus irreducible group-representations and their characters. To do this systematically requires that the local theory intimated earlier in this section has to be developed.

Awaiting this, it is worthwhile to reflect on the intimations in [CFT], in which clues of two different kinds are found. One is already there for the abelian case. In §4.4, it is shown that for $G = GL(1)$ the space of local systems is connected, compact and of (complex) dimension equal to twice the genus of X. So the "variety" of Hecke sheaves should have the same dimension. Apparently this, or rather more than this, has been shown by Laumon and by Rothstein. The relevant statement in [CFT] is formulated in categorical terms not in spectral-theoretical terms, even though the theory is described in [CFT] as the Fourier-Mukai transform. I find an analytic formulation more to my taste. It appears from the title of Rothstein's paper *Connections on the total Picard sheaf and the* KP *hierarchy* that such a formulation may be found there. The variety of Hecke sheaves seems described in terms of functions in §4.5 of [CFT]. These are functions on the elements of (4) of degree 0. The integer n is taken to be 1. An element of degree 0 is a divisor of degree 0, thus a line bundle.

Reflecting on what this must be in the most concrete form, in terms say of $\mathcal{D}$-modules, I recalled the familiar classical theorem that the dimension of the space V of integrals of rational forms $d\omega = f dz$ on the curve that have no logarithmic singularity modulo the space $V_{\rm rat}$ of those dh whose integrals are rational functions is twice the genus, thus $2g$. Consider the $\mathcal{D}$-module defined by

$$dg/dz - fg = 0. \tag{5}$$

It has no singularities in the sense of [CL], and thus no singularities in any reasonable sense. If we replace g by $\hat{g} = ge^{\int h}$, where $\int h$ is rational then the equation for $\hat{g}$ is

$$0 = \frac{d(\hat{g}e^{-\int dh})}{dz} - f\hat{g}e^{-\int dh} = \left(\frac{d\hat{g}}{dz} - h\hat{g} - f\hat{g}\right)e^{-\int dh},$$

or

$$\frac{d\hat{g}}{dz} - (h+f)\hat{g} = 0.$$

So the module associated to $h+f$ is isomorphic to that associated to f and the construction yields a family of $\mathcal{D}$-modules with the correct dimension $2g$. We obtain the fibers given by $H^0(X, \mathcal{K})$ by taking those h that have no poles.

I add a remark or two here that will be relevant later. The space of regular dh is of course a subspace $V_{\rm reg}$ of dimension g in the larger space of dimension $2g$. We may associate to $d\omega$ the sum over its poles $x_1, x_2, \dots$ supposed to have orders $m_1 + 1, m_2 + 1, \dots,$ $m_i > 0$, the divisor $\sum m_i x_i$. Since this is divisor is constant along cosets of V_{reg}, we obtain a map from V/V_{reg} to the Picard variety.

The possibility of a general form of the Fourier-Mukai transform available for all G is raised in §6 of [CFT]. I would like to suggest, once again, that we not look for an equivalence of categories but that, so far as possible, we treat the problem, as for types (1) and (2), as a problem in spectral theory, recognizing that in contrast to types (1) and (2), for type (3) the relevant (cuspidal) spectrum is continuous, while the continuous spectrum (the Eisenstein spectrum) may be almost imperceptible. This is compatible with the existence of a trace formula.[3] I cannot even hazard a guess as to what extent the difficulties described in, say, §6.2 of [CFT] will reveal themselves in an analytic context as factitious. What does appear, and this also seems to be the starting point of §6 of [CFT] which is the transition to Part III of the paper *Conformal field theory approach* is that we have to find a way of constructing the Hecke eigensheaves, apparently in their incarnation as $\mathcal{D}$-modules, thus in a more analytic form.

I observe the contrast between types (1) and (2) on one hand and type (3) on the other. For types (1) and (2) the continuous part of the spectrum is given by the Eisenstein series and has to be constructed. The remaining part, the cuspidal part, is discrete and does not need to be described explicitly. For type (3), there seems to be little in the way of Eisenstein series that is pertinent. On the other hand, the part that might be called cuspidal is continuous, and we have to construct it. We are apparently being urged in Part III of [CFT] and [GT] to turn to contemporary mathematical physics for help.

Before we attempt to understand these two sources, there are a few cautionary or explanatory remarks to make. In contrast to the global forms of the correspondence for types (1) and (2), for which it is not entirely clear what is wanted on the left side, only that it will be difficult to describe with the desired precision, for type (3.a) the left-side is fairly clear, particularly when there is no ramification. The right side, both locally and globally, is more obscure, perhaps sheaves, perhaps $\mathcal{D}$-modules of some sort, sometimes taking account of the large group of isomorphisms, sometimes, especially in an unramified context, resolutely ignoring it. It does seem to me that a focussed examination of the possibilities for an adequate local theory is called for. We have also already observed that for type (3) the strategy of §3 of this note, which is based on the trace formula, is faced with a serious obstacle. We have no way of determining the multiplicity with which the trivial representation is contained in a global object associated to $GL(n)$. It may turn out that for type (3) the trace formula serves different purposes.

Even though I do not yet have information about the progress of Ben-Zvi and Nadler, I have a notion of what to expect from a trace formula. If the local possibility suggested above is in any sense valid, then we can expect that for functions $f = \prod_{v \in V} f_v$, the set V of points on X being finite,

$$\varphi \mapsto \varphi' : \quad \varphi'(g) = \prod_v \int \varphi(gh) f(h) d\mu(h)$$

[3] Apparently, David Ben-Zvi and David Nadler are investigating the possibility of a trace formula for type (3). Nadler has promised me a summary of their conclusions. I am, of course, eager to see it.

is meaningful, the integrals being interpreted as above. Since the spectrum is — or so we suppose — finite-dimensional, although not necessarily compact, it might be possible to give the trace a sense.

The typical example is formed by functions on the line. The trace is

$$\int_{\infty}^{\infty} \hat{f}(x)dx = f(0)$$

Something similar may be possible in the present circumstances, where the spectrum is — if I am not mistaken — largely a vector-bundle over a compact base.

Can we hope that the trace formula can be treated as in [FLN] as an integral over a Steinberg-Hitchin base, in this case, basically

$$\sum_{v \in V} \mathbb{C}((t_v))/t^{n_v} C[[t_v]],$$

where V is finite and t_v a local parameter at V? Even if this is so, what can we do with it? If we had a proposal for a collection of sheaves or $\mathcal{D}$-modules presumed to be complete because attached to a full set of parameters, namely one for each flat ${}^{L}G$ bundle, we might be able to verify their completeness with the help of a trace formula.[4]

6. Gauge theory and reciprocity

I am in no position to understand the language of gauge theories, nor even to judge the extent to which it has any meaning for a mathematician. It is none the less appropriate to ask oneself, as a mathematician, whether there could be not only a genuine mathematical content to the suggestions of a possible relation between the "physical" concepts appearing in say the excursions of Edward Witten and collaborators with their geometrical and physical baggage with the problem of reciprocity in the sense of (3.c), but even some useful suggestions for establishing it in a mathematical sense. In spite of its oneiric quality, the reference [GT], in conjunction with Part III of [CFT] is not such a bad place for a mathematician to begin the attempt to penetrate the constructions of the physicists.

What I find curious, even, although in a weak sense, promising is that the Hitchin base and the associated fibration appear, not in the sense of Ngô but in the original narrower sense. The Hitchin base, even in Ngô's wider sense, is not the same as the similar Steinberg-Hitchin base, but is closely related to it, as in the passage from the Lie group to the Lie algebra in the study of orbital integrals and the fundamental lemma by Waldspurger and others.

This said, I am not yet prepared to describe or assess the somewhat fanciful arguments of [GT], but I do want to say enough that a possible connection to the constructions of the previous section does not seem too far-fetched. I add that — for the moment at least —

[4] Reflecting on these notes as a whole some time (20/2/2011) after I began to write them, I noticed that a few obvious matters were not treated. It should, for example, be possible, and would certainly be useful, to treat, without much difficulty, the theory for $GL(1)$ fully, thus considering not only the unramified theory but the general theory, first for $X = \mathbb{P}^1$, and then in general.

the more I look at [CFT,GT], the more I feel that there is something there. Since for me, reciprocity is something that is only introduced after questions of endoscopy and stability have been resolved and that only operates at the level of stable objects, the introduction in [GT] of endoscopy into the discussion of reciprocity makes me uneasy, but that is a confusion for mathematicians to clarify, if indeed it can be clarified. A too hasty reaction is inappropriate. Stable global characters are — to some extent — factitious objects. The "stable" representations that they represent do not always occur within the action on the space of automorphic forms. So we have no reason to believe that the corresponding stable object of type (3) necessarily appears as a sheaf or $\mathcal{D}$-module on Bun_G. Nevertheless I have not yet examined the pertinent sections §5.4 and 6.3 of [GT] with any care.

I am also not comfortable with the discussion of local issues in [CLG]. The relation to representations of affine Kac-Moody algebras at critical level is, I can well believe, not without significance, but my first impression, on which I should continue to reflect, is that it may be at the level of eigenvalues of the center of the universal enveloping algebra in Harish-Chandra's theory of tempered representations and the Plancherel formula. The center was introduced at the very beginning of that theory and never lost its importance but it was the introduction and construction of the pertinent eigendistributions that was the core element in the analysis. An understanding of the structure of the center was essential but by itself not enough.

It is perhaps best to pass to [GT] over the last section of [CFT], where the connection to the mathematical problem of constructing a parametrized family of solutions to a spectral problem is clearer. Part of the difficulty for me is a result of my unfamiliarity with $\mathcal{D}$ modules. As I write these notes, I come to understand that I need to be steeping myself in several subjects and the appropriate introductory texts, subjects and texts that I had previously happily ignored. Classically, as say in the theory of Eisenstein series, the operators are fixed and only the eigenvalues vary with the parameters. On the other hand, it appears that in the present context the parameters appear in the operators as well. There is a larger dose of geometry in the analysis than for type (1).

For any given value of the parameters, we need a collection of commuting differential operators on Bun_G with $d = \dim \mathrm{Bun}_G$ independent elements and with a common eigenfunction. In [CFT] such families are constructed with the help of the enormous algebras of operators that have been, for me, one of the distressing characteristics of modern mathematical physics. It may be that if I understand them in a more analytic and less algebraic context, I can reconcile myself to them.

This is the advantage of [CFT]. Ramification aside, and as I have observed, the first matter to clarify before attacking ramification is an adequate local theory, the problem is to construct sufficiently many Hecke eigensheaves or, better, the associated $\mathcal{D}$-modules. This is a problem in which some progress in the form of *opers* has been made by Beilinson-Drinfeld. It is explained in [CFT], from which a number of issues become clear. In particular, it is frankly admitted that opers, interesting as they are, yield far too few $\mathcal{D}$-modules. Since the first draft of these notes has to be ready before the lectures are delivered, I have decided to postpone, for the moment, any attempt to understand the geometry of opers.

So other methods are necessary. I have more than a little trouble in following the transition from [CFT], in which the emphasis is on conformal field theory, thus on two-dimensional theories, to [GT], in which four-dimensional theories and branes manifest themselves. I have to try to understand not only what physical notions are available for the construction of the appropriate $\mathcal{D}$-modules, whether they have any mathematical substance, and whether they yield a signicant set of these $\mathcal{D}$-modules, thus one with the dimension of the moduli space of flat ${}^LG$-bundles. As for opers and for the same reasons, I have not yet tried to penetrate the works of Witten and his coauthors. They must certainly be consulted, but later, at leisure.

A notion of gauge theory that is adequate to the purposes of [GT] appears to be available in §4.3 of [QFS]. The report [GT], or at least the section entitled *Enter physics* begins with a gauge theory on a four-dimensional Riemannian manifold M_4 whose lagrangian is provided with a supplemental θ term. The manifold is the product $\Sigma \times X$ of an algebraic curve X, for our present purposes the curve that appears as the base for functoriality or reciprocity of type (3.a), and a Riemann surface Σ with boundary. So a metric on X determines one on M_4. We also need a compact Lie group G_c, which appears to be customarily connected. Whether this is necessary or desirable is not clear to me. The group G_c will be an inner twisting of a quasi-split group G and a twisting of a Chevalley group G'. Presumably the gauge-theoretic duality in [GT] is thought to be related to the duality of §7 for the group G, but I am by no means certain. The notion of a σ-model is defined in [QFS]. As in that book, the theory pertinent in [GT] is a theory that is coupled with a gauge theory. The basic fields are connections A on principal G_c-bundles. If F_A is the curvature of the connection the (Euclidean) action is given by

$$I = \frac{1}{4g^2}\int_{M_4} \operatorname{tr} F_A \wedge *F_A + i\frac{\theta}{8\pi^2}\int_{M_4} \operatorname{tr} F_A \wedge F_A, \tag{6}$$

a formula taken from [GT]. Here F_A is the curvature of A, the operator $*$ is the Hodge star-operator, and tr is the bilinear form on the Lie algebra $\mathfrak{g}$ of G_c, so normalized that the second term is of the form $i\theta k$ with k integral. I confess that at this point I am just copying the original text. What I want is to arrive with as few words and explanations as possible at some inkling of the possible relation between *mirror symmetry* and the reflections of §5.

The complex parameter

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

is introduced in [GT] and, after a passage to a supersymmetric extension, a notion whose precise definition is not initially explained, the following definition is offered: the S-duality of the theory with gauge group G_c and complex coupling τ is equivalent to the theory with gauge group ${}^LG_c$ and coupling constant ${}^L\tau = -1/n_{\mathfrak{g}}$, where $n_{\mathfrak{g}}$ is the lacing number of $\mathfrak{g}$. Unfortunately, I do not yet know what it means for two theories to be equivalent.

After a sequence of explanations that I am not yet capable of understanding, Edward Frenkel, the author of [GT], arrives, following Witten and various co-authors, at the conclusion that

The S-duality of the supersymmetric gauge theories on $\Sigma \times X$ (for particular values of Ψ, thus, more or less, of the parameter τ) becomes mirror symmetry between the topological σ-models with the targets $\mathcal{M}_H(G_c)$ and $\mathcal{M}_H({}^LG_c)$, a statement in which both G and its L-group appear in compact form. Leaving aside the possible meanings of the various intermediate statements and the question of whether they have a genuine mathematical content, we examine the statement in italics. There are three expressions in it patiently waiting to be defined: mirror symmetry; topological sigma models; and the two expressions $\mathcal{M}_H(G_c)$, $\mathcal{M}_H({}^LG_c)$.

There are some necessary elements missing in the statement of the conclusion. There are two types of *twisted two-dimensional σ-models*, the *A-models and the B-models*. In this sentence too a number of still undefined terms appear. In fact, the original theory on $\Sigma \times X$ is modified. According to [GT], "Kapustin and Witten ... study the limit of the gauge theory when X becomes very small. In this limit the theory is described by an effective *two-dimensional* topological field theory on Σ." Since I myself thought that X was fixed, I assume an auxiliary metric is introduced and then multiplied by a constant that goes to 0. The result will certainly depend on X, but is treated, in spite of the ambiguous language, as a theory on the two-dimensional manifold Σ. Apparently the theory can be identified with the twisted topological σ-model with the target manifold $\mathcal{M}_H(G)$. What is as yet only implicit, is that the appropriate theory is given by boundary conditions on Σ. For the mirror symmetry to intervene, these boundary conditions, referred to as "branes" have to be different for G and for ${}^LG$. It turns out that one type, the A-branes, are to be used for G and another type, the B-branes, for ${}^LG$. They cannot be introduced without describing the special natures of the varieties $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^LG)$

At a first innocent glance, it is the manifolds $\mathcal{M}_H(G_c)$ or $\mathcal{M}_H({}^LG_c)$ in which there is something that has grown familiar to specialists in the theory of automorphic forms or representations in recent months, because the manifolds that appear in the proof of the fundamental lemma for Lie algebras given by Ngô are modeled on them. This is not the original fundamental lemma as stated in [STF], but it implies it. We have already seen in §4 that its proof contains elements whose interest, even for the theory of automorphic forms and the trace formula, transcends the theory of endoscopy. The present manifolds differ from those in Ngô's work in two respects: (1) they are varieties over $\mathbb{C}$, not over a finite field; (2) in the proofs of the fundamental lemma, the canonical bundle K_X is replaced by an arbitrary, but fixed, line bundle that has to be correctly chosen. In other words, in the context of Higgs fields and Hitchin moduli spaces there is a constraint on the poles of ϕ not present in the context of the fundamental lemma. It may be that this constraint will disappear when we remove the condition that the sheaves are unramified, thus when we allow irregular singularities.

What the semistable condition entails is not clear to me. A rough intuition is that it amounts to a removal of the cusps, thus an elimination of the need for the partially continuous spectrum expressed by Eisenstein series and their geometric counterpart.

The definitions of these two manifolds are the same, except that they refer to two groups, G and ${}^LG$, that are normally different. The manifold $\mathcal{M}_H(G_c)$ can be defined as an algebraic variety, the moduli space of (semistable) Higgs bundles on X. These are pairs (E, ϕ), where E is a (semistable) G-bundle and ϕ a Higgs field, thus ϕ is a holomorphic section on X of $\mathfrak{g}_E \otimes K_X$,

$$\phi \in H^0(X, \mathfrak{g}_E \otimes K_X). \tag{7}$$

As in the proof of the fundamental lemma, there is a map of $\mathcal{M}_H(G_c)$ to the Hitchin base

$$\mathbf{B} = \oplus H^0(X, K_X^{\otimes(d_i+1)}).$$

The first observation is that ${}^L\mathbf{B}$, thus the Hitchin base for the dual group is isomorphic to $\mathbf{B}$, but only after choosing an invariant bilinear form on a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. This is because, by definition, ${}^L\mathfrak{h}$ is the dual of $\mathfrak{h}$. Moreover, the bilinear form that identifies $\mathfrak{h}$ with its dual is invariant under the Weyl group and, if G is simple, is well-defined up to a scalar. Since this scalar appears again as the automorphism of $\mathbf{B}$ arising from the scalar action on the canonical bundle, it can be realized by an automorphism of one or the other of the two bundles. So the Hitchin bases for $\mathcal{M}_H(G_c)$ and $\mathcal{M}_H({}^LG_c)$ can be identified. This yields a diagram

(8)
$$\begin{array}{ccc} \mathcal{M}_H({}^LG) & & \mathcal{M}_H(G) \\ & \searrow \quad \swarrow & \\ & \mathbf{B} & \end{array}$$

The possibility of identifying the Hitchin bases seems to be unquestionable, not only for the canonical bundle but for any bundle. Before attempting to understand what one might make of this, I return to the other curious properties of these Hitchin moduli spaces.

Both manifolds are hyperkähler manifolds, an essential feature of mirror symmetry. I consult Wikipedia for the definitions. First of all, “a Kähler manifold is a manifold with unitary structure (a U(n)-structure) satisfying an integrability condition. In particular, it is a Riemannian manifold, a complex manifold, and a symplectic manifold, with these three structures all mutually compatible. This threefold structure corresponds to the presentation of the unitary group as an intersection:

$$U(n) = O(2n) \cup GL(n, \mathbb{C}) \cup Sp(2n).\text{”}$$

This definition is paraphrased and the integrability condition made precise in “A Kähler metric on a complex manifold is a hermitian metric on the tangent bundle satisfying a condition that has several equivalent characterizations (the most geometric being that parallel transport induced by the metric gives rise to complex-linear mappings on the tangent spaces). In terms of local coordinates it is specified in this way: if

$$h = \sum h_{i\bar{j}} dz^i \otimes d\bar{z}^j$$

is the hermitian metric, then the associated Kähler form defined (up to a factor of i/2) by

$$\omega = \sum h_{i\bar{j}} dz^i \wedge d\bar{z}^j$$

is closed: that is, d = 0.

A hyperkähler manifold is, in particular, a real manifold whose tangent space carries an action of the quaternion algebra so that, in particular, it carries an action of three anti-commuting elements I, J and K or of $aI + bJ + cK$, $a^2 + b^2 + c^2 = 1$, each one of which defines a complex structure. Those associated to I, J, K can be singled out. According to Wikipedia, they have various properties. First of all, "a hyperkähler manifold is a Riemannian manifold of dimension $4k$ and holonomy group contained in $Sp(k)$ (here $Sp(k)$ denotes a compact form of a symplectic group, identified with the group of quaternionic-linear unitary endomorphisms of an n-dimensional quaternionic Hermitian space). Hyperkähler manifolds are special classes of Kähler manifolds. They can be thought of as quaternionic analogues of Kähler manifolds." Secondly, "a hyperkähler manifold (M,I,J,K), considered as a complex manifold (M,I), is holomorphically symplectic (equipped with a holomorphic, non-degenerate 2-form)." Presumably there is such a form attached to J or K, and indeed to any element $aI + bJ + cK$, $a^2 + b^2 + c^2 = 1$."

The varieties $\mathcal{M}_H(G)$ are, apparently, hyperkähler varieties. So far as I know, this is not so for the more general varieties introduced by Ngô. Certainly the definition of J does not suggest it. The natural complex structure on $\mathcal{M}_H(G)$ as a moduli space is taken to be that defining I The second complex structure, that given by J, seems to be defined for semistable bundles alone. At the moment, it is not clear to me why, nor is it fully clear how the second structure is to be defined. There is a clue in [CFT].. The space $\mathcal{M}_H(G)$ (or $\mathcal{M}_H({}^LG)$) with the complex structure J is defined in a different way. It is defined as the moduli space of semi-stable flat bundles, thus pairs (E, Δ) with E as before, but with Δ a connection on E. The collection of flat bundles is fibred over the collection of bundles without connection. There seems to be no ambiguity about the complex structure on the fibers. They are, as observed at the beginning of §4.5 of [CFT] all the same, just $H^0(X, \mathcal{K})$, but only after we have fixed a point in the fiber.

To clarify the issues just a little, I consider the case that $G = GL(1)$, which we have already examined from a classical point of view, that was again different. The $\mathcal{D}$-modules of (5) were defined by $2g$ complex parameters, but we did not inquire which parameters led to equivalent $\mathcal{D}$-modules, nor did we recall the relation between the parameters f and the de Rham cohomology,[5] or with points on the jacobian or Picard varieties of X. We also did not inquire whether we had exhausted all possible $\mathcal{D}$-modules or how to pass between $\mathcal{D}$-modules and Hecke eigensheaves. These questions have to be answered before we are certain of the parametrization. What is clear is that there is no reason to expect that the relation between the various possible complex parameters, especially those on the side of G and those on the side of ${}^LG$, is holomorphic. The parameters f lie in a $2g$-dimensional complex space, $H^0(X, \mathbb{C})$. So they can be projected onto $H^{1,0}(X, \mathbb{C})$ or $H^{0,1}(X, \mathbb{C})$. Somewhere or other there will be an incompatibility of the various complex structures. I have not yet sorted it out. Since we are dealing with a basic question in the theory of algebraic curves, this is presumably just a matter of putting one's knowledge of the classical theory

[5] As in Grothendieck's *On the de Rham cohomology of algebraic varieties*, where he states that "the theory of "integrals of second kind" is essentially contained in the following very simple Theorem 1 – Let X be an affine algebraic scheme over the field $\mathbb{C}$ of complex numbers; assume X regular (i.e. "non singular"). Then the complex cohomology $H^*(X, \mathbb{C})$ can be calculated as the cohomology of the algebraic de Rham complex (i.e. the complex of differential forms on X which are "rational and everywhere defined").

in order, consulting where it is lacking some standard texts. What is clearly suggested is that we should not be looking for a parametrization of the $\mathcal{D}$-modules/Hecke-eigensheaves with their natural complex structure by flat ${}^LG$-bundles with their natural complex structure in the expectation that the two will be compatible. Apparently, we are faced with this issue in the constructions of Witten et al. That is an encouraging sign.

In the present context, we have the following assertion of Frenkel, "There are two types of two-dimensional σ-models: A-model and B-model. The former depends on the symplectic structure on the target manifold and the latter depends on the complex structure." For both G and LG the two choices of model are available. Apparently, there are two choices of pairing, interchanged on interchanging G and ${}^LG$. The paired elements for the first are $(\mathcal{M}_H(G), \omega K)$, where ω_K is either the symplectic structure associated to the third complex structure, that denoted by K, or the Kähler form associated to it — they are presumably the same, but this is not yet clear to me — and $(\mathcal{M}_H({}^LG) = \mathcal{Y}({}^LG), \omega_J)$, where ω_J is the Kähler form associated to J.

We have not tried to define σ-models. The definition is given at the very beginning of [QFS] and seems clear enough. What we need in order to understand fully Frenkel's exposition is the notion of an *effective two-dimensional topological field theory on* Σ and of a *topological σ-model.* Once again, I tried Wikipedia and found that, "A topological quantum field theory (or topological field theory or TQFT) is a quantum field theory which computes topological invariants." I also found that, "The known topological field theories fall into two general classes: Schwarz-type TQFTs and Witten-type TQFTs. Witten TQFTs are also sometimes referred to as cohomological field theories."

This is the abstract from Witten's paper *Topological Sigma Models*

Abstract. *A variant of the usual supersymmetric nonlinear sigma model is described, governing maps from a Riemann surface Σ to an arbitrary almost complex manifold M. It possesses a fermionic BRST-like symmetry, conserved for arbitrary Σ, and obeying $Q^2 = 0$. In a suitable version, the quantum ground states are the* $1+1$ *dimensional Floer groups. The correlation functions of the BRST-invariant operators are invariants (depending only on the homotopy type of the almost complex structure of M) similar to those that have entered in recent work of Gromov on symplectic geometry. The model can be coupled to dynamical gravitational or gauge fields while preserving the fermionic symmetry; some observations by Atiyah suggest that the latter coupling may be related to the Jones polynomial of knot theory. From the point of view of string theory, the main novelty of this type of sigma model is that the graviton vertex operator is a BRST commutator. Thus, models of this type may correspond to a realization at the level of string theory of an unbroken phase of quantum gravity.*

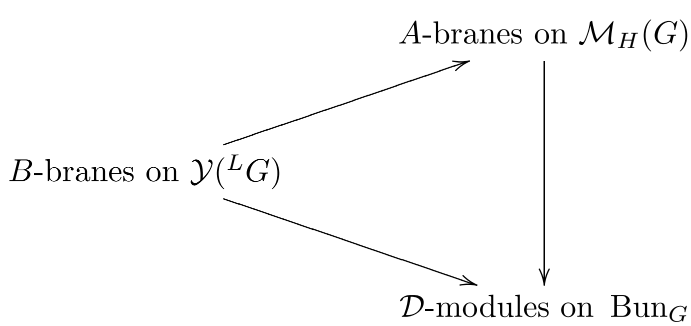
I decided that if I were to continue to try to understand whether there was something on which I, as a mathematician, could put my hands, I would, for now, have to skip a few paragraphs in [GT] and go on to the more structural explanations. Indeed for the moment, I am pretty much defeated by the later sections of [GT]. This does not mean that there is nothing to be learned from the papers on which it is reporting, just that it seems that in order for a simple mathematician to profit from the suggestions of Witten and coauthors, the first and best step will be to turn to their papers and to examine specific examples, taken to be as simple as possible, in hopes of understanding — in as concrete a fashion as possible — their possible mathematical, thus geometrical and analytic, content.

It is best, I think, to remain in the unramified context. There will be issues in common — global issues — that will appear in purer form in the unramified theory. The supplementary issues that arise with ramification are, one can hope, principally local. Recall what we want in the unramified context. We want to parametrize either the family of Hecke eigensheaves for a given group G or the associated family of $\mathcal{D}$-modules, the parameters being flat ${}^LG$-modules. I do not know whether questions of stability or endoscopy arise. There are examples in [GT] that are asserted to be pertinent, but am not yet in a position to examine them. With types (1) and (2) the question of stability and L-packets arises only on the right side of the correspondence. So I am startled by the language of §6.1 of [GT].

An important question that is not, so far as I can see, broached in [GT] is whether the construction of eigensheaves parametrized by flat ${}^LG$-bundles yield all eigensheaves, either in the unramified context or in general. Here we have to be careful with the word “all”, if for no other reason than because of the possibilities arising from “Eisenstein eigensheaves”. We know that the answer to the corresponding question for type (1) has a negative answer. For the geometric theory even for the group $GL(1)$, the statement of Theorem 3 of [CFT] contains an arrow only in one sense. It has been already observed at the end of the preceding section that a trace formula could perhaps be used to establish that, in the geometric theory, the image of the flat ${}^LG$-bundles contains all possible eigensheaves except for those with an Eisenstein element.

For those who understand the physical theories, the key to the reciprocity seems to be the commutative diagram (9). The adjective holonomic does not appear in [GT]. I suppose it is pretty much superfluous for curves. Moreover we have not yet introduced A-branes or B-branes. It appears to be the A-branes and the B-branes that manifest themselves in the physical-geometrics or geometrical physics context of Gukov/Witten or Kapustin/Witten, whereas it is the B-branes and the $\mathcal{D}$-modules that are pertinent for the correspondence. So to pass from the physical arguments, which are supposed the upper skewed arrow, to the correspondence, which would be the lower skewed arrow, one needs the vertical arrow of the diagram. According to [GT], this is established in the papers [Na] and [NZ].

(9)



One can hope that these papers can be read in the customary mathematical manner. To understand the two skewed arrows, one will have to know what a “brane” is. I think it is best for the moment simply to quote the explanations of [GT]. This gives the flavor.

"Branes in two-dimensional sigma-models are certain generalizations of boundary conditions. When writing path integrals for maps $\Phi : \Sigma - M$, where Σ has a boundary, we need to specify the boundary conditions for Φ on $\partial\Sigma$. We may also "couple" the sigma model to a quantum field theory on $\partial\Sigma$ (that is, modify the action by a boundary term),... 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 a B-model, and a flat unitary vector bundle on M' in the case of an A-model."

This brief explanation gives me, at least, some hope that with sufficient effort I may one day understand the issues in a concrete way. The sentences that follow are, nevertheless, a little disquieting for someone who prefers to be sparing in his appeals to categories.

"More generally, the category of branes in the B-model with a complext 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 Δ is a flat unitary vector bundle on L. However, it is also expected to contain more general objects..."

I, myself, am not yet prepared to add to these remarks. With these definitions, the upper skewed arrow of diagram (9) is an expression of what is referred to as *Homological mirror symmetry* for the underlying hyperkählerian manifolds $\mathcal{M}_H(G)$ and $\mathcal{M}_H({}^LG)$. Whether there is a "mirror symmetry" for other pairs or whether it is limited to pairs of Hitchin moduli spaces, I do not know!

7. Establishing reciprocity.

This section comes at the end because I have nothing to write. It is reciprocity for the theories (1.a) and (2.a) with which we are concerned here. I have suggested in [H] and in [SP] that one might combine the seeding provided by the Taniyama group with deformation arguments from the theory of p-adic automorphic forms of the sort used in the proof of Fermat's theorem. I have not tried to pursue this, and am not entirely easy about the attempts I have seen to develop the deformation theory. The insistence on Galois representations of odd type troubles me, but that may be simply a lack of understanding on my part. I should perhaps make a more serious attempt to understand what has been achieved.

It may or may not have escaped the reader's attention that for type (3) it is functoriality that is neglected, both locally and globally. It is not necessarily a consequence of reciprocity, since we have no assurance that reciprocity — which is unlikely to be entirely reciprocal — passes from the left to the right and may not be surjective. I have suggested none the less that there are some grounds for suspecting that it may be essentially so, not only for type (3) but also for type (2). Unfortunately, almost all local problems for type (3) have been neglected, so that there is much room for uncertainty as to the nature of the local reciprocity.

The divergence between the arguments that suggest themselves for type (1) and those that suggest themselves for type (3) is intriguing. Are both types valid for type (2)?

For type (1), global reciprocity is almost certainly at a different level than for type (3), with type (2) perhaps lying somewhere in the middle. As stressed in diagram (1), the goal is to tensor some mock Tannakian structure over the field F, for example over $\mathbb{Q}$, with $\mathbb{C}$ and to show that the result is contained in the mock Tannaka category over $\mathbb{C}$ on the right defined by functoriality. Since there are major obstacles to the construction of the middle object, whether motives or some larger category that allows for the flat vector bundles appearing in the theory of Shimura varieties, the current impulse is to attempt to finesse the problem with the help of the ℓ-adic (or p-adic) representations. There is a feeling even among those who have seriously reflected on the matter that this passage over ℓ-adic (or p-adic) representations is inevitable. This feeling may or may not be correct. I myself am inclined to take it seriously, although I have been tardy, for lack of time rather than of good intentions, in learning not only the relevant p-adic representation theory but also the theory of p-adic automorphic forms. There is also a — in my view somewhat lamentable — tendency to ignore the object in the middle and to try to pass directly between automorphic representations and ℓ-adic representations.

8. The simplest case of the geometric theory.

The simplest case of the geometric theory is the group $G = GL(1)$. I begin with the local theory. The local group is abelian, the product of $\mathbb{C}^\times$ and truncated power series

$$(10) \qquad 1 + az + bz^2 + cz^3 + \cdots = 1 + a_1 z + a_2 z^2 + a_3 z^3 + \ldots, \quad (\mathrm{mod}\ z^{n+1}), \quad n \geq 0$$

where I have allowed myself two choices of notation, the first for n small, the second for n large. Since this group is abelian, it is pretty clear that it will have to be the direct sum of n copies of the complex line.

For $n = 1$, the group is $\{1 + az\}$ with multiplication $(1 + az)(1 + bz) = 1 = (a + b)z$, thus the additive group of the line. For $n = 2$, we have a subgroup $\{1 + bz^2\}$, also isomorphic to the additive group. A second factor is $\{1 + az + \varphi(a)z^2\}$ with $\phi(a) = 1/2$. It is easily verified that

$$(1 + az + \frac{a}{2}z^2)(1 + bz + \frac{b}{2}z^2) = 1 + (a + b)z + \frac{(a + b)^2}{2}z^2 \quad (\mathrm{mod}\ z^3).$$

These simple calculations suggest the obvious way to describe the multiplicative group of elements (10) as the additive group of a vector space. We write

$$1 + az + bz^2 + cz^3 + \cdots = \exp(\alpha z + \beta z^2 + \gamma z^3 + \cdots + \epsilon z^n)$$

and use $\alpha, \beta, \gamma, \ldots, \epsilon$ as the coordinates in the vector space.

For the group $GL(1)$ the space of automorphic forms in the geometric context is easy to describe. First of all, the analogue of the space $F^\times\backslash\mathbb{A}_F$ is best taken in ascending pieces. Choose a finite set $\mathfrak{d}\subset X$ of distinct points in X. Let $F_\mathfrak{d}^\times$ be the set of elements in $F^\times$ whose zeros and poles all lie in $\mathfrak{d}$.

Set

$$\mathbb{A}_\mathfrak{d}^\times=\prod_{x\in\mathfrak{d}}F_x^\times.$$

Finally let D be a positive divisor $D=\sum_{x\in\mathfrak{d}}a_x$, $a_x\geq 0$. Then the appropriate space $\mathbf{P}_D$ is almost

$$F_\mathfrak{d}^\times\backslash\mathbb{A}_\mathfrak{d}^\times/\prod_{x\in\mathfrak{d}}G_{a_x}(\mathcal{O}_x).$$

It is, however, better as in [FLN] to choose an element I in $\mathbb{A}_\mathfrak{d}^\times$ whose associated divisor is of degree 1 and to replace $F_\mathfrak{d}^\times$ by $\tilde{F}_\mathfrak{d}^\times=I^\mathbb{Z}F_\mathfrak{d}^\times$ in the definition.

$$\mathbf{P}_D=\tilde{F}_\mathfrak{d}^\times\backslash\mathbb{A}_\mathfrak{d}^\times/\prod_{x\in\mathfrak{d}}G_{a_x}(\mathcal{O}_x).$$

The point is that in the present context as in other contexts we are sometimes only concerned with divisors of degree 0, but that the ordinary theory of automorphic forms or representations does not accomodate this restraint, except in the artificial way just described.

We introduce also for each divisor $D'=\sum_{x\in\mathfrak{d}}a'_x(x)$, $a'_x\geq 0$ the subset $\mathbb{A}_\mathfrak{d}^{D'}$ of $\prod_{x\in\mathfrak{d}}g_x^\times$ in $\mathbb{A}_\mathfrak{d}$ for which $O(g_x)=a'_x$. Recall that g_x is a formal Laurent series, so that the order $O(g_x)$ of the zero of g_x is a well-defined integer. There is a finite set $\mathfrak{D}$ of D' such that

$$\mathbf{P}_D=\cup_{D'\in\mathfrak{D}}\mathbb{C}^\times\backslash\mathbb{A}_\mathfrak{d}^{D'}/\prod_{x\in\mathfrak{d}}G_{a_x}(\mathcal{O}_x). \tag{12}$$

Here $\mathbb{C}^\times$ is to be treated as a subset of $F^\times$, itself contained in $G(F_x)$ for all x, thus, with the diagonal imbedding, as a subset of $\prod_{x\in\mathfrak{d}}G_{a_d}(\mathcal{O}_x)$. Each of the components in (12) is, as a set on which $\mathbb{C}^\times\times\prod_{x\in\mathfrak{d}}G_{a_d}(\mathcal{O}_x)$ acts and is to be identified with

$$\mathbb{C}^\times\backslash\prod_{x\in\mathfrak{d}}G_0(\mathcal{O}_x)/G_{a_x}(\mathcal{O}_x), \tag{13}$$

a finite-dimensional abelian Lie group whose structure we have identified. It is a product of $d-1$ copies of $\mathbb{C}^\times$ and of $|D|$ copies of $\mathbb{C}^+$, thus of the additive group $\mathbb{C}$. So, as in the remarks at the end of §5, the representation of $\prod_{x\in\mathfrak{d}}G_{a_d}(\mathcal{O}_x)$ and of $\mathbb{A}_\mathfrak{d}^\times$ on the functions on $\mathbf{P}_D$ has a spectral decomposition and a trace.

The elements of the spectral decomposition are given by compatible families of characters on the groups $\mathbf{P}_D$. If a character is defined on this group for one D, it can then be extended in a single way to $\mathbf{P}''_D$ for all $D''\geq D$. This is what is meant by compatible.

To accommodate the language of CFT these eigenfunctions have to be interpreted as $\mathcal{D}-$*modules* or as the associated sheaves. We have also to understand the connection with flat line-bundles. We have already prepared the way for the associated $\mathcal{D}$-modules. It is perhaps best to precede the discussion for general curves by a discussion for $\mathbb{P}^1$, where there are is only one flat line-bundle, the trivial bundle, and only one nonsingular connection on it, the trivial connection. Then all the interest is in the singularities. On the left-hand side, the ${}^LG$-side, there are singular connections. For a general X and the trivial line-bundle, a connection is in a local coordinate z, in a neighborhood of $z = z_0$ given by

$$\frac{df}{dz} = A(z)f, \quad f(0) = 1. \tag{14}$$

One writes, as in Coddington-Levinson $A(z) = (z - z_0)^{-\mu-1}\tilde{A}(z)$. If $\mu \leq -1$, then the solution of the equation is holomorphic at z_0, and there is no singularity there. Otherwise there is a singularity.

We can also allow singularities. The condition $f(0) = 1$ is then usually inappropriate because f will have singularities. For the trivial bundle on $\mathbb{P}^1$, there is everywhere but at ∞ a natural parameter z. At infinity, the parameter is $w = 1/z$. As a equation in w, the equation (14) becomes

$$-\frac{1}{z^2}\frac{df}{dw} = \frac{df}{dw}\frac{dw}{dz} = A\left(\frac{1}{w}\right)f,$$

so that at ∞, $A(z)$ is replaced by $-A(1/w)/w^2$. Write $f(z) = \exp(h(z))$. The equation (14) is

$$\frac{dh}{dz} = A(z), \qquad h(z) = \int A(z)dz$$

so that the nature of the singularity is clear. If

$$A(z) = \frac{\alpha_m}{(z - z_0)^m} + \ldots \frac{\alpha_2}{(z - z_0)^2} + \frac{\alpha_1}{z - z_0} + B(z),$$

with B regular at $z = z_0$, then

$$h(z) = -\frac{\alpha_m}{(m-1)(z - z_0)^{m-1}} - \cdots - \frac{\alpha_2}{(z - z_0)} + \alpha_1 \ln(z - z_0) + \int B(z), \tag{15}$$

where the final term is again regular. At infinity,

$$A(z) = \alpha_{m-2}z^{m-2} + \cdots + \alpha_1 z + \alpha_0 + \frac{\alpha_{-1}}{z} + O(\frac{1}{z^2}).$$

This gives the analogue of (15) at infinity. Notice that the connection ceases to operate at a singular point. We cannot insist that f takes a finite value, in particular 1.

In any case, if we demand that $A(z)$ and its analogue are locally Laurent series with finite principal part, then $A(z)$ is a rational function. This yields all algebraic connections on the trivial bundle on $\mathcal{P}^1$. Our task is now to construct the corresponding sheaves, the corresponding $\mathcal{D}$-modules, and the corresponding automorphic form (or representation) in the sense of an eigenfunction on $\mathbf{P}_D$. If we fix the singular points $z_1, \ldots z_n$and the degree

m_i of $A(z)$ at each of them, then we have $\sum_i m_i$ with only one constraint. The sum of the residues of A is 0.

For the curve $\mathbb{P}^1$, there is only one component in (12) and it can be taken to be (13). So $\mathbb{A}_{\mathfrak{d}}^\times$ is the product of $F_{\mathfrak{d}}^\times$ with $\prod_{x\in\mathfrak{d}} G_{a_d}(\mathcal{O}_x)$. The intersection of these two groups is $\mathbb{C}^\times$. Thus any character of (13) can be extended in a unique way to a character of $\mathbf{P}_D$. Thus the spectral decomposition of the space of automorphic forms with the ramification limited or prescribed by D is given by the set of characters of (13). These are, of course, just characters of $\prod_{x\in\mathfrak{d}} G_0(\mathcal{O}_x)/G_{a_x}(\mathcal{O}_x)$, constrained by the condition that they are 1 on the diagonally imbedded $\mathbb{C}^\times$.

Recall that an element of any one of the factors may be written as

$$\exp(\gamma_0 z + \gamma_1 z + \cdots + \gamma_{a_x} z^{a_x} + \dots),$$

where the expansion of the exponential is taken modulo z^{a_x+1}. The first coefficient is only determined modulo $2\pi i$. So (unitary) characters of the factor are of the form

$$\exp(\gamma_0 + \gamma_1 z + \cdots + \gamma_{a_x} z^{a_x} + \dots) \mapsto \exp\left(i\,\mathrm{Re}(\sum_0^{a_x} \gamma_i\beta_i)\right), \tag{16}$$

where $\beta_0, \dots, \beta_{a_x}$ are almost arbitrary complex numbers. Recall that z is here just a symbol. The element of $\mathbf{P}_D$ is given by $\gamma_0, \gamma_1, \dots$. The coefficient γ_0 is only determined up to multiples of $2\pi i$. In order that the character (16) be well defined, we need

$$2\pi\,\mathrm{Im}\,\beta_0 \in 2\pi\mathbb{Z}. \tag{17}$$

so that $\mathrm{Im}\,\beta_0 \in \mathbb{Z}$. To specify the place x to which the coefficients are attached, we write $\beta_i(x)$. Because we divide by the diagonally imbedded $\mathbb{C}^\times$, we also need

$$\sum_{x\in\mathfrak{d}} \beta_0(x) = 0,$$

where I have indicated explicitly a dependence on $x \in \mathfrak{d}$.

We have not done very well with the notation, but if we overlook that we see that the parameters of the connection defined by A and the characters are the same. For both, we specify a certain number of points. At each point we introduce the fractional part of A as

$$\frac{\alpha_{m_x}(x)}{(z-z_0)^m} + \cdots + \frac{\alpha_1(x)}{z-z_0}, \tag{18}$$

where I have added subscript and a parameter to emphasize that the number of terms and their values depend on x. For the characters, we have at each point the character given by $\beta_0(x), \dots, \beta_{a_x}(x)$. What I claim is that these parameters correspond. Since

$$\sum_x \beta_0(x) = \sum \alpha_1(x) = 0,$$

we take $m_x = a_x + 1$, $\beta_i(x) = a_{i+1}(x)$. This is reasonable and means that $m_x = a_x + 1$.

The only thing to clarify is (17). There is a constraint implicit in (15). For a flat connection with singularity, we demand, as we would for a connection without singularity that f be single-valued. This condition appears no doubt in the literature, but I have not come across it in the literature cited, perhaps because the emphasis is generally on the unramified case, perhaps for other reasons. It means that $2\pi\alpha_1 i$ is an integral multiple of $2\pi i$, so that α_1 is integral. This requires that α_1 in (15) be an integer.[6] So the relation between characters and integrals seems to be that $\alpha_{i+1}(x) = \pm\beta_i(x)$, where the sign is the same for all i and all x, but, for the moment at least, undetermined.

Observe that the constraint on α_1 is stronger than that on β_0. This conforms to what we have seen above. The automorphic or right side of the correspondence contains more than the motivic or Galois side on the left! So it is at best an injection from one to the other.

Thus at least for $X = \mathbb{P}^1$, we have introduced a correspondence in the original sense, that of automorphic representations, between ${}^LG$-bundles with connections — implicit here is the standard one-dimensional representation of ${}^LG = GL(1)$ — and automorphic characters — in the sense given to them in §5. The equation (14) is an extension of the equation (5) in which singularities are allowed.[7] So the $\mathcal{D}$-modules are already there, although they are defined in the context of ${}^LG$ not in that of G. Moreover they are not modules on Bun_G or on $\mathrm{Bun}_{{}^LG}$ but on X, so that they are not what is sought in [CFT]. Before turning to the correspondence in the sense of those notes, we examine briefly the case of a curve of arbitrary genus.

In fact the relation between connections, of this or that type, and the characters on the various sets $F_{\mathfrak{d}}^\times\backslash\mathbb{A}_{\mathfrak{d}}^\times$ is not what the a reading of [CFT] suggested to me. Suppose that in (15), which refers to the local coordinates at one point x, the higher order terms are 0, so that only $\alpha_1 = 2\pi n$ matters. We are dealing with a connection flat at x — in my understanding of the term — at x if and only if $n = 0$. On the other hand, if we modify the bundle by agreeing that a nonzero section is exactly of the form az^m, $a \neq 0$ and m a given integer, we are dealing with a connection flat at x if and only if $m = n$. On the right-hand side, the side of "idèle-class characters", the ramification is determined by the alternative $n = 0$ or $n \neq 0$, an alternative that is absolute. On the left-hand side, the side of "Galois representations" it loses to some extent its absolute character. For a curve of genus 0, we can modify the bundle (globally) in the way just described provided the sum of the various n_x that intervene is 0. For an arbitrary curve, provided that $\sum_x n_x \cdot x$ is a divisor linearly equivalent to 0.

For the field of rational functions, every adelic element is the product of a an element of $\tilde{F}^\times = I^{\mathbb{Z}}F^\times$ and an adelic element that is a unit everywhere, the two factors being unique up to an element of $\mathbb{C}^\times$. So an "idele-class character" is expressible uniquely as a character of $\prod_{x\in\mathfrak{d}} G_{a_x}(\mathcal{O}_x)$, trivial on the diagonal $\mathbb{C}^\times$ in $\prod_{x\in\mathfrak{d}}\mathbb{C}^\times$, the finite set $\mathfrak{d}$ being taken sufficiently large. There is no such easy decomposition in general so that characters are more difficult to describe.

[6] I am not entirely easy about this condition.

[7] Notice that the f of (5) has become the A of (14) and the g of (5) is now f.

As before we consider finite sets $\mathfrak{d} = (x_1, \dots)$ of distinct points and divisors $D = \sum_{x\in\mathfrak{d}} a_x$, $a_x \geq 0$, but we now fix the integers a_x and the cardinality of $\mathfrak{d}$, but we let the points $x_1, \dots$ vary. We have a homomorphism of the collection of "automorphic characters" to a direct limit

$$\varinjlim(\mathbb{C}^\times \backslash \prod_{x\in\mathfrak{d}} G_0(\mathcal{O}_x)/G_{a_x}(\mathcal{O}_x))\hat{}, \tag{19}$$

where the groups are groups of characters the limit is taken on increasing $\mathfrak{d}$ and increasing a_x. All arrows in the direct limit are injective.

In addition, we have the kernel of (19), which is the subgroup of unramified automorphic characters, which is, because of the presence of $I^{\mathbb{Z}}$ the character group of the group of divisors of degree 0 modulo principal divisors, thus the dual in the topological sense of $\mathrm{Pic}^0(X)$. So the group B of "idele-class characters" is imbedded in an exact sequence

$$\{1\} \to A \to B \to C \to \{1\},$$

where C is the group (19) and A is the group of unramified characters.

If we have a character of

$$\prod_{x\in\mathfrak{d}} G_0(\mathcal{O}_x)/G_{a_x}(\mathcal{O}_x)$$

then we introduce the fractional parts (18) as before. They serve as its parameters. For x not in $\mathfrak{d}$, they are taken to be 0. There may or may not be a function on X with these principal parts. So we introduce an auxiliary set of points $\mathfrak{d}'$, disjoint from d and auxiliary indeterminates $a_1(x)$ at $x \in \mathfrak{d}'$. If $\mathfrak{d}'$ contains at least 2d parameters, we can find a function with the correct principal parts outside of $\mathfrak{d}'$, but with — at first glance — pretty much uncontrollable principal parts at the points of $\mathfrak{d}'$ except that $a_i(x) = 0$, $i > 1$. In any case for $\mathfrak{d}'' = \mathfrak{d} \cup \mathfrak{d}'$, we have a collection of principal parts that are arbitrary on $\mathfrak{d}$

It is best to begin by clarifying for ourselves the notion of line bundles with connection. We represent line bundles on X as modifications of the trivial bundle. There is a redundancy in this representation. A modification, thus the decision to allow a pole of at most a given order at a finite number of points and to demand a zero of a certain order at certain, is given by a divisor. Two divisors that are linearly equivalent, give rise to isomorphic bundles. Sections of one are sent to sections of the other by multiplication with a fixed meromorphic function. Connections on the bundle are just meromorphic forms, thus connections on the trivial bundle, but with a different notion of admissibility and, in a particular a different notion of singularity.

I repeat these observations giving them a slightly different cast. I begin locally. The modification means that a section of the modified bundle is $f' = gf$, where g is fixed and f is a section of the original bundle. The sections f and f' do not have the same singularity at the points where g has a zero or pole. A connection $f \to df/dz + Af$ is replaced by

$$f' \to \frac{df'}{dz} + (A - 1g\frac{dg}{dz})f',$$

so that A is modified to

$$A' = A - \frac{1}{g}\frac{dg}{dz}.$$

. The principal part of the form $A'dz$ is modified by the subtraction of dg/g. To a given representation of a bundle in this way (any bundle) and to any realization of the collection as the principal parts of a function (unique up to an additive constant), we have a possible "motivic" or "Galois" object in our context. There are two things to do. Match the parameters on the two sides and then interpret the matching.

The matching of principal parts is clear. The only surprising discovery is that on the left-hand side, they do not all occur.This appears to be a confirmation of the existence of a phenomenon to which we have already drawn attention. Not all parameters on the motivic side appear on the automorphic side! It is not a very deep restriction As we saw earlier, at a point the connection has the form

$$A'(z) = \frac{\alpha_m}{(z-z_0)^m} + \cdots + \frac{\alpha_1}{z-z_0} + \ldots.$$

As we turn around the point on the path determined by connection the function is multiplied by $\exp(2\pi\alpha_1)$. The connection is therefore not well defined unless α_1 is an integer.

On the other hand the bundle is given by modifying the trivial bundle by the introduction of a finite, but equal, number of zeros and poles. We may even permit an intersection $\mathfrak{d} \cap \mathfrak{d}'$ that is not empty. This modification entails a change in $A(z)$ in (14), namely $f(z)$ is replaced by $g(z)$ with $f(z) = (z-z_0)^a g(z)$, where a is an integer, positive or negative according as to whether we demand a 0 or allow a pole at z_0. Then

$$\frac{df}{dz} = a(z-z_0)^{a-1}g(z) + (z-z_0)^a\frac{dg}{dz},$$

so that

$$\frac{dg}{dz} = \left(A(z) - \frac{a}{z-z_0}\right)g(z) = B(z)g(z), \qquad B(z) = A(z) - \frac{a}{z-z_0} \tag{20}$$

So the coefficient of the residual term is modified by an integer. Thus the integrality condition on the coefficient α_0 is preserved under modifications.

I also recall that there is a familiar classical auto-duality: between the group of classes of line bundles of degree 0 on a curve and its topological dual, each in one way or another a realization of the Picard group $\mathrm{Pic}^0(X)$. I shall return to these questions, especially the matching of principal parts, again later.

So we have two sets, actually two abelian groups. That is clear for the group of "idele-class characters". We can also add principal parts and take the tensor product of two line bundles, so that we do in fact have a group on the "Galois" side. They contain isomorphic subgroups. The group of unramified characters on the right; the group of all line bundles of degree 0 on the left, with the principal part taken to be 0. Notice that the corresponding connections are not nonsingular! They may have a principal part of degree 1 with an integral coefficient. The sum of these coefficients is constrained to be 0. They also have related quotients, one is a subgroup of the other. The group (19) on the one hand; on the other the group formed by the admissible principal parts, thus the sum of the residues must be 0 and the residues integral. The second may, as we have seen, be regarded as a subgroup of the first.

There is, however, something to be remarked here. The division on the left by divisors of functions has to performed on the two factors, namely both on the line bundles of degree 0 realized as modifications of the trivial bundle and on the principal parts, as in (20). If we confine ourselves to principal parts with integral residues, there is a way to remove this ambiguity. We modify — if possible! — the bundle in such a way that, on using (20), we reduce the residues to 0. This done, the bundle can no longer be modified. What we are left with are integrals of the second kind. Since we are confining ourselves to principal parts with integral residues, these residues define a divisor. If we modify the integral residues, we have to modify the bundle. This means only that we introduce a meromorphic function and replace sections by their products with this function. So the bundle together with the residue is determined by two things: the linear equivalence of the divisor of a meromorphic section of the bundle, and within the class of equivalent bundles, the meromorphic function used to pass from one fixed one (with residues taken to be 0) to another, the same bundle but new residues, subject to the constraint of being given by the zeros and poles of a meromorphic function.

I repeat this, because it has taken me a certain time to arrive at the necessary clarity. For a given X, we start with the trivial bundle. All other bundles $\mathcal{L}$ are to be regarded as obtain from it by a modification. Thus $\mathcal{L}$ is, to some extent, to be thought of a the trivial bundle, but with a modified notion of zero and pole. The second object is a form ω, whose principal part is, from one point of view, the principal part as a form on the trivial bundle, from another point of view, the principal part on the modified bundle. Thus if the modification is by the divisor $\mathfrak{a}$ — always taken to be of degree 0 — and a section f (perhaps local) has a divisor $\mathfrak{d}_f$ as a function, then its divisor as a section is $\mathfrak{d}_f - \mathfrak{a}$. If we are dealing with local sections, the divisor is to be interpreted locally. If the first components of the two pairs $(\mathcal{L}, \omega)$ and $\mathcal{L}', \omega')$ are isomorphic, then the two divisors $\mathfrak{a}$ and $\mathfrak{a}'$ defining the bundles are linearly equivalent. We have $\mathfrak{a}' - \mathfrak{a} = \mathrm{div}(h)$. The function h is determined up to a constant. The pairs are equivalent if in addition

$$\omega' - \omega = \frac{dh}{h}.$$

I point out once again that for all three forms in this relation the residues at any point are integral.

Clarity begins to arrive.

At this point I began to appreciate that there is a better way of viewing matters if one uses the theorem attributed to Weil on p. 242 of [GH]. Indeed the explanations in the preceding lines have been modified so that we can more readily profit from the insight it offers. We will need not only the theorem itself, but some improvements on it, which I record in the order in which they forced themselves on my attention.

Theorem. *Suppose f and g are meromorphic functions on the compact Riemann surface such that the set of zeros and poles of f is disjoint from that of G. Then*

$$\prod_p f(p)^{\mathrm{ord}_p(g)} = \prod_p g(p)^{\mathrm{ord}_p(f)}.$$

Using this theorem, we associate to g a character of the ideles equal to 1 on $F^\times$. If a is an idele, we may write $a = fb$, where f is a function and b has its zeros and poles outside the set of zeros and poles of g. Then

$$\prod_x \frac{b(x)^{\mathrm{ord}_x(g)}}{g(x)^{\mathrm{ord}_x(b)}} \tag{21}$$

is, by the theorem, independent of the choice of b and defines an idele-class character ξ_g. So there was a mistake in the point of view. We should not be representing our elements locally by the order of their zero and by an element of $G_0(\mathcal{O}_x)/G_a(CalO_x)$ but by their residue modulo $G_1(\mathcal{O}_x)$, thus by a factor $c(z - z_0)$, $c \in \mathbb{C}$. On reflection, this has an obvious advantage. It is independent of the choice of uniformizing parameters.

Consider the restriction of ξ_g to those ideles that are units everywhere. I claim that it is trivial if and only if g is a constant function. According to (21), $\xi_g : b \mapsto \prod b(x)\,\mathrm{ord}_x(g)$. if b is a unit everywhere. This can be 1 for all b if and only if g has neither zeros nor poles. On the other hand, we see also that we only obtain those characters on $\prod G_0(\mathcal{O}_x)/G_1(\mathcal{O}_x)$ associated to divisors that are linearly equivalent to 0. Notice moreover that we do not obtain unitary characters. For unitary characters, we have to take $\chi_g = \xi_g/|\xi_g|$. It is unitary characters with which we are concerned, but the principles are the same $g \mapsto \chi_g$ is injective — or rather χ_g is trivial if and only if g is a constant — and the constraint on the image — in the notation of (16) is that the divisor —

$$\sum_x \frac{\beta_0(x)}{2\pi}$$

is linearly equivalent to 0. The number $\beta_0(x)/2\pi$ is the order of the function g at x. So the group of automorphic representations or (unitary) characters that are at most tamely ramified, thus with all $\beta_i(x)$, $i > 0$, equal to 0, and that satisfy the integrality condition (17)— and with the supplemental constraint that they are 1 on $I^{\mathbb{Z}}$ — contains the product of two subgroups with only the trivial character in common: the group of unramified characters trivial on $I^{\mathbb{Z}}$ and the group of characters χ_g. The second group is very large but discrete. The quotient of the full group by this subgroup is isomorphic to $\mathbb{Z}^{2g}$, where g is the genus of the curve. (I apologize for the conflicting notations!)

I observe that I am trying to describe the nature of the possible reciprocity in this context: geometric theory and $G = GL(1)$. With the arithmetic case, for which the global field is a number field, even $\mathbb{Q}$, in mind, we are prepared that the "motivic" side is smaller than the "automorphic". The missing tamely ramified characters are a sign of this. In fact, too strong a sign. We can do better than the Weil theorem, and with the same proof, taken from [GH], whose authors observe that the Weil theorem is, in any case, just one of several possibilities

Suppose we have a differential form ω on X whose singularities are all at most of first order, thus of the form a_x/x. We take the a_x to be integral. We want then to associate to ω a character on the idele-class group. We follow the notation, in so far as it is appropriate, on p. 229 and on pp.242-243 of [GH]. In particular, we follow it for the usual description of cycles $\delta_1, \ldots, \delta < g, \delta_{g+1}, \ldots, \delta_{2g}$ that display the surface as a planar polygon with sides identified. We have a function f with poles and zeros at p_i and a form ω (the form dg/g of [GH]) with first-order poles at q_j, the p_i and the q_j lie in the interior of the planar region and we join the p_i to a common point p on the boundary on curves avoiding the q_j. The sets $\{p_i\}$ and $\{q_j\}$ are taken to be disjoint. What we want to prove is that

$$\lambda = \sum_j \mathrm{res}_{q_j}(\omega) \log f(q_j) - \sum_i \mathrm{ord}_{p_i}(f) \int_p^{p_i} \omega \in 2\pi i \mathbb{Z}. \tag{22}$$

Since $\sum_i \mathrm{ord}_{p_i}(f) = 0$ and $\phi(p_i) = \ln \int_p^{p_i} \omega$ is well-determined up to a constant independent of p_i, the possible ambiguities, for example in the choice of the base-point p, have no effect on this relation.

Think of the set $\{p_i\}$ with the various ord_{p_i} as defining a finite collection of pairs of points on X, the collection $\{a_j, b_j\}$, where ord_{p_i} is the number of times p_i occurs as an a_j minus the number of times it occurs as a b_j. The second sum in (22) may be interpreted as the sum of the integrals of ω along paths from a_j to b_j, the paths being chosen to conform to the geometry of the planar representation.

The relation (22) once proved we can associate to ω an idele-class character, by taking any idele and writing it as the product $f_1 f_2$ of a principal idele f_1 and an idele f_2 that has neither zero nor pole at the places q_j and then sending the idele to $\exp \lambda$, where λ is defined by (22) with $f = f_2$. The result will be well defined.

As in [GH], we introduce the form $\varphi = \log(f)\omega$, where we have substituted for $dg/\ln g$, and integrate over the boundary of the region. By the residue theorem, this integral is given by

$$\begin{aligned} \int_{\partial\Delta'} \varphi &= 2\pi i \sum_{q_j} \mathrm{Res}_{q_j}(\varphi) \\ &= 2\pi i \sum_{q_j} \mathrm{res}_{q_j}(\omega) \log f(q_j). \end{aligned}$$

We collect terms as in [GH]. First of all, for identified pairs p, p' on the arc δ_i and on the inverse arc δ_i^{-1},

$$\log f(p') = \ln f(p) + \int_{\delta_{g+i}} d \log f,$$

so that

$$\int_{\delta_i+\delta_i^{-1}} \varphi = \left(\int_{\delta_i} \omega\right)\left(-\int_{\delta_{g+i}} d\log f\right).$$

In the same way,

$$\int_{\delta_{g+i}+\delta_{g+i}^{-1}} \varphi = \left(\int \delta_{g+i}\omega\right)\left(-\int_{\delta_i} d\log f\right).$$

Moreover for identified points $p \in \alpha_i$ and $p' \in \alpha$,

$$\log f(p') - \log f(p) = -2\pi i \,\mathrm{ord}_{p_i}(f).$$

so that[8]

$$\int_{\alpha_i+\alpha_i^{-1}} \varphi = 2\pi i \,\mathrm{ord}_{p_i}(f) \int_{s_0}^{p_i} \omega = 2\pi i \phi(p_i).$$

As in [GH], the conclusion is that

$$2\pi i \sum_j \left(\mathrm{res}_{q_j}(\omega) \log f(q_j) - \sum_i \mathrm{ord}_{p_i}(f) \int^{p_i} \omega\right)$$

is equal to

$$\sum_{k=1}^{g} \left(\int_{\delta_k} d\log f \cdot \int_{\delta_{g+k}} \omega - \int_{\delta_k} \omega \cdot \int_{\delta_{g+k}} d\log f\right). \tag{23}$$

In (23), all four integrals are of functions all of whose residues are integral and all integrals are over closed curves. The conclusion is, as in [GH]¡ that the sum is an integral multiple of $(2\pi i)^2$. The relation (22) follows.

Advantage. Since they seem to be difficult for me to retain, I want to record the reasons for passing from the collection g, or better dg/g, to differentials ω with only poles of order 0 and integral residue. Each dg/g is a possible, but there are more ω. Indeed, the possible dg/g are limited by the condition that $\sum a_x x$, where a_x is the residue of dg/g is linearly equivalent to 0. So we have a constraint that entails the restriction to a subvariety of codimension $2g$, g is the genus of X, in a variety of infinite dimension. There is no such restriction on the divisors $\sum a_x x$ if we allow ω. Indeed the lemma on p.233 of [GH] states that there is no restriction at all on the possible residues of ω, except that the sum be 0. This is a constraint that we impose on the collection a_x in any case, but, at first, for different reasons. It is because characters are to be trivial on the diagonally imbedded $\mathbb{C}$.

[8] In the diagram of [GH], δ_0 is meant to be s_0, an arbitrarily chosen point on the boundary of the planar region. I have denoted it above by p.

Supplement. In fact the assertion (22) can be extended to an arbitrary differential ω and to a function whose zeros and poles are disjoint from the singularities of ω. The number λ is then replaced by

$$\lambda = \sum_j \mathrm{res}_{q_j}\{\omega \log f(q_j)\} - \sum_i \mathrm{ord}_{p_i}(f)\int_p^{p_i}\omega \in 2\pi i\mathbb{Z} \tag{24}$$

The proof is the same as before. The calculation remains valid, but the relation between $\log f$ and the residue of $\omega \log f$ involves the higher order terms of the expansions of f. This allows us to attach to ω a character of the idele-classes. If $h = \prod_x h_x$ is an idele, say of norm 1, thus $\prod_x |h_x| = 1$, it being understood that $|h_x| = 1$ for almost all x, we associate to ω the character defined on elements whose set of zeros and poles is disjoint from the set of zeros and poles of ω, and thus by (24) on all ideles by

$$h \mapsto e^{\lambda}, \qquad \lambda = \sum_j \mathrm{res}_{q_j}\{\omega \log h_{q_j}\} - \sum_i \mathrm{ord}_{p_i}(h)\int_p^{p_i}\omega \tag{25}$$

If the principal part of ω at a given point q with local coordinate z

$$\frac{\alpha_m}{z^m} + \cdots + \alpha_1 z$$

and

$$\log h_q = \beta_0 + \beta_1 z + \ldots,$$

the contribution of q to the first sum of (25) is

$$\alpha_1\beta_0 + \alpha_2\beta_1 + \ldots \alpha_m\beta_{m-1}$$

A curious, even troubling, point. Take the differential ω in (24) to be regular. Then the first sum is 0. As we observed when discussing (22), the second may be interpreted as the integral

$$\sum_j \int_{a_j}^{b_j}\omega, \tag{26}$$

where the divisors $\sum_j a_j$ amd $\sum_j b_j$ are linearly equivalent, but otherwise arbitrary. The collection of paths from a_j to b_j is however strongly constrained. This is just as well, because the ω that can appear in (26) vary over a continuous family, the family of differentials (or integrals) of the first kind. So (26) can be an integral multiple of $2\pi i$ for all ω if and only if it is identically 0.

This is clear enough for curves of genus one. Take the development of the Riemann surface X to be that onto a parallelogram defined by a basis of the lattice defining it. Take a_j and b_j all to be in the plane. If they are all close to each other, then

$$\sum_j \int_{a_j}^{b_j} dz = \sum_j (b_j - a_j) = 0.$$

Somehow the geometry forces us to move the points so as to preserve this relation. My intuition fails me, as it fails me in higher genus.

A simple, but important, point. Suppose we take the form ω to be dg/g. Then for the λ in (24) the value of $\exp\lambda$ is

$$\prod_{q_j} f(q_j)^{\mathrm{ord}_{q_j} g} / \prod_{p_i} g(p_i)^{\mathrm{ord}_{p_i} f}.$$

So, as we already know, Weil's theorem is a particular expression of (24).

The structure of the group of idele classes. We can introduce a filtration with four stages, the elements of the filtration having a variable character in order to accomodate the fact that characters are a product $\prod_x \chi_x$, where χ_x is trivial for all but a finite number of x.

(i)The smallest element C_4 of the filtration is a group

(27) $$\mathbb{C}^\times \prod_{x\in\mathfrak{d}} G_{a_x}(\mathcal{O}_x) \prod_{x\notin\mathfrak{d}} G_0(\mathcal{O}_x),$$

where $\mathfrak{d}$ and the integers $a_x \geq 1$, $x \in \mathfrak{d}$, are taken as large as appropriate. This is also the first quotient. Each character will be 1 on a subgroup of this sort.

(ii) The next element C_3 of the filtration is the collection of groups

(28) $$\mathbb{C}^\times \backslash \prod_{x\in\mathfrak{d}} G_1(\mathcal{O}_x)/G_{a_x}(\mathcal{O}_x) \prod_{x\notin\mathfrak{d}} G_0(\mathcal{O}_x).$$

Since we are dividing by the group (27) the second quotient is

$$\prod_{x\in\mathfrak{d}} G_1(\mathcal{O}_x)/\mathbb{C}^\times \prod_{x\in\mathfrak{d}} G_{a_x}(\mathcal{O}_x).$$

(iii) The next element of the filtration C_2 is

$$\prod G_0(\mathcal{O}_x)$$

and the third quotient

(29) $$\prod_{x\in\mathfrak{d}} G_0(\mathcal{O}_x)/G_1(\mathcal{O}_x)/\mathbb{C}^\times.$$

(iv) The final group is $C = C_1$ and its quotient is a restricted product,

(30) $$\textstyle\coprod_x G(F_x)/I^{\mathbb{Z}} \prod_x G_0(\mathcal{O}_x)$$

We take the last element of the filtration to be the full group $C = C_1$ of all idele-classes.

The first pairing. We want to introduce a pairing on the product $C \times C$ such that:
(i)it is multiplicative in both variables $(b_1 b_2, b_2) = (b_1, b_3)(b_2, b_3)$ and skew-symmetric, $(b_1, b_2)(b_2, b_1) = 1$;
(ii) if b_1 and b_2 are in $F^\times$ then $(b_1, b_2) = 1$;
(iii) if the divisors of f and g are disjoint then

$$(f, g) = \frac{\prod_{x \in \mathfrak{d}_g} f_x(0)^{\operatorname{ord} g_x}}{\prod_{x \in \mathfrak{d}_f} g_x(0)^{\operatorname{ord} f_x}};$$

(iv) if f and g are principal then $(f, g) = 1$.

We have already seen in the comments following formula (21) that there is a unique way to define (f, g) if one or the other of the arguments is principal. The same argument applies in general. Let $f = f_1 f_2$, $g = g_1 g_2$, where f_2 and g_2 are principal and f_1 and g_1 have no poles in common. Then $(f, g) = (f_1, g)(f_2, g_1)(f_2, g_2)$ and all three factors are well-defined by the conditions (ii) and (iii). It is also easily verified that the bilinear form is skew-symmetric.

This pairing has an important property. Suppose $f = \prod_v (c_v)$, $c_v \in \mathbb{C}$. Then for any idele h the value of (f, h) is $\prod_v (c_v)^{\operatorname{ord} g_v}$.

A second pairing This is a pairing between meromorphic integrals and ideles h of degree 0. It is given by $\exp(\lambda)$, where λ is defined by (25), it being understood that h is so chosen that its poles and zeros are disjoint from the singularities. This is possible because we can always choose modify h by a principal idele so that the resulting idele has a divisor disjoint from the set of singularities of ω. The relation (24) guarantees that the value of the resulting $\exp(\lambda)$ is independent of the choice of modification.

The correspondence. It is the theorem credited to Weil, but in its supplementary form, that is the key to constructing the duality. We have effectively identified connections with integrals or differential forms, the notion of principal part being somewhat different in the two cases, because our line bundles $\mathcal{L}$ are modifications of the trivial bundle, so that $\mathcal{L}$ denotes not merely a line bundle but a divisor of degree 0. The modification affects only the residual terms, thus the terms α_1 in (15). The condition that α_1 be integral is not affected by the addition of a form dg/g. It is clear from the Riemann-Roch thbeorem that the higher order terms (in (15), α_i, $i > 1$) in the principal parts of a differential can be freely prescribed. From the lemma on p. 233 of [GH] it is clear that the residues can also be freely prescribed. As observed, we take them to be integral.

The "motivic" side will be defined by a group. This group consists of pairs (h, ω), in which h is an idele of degree 0 taken modulo ideles of the form $\prod h_v$ where the h_v are in $\mathcal{O}_v$ for all v and equal to 1 at the center of the valuation v. Two pairs (h, ω) and (h', ω') are taken to be equivalent if $h' = hg$, $\omega' = \omega - dg/g$. We may associate to a pair (h, ω) a line-bundle, the bundle obtained from the trivial bundle by the modification specified by the divisor of h. The form ω is a form for the trivial bundle and thus, itself unchanged, a form on the modified bundle. Its principal parts are different on the modified bundle because

$$g^{-1}\frac{dgf}{dx} = \frac{df}{dx} + \frac{f}{g}\frac{dg}{dx}.$$

Thanks to the first and the second pairings both factors define forms on ideles of norm 1 modulo principal ideles. Therefore the pair does as well.

$$(h,\omega):\ f \mapsto (h,f)(\omega,f). \tag{31}$$

The definitions are such that this is well-defined. That equivalent pairs give equal characters is a consequence of (24) — if I have managed to achieve correct signs. The pairs (h,ω) form a group. The pairing is bilinear.

Our calculations of the local contributions to (25) show that the arbitrariness of the principal parts, except for the residues, allows us to find an ω whose restriction to the first element of the filtration is arbitrary. So what we have to do is examine the image of the pairs $(h,\omega$ for which ω is a differential of the third kind.

The structure of the space of perhaps singular connections on X. This section is very rough. Although we have already prepared ourselves for describing the duality between two groups — (i) ideles of norm 1 modulo principal idele; (ii) the group of pairs (h,ω) — we have not done it very well. So our discussion here will be clumsy. Thanks to a well-known theorem [Le], the differentials of the second kind have arbitrary singularities. So they can be used to provide for any character of the quotient C_3/C_4. The differentials of the third kind — with integral residues — can, by a theorem in [GH] alread cited, provide any character on the group of elements $\prod_v C^\times$ — a group that generates C_2 modulo C_3 — of the form $\prod c_v \mapsto \prod c_v^{n_v}$, n_v integral and almost all $n_v = 0$. These are the only characters we can expect to obtain. They certainly form a large class.

The group C_1/C_2 is an infinite discrete group and its characters are of the form

$$\prod g_v \to \prod c_v^{\operatorname{ord} g_v}, \qquad c_v \in \mathbb{C}^\times$$

The g_v are arbitrary integers, except that all but a finite number must be 0, but subject to the condition $\sum g_v = 0$. The c_v are arbitrary. These characters are given by elements $h = \prod_v h_v$ with all $h_v \in \mathbb{C}_v$. The full group of h modulo the subgroup of ideles with no zeros and no poles is of course the set of points on the jacobian, but as a discrete(!) group. So the duality is a little bizarre but seems to be otherwise natural.

We have already made various remarks about the h and the ω that appear as the first or second elements of the pairs (h,g). At the cost of repeating myself, I make a few further observations. The line bundle defined by (h) is of degree 0 and may be obtained as a modification of the trivial bundle. The meromorphic one-form has principal parts, but these are generally not arbitrary. By the lemma on p. 233 of [GH], every such form may be written as the sum of a form of the third kind, thus a form whose principal parts contain only first terms $a/(z-z_0)$ and forms of the second kind, thus forms with no residues. The question of whether the forms of the second kind have single-valued integrals is not pertinent here. What is important is that their principal parts are arbitrary. This is an immediate consequence of the Riemann-Roch theorem and the lemma on p.233 of [GH].

The group law is that line bundles are multiplied, thus the modifications, allowing a larger order of pole, or demanding a zero of some order, which are determined by a collection of integers $\{a_x\}$, are multiplied and their coefficients added. The forms are then modified. I was at first inclined to take not the pairs (h, ω) but the pairs $(\mathcal{L}, \omega)$ consisting of a line bundle $\mathcal{L}$ defined by the modification of the trivial bundle given by h and a form ω. If $\mathcal{L}'$ is obtained from $\mathcal{L}$ by the modification given by a meromorphic function g on X and ω' is $\omega - dg/g$, then $(\mathcal{L}', \omega')$ is equivalent to $(\mathcal{L}, \omega)$. Dividing by this equivalence relation, we obtain the $GL(1)$-bundles on X with connection, with singular points that may be regular and irregular. Unfortunately, this does not seem to work. We need h

There are several subgroups to be introduced.There is a subgroup B_2 invariant under the equivalence relation, those pairs for which ω has at most a poles of the first order, thus for which ω is a form of the third kind. This subgroup contains an even smaller subgroup B_3, those for which ω is of the third kind and the modification is zero, so that the bundle itself is trivial. The quotient of the two groups is the group of divisors of degree 0 modulo linear equivalence. There is a second group B_3', the group formed by pairs $(\mathcal{L}, 0)$, with an arbitrary bundle but the connection taken from the trivial bundle. As a connection on $\mathcal{L}$, it will have poles of order 1 at the points of modification. Certainly $B_3 \cap B_3'$ is the trivial group. It contains pairs An element of $(\mathcal{L}, 0)$ equivalent to $(\mathcal{L}_0, dg/g)$ if $\mathcal{L}_0$ is the trivial bundle and g is a function whose divisor determines the modifications defining $\mathcal{L}$.

By the Abel-Jacobi theorem the group C_1/C_2 is, as an abstract group, the jacobian of the curve. As a discrete group, the dual group formed by its unitary characters is very large. As the jacobian of the curve, thus as a $2g$-dimensional torus, the group of characters is much smaller, the direct sum of $2g$ copies of $\mathbb{Z}$.

There still remains the matter of an interpretation more along the lines of [CFT] or [GT], in which Hecke eigensheaves and $\mathcal{D}$-modules on $\mathrm{Bun}_G(X)$ play a more prominent role. Whether these alternative interpretations are better or worse is not yet clear, but it will, I suppose, be worthwhile to have them at hand when attempting to understand the contributions of the physicists and to assess the mathematical insight they offer. I hope that this can be done and that a — certainly much more complicated — discussion of the general group can be carried out in a representation-theoretic context along the lines of this section. So far as I know, the necessary representation theory for the Lie groups that arise is not available. For the moment, however, I have to turn to other matters.

9. Appendix.

Structures.. It may be best to review once again the structural background that one is attempting to establish. There are two slightly different structures, both of a motivic and therefore categorical nature. that appear for each of the six possible fields F, whether local or global. I consider first A-structures, $\mathfrak{A}$, the A suggesting representation theoretic or automorphic. Present in these structures is the collection of L-groups ${}^L G_K$, K being a Galois extension of F, large enough to accomodate the necessary action of $\mathrm{Gal}(K/F)$ on the connected component of ${}^L G$. The objects are pairs $({}^\lambda H, \pi_H^{\mathrm{st}})$. The index λ indicates, as I have observed in various places, that we need more liberty in our choice of G than the condition that G_{der} be simply connected, necessary for the effective use of the trace formula, allows. The class π_H^{st} is a stable conjugacy of automorphic representations of a reductive group H over F of which ${}^\lambda H$ is essentially the L-group. Given an object $({}^\lambda H, \pi_H^{\text{st}})$ and an admissible homomorphism

$$\phi : {}^\lambda H \to {}^L G_K, \tag{32}$$

functoriality associates to it a stable family of representations of automorphic representations or, locally, of irreducible representations of the group $G(\mathbb{A}_F)$ or $G(F)$, the representations being restricted — in principle, but there are theorems to demonstrate — to those introduced by Arthur, thus almost to tempered representations. In particular then, for a given ρ and each representation

$$\rho : {}^L G_K \to GL(n, \mathbb{C}), \tag{33}$$

we have a representation, automorphic or local, attached to $\rho \circ \phi : {}^\lambda H \to GL(n, \mathbb{C})$.

It is perhaps better, at this point, just to take ${}^L G_K = GL(n)$ in (32). We denote, recalling that stable classes for $GL(n)$ contain only a single element, the associated representation by π_ϕ. For the ϕ under consideration we have direct sums $\phi_1 \oplus \phi_2$ and direct products $\phi_1 \otimes \phi_2$, so that the collection of ϕ given in this way as well as the collection of π_ϕ form a Tannaka-like category. It is associated to the diagrams (32).

There is another point that I find not so much strange as difficult to keep in mind. Either locally or globally one single stable class of irreducible representations (L-packet) yields a reductive complex group, and thus all its representations. Another stable class is either related to this class by functoriality, or the two classes have a common image under homomorphisms of L-groups, or they give rise under functoriality to basically disjoint collections of stable L-packets, and thus to basically disjoint substructures of the A-structures. The M-structures are of a completely different kind. The individual elements, namely the individual groups in the associated *motivic* group — usually a kind of product — are difficult to isolate, whereas the other elements, their representations are given. For the A-structures associated to automorphic forms, both kinds of individual elements, first of all the hadronic automorphic representations π^{st} and the associated ${}^\lambda H$, secondly those associated to the ϕ of (32) are (will be!) difficult to construct. For both, the attack suggested in §3 entails a mastery, not yet achieved, of the trace formula.

From this point of view the simple factors of ${}^\lambda H$ represent isolated factors of the corresponding structure, but the abelian part is more elusive. There are many possible relations. Consider the local context, over F_v, for the group $GL(1)$. The irreducible representations are the unitary characters $\pi = \chi$. The L-packets all contain a single-element. So in principle, on the M-side of the correspondence we have a free group generated by the characters. However, the possible relations $\chi_1^{a_1} = \chi_2^{a_2}$ as possible relations among the images under the corresponding ϕ_1 and ϕ_2 in (32) have to be imposed. For a group like $SL(2)$, there are relatively few relations of this kind that arise, certainly not among those π^{st} for which ${}^\lambda H_{\pi^{\text{st}}}$ is $SL(2)$. So the group for the M-structure is of quite a different kind. It is, as I said, best to take a finite number of π^{st} at a time.

I point out that even when ${}^\lambda H$ is abelian, and thus a topological group the group on the side of the M-structures does not, in this picture, acquire a topological character. It is just the dual group, in the sense of abelian topological groups, deprived of its topology.

In general a Tannaka category is associated to a group or to a group-like object and to a field Φ, often but always $\mathbb{Q}$ or a finite extension of it. The field can be extended to $\mathbb{C}$, and this will be the only possibility in the categories constructed by functoriality. The category is constructed from representations of the object by taking their sums and products. Since we can take direct products of groups, we can build more and more complicated objects. Categories of this type were employed, conjecturally, by Grothendieck to decompose algebraic varieties as generators of spaces of cohomology or of L-functions into their fundamental components. This was his notion of *motive*. He considered all motives — or all motives over a given ground field F — at once, so that the groups were enormous. This is not necessary or, in my view, even desirable. We could consider the motives generated by a finite number of varieties, thus, allowing disconnected varieties, by a single variety. The resulting groups would be finite-dimensional algebraic groups, perhaps even reductive, although certainly not always connected. I add, as an aside, that we have already observed that it may be appropriate for the purposes of reciprocity to enlarge the notion of motive used by Grothendieck to include, for example, whatever the vector bundles appearing in the theory of Shimura varieties represent.

I also recall that even the category that Grothendieck tried to define has, in contrast to those to be associated to automorphic forms or representations, the associated field $\Phi = \mathbb{Q}$. When comparing this second type of category, which we may call an M-structure, $\mathfrak{M}$, with the first type, an A-structure, we have to enlarge the field appropriately. Anyhow, after an appropriate enlargement, reciprocity is to be some form of functor, $\mathfrak{M} \to \mathfrak{A}$. The M is meant to suggest *motive* and it is likely that whatever the category is for fields F of the second and third kind it will not be so different from that associated to algebraic number fields.

As I have tried to indicate in these notes, an adequate theory is a long way off, but there are aspects of it that, less forbidding and more accessible than others, it would well be profitable to investigate now. This is, however, no reason for a large number of people to spend a great deal of time or energy on peripheral or, in the long term, fruitless topics. Accessibility is not the only criterion.

******

For the references to my own papers, it is simplest to turn to

http://publications.ias.edu/rpl

References

[BBD] Beilinson, Bernstein, Deligne.

[CL] Earl A. Coddington and Norman Levinson, *Theory of ordinary differential equations.*

[QFS] P. Deligne et al., *Quantum fields and strings: A course for mathematicians.*

[CFT] E. Frenkel, *Lectures on the Langlands program and conformal field theory.*

[CLG] ———, *Langlands correspondence for Loop groups.*
[GT] ———, *Gauge theory and Langlands duality*, Sém. Bourbaki, 61ème année **1010** (2008—2009).
[FG] E. Frenkel and B. Gross, *A rigid irregular connection on the projective line.*
[FLN] E. Frenkel, R. Langlands, and Ngô Bao Châu.
[GH] P. Griffiths and J. Harris, *Principles of algebraic geometry.*
[G] Benedict H. Gross, *Irreducible cuspidal representations with prescribed local behavior.*
[Pr] *Problems in the theory of automorphic forms.*
[STF] *Les débuts d'une formule des traces stable.*
[BE] *Beyond endoscopy.*
[H] *Review of Hida's book.*
[SP] *Reflexions on receiving the Shaw prize.*
[ST] *Singularités et transfert.*
S. Lefschetz, *Algebraic geometry.*
[LP] M. Larsen and R. Pink, *Determining representations from invariant dimensions*, Invent. Math. **102** (1990), 377398.
[N] Ngô Bau Chaô, *lemme fondamental.*
[Na] D Nadler, *Microlocal branes are constructible sheaves.*
[NZ] D. Nadler and E. Zaslow, *Constructible sheaves and the Fukaya categories.*
[S] Jean-Pierre Serre, *Groupes algébriques et corps de classes.*
[Wa] Jean-Loup Waldpsurger.
[WS] Wang Song, *http ://www.mcm.ac.cn/people/ws.aspx.*