Pacific Journal of Mathematics

A PROLOGUE TO "FUNCTORIALITY AND RECIPROCITY" PART I

ROBERT LANGLANDS

Volume 260 No. 2 December 2012

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 2, 2012
dx.doi.org/10.2140/pjm.2012.260.583

A PROLOGUE TO "FUNCTORIALITY AND RECIPROCITY" PART I

ROBERT LANGLANDS

In memoriam — Jonathan Rogawski

Contents of Part I

Contents of Part II

1. Introduction

The first of my students who struggled seriously with the problems raised by the letter to André Weil of 1967 was Diana Shelstad in 1970–74, at Yale University and at the Institute for Advanced Study, who studied what we later called, at her suggestion and with the advice of Avner Ash, endoscopy, but for real groups. Although endoscopy for reductive groups over nonarchimedean fields was an issue from 1970 onwards, especially for SL(2), it took a decade to arrive at a clear and confident statement of one central issue, the fundamental lemma. This was given in my lectures at the École normale supérieure des jeunes filles in Paris in the summer

MSC2010: 11FXX.
Keywords: functoriality, reciprocity, Langlands program.

of 1980, although not in the form finally proved by Ngô, which is a statement that results from a sequence of reductions of the original statement. The theory of endoscopy is a theory for certain pairs (H, G) of reductive groups. A central property of these pairs is a homomorphism $\phi_G^H : f^G \mapsto f^H$ of the Hecke algebra for $G(F)$ to the Hecke algebra for $H(F)$, F a nonarchimedean local field, and the fundamental lemma was an equality between certain orbital integrals for f^G and stable orbital integrals for f^H.

Clozel observed at an early stage that, with the help of the trace formula, it was sufficient to treat the case that f^H and f^G were both the unit element in the respective Hecke algebras. Other more difficult reductions came later, but in the seventies it was the fundamental lemma in a raw form, but for specific groups, that I proposed as a problem to a number of students who worked with me at the Institute for Advanced Study, although their formal advisors were elsewhere because the IAS had no graduate program: first Robert Kottwitz in 1976–77, Tate's student at Harvard; then Jonathan Rogawski, who received his degree in 1980; and later Thomas Hales, in the mid-eighties. Nicholas Katz was the formal advisor of both Rogawski and Hales. The experience was perhaps not entirely a happy one for at least two of the three students, but all survived to thrive as mathematicians. Jon left us too soon.

He had come to me on arriving at Princeton from Yale thanks to the advice of Serge Lang. It was Jon's ambition to become a number theorist, an ambition he ultimately realized, but the fundamental lemma for SU(2, 1) looked to him, with reason, to be far from real number theory. I think he would rather have proved the lemma wrong for SU(2, 1), abandoned the whole project, and gone on to something where elliptic curves figured more prominently. Fortunately, in my view, he never found a semisimple element for which the desired equality was false, proved the lemma for this group, and went on to write an extremely useful treatise on SU(2, 1), *Automorphic representations of unitary groups in three variables*, with very instructive examples of endoscopy, and then spent a good part of the remainder of his life with automorphic forms as an expression of the theory of numbers. Unfortunately, I never had an opportunity to discuss with him the very sophisticated, and very difficult, subsequent development of the fundamental lemma as a central element in the analytic theory of automorphic forms at the hands of Kottwitz, Waldspurger, Ngô and many others.

Indeed, we lived on opposite coasts of North America, and met only rarely, so that we never had an occasion to share our views on the changing face of the theory of automorphic forms in the years after late sixties when functoriality first appeared, together with some indications of reciprocity, or after the seventies, when the trace formula began to be used more systematically in the study of automorphic forms and of Shimura varieties. We were together for a conference on Picard modular surfaces in Montreal in the late eighties, at which his book was a central reference, but we were both too busy to have much time for conversation. Moreover, the subject was changing around both of us: the geometrical theory was growing at an astronomical rate; the genuinely arithmetical applications, such as Fermat's theorem were utterly unexpected; and the trace formula was being developed by Arthur not merely as an occasional tool but as an elaborate theory crying out for applications. In spite of many remaining points that are both obscure and difficult, many more, and more important, applications are in the offing.

That there are common threads running through this material and many later contributions, often referred to in the bulk as the Langlands program, was generally accepted, but, as I found when attempting — to some extent to indulge my vanity, because of the label, and to some extent for sentimental reasons, for it is also related to a number of topics that appealed to me in my early years as a mathematician but that I had never actively investigated — to acquire some understanding of the scope of the program at present, there is a great deal of confusion: the central issues are not always distinguished from the peripheral; partial results obtained by methods that are almost certainly dead-ends are offered with a frequently misplaced satisfaction; many suggestions are facile and, in my eyes, more than doubtful. Some of these shortcomings reflect the failings of our current mathematical culture; others may be inevitable in any cooperative intellectual effort. They are nonetheless troubling and, for the incautious, misleading. Some coherent reflection on the topic, its goals, its limitations at present, and achievements so far, is necessary. It is also difficult.

To write at this point a synopsis of the subject would be premature. Too much is left to do and my command of the material is inadequate. Nevertheless, I am trying to describe the goals of the theory and the methods with which they might be achieved — for my own satisfaction first of all, but secondly because the subject of automorphic representations and their applications appears to me central. As I attempted to explain in the essay "Is there beauty in mathematical theories?" [ND], it is the natural issue of several major currents in pure mathematics of the past two centuries: algebraic number theory; algebraic geometry; group representations — as created by Frobenius, Weyl and Harish-Chandra; and even a dollop of topological ideas, such as perverse sheaves. There is a speculative element in this attempt, and I try to be clear about it when the occasion arises. Nevertheless, the intention is to offer, when I can, possibilities that are not, in my view, impasses and that will lead to a theory at the level of its historical origins. If some results, even, or especially, much-acclaimed or important results, are not mentioned, it may because I see them as leading ultimately nowhere, not as an absolute conviction — absolute convictions are seldom useful — but as a suspicion; but it may also be because they refer to issues like endoscopy or the fundamental lemma, which are basic and important, but for reasons that are tactical more than strategic. Unfortunately, inadequate as it will be, there was no question of completing this description in time for it to appear in the present collection; there are far too many questions and difficulties on which I have hardly begun to reflect. At my age the future offers an uncertain quantity of time, so that whatever success I have will certainly be limited. Nevertheless, this memorial volume is an opportunity to describe and explain in a provisional and, at this time, necessarily incomplete form not only what I mean, in the context of the Langlands program — even in that part of it that owes little or nothing to me — by the two words *functoriality* and *reciprocity* — concepts that are maturing only slowly and in whose development Jon participated — but also how I expect them to be given a clear mathematical content. I apologize, once and for all, for the large tentative element that still remains not only in this prologue but also in the longer, more substantial text "Functoriality and Reciprocity" that it anticipates.

It is best to begin with a rough description of some basic concepts, concepts which it would be idle at this point to formulate too precisely, but which help in appreciating the structure of the theory we are attempting to construct. To introduce the notions of functoriality and reciprocity we need a crude notion of a mock Tannakian category: a generalization of the notion of the category of representations of a group.

Take, as an introductory example, G to be a group, for example, to be as simple as possible, a finite group. Suppose we have a family $\mathcal{G}^{\text{range}}$ of groups and homomorphisms between then, for example the family $\{\mathrm{GL}(n, K) \mid n = 1, 2, \ldots\}$, where K is a field, say the field of complex numbers. Consider the collection of homomorphisms $\varphi : G \to \mathrm{GL}(n, K)$ from G to an element of $\mathcal{G}^{\text{range}}$. This is a Tannakian category in the sense of, for example, [M], a very simple one. The morphisms are given by $\varphi \to \phi \circ \varphi$, where ϕ is an algebraic homomorphism from $\mathrm{GL}(n, K)$ to $\mathrm{GL}(n', K)$.

A more sophisticated choice for $\mathcal{G}^{\text{range}}$ would be the collection ${}^L\mathcal{G}(K/F)$ of all L-groups ${}^LG$ defined for a given extension K/F (see [BC]). The possible choices for these Galois extensions will be described later. The objects of the category will be pairs, the first element of which is an L-group ${}^LG$ in ${}^L\mathcal{G}(K/F)$ and the second an object whose nature depends on F. What is important is that these objects behave functorially: given a pair with first element ${}^LG$ and a homomorphism $\phi : G \to G'$ in ${}^L\mathcal{G}(K/F)$, thus a homomorphism for which the diagram

$$\begin{tikzcd} {}^LG \arrow[rr, "\phi"] \arrow[dr] & & {}^LG' \arrow[dl] \\ & \mathrm{Gal}(K/F) & \end{tikzcd} \tag{1.1}$$

is commutative, there is an image — of the given pair with first element ${}^LG$ — whose first element is ${}^LG'$. We have, for the moment, to be coy about the second element of the pairs because its nature depends on the nature of the field F, whether it is local or global, a field of algebraic numbers or the function field of an algebraic curve.

If K/F is given, the principal property of a mock Tannakian category — and the word "mock" is there to allow a certain latitude and a certain imprecision — is that there is a group $\mathfrak{G}$, usually not a group in ${}^L\mathcal{G}$ although it will have to be provided with a homomorphism $\mathfrak{G} \to \mathrm{Gal}(K/F)$, such that for any ${}^LG \in {}^L\mathcal{G}(K/F)$ the set of pairs with first element ${}^LG$ may be identified with — or, better, parametrized by — the homomorphisms $\varphi : \mathfrak{G} \to {}^LG$ for which the diagram

$$\begin{tikzcd} \mathfrak{G} \arrow[rr, "\varphi"] \arrow[dr] & & {}^LG' \arrow[dl] \\ & \mathrm{Gal}(K/F) & \end{tikzcd} \tag{1.2}$$

is commutative. The existence of this group is closely related to functoriality, usually by no means evident, and in many cases of great interest it is not yet established, although it is expected that it can, in the simpler local contexts, be identified with groups familiar from mid-twentieth century algebraic number theory or from geometry. For global fields, the group cannot be known except through the category it describes. We are striving for a notion that is, in one way, more general than that of a Tannakian category formulated in [M] and, in another, less broad. It is certainly at the moment much less precise.

Functoriality in the L-group appeared first, although not with that name and not so clearly circumscribed as at present, in my letter to Weil of 1967 [LW]. There is a reciprocity — of which, even today, only a small part has been realized — already implicit in functoriality, but a general form of reciprocity only appeared in connection with Shimura varieties, first with the theorem of Eichler–Shimura, but later in a bolder form, once the relation between the cohomology of general Shimura varieties and the discrete series of Harish-Chandra was clarified. The appropriate expression of its general form is one of the central issues of the program, an issue that, as will be intimated later in this prologue, not yet fully resolved, even conjecturally. Indeed in spite of the spectacular success of Wiles with the conjecture for elliptic curves to which the names of Taniyama and Shimura are attached, it is hardly broached. That the initial expression of reciprocity was the Artin reciprocity law for abelian characters of the Galois group of number fields, which itself had its historic source in the quadratic reciprocity law is, however, clear. In the letter to Weil, it was global functoriality that manifested itself as a possible strategy for the analytic continuation of automorphic L-functions. It can still be regarded as the only genuinely promising method for attacking this problem. Local functoriality only manifested itself later, as the consequences and the possibility of more precise formulations of global functoriality began to appear.

In a general context, in which all avatars of the original "program" are to be embraced not by an absolutely uniform collection of definitions and theorems, but by structures which do bear a family resemblance to each other, the origins are a less useful source of understanding than a few general concepts. There are three different theories that find a place in the program, each with a global and a local aspect. By and large, the local theory is a prerequisite for the global. Each of the three is attached to a different type of field. Globally these are (i) algebraic number fields of finite degree over $\mathbb{Q}$; (ii) function fields of algebraic curves over a finite field; (iii) function field of a complete nonsingular curve over $\mathbb{C}$. The corresponding local fields are (i) real, complex and p-adic local fields; (ii) fields of Laurent series over a finite field; (iii) fields of Laurent series over $\mathbb{C}$. The second type of field, a kind of poor relative, usually ignored, shares properties with both the first and the third, themselves quite different from each other in the details although with a common structure.

Let, for example, $F = F_v$ be a local field, for the moment the completion of a number field, and G a reductive group over F_v. Consider the collection of irreducible representations of $G(F_v)$. These are usually infinite-dimensional. The theory of irreducible representations of $G(F_v)$ is a theory that began with Dedekind, Frobenius, and Schur, and whose current structure, the structure with which we shall be dealing, owes an enormous debt to Élie Cartan, Weyl, and Harish-Chandra, but it is nonetheless a theory that is far from complete. We know more for $F = \mathbb{R}, \mathbb{C}$ than for a nonarchimedean F, but the theory appears to be similar for all.

An important observation is that the theory of which we speak is, for a given F_v, not unique. There are several possibilities. First of all, the representations being infinite-dimensional, there are technical constraints, discussed in all the standard texts: they are to be admissible. The notion of equivalent representation has also to be formulated carefully. That demands a good deal of understanding of the structure of the group, its subgroups, and its Lie algebra, all of which I take for granted. Secondly the theories for different groups should be fused. There are distinguished reductive groups: the quasisplit groups, even perhaps, in a pinch, the quasisplit simply connected groups. The reduction of the representation theory for general reductive groups to the theory for quasisplit groups is a part of endoscopy, for which the famous fundamental lemma is necessary and which is absolutely essential for the representation theory of reductive groups over local fields, those of the first two types and probably those of the third type as well. This reduction I do not emphasize; I take it for granted, simply confining myself to quasisplit groups. Moreover, even for quasisplit groups there is another consequence of the yet only very incompletely developed endoscopy that we accept: classes of irreducible representations and their characters are not the objects to which functoriality and reciprocity apply. It is stable classes of representations and, at least for fields of the first or second kinds, their characters that are pertinent. The local theory for function fields over $\mathbb{C}$ is only available in a nascent form, and it is by no means certain that characters have a role to play. It is hard — at the moment — to imagine that their equivalent will not appear. For local fields of the first two types, a stable class π^{st} consists of finitely many equivalence classes of representations and the character of a stable class is a sum

$$\chi_{\pi^{\mathrm{st}}} = \sum_{\pi \in \pi^{\mathrm{st}}} \alpha_\pi \chi_\pi ,$$

where the coefficients are often, perhaps always, integers and these stable characters are not merely class functions, which is what we normally expect from characters, but functions on stable conjugacy classes, stable conjugacy meaning — essentially — conjugacy in $G(F^{\mathrm{sep}})$, F^{sep} being the separable algebraic closure of F, of two elements in $G(F)$. In their full generality both functoriality and reciprocity are predicated on complete theories of endoscopy. Although we are far from possessing such theories, there are many questions related to the two notions on which we can reflect at present.

In order to describe the mock Tannakian categories that are of concern to us, we need to fix a field, either global or local. The first element of a pair is then the L-group ${}^LG$ associated to a reductive group G over F or to the quasisplit group associated to it. They are the same. The second, about which we have been until now reticent, is a stable conjugacy class π^{st} of irreducible representations of $G(F)$, if F is local, or of automorphic representations $G(\mathbb{A}_F)$, if F is global. As already observed, it is best to take G itself quasisplit, referring the rest to endoscopy.

The second element introduces new subtleties. Suppose, for example, that F is a local field and that we are dealing the first of the three types, so that F is the completion of a number field, even $\mathbb{R}$ or $\mathbb{C}$. Then there are several categories of irreducible admissible representation that can — and must — be distinguished: all; unitary; tempered; the Arthur class. For each of these classes, in so far as it is in the present context of any interest at all, there will be a mock Tannakian category, each a slight modification of the others. If we are dealing with all representations then in the semidirect product defining the group ${}^LG = \hat{G} \rtimes \mathrm{Gal}(K/F)$, the connected component $\hat{G}$ is the group of complex points of a reductive group. It is not clear that it is appropriate to consider the category of unitary representations in the context of functoriality. They are, as a class, recalcitrant, and it is very likely that only the Arthur class, of which tempered representations, which are unitary, form an important subclass, is needed. So it may be best to exclude the class of unitary representations as such. For tempered representations, $\hat{G}$ is taken to be the unitary form of $\hat{G}$ — with no change in notation. For the parametrization of the Arthur class, ${}^LG$ is presumably replaced by $\mathrm{SL}(2, \mathbb{C}) \times {}^LG$, but here again it is best to impose some growth conditions on the characters and some unitary condition on the parameters, thus, as it turns out, some growth conditions, not yet understood or formulated, on the matrix coefficients of the representation. The class of all representations has obscure aspects that remain unsettled. We can classify its elements, so that we have a notion of L-packet for them, but so far as I know, there is no stable theory available, even over $\mathbb{R}$; there are L-packets but, at this moment, no stable characters. They may not exist outside the Arthur class. Over nonarchimedean local fields, our ignorance is even more thorough. When the groups ${}^LG$ are replaced by their unitary form, the conditions on the homomorphisms between them are modified as well. For example, for tempered L-packets, homomorphisms from $\hat{H} \rtimes \mathrm{Gal}(K/F)$ to $\hat{G} \rtimes \mathrm{Gal}(K/F)$ restricted to $\hat{H}$ are homomorphisms from a compact group to a compact group.

In the simple example we gave of a Tannakian category, each morphism $\varphi : G \to \mathrm{GL}(n, K)$ represented something, namely itself, a linear representation of G. Composed with $\phi : \mathrm{GL}(n, K) \to \mathrm{GL}(m, K)$ it represents a second linear representation, this time m-dimensional. In the more general mock Tannakian categories, like those associated to ${}^L\mathcal{G}(K/F)$, and we may as well restrict our attention to it, each object is a pair, the first element of which is the L-group ${}^LG$, which determines and is uniquely determined by the corresponding quasisplit group G over F. There is also a second element, a stable conjugacy class π^{st} of irreducible representations $G(F)$, whose type must, as observed, be specified, whether all, tempered, or in the Arthur class. The geometric theory, as described in [CFT], is still immature, so that the possibilities for this second element are even less clear. We shall come to the geometric theory, with its many unresolved questions, later. For now it is best to concentrate on the arithmetic theory.

In order to be able to discuss reciprocity we need, whether at a local or global level, a group $\mathfrak{A}$ such that for a given group G or ${}^LG$ the stable conjugacy classes of irreducible or, globally, irreducible automorphic representations are represented by homomorphism of $\mathfrak{A}$ to ${}^LG$. This possibility was already mentioned, and it was intimated that to prove the existence of $\mathfrak{A}$, it was necessary to prove first that to any homomorphism $\phi : {}^LG_1 \to {}^LG_2$, there was associated a map $\Pi_\phi : \pi^{\mathrm{st}} \to \pi_\phi^{\mathrm{st}}$ of L-packets for G_1 to L-packets for G_2. This possibility I refer to as functoriality or, at more length, functoriality in the L-group.

Once functoriality in the L-group is proved, we shall be on the road to the proof of the existence of $\mathfrak{A}$, locally or globally, and for each kind of field. We have, as will be explained, to envisage different kinds of proof for the various types of fields. Before attempting to describe the possible nature of these proofs, I comment on the second principal topic of this prologue and of the essay to follow it: reciprocity.

We have suggested that the existence of groups $\mathfrak{A} = \mathfrak{A}_F = \mathfrak{A}_{K/F}$, one associated to each field F of the six various types of field, and, to be pedantic, to each sufficiently large Galois extension K of F, was the appropriate classification of representations either locally or globally. Indeed there are other constraints that have to be taken into account: first of all, whether the representations considered are tempered, of Arthur type, or, globally, of Ramanujan type, which entails, even for the same field, the introduction of more than one $\mathfrak{A}_F$; secondly, and this is important only in order not to be forced to pass to senselessly large inverse limits, we should consider the stable classes of representations generated by a finite set. This provides us with one ingredient of reciprocity. The other has been provided, at least partially, by two mathematicians: Galois in the early nineteenth century and Grothendieck in the late twentieth century. Galois groups and their importance are well understood; Grothendieck's notion of motivic Galois group is not well understood and not yet even in a satisfactory form. One task for mathematicians in the coming decades is to discover a better form. Whatever else, these are groups $\mathfrak{M} = \mathfrak{M}_F = \mathfrak{M}_{K/F}$ attached to the various fields. It can be said once again that they are only known through the objects they describe. Over local fields these groups are familiar, especially those for the fields of the first two types, and are known as Weil(–Shafarevich) groups. Globally, however, they are not and the function of reciprocity is to provide some understanding of them. It will be expressed as a homomorphism from $\mathfrak{A}_{K/F}$ to $\mathfrak{M}_{K/F}$, so that it attaches a representation of the former group to one of the latter. Reciprocity is of course the most abstruse, the most profound, and the most difficult of the topics discussed in this prologue and in the essay to follow. I do not expect to have much useful to write. So far as I can tell, we do not understand motives, not even hypothetically, and any real understanding of them requires the solution of major problems in algebraic and diophantine geometry. It would be presumptuous for me even to attempt to describe them at this moment.

I am not certain how it is best to refer to the various groups $\mathfrak{A}$ and $\mathfrak{M}$, in either their local or global forms. For lack of anything better, I shall refer to automorphic and motivic galoisian groups, the adjective *galoisian* indicating that the group describes some other kind of algebraic structure or is defined by it. It may be useful to observe immediately that, in the arithmetic theory, the relation between $\mathfrak{A}$ and $\mathfrak{M}$ is inevitably reflected in an important analytic object associated to irreducible representations and automorphic representations on one hand and motives on the other: L-functions. A homomorphism from $\mathfrak{A}$ to $\mathfrak{M}$ entails a mapping from *complex* representations of $\mathfrak{M}$ to *complex* representations of $\mathfrak{A}$. The definitions on the motivic side are delicate because of the intervention of ℓ-adic-representations. An ℓ-adic representation is not, at least not without further ado, a complex representation. Useful and important as ℓ-adic representations are — they are indeed indispensable —

some reflection is necessary before incorporating them into statements of reciprocity. I find that this preliminary reflection is often missing.

Further discussion of these questions will appear in the essay itself. What is important at present, especially for a reader who may not appreciate the need for the development of sound general concepts, is some understanding of how the general concepts are incorporated into the search for proofs. I begin with a brief list of the necessary steps, employing sometimes notions that have yet to be introduced.

(i) *The local theory over the real field.* What is needed is, first of all, to complete the theory for real groups developed by Harish-Chandra. This means, first of all, a theory of the Arthur class, and secondly a theory of stable transfer.

(ii) *The local theory over nonarchimedean fields.* It is again a matter of completing the theory created by Harish-Chandra, but, as he knew, he left the theory for p-adic fields in a form in which much that he had established over $\mathbb{R}$ was not yet available. Not only is there no theory for the Arthur class and no theory of stable transfer over p-adic fields, there is also no adequate tempered theory.

(iii) *The global theory for algebraic number fields.* In my view, which may not be unanimously shared, the only possibility is to pursue the suggestions of [FLN; BE; ST]. This is no easy in matter. It requires the local theories of (i) and (ii). Globally, it demands a completion of the analytic beginnings of [ST], thus some way of transforming the limits that appear in [ST] into a useful form. Efforts with some promise are being made, although not by me. I am keeping my fingers crossed that they succeed. These will be, at first, results only for $G = \mathrm{PGL}(2)$, but it is possible that they will substantially strengthen our confidence in the trace formula as the route to global functoriality. Moreover the creation of the theories of (i) and (ii) will make it possible to pursue the general global theory effectively. For $\mathrm{PGL}(2)$, there are two bench marks: (a) the second symmetric power and dihedral representations; (b) the fourth symmetric power and quaternionic representations. The second of these bench marks, if reached, would, I believe, encourage the search for concrete methods of counting fields with specific properties in a way that can be compared with the results reached analytically with the trace formula. This may more closely resemble the class field theory of the first half of the twentieth century than of the second.

(iv) *The local geometric theory.* This is the local theory for the field of Laurent series over $\mathbb{C}$. The fascination of contemporary mathematicians with sheaves has, on the one hand, encouraged the development of the local and global theories, but only in the context of spherical functions, which were also of considerable importance in the early years of representation theory for semisimple groups. It has, at the same time, inhibited the development of a theory with ramification, although not entirely [FG]. If this were remedied, the theory would be richer. The structure of a complete local theory is by no means evident, although there are some intimations of the form to be taken by reciprocity, or, better, of the form of the galoisian group $\mathfrak{A}_{\text{geom}}$.

(v) *The global geometric theory.* This is a theory strongly related to the theory of abelian integrals on one hand and the theory of ordinary differential equations with singularities on the other. As with the local theory, the contributions of algebraic geometers, among them Drinfeld, and of mathematical physicists, among them Witten, to the theory have greatly enriched it, but we do prefer a mathematical theory that includes ramification. The best I will be able to do in this prologue are some, with any luck instructive, observations not about reciprocity in a geometric context, where it may not exist, but about the new features that its relation to field theories reveal. I hope that, before coming to this part of the essay itself, I shall have acquired more familiarity and more understanding of the contributions of the mathematical physicists and the geometers.

(vi) *The p-adic theory and diophantine geometry.* These, or rather reciprocity, which can be considered the link between them and the theory of automorphic forms, have to be postponed to the second half of the prologue. It is not clear that, even with the longer period of time available to me for its preparation, I shall be able to write anything useful about these topics. I do hope, at least, to make the stakes clear.

I have no doubt that a lot of reflection will be necessary before the problems presented by (vi) can be broached in any serious way. Deep, quiet reflection over many years may be an indispensable preliminary. My thesis in this prologue and in the essay is that we have, nevertheless, enough tactical understanding to attack the other five problems successfully on a broad front now. Immediate victory is unlikely, but steady advances are not.

2. The local theory over the real field

For reasons connected with the Ramanujan conjecture and its generalizations and with the theory of Eisenstein series, the tempered irreducible representations of $G(F_v)$, F_v a local field, in particular, $F_v = \mathbb{R}$, are not adequate for the modern theory of automorphic forms. There is a larger class of irreducible representations needed that we have already introduced as the Arthur class. The simplest such representation is the trivial one-dimensional representation, which is present for every G and an important factor in the global analytic theory. We have also observed that the local group $\mathfrak{A}_{\mathbb{R}}$ for tempered representations is known to be the Weil group $W_{\mathbb{C}/\mathbb{R}}$, of which we then admit only representations with relatively compact image. The local group for Arthur classes over $\mathbb{R}$, at least for the analogues of tempered representations in the context of Arthur classes, is the group $\mathrm{SL}(2, \mathbb{C}) \times W_{\mathbb{C}/\mathbb{R}}$, or, perhaps better $\mathrm{SU}(2) \times W_{\mathbb{C}/\mathbb{R}}$, but they give equivalent results, and it is better to use the first, for which the Jacobson–Morozov theorem is more easily stated. It is also more concise, if less precise, to use the notation $W_{\mathbb{R}}$ rather than $W_{\mathbb{C}/\mathbb{R}}$ for the Weil group of $\mathbb{R}$.

When attempting to formulate the missing spectral theory for the Arthur packets, we will need to be aware of the need when applying the trace formula for a stable transfer of L-packets. Some very simple cases of this transfer were examined in [ST], but no general theory is available even over $\mathbb{R}$. It is closely related to the stable character for Arthur packets for a (quasisplit) group G that Arthur introduced with the packets in [A1] and their transfers from one group to another are presumably functorial with respect to homomorphisms from ${}^LH$ to ${}^LG$. There is, by the way, no need to introduce any kind of unitary constraint on these homomorphisms: if the image of φ is relatively compact, then so is the image of $\psi \circ \varphi$.

Our focus at the moment is the theory for the real field, which implicitly includes the theory for the complex field. Harish-Chandra's theory for tempered representations, which is the special case of

$$\varphi = \sigma \times \psi : \mathrm{SL}(2, \mathbb{C}) \times W_{\mathbb{R}} \to {}^LG, \qquad W_{\mathbb{R}} = W_{\mathbb{C}/\mathbb{R}}, \tag{2.1}$$

for which σ is trivial, will be in so far as possible the model. It will certainly be used. It is a spectral theory, thus an analytic theory, but it differs from the usual spectral theory. The space of functions to be decomposed is $L^2(G(\mathbb{R}))$, but, as on the line, it is really a more subtle space that is to be decomposed, a Schwartz space. The eigenfunctions or eigendistributions to be employed are invariant under conjugation, thus characters, which are tempered distributions on the Schwartz space. So there is a passage in the theory, not present in, for example, Fourier analysis on the line. From functions on the group, through orbital integrals, to functions on the semisimple conjugacy classes, which for a reductive group is itself a space easily enough described in terms of Cartan subgroups. There is also a passage backwards from distributions on the center to conjugation-invariant distributions on the group and then, by integration on parameters and convolution with functions on the maximal compact subgroup, to functions in the Schwartz space. This means, incidentally, that at this stage it is best not to work with stable packets but with the appropriate classes of irreducible representations, referred to by Harish-Chandra as tempered, those whose matrix-coefficients lie in the Schwartz space. All characters satisfy differential equations, differential equations whose solutions can be concretely described in terms of exponential functions, growth conditions, and jump conditions. Harish-Chandra recognized this. He recognized also, after many years of reflection, that this was all he needed to construct a complete spectral theory for tempered characters. For a more detailed description of Harish-Chandra's representation theory for real groups, I refer to Varadarajan's introduction to his collected works [HC].

In the homomorphism (2.1) σ has a different character than ψ. Only its conjugacy class under interior automorphisms of $\hat{G}$ is pertinent and these are finite in number and correspond to conjugacy classes of unipotent elements in $\hat{G}$ or nilpotent elements in its Lie algebra $\hat{\mathfrak{g}}$. It is usual to study the homomorphisms with a fixed σ and the associated class of representations $\Pi_\sigma(G)$ as a unit, it being understand that the image of ψ is relatively compact. For example, if σ is trivial, we are dealing with the class of tempered representations. To a pair consisting of a homomorphism

$$\phi = \sigma \times \psi : \mathrm{SL}(2, \mathbb{C}) \times {}^L H \to {}^L G$$

and a homomorphism

$$\varphi_H = \sigma_H \times \psi_H : \mathrm{SL}(2, \mathbb{C}) \times W_{\mathbb{C}/\mathbb{R}} \to {}^L H,$$

in which ψ has relatively compact image, we can associate

$$\varphi_G = \sigma_G \times \psi_G : \mathrm{SL}(2, \mathbb{C}) \to {}^L G,$$

in which σ_G is the composite of $\sigma_G \times \sigma_H \circ \phi_H$ with the diagonal imbedding $\mathrm{SL}(2, \mathbb{C}) \to \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})$ and $\psi_G = \psi \circ \psi_H$.

If we had the theory of stable characters for each Arthur class envisaged in [A1], then we would have the mapping that assigns to the stable character of $\pi^{\mathrm{st}}_{\varphi_H}$ on $H(\mathbb{R})$ the stable character of $\pi^{\mathrm{st}}_{\varphi_G}$ on $G(\mathbb{R})$. A grave question, or rather a question central for the trace formula, arises! Is there, for a given ϕ a dual mapping — or, better, correspondence because it will not be single-valued — from smooth functions f^G with compact support on $G(\mathbb{R})$ to smooth functions f^H with compact support on H, thus $f^G \to f^H$, such that

$$\int_{H(\mathbb{R})} f^H(h) \pi^{\mathrm{st}}_{\varphi_H}(h)\, dh = \int_{G(\mathbb{R})} f^G(g) \pi^{\mathrm{st}}_{\varphi_G}(g)\, dg,$$

for all ψ? This question was broached for a very special case in [ST]. It would be premature to attempt to discuss it further here. It is necessary to understand the transfer of stable characters. For this, the first step is to ask what must be done to establish the existence of $\pi^{\mathrm{st}}_{\varphi_H}$.

The stable character will be an eigendistribution, and thus, by an important theorem, an eigenfunction of the center of the universal enveloping algebra with eigenvalues that are given because the definition of Arthur prescribes one representation in it — or rather one stable packet in the sense of the Langlands classification of all irreducible representations, namely, the packet π^{st} associated, as in [A1], to the homomorphism

$$\phi_w = \sigma_G\left(\begin{pmatrix} |w|^{1/2} & 0 \\ 0 & |w|^{-1/2} \end{pmatrix}\right) \cdot \psi_G(w) \tag{2.2}$$

of $W_{\mathbb{R}}$ into ${}^LG$. The infinitesimal character of a representation with Langlands parameter given can be readily calculated from the parameter. So we know the infinitesimal character corresponding to π^{st} and thus that corresponding to the associated — conjecturally — Arthur packet. We can safely assume that all representations in it have the same infinitesimal character, for otherwise we will have no theory. We can study the papers of Harish-Chandra to learn how to calculate all possible eigenfunctions χ associated to this infinitesimal character. On each connected component of the regular elements in each Cartan subgroup T they have the form

$$\chi(t) = \frac{\sum_{w\in\Omega} a_w \exp(w\lambda X)}{|\Delta(t)|}, \tag{2.3}$$

where Ω is the Weyl group, X lies in the Lie algebra of $\mathfrak{t}$ of T, $t = \exp X$, λ is a complex linear form on $\mathfrak{t}$ and the a_w are complex constants. The function $|\Delta(t)|$ is defined as usual by a product of differences of roots. There are constraints attached to χ by the parameter φ_G, constraints studied carefully by Harish-Chandra when σ_G is trivial, thus when the packet is tempered. The constraints, basically on the growth of the function (2.3) and on the propagation of the constants across the subvarieties of singular elements, can be studied as for the tempered characters, although there will be complications that I am in no position to anticipate. They will have to be determined by experience and by the study of some low-dimensional cases.

The existence of the transfer $f^G \to f^H$ and its properties is not the only local problem raised in [ST] in connection with the combined use of the trace formula for G and Poisson summation formula on the Steinberg–Hitchin base. It was also necessary to understand the singularities of θ^G. Both questions are related to the asymptotic behavior of orbital integrals and stable orbital integrals $\mathrm{Orb}(\gamma_{\mathrm{st}}, f^G)$. I have not tried to reflect on them in any serious way.

3. The local theory over nonarchimedean fields

The problems are the same as for the real field; the difficulties are different. I — and, I suspect, many other people — find ourselves here face-to-face with our own ignorance, not just in one domain, but in several. Over both fields we are dealing with problems for characters. Over the real numbers, characters are solutions of a system of holonomic differential equations. Such systems are intimately related to perverse sheaves. In particular, for the complex numbers, the relation between functions and perverse sheaves is mediated by differential equations and belongs to a system of reference familiar to all mathematicians. For representation theory, the real field is more important, but we can overlook that for the sake of the analogy. Over nonarchimedean fields, characters are functions, but there is, as yet, no convenient characterization of them. We have to appeal to the original definition of Dedekind–Frobenius. There are perverse sheaves over these fields and, apparently, perverse sheaves on varieties over finite fields yield functions through the trace of Frobenius — or of the inverse Frobenius. So, if we are willing to overlook the difference between p-adic fields and finite fields, we have parallel constructions for the real field and for nonarchimedean fields

$$\begin{array}{ccccc} \text{perverse sheaves} & \longrightarrow & \text{differential equations and functions} & \longrightarrow & \text{characters} \\ \text{perverse sheaves} & \longrightarrow & \text{trace of Frobenius and functions} & \longrightarrow & \text{characters} \end{array}$$

The trick will be to discover how the perverse sheaves on the second line are to be defined and how they are to be calculated

In the theory of Harish-Chandra ([HC]), whether over an archimedean or over a nonarchimedean field, at least one of characteristic zero, the characters are distributions on $G(F)$ given by functions, or, more precisely, by the product of invariant, but singular, functions with the Haar measure. Over the real or complex field, these singular functions, as distributions, satisfy differential equations, which are — in some sense — holonomic. Since the distributions are invariant, the functions can be considered as functions on the (regular, semisimple) conjugacy classes, and the problem faced and solved by Harish-Chandra was to translate the differential equations satisfied by the characters into jump-conditions for these singular functions. For nonarchimedean fields, there will presumably be similar problems, but I am still uncertain of their nature and certainly in no position to attempt to solve them.

I content myself with a few remarks, influenced, but in no very precise way, by [Wa]. I have no grounds for taking them very seriously, nor do I have any genuine understanding of the necessary algebraic geometry. My goal is to complete the Harish-Chandra theory by finding a handle on the explicit forms of the characters over nonarchimedean fields for tempered representations and, more generally, for characters of representations in the Arthur class; my immediate question is whether, with this precise goal in mind, it is worthwhile to learn the theory of perverse ℓ-adic sheaves. We shall need sheaves on the Cartan subgroups of G and the functions are to be obtained by the traces of the Frobenius on the ℓ-adic cohomology of the fibers.

We also need to convert varieties over a nonarchimedean local field F_v of characteristic 0, or rather schemes over $\mathbb{O}_v$ with residue field κ_v. Let q be the number of elements in κ_v. A preliminary study of [Ha] suggests that Witt vectors are the appropriate instrument for this. The elements of $\mathbb{O}_v$ or, more generally, of the analogous ring $\bar{\mathbb{O}}_v$ in the maximal unramified extension $\bar{F}_v$ of F_v, can be written as series $(x_0, x_1, \dots)$ with coefficients in κ_v or $\bar{\kappa}_v$. This applies to the equations defining any scheme being considered. If the scheme A lies in an n-dimensional space $(X_1, \dots, X_m)$, and if we truncate the coefficients of the equations and of the variables after the mth variable, we obtain equations in $(m+1)n$ variables and schemes A_m, $m = 0, 1, 2, \ldots$. There is a morphism $A_{m+1} \to A_m$, $m = 0, 1, \ldots$. We might guess that for large m this will usually be smooth with fiber the n-dimensional affine space.

If there is a perverse ℓ-adic sheaf on the scheme being considered, we can think of restricting it to A_{m+1} and to the fibers of the morphism. If this restriction is a constant sheaf, just the pull-back of the restriction to the base point of the fiber, then the restriction has cohomology with compact support only in dimension 0. So the summation over the points in the fiber of the trace of the Frobenius is q^n. There is, however, something to remember. Although the character is a function, it always appears multiplied by a measure, either the Haar measure on the group or, if we pass to an integral not against a function f on $G(F)$, but against the orbital integral of F a measure on the Steinberg–Hitchin base or on a Cartan subgroup. The two are locally equivalent. A is either this base or the Cartan subgroup — give or take some singular subvarieties. The measure of a point on A_{m+1} is, up to a constant, $1/q^n$ times the measure of its image, so that the factors q^n and $1/q^n$ cancel each other. When passing to the Steinberg–Hitchin base, we multiply the character by $|\Delta(t)|$ and the measure is the measure on the Cartan subgroup, for which conventions have been established in [FLN]. The remaining factor in the measure is implicit in the orbital integrals.

I am tempted to think that the road to follow is already blazed in the literature. The theory over $\mathbb{R}$, with the Borel–Weil–Bott theorem, the homological realization of the discrete series verified by W. Schmid, Harish-Chandra's analytic construction of tempered representations from the discrete series, and the proof by Deligne–Lusztig of a conjecture of Macdonald, all point in the same direction: first introduce the characters of tori in a form adapted to the use of perverse sheaves, then combine it with some twisted form of parabolic induction — which can be formulated I suppose, in the context of perverse sheaves. This is not a matter of an effort lasting a few days or a few weeks, but unless the basic idea of using truncated Witt vectors is misguided, a careful study of the works mentioned should allow some progress.

I confess that I have never attempted, even in a modest experimental way, to examine the possibilities or to understand the initial difficulties when attempting to transpose the constructions in [DL] to a nonarchimedean context using truncated Witt vectors. To begin would be easy enough, as the only difficulty is to find the time, but the possible virtues of these secondary constructions was not apparent to me until I began to write this prologue and the essay on functoriality and reciprocity to follow, both a continuation of the reflections begun in [FLN; BE; ST]. In the following section, I simply take the existence of the necessary local theory as established. A good deal of it, not necessarily in the most suggestive form, is certainly available for $G = \mathrm{SL}(2)$ for which the global analytic and arithmetic problems are already daunting and well worth investigating.

4. The global theory for algebraic number fields

There are two aspects to the continuing reflections on the methods suggested in [BE]: the formal or structural aspects and the analytic aspects. The latter are extremely difficult. Ali Altuğ has been reflecting on them and I leave it to him to present, when it is appropriate, his conclusions. I concentrate on the former. The principal goal, indeed the overriding goal, is to establish functoriality and its consequences with the help of the trace formula and Poisson summation. The objects studied are the automorphic L-functions $L(s, \pi, \rho)$ associated to an automorphic representation π or, better, a stable class of automorphic representations π^{st} that contains π. It is their analytic properties that need to be studied, especially near $s = 1$ or in the half-plane $\mathrm{Re}\, s > 1$.

There are two possibilities: examine $L(s, \pi, \rho)$ itself or examine its logarithmic derivative. Although the logarithmic derivative contains in clearer form the pertinent information, the function $L(s, \pi, \rho)$ is the more accessible. So there is a difficult passage, as with the prime number theorem, from its study to that of $-L'(s, \pi, \rho)/L(s, \pi, \rho)$. This I leave, at least for the moment, to others and concentrate on the properties of $-L'(s, \pi, \rho)/L(s, \pi, \rho)$ that one hopes can be established and that lead to functoriality.

We anticipate a complete form of endoscopy, which is certainly available in some simple and instructive cases. With an appropriate choice of test functions, the stable trace formula leads to sums

$$\sum_{\pi^{\mathrm{st}}} m_{\pi^{\mathrm{st}}} \left\{ \prod_{v \in S} \mathrm{tr}\, \pi_v^{\mathrm{st}}(f_v) \left\{ \sum_{v \notin S} \ln L_v(s, \pi_v, \rho) \right\} \right\}. \tag{4.1}$$

Here S is an arbitrary finite collection of places, containing the infinite places, Each f_v is a smooth function with compact support, taken otherwise arbitrary, and ρ is an algebraic representation of ${}^L G$. There are loose ends, some terms missing, and some imprecision in the formula (4.1), but none of this is a serious issue for us here. The sum itself is over stable classes of representations unramified outside of S. So it is likely that only H that are unramified outside of S are pertinent.

The first serious issue is related to the generalized form of Ramanujan's conjecture and Arthur L-packets. The global L-packets are expected to be related to homomorphisms

$$\phi = \sigma \times \psi : \mathrm{SL}(2, \mathbb{C}) \times {}^\lambda H \to {}^L G, \tag{4.2}$$

where, for the present purposes, we can quite comfortably write $\mathrm{SL}(2, \mathbb{C}) \times {}^L H$, the necessity of modifying the L-groups ${}^L H$ slightly to ${}^\lambda H$ for technical homological reasons being one of the minor nuisances that plague the subject. Our principal purpose is to establish that the stable classes of automorphic representations can be written as a sum over the functorial transfers associated to (4.2) of the stable tempered automorphic representations of $H(\mathbb{A}_F)$. There will be ambiguities to be clarified.

One stable class π_G for G may be the image of several π_H. This is why we appeal in the essay to the notion of a hadronic or thick class introduced in [LSP]. We use only the transfers associated to classes that are hadronic. It is then to be expected that each class π_G^{st} is associated to a unique pair $\mathrm{SL}(2,\mathbb{C}) \times {}^L H$, although we will have to allow different ψ, for the reasons that can be inferred from [LP], and perhaps even different σ, although this is unlikely.

In the discussion of local packets and local transfers, we made it clear that the transfers associated to (4.2) are of tempered representations of H to representations of G that are tempered if and only if σ is trivial. The σ-factor is otherwise a measure of the extent to which the local images $\pi_{G,v}$ are not tempered. This is measured by the eigenvalues of

$$\rho\left(\sigma\left(\begin{pmatrix} q^{1/2} & 0 \\ 0 & q^{-1/2} \end{pmatrix}\right)\right)$$

in the various representations ρ of ${}^L G$. We want to sort the representations appearing in (4.1) according to type. This means we take the sum over pairs $\mathrm{SL}(2,\mathbb{C}) \times {}^L H$ and over conjugacy classes of ϕ but only include for a given such pair — if we include ϕ, such triple — hadronic π_H. Such a sum introduces multiplicities. It is natural to assume, and the evidence, such as it is, supports the assumption, that they are accommodated by the multiplicity with which various representations of $G(\mathbb{A}_F)$ occur in the space of automorphic forms.

So (4.1) should be equal to a sum, implicitly over triples $(\phi, \mathrm{SL}(2,\mathbb{C}) \times {}^L H)$,

$$\sum \left\{ \sum\nolimits_H^{\mathrm{temp}} \left\{ \prod_{v \in S} \operatorname{tr} \pi_v^{\mathrm{st}}\bigl(f_v^H\bigr) \left\{ \sum_{v \notin S} \ln L_v(s, \pi_v, \rho) \right\} \right\} \right\} \tag{4.3}$$

It is understood that at each place f_v^H is the transfer in the sense of the previous sections of f^G. I have left out any reference to multiplicities on the assumption, made largely for the purposes of simplicity, that the multiplicities are largely caused by multiple homomorphisms ϕ. Any necessary corrections can be made when proofs have been found. What is important at the moment is to be clear about the structure proposed for the proof. For (4.1) there is a formula, the trace formula. In (4.3), the outer sum is over triples, the first inner sum, $\sum_H^{\mathrm{temp}}$ is over the stable tempered automorphic representations of $H(\mathbb{A}_F)$. We can assume by induction that for each triple, except the triple with $H = G$, thus with ϕ trivial on SL(2) and the identity on G itself, we have a formula for the inner sum $\sum_H^{\mathrm{temp}}$. This would yield a formula for the remaining term of the inner sum,

$$\sum\nolimits_G^{\mathrm{temp}} \Big\{ \prod_{v \in S} \operatorname{tr} \pi_v^{\mathrm{st}}\big(f_v^G\big) \Big\{ \sum_{v \notin S} \ln L_v(s, \pi_v, \rho) \Big\} \Big\}, \tag{4.4}$$

except that we would not know that the only automorphic representations that are not the image of a hadronic tempered automorphic representation with respect to some ψ with $H \neq G$ are themselves hadronic and tempered. However, in the terms of the sum (4.4) the first factor $\prod_{v \in S} \operatorname{tr} \pi_v^{\mathrm{st}}(f_v^G)$ is essentially arbitrary and serves to distinguish one π from another. So an understanding of (4.3) is essentially an understanding of the logarithmic derivative

$$\frac{d}{ds} \ln L_v(s, \pi_v, \rho) = \frac{L_v'(s, \pi_v, \rho)}{L_v(s, \pi_v, \rho)}. \tag{4.5}$$

The analytic problem is to show, with the aid of the formula for (4.4) just described, that it is holomorphic to the right of $\operatorname{Re} s = 1$ for every ρ. This implies not only that the representation is tempered but that it is hadronic. This problem is central, very serious, and, in my view, it will be a matter of decades before it is solved in any generality. The method suggested in [FLN; ST] was to use the Poisson summation formula on the Steinberg–Hitchin base, but the hard questions were not broached.

Although it is premature to make too much of a fuss of the notion of hadronic representation, one observation is in order. If π_G^{st} is the image of π_H^{st} under the functorial transfer associated to ϕ in (4.2). Then

$$L(s, \phi_G^{\mathrm{st}}, \rho) = L(s, \phi_H^{\mathrm{st}}, \rho \circ \psi). \tag{4.6}$$

The representation $\rho \circ \phi$ of $\mathrm{SL}(2, \mathbb{C}) \times {}^L H$ decomposes into a direct sum $\bigoplus_n \tau_n \otimes \rho_n$, where τ_n is the irreducible representation of $\mathrm{SL}(2, \mathbb{C})$ of degree n, which can be any positive integer. The L-function (4.6) is then given by

$$\prod_{n=1}^{\infty} \prod_{j=0}^{n} L(s - 2j + n, \pi_H^{\mathrm{st}}, \rho_n).$$

Each representation ρ_n is a direct sum of irreducible representations $\bigoplus_{i=1}^{m_n} \rho_{n,i}$. To show that (4.4) is holomorphic for $\operatorname{Re} s > 1$ for every choice of S and every choice of the functions f_v, $v \in S$ is to show that $\dim \rho_n = 0$ for $n > 0$ and that for all ρ none of the representations $\rho_{0,i}$ is trivial. It may be appropriate to remind ourselves at this point that the L-groups that appear are defined with respect to any extension K/F, which can be arbitrarily large. Since H itself may be the group $\{1\}$, we will be dealing, in particular, with those representations that are attached to homomorphisms of the Galois group into ${}^L G$.

The solution of these analytic problems, even for very specific low-dimensional questions, for example, the existence of automorphic representations associated to quaternionic representations, can entail at least partial answers to the arithmetic questions raised in §2 and §3 and to their global forms. I, myself, find that these questions and their answers add considerably to the appeal of the algebraic theory of numbers [D; JY]. The Dedekind paper [D], which we shall review in the next section, is particularly charming. The arithmetic problems to be confronted and solved in the course of establishing global functoriality are nevertheless every bit as formidable, if not more so, than the analytic problems.

5. Classical algebraic number theory

There are two very different aspects of the construction of global functoriality proposed in the previous section: analytic and arithmetic. The analysis does not end with the introduction of the Steinberg–Hitchin base and the use of the Poisson summation formula; as in [FLN; ST], considerably more is needed. I hope that this will be explained in Altuğ's forthcoming thesis. As just intimated, there will also be arithmetic problems. In §4 it was explained that we expect, for a given G and each representation ρ of its L-group, thanks in part to the trace formula and Poisson summation, to be able to express the sum of the logarithmic derivatives of the L-functions $L(s, \pi, \rho)$ as a sum over imbeddings $\phi : \mathrm{SL}(2, \mathbb{C}) \times {}^L H \to {}^L G$, and in particular, with this in hand, to examine the asymptotic behavior of this sum as $s \to 1$. This will be complicated, because, for example, the L-groups ${}^L H = \hat{H} \rtimes \mathrm{Gal}(K/F)$ and ϕ can reduce to an imbedding of a Galois group in ${}^L G$. As a result the proposed analysis entails an understanding of such imbeddings. For abelian class field theory, this becomes an understanding of, say, the cyclic extensions of a given degree of the base field F. For the group GL(2) or P GL(2), it becomes an understanding of the imbeddings of Galois groups in GL(2) or SL(2). If ρ is the fourth symmetric power of the defining representation of SL(2), the most interesting possibility is that $\mathrm{Gal}(K/F)$ is imbedded as the quaternion group. Such extensions were studied, as observed in the previous section, not so long ago by Jensen and Yui, to whose paper my attention was drawn by Anthony Pulido. They were influenced by an earlier paper of Reichardt ([Re], see also [Ri]). There is a much earlier, more concrete paper by Dedekind ([D]), that it is worthwhile to review briefly, because, or so it seems to me, algebraic number theory in the, often concrete, style of Dedekind was abandoned after the Second World War, with the mounting enthusiasm in the USA and elsewhere for the more formal, more abstract styles of Artin and Chevalley. It may be that a successful attack in the spirit of §4 will demand a return to Dedekind.

The focus in Dedekind's paper is on quaternion extensions of $F = \mathbb{Q}$. Following his notation, let the quaternion group be formed by 1, a central element ϵ, $\epsilon^2 = 1$, and elements $\alpha, \beta, \gamma, \epsilon\alpha, \epsilon\beta, \epsilon\gamma$, $\alpha^2 = \beta^2 = \gamma^2 = \epsilon$, $\alpha\beta = \gamma = \epsilon\beta\alpha$. Any such extension contains a biquadratic extension, the fixed field H of ϵ. This is a field of the form $\mathbb{Q}(\sqrt{a}, \sqrt{b}, \sqrt{c})$, $c = ab$. Dedekind takes $a = 2$, $b = 3$, which pretty much leads to the minimal ramification of the field, which, as it turns out, has to be real. This is very convenient in connection with the trace formula. The field H is the only biquadratic field unramified outside of $\{2, 3\}$ and, using spherical functions outside of $\{2, 3\}$, in particular at the infinite place, we exclude all representations π with ramification outside this set. So our comparison will be very focussed. The field Ω whose Galois group is the quaternion group will be of the form $\Omega = H(\omega)$, $\omega^2 = \mu \in H$ and the problem is to determine those ω that lead to an Ω also unramified outside of $\{2, 3\}$. A helpful feature that simplifies the constructions, but that is not present in general, is that the class number of H is one. We shall verify this later.

We begin with some other considerations, more generally valid. After some hesitation, I chose to follow Dedekind's convention of writing the action of the Galois group on the right. This is, otherwise, inconsistent with the notation of the paper, but without it the comparison with Dedekind's paper is awkward. The elements $\omega\alpha$, $\omega\beta$, $\omega\gamma$ all lie in Ω and their squares all lie in H. Since they themselves do not lie in H, they all lie in $H\omega$. As a result, we obtain,

$$\omega\alpha = u\omega, \quad \omega\beta = v\omega, \quad \omega\gamma = w\omega, \qquad u, v, w \in H. \tag{5.1}$$

Moreover, $\omega\epsilon = -\omega$. Applying α to the first of the equations (5.1), we obtain $-\omega = \omega\epsilon = u\alpha u\omega$ or, as Dedekind writes, $u\alpha = -u^{-1}$. There is a collection of similar equations, verified in the same way, that Dedekind writes as

$$\begin{array}{lll} u\alpha = -u^{-1}, & u\beta = wv^{-1}, & u\gamma = -vw^{-1}, \\ v\alpha = -wu^{-1}, & v\beta = -v^{-1}, & v\gamma = uw^{-1}, \\ w\alpha = vu^{-1}, & w\beta = -uv^{-1}, & w\gamma = -w^{-1}, \\ \mu\alpha = \mu u^2, & \mu\beta = \mu v^2, & \mu\gamma = \mu w^2. \end{array} \tag{5.2}$$

If μ is replaced by $\mu\nu^2$, the extension does not change and u, v, w can be replaced by $u\nu\alpha/\nu$, $v\nu\beta/\nu$, $w\nu\gamma/\nu$. This allows simplifications. For example, Dedekind observes that if the class number of H is one, then we can assume that μ is integral and not divisible by the square of any prime ideal. This is then also true of its conjugates $\mu\alpha$, $\mu\beta$, and $\mu\gamma$, so that, thanks to the last line in (5.2), u, v, w must all be units. As a consequence, if μ is divisible by any prime ideal π, it is divisible by all the conjugates of that ideal. If, therefore, p is the prime number in $\mathbb{Q}$ that π divides and if p does not divide the discriminant, then p divides μ. Dedekind concludes that μ must be the product of a natural number m, a unit, and perhaps powers of the generators of the prime divisors of the discriminant. The pertinent information for our particular H is (i) its discriminant is $48^2 = 2^8 3^2$; (ii) the ideal

$(3) = (\sqrt{3})^2$ and $\sqrt{3}$ is a prime in H; (iii) the ideal (2) is the fourth power of the ideal $(1+\eta)$, $\eta = (1+\sqrt{3})/\sqrt{2}$, with $\eta^2 = 2+\sqrt{3}$, $\eta^{-2} = 2-\sqrt{3}$; (iv) the fundamental units in H are $a = 1+\sqrt{2}$, η, $\tau = \sqrt{2}+\sqrt{3}$, with inverses, $\sqrt{2}-1$, $(\sqrt{3}-1)/\sqrt{2}$, $\sqrt{3}-\sqrt{2}$. The possibilities for μ are therefore

$$\mu = \pm m a^{e_1}\eta^{e_2}\tau^{e_3}(1+\eta)^{e_4}(\sqrt{3})^{e_5}, \tag{5.3}$$

in which each e_i, $i = 1, \dots, 5$, is 0 or 1 and m is a natural number prime to 6 and a product of primes. Not all possible values of the exponents are admissible. Examining (5.3) on the basis of (i)–(iv), Dedekind arrives at the conclusion that $e_1 = e_2 = 1,\ e_3 = e_4 = 0,\ e_5 = 1$. As a consequence

$$\mu = \pm m a\eta\sqrt{3}. \tag{5.4}$$

I repeat his calculations. It is necessary to calculate $\mu\alpha/\mu$, $\mu\beta/\mu$, $\mu\gamma/\mu$ and to demand that they all be squares. For this, following Dedekind, we compute the Galois action on each factor of (5.3). We repeat that

$$\begin{aligned}
&(\sqrt{2}, \sqrt{3}, \sqrt{6}, \omega)\alpha = (\sqrt{2}, -\sqrt{3}, -\sqrt{6}, u\omega),\\
&(\sqrt{2}, \sqrt{3}, \sqrt{6}, \omega)\beta = (-\sqrt{2}, \sqrt{3}, -\sqrt{6}, v\omega), \qquad \omega\epsilon = -\omega, \quad \mu\epsilon = \mu,\\
&(\sqrt{2}, \sqrt{3}, \sqrt{6}, \omega)\gamma = (-\sqrt{2}, -\sqrt{3}, \sqrt{6}, w\omega).
\end{aligned}$$

The Galois action on the units is given by

$$\begin{array}{lll}
a\alpha = a, & a\beta = -a^{-1}, & a\gamma = -a^{-1},\\
\eta\alpha = -\eta^{-1}, & \eta\beta = -\eta, & \eta\gamma = \eta^{-1},\\
\tau\alpha = -\tau^{-1}, & \tau\beta = \tau^{-1}, & \tau\gamma = -\tau.
\end{array} \tag{5.5}$$

The first line follows from $(1+\sqrt{2})(1-\sqrt{2}) = -1$. The action of the Galois group takes η to $\pm(1\pm\sqrt{3})/\sqrt{2}$ and

$$\frac{1+\sqrt{3}}{\sqrt{2}}\cdot\frac{1-\sqrt{3}}{\sqrt{2}} = -1,$$

thus η to $\pm\eta^{\pm 1}$. This is the second line. Since $(\sqrt{3}+\sqrt{2})(\sqrt{3}-\sqrt{2}) = 1$, the Galois group also takes τ to $\pm\tau^{\pm 1}$. This is the third line.

The action of the Galois group on $\sqrt{3}$ is given by $(\sqrt{3})\alpha = -\sqrt{3}$; $(\sqrt{3})\beta = \sqrt{3}$; $(\sqrt{3})\gamma = -\sqrt{3}$. The second line yields immediately a first form for the Galois action on the supplementary prime $1+\eta$ that divides 2,

$$(1+\eta)\alpha = -\eta^{-1}(1-\eta), \quad (1+\eta)\beta = (1-\eta), \quad (1+\eta)\gamma = \eta^{-1}(1+\eta). \tag{5.6}$$

Each of these numbers are units. It only remains to express them as products of powers of the fundamental units times $1+\eta$.

$$\frac{1-\eta}{1+\eta}=\frac{\sqrt{2}-1-\sqrt{3}}{\sqrt{2}+1+\sqrt{3}}=\frac{(\sqrt{2}-1-\sqrt{3})^2}{2-(1+\sqrt{3})^2}=-\frac{2-2\sqrt{2}(1+\sqrt{3})+(1+\sqrt{3})^2}{2+2\sqrt{3}}.$$

Multiplying numerator and denominator by $1-\sqrt{3}$, we obtain

$$-\frac{2(1-\sqrt{3})-2\sqrt{2}(1-3)-2(1+\sqrt{3})}{2(1-3)}=-(\sqrt{2}-\sqrt{3})=-\tau^{-1}.$$

Thus the three numbers (5.6) are $1+\eta$ times, respectively,

$$\eta^{-1}\tau^{-1}, \quad -\tau^{-1}, \quad \eta^{-1}.$$

From these relations, we conclude with Dedekind that

$$\mu\alpha=\pm m a^{e_1}(-\eta)^{-e_2-e_4}(-\tau)^{-e_3-e_4}(1+\eta)^{e_4}(-\sqrt{3})^{e_5},$$

the sign being the same as in (5.3), and that

$$u^2=\frac{\mu\alpha}{\mu}=(-1)^{e_2+e_3+2e_4+e_5}\eta^{-2e_2-e_4}\tau^{-2e_3-e_4}. \tag{5.7}$$

For this to be a square it is necessary and sufficient that e_4 be 0 and $e_2+e_3+e_5\equiv 0 \pmod 2$. Further conditions are given by $\mu\beta/\mu$. Since

$$\mu\beta=\pm m(-a)^{-e_1}(-\eta)^{e_2}(\tau)^{-e_3-e_4}(1+\eta)^{e_4}$$

and $e_4=0$, the quotient $v^2=\mu\beta/\mu$ is

$$(-1)^{e_1+e_2}a^{-2e_1}\tau^{-2e_3}. \tag{5.8}$$

For this to be a square $e_1+e_2\equiv 0 \pmod 2$. Thus $e_1=e_2$.

The first two of the equations in the last line of (5.2) imply the third. They imply as well that Ω is a quadratic extension of H, Galois over $\mathbb{Q}$. They do not imply that Ω is a quaternion extension of $\mathbb{Q}$. For that we need the earlier lines, which assure us that this is so. Dedekind uses the first two of the three diagonal equations, which must imply all nine equations because the first completely defines the action of α on Ω and the second that of β. Consider the first diagonal equation. The number u is the square root of (5.7). The information at our disposition yields

$$u=(\pm)'\eta^{-e_1}\tau^{-e_3},$$

where, following Dedekind, we have explicitly indicated with a prime that the sign appearing here is not the sign in (5.3). The first diagonal equation yields

$$\begin{aligned}u\alpha=(\pm)'(-\eta^{-1})^{-e_1}(-\tau^{-1})^{-e_3}&=(\pm)'(-1)e_1+e_3\eta^{e_1}\tau^{e_3}\\&=-u^{-1}=-(\pm)'\eta^{e_1}\tau^{e_3},\end{aligned}$$

from which we conclude that $e_1 + e_3 \equiv 1 \pmod 2$. This implies that $e_5 = 1$ and that (e_1, e_3) is either $(1, 0)$ or $(0, 1)$. Dedekind settles the matter with the second diagonal equation.

The element v is the square root of (5.8), $v = (\pm)'' a^{-e_1}\tau^{-e_3}$ and, thanks to (5.5),

$$\begin{aligned} v\beta = (\pm)''(-a)^{e_1}\tau^{e_3} &= (\pm)''(-1)^{e_1}a^{e_1}\tau^{e_3} \\ &= -v^{-1} = -(\pm)''a^{e_1}\tau^{e_3}. \end{aligned}$$

We infer that $e_1 = 1$, and therefore that $e_3 = 0$, arriving finally at Dedekind's conclusion (5.4).

Dedekind does not offer any hints for the verification that the class number is one. So we apply the standard theorems. Since there are a number of other points about the field H to be verified, we postpone this until the end of the section and explain first the pertinence of the quaternionic fields to the study of the trace formula and its applications.

There are two tests that may be undertaken to persuade oneself of the validity of the strategy proposed in §4 and of Altuğ's analytic development of the necessary analysis. He, himself, has begun to reflect on them. The two tests are: the application to dihedral automorphic representations and the possible application to quaternionic representations. The interest is less in the results and more in the conviction to be obtained that the methods proposed, although difficult, are sound. As explained in §4, the method, as so often with L-functions, is focussed on the behavior of $-dL(s, \pi, \rho)/ds$ as $s \searrow 1$, or, rather, assuming for simplicity that $F = \mathbb{Q}$, on that of

$$\sum_{\pi}\Bigl\{\operatorname{tr}\pi(f^{\rho}_{\infty}) + \sum_{p}\sum_{n=1}^{\infty}\frac{n\ln p}{p^{ns}}\operatorname{tr}\pi_v(f^n_v)\Bigr\}, \tag{5.9}$$

where outside of a finite set S of places v, the functions f^n_v are chosen to be spherical functions such that

$$\operatorname{tr}\pi^n_v = \rho(A^n(\pi_v)),$$

where $A(\pi_v)$ is the Frobenius–Hecke class attached to π_v. The representation ρ is a representation of ${}^LG$. The development of the stable trace formula described in §4 allows for the removal of all nontempered π from (5.9), thus of all stable π whose parameter contains a nontrivial SL(2) component. It is understood that these have been removed, so that the remaining sum has no singularities to the right of $\operatorname{Re} s = 1$. It will be part of the analysis to show this! It is moreover expected, and will have to be shown, that only those π associated to a homomorphism $\phi : {}^{\lambda}H \to {}^LG$ whose image is a proper subgroup of ${}^LG$ will contribute to the pole at $s = 1$. The representations of $G(\mathbb{A}_F)$, thus of $H(\mathbb{A}_F)$, are to be understood inductively. For the two tests, we take $G = \mathrm{PGL}(2)$ or, but that would be slightly more elaborate, GL(2). The L-groups are SL(2) or GL(2) or, better because we must consider all possible H, $\mathrm{SL}(2) \times \mathrm{Gal}(K/\mathbb{Q})$ or $\mathrm{GL}(2) \times \mathrm{Gal}(K/\mathbb{Q})$. We consider only the first.

If $\pi = \pi_G$ is the image of π_H, then the contribution of π to the sum (5.9) is $1/(s-1)$ times the multiplicity of the trivial representation of ${}^\lambda H$ in $\rho \circ \phi$. To avoid redundancy, we always suppose π_H is hadronic. In particular, if ρ is irreducible and nontrivial, as we may as well suppose, there is no contribution from any hadronic π_G. For P GL(2) or GL(2), say GL(2) because this allows a simpler notation, this means any one of the following three possibilities. First ${}^\lambda H = {}^L H$, $H = \mathrm{GL}(1) \times \mathrm{GL}(1)$. For the second there is a quadratic extension E of $\mathbb{Q}$, H is the two-dimensional torus obtained from GL(1) by restriction of scalars from E to $\mathbb{Q}$, ${}^\lambda H = {}^L H$, and

$$\tag{5.10} \begin{aligned} &\phi : (a, b) \rtimes 1 \mapsto \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \qquad a, b \in \mathbb{C},\ 1 \in \mathrm{Gal}(E/\mathbb{Q}), \\ &\phi : (1, 1) \rtimes \sigma \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \end{aligned}$$

where $\mathrm{Gal}(E/\mathbb{Q}) = \{1, \sigma\}$. The representation π_H is attached to a homomorphism φ of the global Weil group

$$\{1\} \to E^\times \backslash I_E \to W_{E/F} \to \mathrm{Gal}(E/F) \to \{1\}$$

into ${}^L H$ and this homomorphism is defined by a character χ of $E^\times \backslash I_E$. If w is a fixed element in $W_{E/F}$, $w \notin I_E$, then

$$\tag{5.11} \varphi : \begin{cases} \alpha \in I_E \mapsto (\chi(\alpha), \chi(\sigma\alpha)) \rtimes 1, \\ \qquad w \mapsto (\chi(w^2), 1) \rtimes \sigma. \end{cases}$$

The third possibility is that H=$\{1\}$. There is overlapping of all three cases.

The first case leads to noncuspidal representations and is thus understood. The third case is most interesting when we take ${}^\lambda H$ to be a Galois group, especially when this group is tetrahedral, octahedral or icosahedral. We have not reached the stage where they can be treated by the methods under discussion. The overlapping occurs when the Galois group is a finite dihedral group and, in particular, when it is a quaternion group.

Consider first the case (ii) and let ρ be the $2n$-th symmetric power of the defining representation of GL(2). Then the L-function $L(s, \pi_G, \rho) = L(s, \pi_H, \rho \circ \phi)$ will be the product of the L-function of $\mathbb{Q}$ associated to the character $\chi | I_{\mathbb{Q}}$ and the L-functions of the field E associated to the characters $\chi^2, \ldots, \chi^n$. The first of these functions has a pole of order 1 at $s = 1$ if and only if $\chi | I_{\mathbb{Q}}$ is trivial. So for any natural number n, these functions will contribute a pole at $s = 1$ to the sum (4.3), in particular for $n = 1$. For the results achieved by the method of §4 for $n = 1$, it will be best to refer to Altuğ's thesis. Since they concern functoriality only for the group GL(2) and a torus H, a case that can be treated in the context of the Hecke theory, they may not convince the sceptical, however interesting they may be for those whose concern is with functoriality and its consequences in general. It is only for exceptional n and very exceptional H and ϕ that further poles appear.

One possibility is that H is associated to a quadratic extension, χ is exactly of order 4, $\chi(\alpha)$ is not identically equal to $\chi(\sigma\alpha)$, and $\chi(w^2) = -1$. Of course, π_H is then not hadronic, but π_G is also associated to another group H, the group $H = \{1\}$, and extensions $K/\mathbb{Q}$ with $\mathrm{Gal}(K/\mathbb{Q})$ isomorphic to the image of the original ${}^L H$ under $\phi \circ \varphi$. This image is a group of order 8, isomorphic to the group

$$\left\{ \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 0 & a \\ -\bar{a} & 0 \end{pmatrix} \right\}, \quad a^4 = 1.$$

This is the quaternion group imbedded in SL(2). In addition to the irreducible representation this yields, the group has four one-dimensional representations, the trivial representation and the three nontrivial characters of the group divided by its center, which is, of course, $\pm I$. The even symmetric powers of the two-dimensional representation are clearly the direct sums of characters. Since the quaternion group has a group of outer automorphisms of order three, $\alpha \to \beta \to \gamma \to \alpha$, all three nontrivial characters appear with the same multiplicity μ and the trivial representation then appears with multiplicity $\nu = 2n + 1 - 3\mu$. For $n = 1$, $\mu = 1$, $\nu = 0$; for $n = 2$, $\mu = 1$, $\nu = 2$. This means that if ρ is the fourth symmetric power and π_G is the image of the trivial representation of π_H, then $L(s, \pi_G, \rho)$ has a pole of order two at $s = 1$. This will not be so for hadronic π_G, nor for other dihedral π_G, nor for tetrahedral, octahedral, or icosahedral π_G. For these, as observed in [BE], the exceptional poles at $s = 1$ begin only with higher values of n.

Thus the method of §4, if it is to work at all, must detect the quaternionic representations — and only the quaternionic representations — by the extra pole for $n = 4$. Although this leads to no new number-theoretical conclusions, it would be a very important indication of the promise of the method. It would also be a sign that the investigations of Dedekind or Jensen–Yui and the other authors have to be pursued, perhaps along the lines suggested in [ST], perhaps in other ways.

When applying the method, we usually fix a bound for the ramification of the representations π that we wish to consider. This is done by choosing $f = \prod_v f_v$ to be a spherical function outside of a finite set S of places and then choosing the f_v, $v \in S$, with appropriate restrictions. Their precise description is limited by one's understanding of the local harmonic analysis and arithmetic. In the present case, we might want to take $m = 1$ in (5.4), which restricts the ramification to 2 and 3, where it can be the minimum that permits the quaternion group to appear. Ramification for nonabelian representations of the Galois group and for representations of a general reductive group over a local field demands, of course, a more sophisticated, more technical examination, than necessary for abelian Galois representations or for the group GL(1).

To complete our report on Dedekind's paper, we have still to deal with some details of the structure of H, which can be obtained in two steps: (i) by the adjunction of $\sqrt{3}$ to obtain $H_1 = \mathbb{Q}(\sqrt{3})$; (ii) by the subsequent adjunction of $\sqrt{2}$, $H = H_1(\sqrt{2})$. Unfortunately, I am not so familiar with such calculations as Dedekind.

The discriminant of $H_1/\mathbb{Q}$ is $2^2 3$ and the ideal $\sqrt{3}$ is clearly unramified in H_2 where it does not split. So, by the usual formulas for the differents and discriminants of fields obtained by repeated extensions, the contribution of 3 to the discriminant of $H/\mathbb{Q}$ is 3^2. The two numbers a and τ are clearly integral. Moreover $\eta^2 = 2+\sqrt{3}$, so that η is also integral. It follows from (5.5) that all three of these numbers are units and from (5.6) that

$$N_{H/\mathbb{Q}}(1+\eta) = -\eta^{-2}(1-\eta^2)^2 = -\frac{(1+\sqrt{3})^2}{2+\sqrt{3}} = -2$$

is a unit times the fourth power of $1+\eta$.

The only other prime dividing the discriminant is 2. Let Z_2 be the 2-adic integers. Since the powers $(1+\eta)^j$, $j = 0, 1, 2, 3$ form an integral basis over $\mathbb{Z}_2$ of $H \otimes \mathbb{Z}_2$, we can calculate the power of 2 in the discriminant as $\prod_{i \neq j}(\eta_i - \eta_j)$, where η_i, $i = 1, 2, 3, 4$ are the conjugates of η, namely $\eta, -\eta, \eta^{-1}, -\eta^{-1}$. The result is $\pm(\eta^2 - \eta^{-2})^4$ and

$$16(\eta^2 - \eta^{-2})^4 = ((2+\sqrt{3}) - (2-\sqrt{3})^4 = 16^2 \cdot 3^2 = 2^8 \cdot 3^2.$$

This gives the correct result not only for 2 but also for 3. As a consequence, $\{1, 1+\eta, (1+\eta)^2, (1+\eta)^3\}$ or $\{1, \eta, \eta^2, \eta^3\}$ is an integral basis for the ring of integers in H.

Dedekind observes — discretely and without comment — that $\eta^2 = 2+\sqrt{3}$ and $\tau^2 = 5+2\sqrt{6}$. This is very useful information. The three quadratic subfields of H are $E_1 = \mathbb{Q}(\sqrt{2})$, $E_2 = \mathbb{Q}(\sqrt{3})$, $E_3 = \mathbb{Q}(\sqrt{6})$. For the units in E_1, the two basic hyperbolas are $x^2 - 2y^2 = \pm 1$. The units of positive norm are contained in $x^2 - 2y^2 = 1$ and generated, up to sign, by $3 \pm 2\sqrt{2}$, themselves the square of $1 \pm \sqrt{2}$. So a is a fundamental unit of E_1. For E_2, the corresponding hyperbolas are $x^2 - 3y^2 = \pm 1$. The units of positive norm are generated, again up to sign, by $2 \pm \sqrt{3} = \eta^2$. There are none with negative norm. For E_3 the hyperbolas are $x^2 - 6y^2 = \pm 1$ with points $5 \pm 2\sqrt{6}$, thus τ^2, $(\tau\alpha)^2$. They generate the units of positive norm. There are again no units of negative norm.

Consider a unit $x = x_1$ and its conjugates, $x_2 = x\alpha$, $x_3 = x\beta$, $x_4 = x\gamma$. Thus $|x_1| \cdot |x_2| \cdot |x_3| \cdot |x_4| = 1$, $x_1x_2x_3x_4 = \pm 1$. Since x_1x_3 is a unit in E_2, it is up to sign an even power of η. Thus, dividing x by an appropriate power of η we can conclude that $x_1x_3 = \pm 1$. Then, of course, $x_2x_4 = \pm 1$ as well. Now we divide by a power of

τ to obtain $x_1x_4 = \pm 1$, but without affecting the value of x_1x_3. As a result

$$\pm 1 = x_1x_2x_3x_4 = \pm\frac{x_2}{x_1} \quad \text{and} \quad \frac{x_1}{x_2} = \pm 1,$$

so that $x_1\alpha = \pm x_1$. If $x_1\alpha = x_1$, then x_1 is in E_1 and up to sign a power of a. Otherwise, $x_1 = y\sqrt{3}$, $y \in E_2$. Since 3 remains prime in E_2 and x_1 is a unit, this is impossible. We conclude that, as affirmed by Dedekind, a, η, and τ generate, up to sign, the group of units of H. Dedekind's example is marvelously simple!

Unfortunately, I am not familiar enough with Dedekind's style to know how he would have established that the class number of H is one. It follows readily enough from standard theorems. Dedekind's argument would have been more elegant. According to a familiar theorem [He, Satz 96], if there is a prime ideal in H that is not principal, there is one with norm less than or equal to the square root of the discriminant of H, thus $2^4 \cdot 3 = 48$.

The field H is a composite of two quadratic fields the class field associated to the group of ideles multiplicatively congruent to 1 or 7 modulo 8, 1 modulo 3, and positive. So there are four classes of primes different from 2 and 3. According to the law of quadratic reciprocity, they are distinguished by their residues modulo 3 and 8. First of all, in the field $\mathbb{Q}(\sqrt{2})$ the decomposition is:

(i) If $p \equiv 1, 7 \pmod 8$ then p splits.

(ii) If $p \equiv 3, 5 \pmod 8$ then p does not split.

In the field $\mathbb{Q}(\sqrt{3})$:

(i) If $p \equiv 1 \pmod 3$ and $p \equiv 1 \pmod 4$ or $p \equiv 2 \pmod 3$ and $p \equiv 3 \pmod 4$ then p splits.

(ii) If $p \equiv 2 \pmod 3$ and $p \equiv 1 \pmod 4$ or $p \equiv 1 \pmod 3$ and $p \equiv 3 \pmod 4$ then p does not split.

In the field $\mathbb{Q}(\sqrt{6})$, if $p \equiv 1 \pmod 8$ and $p \equiv 1 \pmod 3$, if $p \equiv 7 \pmod 8$ and $p \equiv 2 \pmod 3$, if $p \equiv 3 \pmod 8$ and $p \equiv 1 \pmod 3$, or if $p \equiv 5 \pmod 8$ and $p \equiv 2 \pmod 3$ then p splits, otherwise it does not. From this, we determine immediately the nature of the decomposition in H, whether a prime different from 2, 3 splits into 1, 2 or 4 primes. It splits into four if and only if it splits into two in the three intermediate fields.

According to the theorem cited, all we need do is show, first of all, that every prime ideal $\mathfrak{p}$ of norm p in one of the three fields E dividing a prime $p > 3$ in $\mathbb{Q}$ and with $\mathrm{N}\,\mathfrak{p}$ less than or equal to the square root of the discriminant of E is principal and, secondly, that every prime ideal of norm p in H dividing a prime $p > 3$ and with $\mathrm{N}\,\mathfrak{p}$ less than or equal to the discriminant of H is also principal. For the first type this is hardly necessary, but the results are as follows.

(1) For $\mathbb{Q}(\sqrt{2})$, the discriminant is 8 and there are no such primes.

(2) For $\mathbb{Q}(\sqrt{3})$, the discriminant is 12 and the only pertinent prime seems to be 11. Since $\mathrm{N}(1+2\sqrt{3}) = -11$, $\mathfrak{p} = (1+2\sqrt{3})$ is a prime of norm 11.

(3) For $\mathbb{Q}(\sqrt{6})$, the discriminant is 24. Of the primes 5, 7, 11, 13, 17, 19, 23, only 5, 19, 23 seem to satisfy the necessary conditions. We have $N(1+\sqrt{6}) = -5$, $N(5+\sqrt{6}) = 19$, $\mathrm{N}(1+2\sqrt{6}) = -23$.

(4) There are many primes less than or equal to 48, namely

$$5, 7, 11, 13, 17, 19, 23, 31, 37, 41, 43, 47,$$

but very few with the correct congruence properties, namely that split completely in H. For this we need either $p \equiv 1 \pmod 8$, $p \equiv 1 \pmod 3$, or $p \equiv 7 \pmod 8$, $p \equiv 2 \pmod 3$. Thus $p = 23, 47$ seem to be the only possibilities. We have to show that each of them factors in H into the product of four distinct principal prime ideals. It is enough to show that each of them is the norm of an element in H.

We can factor each of them in the three quadratic fields.

$$\begin{array}{lll} \mathrm{N}(5-\sqrt{2}) = 23; & \mathrm{N}(2-3\sqrt{3}) = -23; & \mathrm{N}(1-2\sqrt{6}) = -23, \\ \mathrm{N}(7-\sqrt{2}) = 47; & \mathrm{N}(1-4\sqrt{3}) = -47; & \mathrm{N}(7-4\sqrt{6}) = -47. \end{array}$$

Because we have so much information about η, it is convenient — and sufficient — to establish that the central element in each of these rows is, up to a unit, the norm in $E = \mathbb{Q}(\sqrt{3})$ of an element u in H, which is $E(\eta)$, because $\eta^2 = 2+\sqrt{3}$. The field E is the fixed field of β. Thus, if we can find one $u = a + b\eta$ such that

$$\mathrm{N}_{H/E}(u) = u \cdot u\beta = a^2 - b^2(2+\sqrt{3}), \qquad a, b \in E. \tag{5.12}$$

differs from $2-3\sqrt{3}$ by a unit in E and another such that it differs from $1-4\sqrt{3}$ by another unit, then our task will be complete. So ran my first reflections.

I thought it would be necessary to attack the problem systematically, by a careful analysis that would determine where the numbers whose norm was ± 23 or ± 47 were to be found. The field H seemed to be singularly adapted to the necessary calculation. Consider the absolute values of the numbers a, η, τ and of their conjugates in the order: the number itself, then its conjugate under α, β, γ in that vertical order. The first column is supplementary, $x > 0$, $x \neq 1$.

$$\begin{array}{cccc} 1 & a & \eta & \tau \\ x & |a| & |\eta| & |\tau| \\ x & |a| & |\eta|^{-1} & |\tau|^{-1} \\ x & |a|^{-1} & |\eta| & |\tau|^{-1} \\ x & |a|^{-1} & |\eta|^{-1} & |\tau|. \end{array} \tag{5.13}$$

Taking the logarithms, we obtain a matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \ln x & 0 & 0 & 0 \\ 0 & \ln|a| & 0 & 0 \\ 0 & 0 & \ln|\eta| & 0 \\ 0 & 0 & 0 & \ln|\tau| \end{pmatrix} \tag{5.14}$$

The first matrix is up to a factor an orthogonal matrix with inverse,

$$\frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

I wrote this down, looked at it, thought of the effort that a further systematic analysis would require, and decided I would just resort to Mathematica and calculate the norms of a few numbers in H in order to have a better feel for the sizes entailed. To my surprise and delight, because I was growing very fatigued, among the first ten norms generated appeared both -23 and -47.

$$\mathrm{N}(1-\eta^2-\eta^3) = -23; \quad \mathrm{N}(1-\eta-\eta^3) = -47.$$

6. Reciprocity

The meaning of reciprocity, as it appears in this prologue, is somewhat uncertain and variable. This appears to be inevitable. Although I have attempted to confine it to a relation between a group of $\mathfrak{A}$ and a group $\mathfrak{M}$, it sometimes appears to be simply a description of a group, either a motivic group or, more often, an automorphic group. This is, to a large extent, because the traditional Weil group already incorporates both aspects: (i) the multiplicative group of the field or of the idele classes as a carrier of characters; (ii) the Galois group as a description of finite extensions of the base field F, thus as a description of motives of dimension 0. Moreover, although reciprocity has a certain universality, it appears under more than form and this form adapts itself to the circumstances, local or global, geometric or arithmetic, and is, as a consequence, somewhat protean.

The Ramanujan conjecture in its general form — if properly interpreted, even in its classical form — is a statement about the local factors of automorphic representations $\pi = \otimes\pi_v$ and their Arthur parameters. We have not had occasion to comment on the local form of the group $\mathfrak{A}$ in an Arthurian context over nonarchimedean fields F_v. It appears to be $\mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C}) \times W_{F_v}$. As for archimedean fields, the first factor determines whether the representation is tempered or not and, if it is not, the nature of its failure to be tempered: determined by the asymptotic behavior of characters or matrix-coefficients. The second $\mathrm{SL}(2,\mathbb{C})$ does not appear for archimedean fields and is present to accommodate the needs of local reciprocity, which some authors satisfy by introducing the Weil–Deligne group WD.

We can introduce into the global arithmetic theory a formal but suggestive diagram:

$$\text{(6.1)}\quad \ell\text{-adic representations} \xleftarrow{\otimes\bar{\mathbb{Q}}_\ell} \text{motives}/F \xrightarrow{\otimes\mathbb{C}} \text{automorphic representations,}$$

The Weil–Deligne group is introduced in the context of ℓ-adic representations, thus on the left; the motivic groups, at least those introduced by Grothendieck can be considered, for the present purposes, as being defined over $\mathbb{Q}$; the group $\mathfrak{A}$ is defined over $\mathbb{C}$. So an arrow from the extreme left to the extreme right is, without further explanations, not immediately at hand. The further explanations necessary are not, given the theorems and conjectures currently at hand, particularly difficult. The Weil–Deligne group has two disadvantages: (i) it introduces isomorphisms between fields that have no natural relation to each other, namely ℓ-adic fields and the complex field; (ii) it introduces classes of representations that are not semisimple. Neither of these is overwhelming, but both are unnecessary and clumsy. It is best to introduce a second SL(2) factor, either in the local $\mathfrak{A}$ or in the local $\mathfrak{M}$. This second factor is not present over $\mathbb{R}$ or $\mathbb{C}$.

Our immediate task, however, is to introduce the appropriate local structures on the left-hand side, for which all we have at hand are the ℓ-adic representations. We begin with them in their local form, taking the necessary material from [T]. Suppose that the local field $F = F_v$ is nonarchimedean with residue characteristic p and $\ell \neq p$. The theory of p-adic representations is more difficult and certainly pertinent, but not for this article.

In [T] a Frobenius element Φ is an element of the Galois group such that $\Phi^{-1}x = x^q$ on the residue field. I follow this convention. The elements of the Galois group that concern us are those that can be written as a product $\Phi^n\iota$, where ι lies in the inertia group. They form a dense subgroup, to be identified with the Weil group, of the Galois group. There is a homomorphism of the inertia group onto $\prod_{p\neq\ell}\mathbb{Z}_\ell$. Let it send ι to $\prod_{\ell\neq p} t_\ell(\iota)$. In [T] the notion of an ℓ-adic WD-representation or a representation of the Weil–Deligne group on a finite-dimensional ℓ-adic vector space is introduced. A “representation” of this group is not a true representation, it is a pair (r, N), where N is a nilpotent transformation of a finite-dimensional vector space V over $\bar{\mathbb{Q}}_\ell$ and r a representation of the Weil group on the same space. The representation r is to be continuous, so that its kernel is open in W. Moreover

$$r(\Phi)Nr(\Phi^{-1}) = q^{-1}N, \tag{6.2}$$

a condition imposed for every choice of Φ. Thus N commutes with the inertia group.

There are supplementary conditions that can be imposed. One, that the Zariski closure G_r of the image of r is reductive, seems especially important. We impose it. As a consequence the image $r(\Phi^n\iota)$ of any element of the Weil group is semisimple. Other conditions of a topological or analytical nature are of lesser importance and we leave them aside for the moment. There is then a second representation associated to the pair (r, N),

$$\rho : \Phi^n\iota \mapsto r(\Phi^n\iota)\exp(t_l(\iota)N).$$

Clearly ρ determines r and N, but it is the pair to which we attach here the most importance, not the ℓ-adic representation ρ.

The restriction of r to the inertia group is defined by a representation of a finite quotient of this group. There is, consequently, an integer $m \neq 0$ such that $r(\Phi^m\iota) = r(\iota\Phi^m)$ for all ι in the inertia group. If $r(\Phi)$ is equal to $\Phi_{\mathrm{ss}}\Phi_{\mathrm{un}} = \Phi_{\mathrm{un}}\Phi_{\mathrm{ss}}$, with Φ_{ss} semisimple and Φ_{un} unipotent, then

$$\Phi_{\mathrm{ss}}^m r(\iota) = r(\iota)\Phi_{\mathrm{ss}}^m, \quad \Phi_{\mathrm{un}}^m r(\iota) = r(\iota)\Phi_{\mathrm{un}}^m,$$

for all ι in the inertia group. The second equation implies that $r(\iota)\Phi_{\mathrm{un}} = \Phi_{\mathrm{un}}r(\iota)$ for all ι. Consequently we can introduce a new representation r_{ss} such that $r_{\mathrm{ss}}(\iota) = r(\iota)$ for all ι, while $r_{\mathrm{ss}}(\Phi) = \Phi_{\mathrm{ss}}$. The representation r_{ss} is a canonical semisimplification of r. This is a representation of the Weil group.

It may seem idle, but we want to be able to replace the homomorphism r by a homomorphism of the WD-group to any L-group ${}^LG$, taken not over $\mathbb{C}$ but over $\bar{\mathbb{Q}}_\ell$. This is easily done. The definitions are identical to those just given for ${}^LG = \mathrm{GL}(n, \bar{\mathbb{Q}}_\ell)$. We shall continue in this vein for it allows us to consider homomorphisms of the categories we construct into the category of (algebraic) representations of ${}^LG(\bar{Q}_\ell)$. We retain the assumption that the Zariski closure of the image of the Weil group in ${}^LG$ is reductive, not forgetting the necessary compatibility with projections to finite Galois groups.

The construction of the WD-group is, unfortunately, clumsy and misleading, because it permits a passage to quotients by kernels of the transformation N. In order to avoid this possibility, we appeal to the Jacobson–Morozov lemma as formulated in [K], but we use it not over $\mathbb{C}$, rather over $\bar{\mathbb{Q}}_\ell$. Let N be a nilpotent element in the Lie algebra ${}^L\mathfrak{g}$ of a reductive group ${}^LG$. The superscript on ${}^L\mathfrak{g}$ serves a largely mnemonic function. There exists an $X \in \mathrm{ad}\, N({}^L\mathfrak{g})$ such that $[X, N] = 2N$. In addition, for each such X there exists a unique N' such that $[X, N'] = 2N'$, $[N, N'] = X$. The algebra $\mathfrak{s} = \{N', X, N\}$ is therefore isomorphic to the algebra $\mathfrak{sl}(2)$. Let σ be the isomorphism

$$\sigma : \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mapsto N, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mapsto X, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \mapsto N'$$

of the algebra $\mathfrak{sl}(2)$ with $\mathfrak{s}$.

For us, N is the element in the equation (6.2). Corresponding to this equation, there is a character of the Galois group Gal_F, such that

$$r(\iota)Nr(\iota)^{-1} = \chi(\iota)N, \quad \forall g \in \mathrm{Gal}_F .$$

On the inertial group $\chi(\iota) = 1$.

Let $\mathfrak{h}$ be the centralizer of the image of the inertial group in the algebra ${}^L\mathfrak{g}$. The algebra $\mathfrak{h}$ is reductive because, by hypothesis, the image of r is reductive. We apply the Jacobson–Morozov theorem to the algebra $\mathfrak{h}$ and the element $N \in \mathfrak{h}$. Let H be the connected component of the identity in the centralizer of the inertia group in ${}^LG$ and S the connected subgroup of H corresponding to $\mathfrak{s}$. The group S has a unique Cartan subgroup, isomorphic to the multiplicative group of the field $\bar{\mathbb{Q}}_\ell$, whose Lie algebra contains X and this subgroup contains an element P such that $\mathrm{Ad}(P)(N) = q^{-1}N$, $P = q^{-X/2}$. Define ψ by the relation

$$\psi(\Phi^n\iota) = P^{-n}r(\Phi^n\iota). \tag{6.3}$$

The set of products $\Phi^n\iota$, $n \in \mathbb{Z}$, is of course the Weil group and ψ is a representation of it. Since $\psi(\iota) = r(\iota)$,

$$\mathrm{Ad}(\psi(\iota))N = N, \quad \mathrm{Ad}(\psi(\iota))X = X, \quad \mathrm{Ad}(\psi(\iota))N' = N'. \tag{6.4a}$$

Moreover,

$$\mathrm{Ad}(\psi(\Phi))N = N, \quad \mathrm{Ad}(\psi(\Phi))X = X. \tag{6.4b}$$

Consequently, $\mathrm{Ad}(\psi(\Phi))N'$ satisfies the conditions of the theorem of Jacobson–Morozov, so that $\mathrm{Ad}(\psi(\Phi))N' = N'$. As a consequence, rather than a representation of the WD-group in the sense given to it in [T] and other sources, we may use the representation (σ, ψ) of the thickened Weil group $\mathbb{W}$. I prefer this. Of course, σ has to be interpreted as a representation of the group, rather than of the algebra, and we have to replace $\mathbb{C}$ by $\bar{\mathbb{Q}}_\ell$. There is nothing to be done about this. It can be effected by an imbedding $\bar{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$, disturbing but in the nature of things. We can, if we prefer, rather take the thickened Weil group not over $\mathbb{C}$ but over $\bar{\mathbb{Q}}_\ell$.

This possibility raises many questions. Since we do not yet have a complete theory of the representations of reductive groups over nonarchimedean local fields, we do not have a parametrization of the various classes, tempered, arbitrary, or the class introduced by Arthur. Moreover, even over archimedean fields there is, so far as I know, no clear indication, even at the speculative level, that there will be a stable theory for arbitrary irreducible representations. On the other hand, the classification of tempered representations, over $\mathbb{R}$ or, presumably, any other local field, will certainly demand a constraint of relative compactness on the image of ψ in, for example, (2.1), and this condition is not one that is invariant under an imbedding of $\bar{\mathbb{Q}}_\ell$ in $\mathbb{C}$. So there is room for reflexion on the local form of (6.1).

We take the imbeddings $\bar{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and $\bar{\mathbb{Q}} \hookrightarrow \bar{\mathbb{Q}}_\ell$ as given; so the algebraic closures of $\mathbb{Q}$ in the two fields are identified. The considerations that follow would lead to definitions that are independent of this identification.

There is a distinguished subgroup of $\mathbb{Q}$, the group S_1 of algebraic numbers β all of whose conjugates have absolute value 1 in $\mathbb{C}$. We introduce, at a local level, the set of parameters ψ, or (σ, ψ), or $(\sigma_1, \sigma_2, \psi)$ such that there is a homomorphism

$$\xi : \mathrm{GL}(1) \to {}^L G \tag{6.5}$$

for which:

(i) The image of ξ commutes with the image of ψ and, if appropriate, the image of σ or $\sigma_1 \times \sigma_2$.

(ii) For every Frobenius element Φ, every element ι of the group of inertia, every integer m, and for every (algebraic) representation ρ of ${}^L G$ all the eigenvalues of $\rho(\xi(|w|^{-m/2})\psi(\Phi^m \iota))$ lie in S_1.

If ξ exists it is determined by r so that when there is no danger of misunderstanding it need not be explicitly given. The distinction between ψ and r is somewhat pedantic. In (6.4a) and (6.4b), r is given and ψ depends on the choice of a square root of q; the representation r does not. In (ii) we are, in effect, making a further modification of ψ, leading to a further dependence on the choice of the square root $q^{1/2}$. The second condition does not, however, depend on the choice of the square root $q^{1/2}$. We must, nevertheless, take care that no implicit dependence on this choice occurs in other definitions, for example, in the L-functions associated to ℓ-adic representations. This would be a different dependence than that entailed by the simultaneous imbeddings of $\bar{\mathbb{Q}}$ in $\bar{\mathbb{C}}$ and $\bar{\mathbb{Q}}_\ell$. In a global context, ξ would first be given and then the various conditions would be satisfied by the local restrictions and this fixed ξ.

For the field $\bar{\mathbb{Q}}_\ell$ — and for ${}^L G = \mathrm{GL}(n)$ — these are the parameters that, because of the last Weil conjecture, yield the ℓ-adic representations with which we are principally concerned. According to the yet to be established local parametrization — for tempered and nontempered representations — they would correspond not only to tempered representations — for ξ trivial — but often to nontempered representations. Thus, it is entirely appropriate to introduce *weighted* parameters locally as well as the attendant global modifications. At a nonarchimedean place the local parametrization consist of pairs $\{\sigma \times \psi, \xi\}$ that satisfy the conditions described, or, for Arthur parameters, pairs $\{\sigma_1 \times \sigma_2 \times \psi, \xi\}$. At an archimedean place, they would just be pairs $\{\psi, \xi\}$ or $\{\sigma \times \psi, \xi\}$, the first condition remaining unchanged, but the second being replaced by the condition that the eigenvalues of $\rho\left(\xi(|w|^{-1})\psi(w)\right)$ have absolute value 1. The local form of the relation (6.1) would no longer be mediated by motives. It would be

(6.6) weighted ℓ-adic representations $\longrightarrow$ (automorphic) representations.

The arrow is now independent of the imbedding $\bar{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$, at least in so far as it is compatible with the identification of $\bar{\mathbb{Q}} \subset \mathbb{C}$ and $\bar{\mathbb{Q}} \subset \bar{\mathbb{Q}}_\ell$. All this, in view of the scarcity of general results, has a very pedantic air, but, for my own sake, I find it useful to have a clear notion of the goals. They are not always clearly understood or formulated. Locally, what is often wanted — apart from careful, appropriate definitions — is simply an independent description of $\mathfrak{A}$ or $\mathfrak{M}$, in terms of familiar objects: the Weil group, the Galois group, or differentials. Globally, at least for arithmetical fields, it is a matter of proving the existence of both $\mathfrak{A}$ and $\mathfrak{M}$, deciding what their relation is and proving it. All three are major problems. For the global geometric theory, it is not clear to me at the moment, whether it is a description of $\mathfrak{A}$ in classical terms that is wanted, or whether there is a motivic group $\mathfrak{M}$ to be introduced and a relation of $\mathfrak{A}$ and $\mathfrak{M}$ to be discovered. The following two sections suggest that there is no $\mathfrak{M}$ in the geometric theory, but they are hardly conclusive.

I add that the supplementary Arthur parameters may not play a role in the correspondence (6.6). It appears to me that the image is likely to consist of objects whose supplementary Arthur parameter is trivial, so that the homomorphism of groups, which is from the right-hand side to the left, will be trivial on the SL(2) component of the global automorphic $\mathfrak{A}$.

There are many relatively simple examples of the various parameters that it might be appropriate to introduce here: (i) for the Arthur parameters, the conjecture of Jacquet for SL(n) proved by Mœglin and Waldspurger; (ii) for the second SL(2) parameter, the reciprocity for elliptic curves with nonintegral j-invariant, a very important and very early example in the development of a general reciprocity. Although they are well-known, they belong in any introduction to the theory. This is none the less only the prologue to an introduction. So I omit them.

7. The geometric theory for the group GL(1)

For the local arithmetic theory, we can identify the group $\mathfrak{A}_F$ as the Weil group or — for Arthur packets or if the local field is nonarchimedean — as a modified form of the Weil group, but we are not yet able to supply the necessary proofs. For the local geometric theory the abelian quotient, thus the group appropriate for $G = \mathrm{GL}(1)$, of the local group $\mathfrak{A}_F$ is readily identified, although the definitions are somewhat forced. The description of this quotient for the global geometric theory can be deduced, as we shall describe, from the classical theory of abelian integrals on a Riemann surface. The description of the abelian quotient of $\mathfrak{A}_F$ suggests, both locally and globally, a definition of $\mathfrak{A}_F$ itself, but as I discovered, one is faced almost immediately with the need for theories that have yet to be developed. I stress that, although the concepts emphasized here differ in some ways from those preferred by Edward Frenkel and have been influenced as well by the proof of a theorem of Weil, the initial impulse has been taken from his writings.

I should perhaps confess as well that, although the references [CFT; CLG; GT], from which I profited considerably, were, together with a letter from their author, my introduction to the geometric theory, my impulses, aesthetic and mathematical, are more analytic, less formal, perhaps less geometric, than those of their author. Even though I have not yet succeeded in exploiting the analytic possibilities of the theory, I do want to draw them to the reader's attention.

For the geometric theory, the local field at a point x is the field F_x of formal Laurent series

$$f(z)=\sum_{n=k}^{\infty} a_n z^n, \qquad k\in\mathbb{Z}.$$

In the present context reciprocity — not the correct word in this context, in which our goal is simply the description of the local automorphic galoisian group $\mathfrak{A}_{F_v}$ — is, at least at first, simply a matter of expressing the characters of $F_x^\times$, or rather the group formed by these characters, in some appealing arithmetic or geometric manner. We must of course fix the choice of characters — unitary, nonunitary, holomorphic, whatever.

The local group $F_x^\times$ is abelian with two particularly important subgroups,

$$\begin{aligned}\mathbb{O}_x^\times &= \{a+bz+cz^2+\cdots \mid a\neq 0\},\\ \mathbb{O}_x^+ &= \{1+bz+cz^2+\cdots\},\end{aligned} \tag{7.1}$$

and $\mathbb{O}_x^\times=\mathbb{C}^\times\mathbb{O}_x^+$, $\mathbb{O}_x^+\backslash\mathbb{O}_x^\times\simeq\mathbb{C}^\times$, $\mathbb{O}_x^\times\backslash F_x^\times\simeq\mathbb{Z}$. The elements of the group $\mathbb{O}_x^+$ are best written in exponential form

$$\exp(\alpha_1 z+\alpha_2 z^2+\cdots). \tag{7.2}$$

The characters of $\mathbb{O}_x^\times$ are, for our purposes, most conveniently given by the residues of the differential forms defined by the product of the logarithm of

$$f=\alpha_0\exp(\alpha_1 z+\alpha_2 z^2+\cdots), \qquad \alpha_0\neq 0,$$

and a given local differential form

$$\omega=\frac{\beta_{-k+1}}{z^k}+\frac{\beta_{-k+2}}{z^{k-1}}+\cdots+\frac{\beta_{-1}}{z^2}+\frac{\beta_0}{z}+\sum_{j=1}^{\infty}\beta_j z^{j-1}, \quad \beta_j\in\mathbb{C}\ \forall j, \quad \beta_0\in\mathbb{Z}, \tag{7.3}$$

although the coefficients β_j, $j > 0$, which are redundant and present only in anticipation of the global theory, do not affect the pairing.

$$(7.4)\quad (\omega, f) = (\omega, f)_x = \exp(i\,\mathrm{Re}(\mathrm{res}\,\omega \ln f)),$$
$$\mathrm{res}\,\omega \ln f = \beta_0 \ln \alpha_0 + \sum_{i=1}^{k-1} \alpha_i \beta_{-i}.$$

There is a second pairing implicit in (7.4), obtained on replacing $\mathrm{Re}(\mathrm{res}\,\omega \ln f)$ by the real linear form $\mathrm{Im}(\mathrm{res}\,\omega \ln f)$. It is understood that both are to be used, alone and in products. If $\beta_0 = 0$ one of them is redundant, since $\mathrm{Im}(\mathrm{res}\,\omega \ln f) = \mathrm{Re}(\mathrm{res}(-i\omega \ln f))$. If $\beta_0 \neq 0$, $i\omega$ is not admissible, because $i\beta_0 \notin \mathbb{Z}$. These characters are unitary. There is another possibility,

$$f \mapsto \exp(\mathrm{Re}(\omega \ln f)).$$

These characters are not unitary, but are pertinent in a more geometric theory, like that of [CFT], if ramification is admitted. We keep them in mind, because the two theories, analytic and geometric, are conceived as parallel to each other.

The group $\mathbb{O}_x^+$ is an infinite-dimensional complex vector space, the inverse limit of finite-dimensional vector spaces. Its dual space is taken to be a direct limit not of the complex dual spaces of the distinguished finite-dimensional spaces defined by the inverse limit, but of the distinguished real linear forms defined by the real and imaginary parts of the complex forms. Since $\mathrm{Im}(\omega, f) = \mathrm{Re}(-i\omega, f)$, this leads to a real vector space of dimension twice — and not four times — the complex spaces from which they arise. The dual space of $\mathbb{C}^\times = \mathbb{O}^+\backslash\mathbb{O}^\times$ is taken to be $\mathbb{R} \times \mathbb{Z}$, $\alpha \to \alpha^m \bar{\alpha}^n$, $m + n \in \mathbb{R}$, $m - n \in \mathbb{Z}$. Here, however, we need to use both the real and the imaginary parts of (ω, f), thus $\beta_0 \,\mathrm{Re} \ln \alpha_0$ and $\beta_0 \,\mathrm{Im} \ln \bar{\alpha}_0$, because β_0 is constrained to be integral, in particular, real. The pairing (ω, f) is linear in ω and multiplicative in f.

There seems to be no natural or unique way to extend this identification of the space Ω_x of local differential forms ω_x at x, implicitly taken modulo their regular parts and modulo the identification described, with the character group of $\mathbb{O}_x^\times$ to a concrete identification of $F_x^\times$, thus no way to incorporate naturally the dual of $\mathbb{O}_x^\times\backslash F_x^\times \simeq \mathbb{Z}$. This dual can be taken to be $\mathbb{C}^\times$. There will be a commutative diagram

$$(7.5)\quad \begin{tikzcd} \{1\} \arrow[r] \arrow[d] & \mathbb{C}^\times \arrow[r] \arrow[d] & \widetilde{\Omega}_x \arrow[r] \arrow[d, "\tilde{\eta}"'] & \Omega_x \arrow[r] \arrow[d, "\eta"'] & \{1\} \arrow[d] \\ \{1\} \arrow[r] & \mathbb{C}^\times \arrow[r] & \mathrm{Char}(F_x^\times) \arrow[r] & \mathrm{Char}\,\mathbb{O}_x \arrow[r] & \{1\}, \end{tikzcd}$$

in which the kernel of $\tilde{\eta}$ is equal to the kernel of η, but a natural precise description is not available. To split the extension, a local parameter must be chosen.

This said and the necessary precisions kept in mind, we have defined the local group $\mathfrak{A}$ for the abelian geometric theory in terms of local differentials with singularities. For the global theory on a Riemann surface X, there are not only global differential forms, there are a number of supplementary objects, whose purpose took me some time to recognize. I brought with me from the arithmetic theory a notion of an automorphic form as a function on $G(F)\backslash G(\mathbb{A}_F)$. I shall return to the notion for a general group in §8. For now, I recall it for GL(1). It is the quotient I_F of the restricted districted product $\mathbb{I}_F = \prod_{x\in X} F_x^\times$ by the diagonally imbedded $F^\times$. There is a filtration

$$\{1\} \subset I_F^{\mathrm{tr}} \subset I_F^{\mathrm{unr}} \subset I_F^0 \subset I_F$$

of the group of ideles $I_F = F^\times\backslash\mathbb{I}_F$ of the group of idele classes, with

$$I_F^{\mathrm{tr}} = \prod_{x\in X} \mathbb{O}_x^+, \qquad I_F^{\mathrm{unr}} = \mathbb{C}^\times \backslash \prod_x \mathbb{O}^\times,$$
$$I_F^0 = F^\times\backslash\mathbb{I}_F^0 = \Big\{x = \prod f_x \in \mathbb{I}_F \,\Big|\, \sum_x \mathrm{ord}_x(f_x) = 0\Big\}.$$

The quotients are

$$I_F^{\mathrm{tr}}\backslash I_F^{\mathrm{unr}} = \mathbb{C}^\times \backslash \prod_x \mathbb{C}^\times,$$

where $\mathbb{C}^\times$ is diagonally imbedded,

$$I_F^{\mathrm{unr}}\backslash I_F^0 = \Big\{(n_x) \in \bigoplus_x \mathbb{Z} \,\Big|\, \sum_x n_x x = \mathrm{div}(f),\ f \in F^\times\Big\} \Big\backslash \Big\{(n_x) \in \bigoplus_x \mathbb{Z} \,\Big|\, \sum_x n_x = 0\Big\},$$

thus the group of divisors of degree 0 modulo principal divisors, and $I_F^0\backslash I_F = \mathbb{Z}$. The idele-class characters in the geometric theory are continuous functions on I_F equal to 1 on a subgroup $\prod_{x\notin S} \mathbb{O}_x^\times$, S a finite set of points in X, and on $\prod_{x\in S} F_x^\times$ to a product of the local characters already introduced.

These characters, or these automorphic forms, certainly need to be considered, but the geometric theory takes a broader view that it took me a good deal of time to appreciate and to reconcile with my simple ideas. The pertinent clue lies in the statement of Theorem 3 of §3.8 of [CFT].

Assertion. *For each irreducible rank n local system E on X there exists a perverse sheaf* Aut_E *on* Bun_n *which is a Hecke eigensheaf with respect to E. Moreover,* Aut_E *is irreducible on each connected component* Bun_n^d.

For the moment, I take $G = \mathrm{GL}(1)$, thus G to be not a general reductive group, and not $\mathrm{GL}(n)$, with n arbitrary as in the assertion, but with $n = 1$, and try to understand the meaning of this assertion. Among other things, it will be important to be clear, as soon as the initial explanations are concluded, about the nature of the difference between automorphic forms in the naive, but legitimate sense taken from the arithmetic theory, even those that are eigenfunctions of the Hecke operators, and a Hecke eigensheaf, at first when there is no ramification. It is an immediate result of Diagram I that for the group GL(1) a local system is just a coset in $H^1(X,\mathbb{Z})\backslash H^1(X,\mathbb{C})$. This is not emphasized in [CFT] or even mentioned, perhaps because the emphasis is on nonabelian groups. Handicapped by my inexperience, I am often, in the theory of algebraic curves, at a loss to distinguish theorems from definitions. Our first task will be to acquire some concrete understanding of the Assertion for $n = 1$ and to introduce its geometric counterpart. I do find it convenient, when reflecting on the Assertion and its analytic counterpart, to fix in mind the dimensions that appear. We shall see, for example, that the line bundles on X are parametrized by a $2g$-dimensional torus and the local systems attached to a given line bundle by a $2g$-dimensional real vector space, so that in the Assertion the possibilities are parametrized by the quotient of a $4g$-dimensional vector lattice by a $2g$-dimensional lattice. Such information assures a failing memory that nothing has been forgotten and nothing counted twice.

The quotient of I_F by $\prod_x \mathbb{O}_x^\times$ is the group of divisors on the nonsingular algebraic curve X for which the global theory is to be developed, taken modulo linear equivalence; it can be given the structure of an algebraic variety. The connected component of this variety, formed by the divisors of degree 0, is then the jacobian of X, which could be identified with the moduli space $\mathcal{P}^0$ of line bundles of degree 0 on X, but we do not do so. The full group is the Picard variety $\mathcal{P}$ itself, which can be identified with the quotient $F^\times \prod_x \mathbb{O}_x^\times\backslash\mathbb{I}_F$, but once again it is convenient to distinguish them.

We can be more precise. Let g be the genus of X. We introduce a complex vector space Ξ of dimension g, the dual space of the space of differential forms of the first kind on X and a lattice Δ in Ξ, given by the complex linear forms

$$\omega \mapsto \int_\delta \omega, \tag{7.6a}$$

$\delta \in H_1(X,\mathbb{Z})$, thus, more informally, but more instructively, δ being a closed curve on X. We introduce as well the real dual space $\hat{\Xi}$ of Ξ (sometimes identified with the space of conjugate linear complex-valued forms, but often with the space of complex linear forms) on Ξ by sending the conjugate linear form μ, $\mu(\alpha x)=\bar{\alpha}\mu(x)$, to $\mathrm{Re}(\mu)$ and the lattice $\hat{\Delta}$ defined by $\hat{\delta}\in\hat{\Delta}$ if and only if $\mathrm{Re}(\hat{\delta}(\delta)) \in 2\pi\mathbb{Z}$ for all $\delta\in\Delta$. It is difficult to distinguish Ξ and $\hat{\Xi}$ or Δ and $\hat{\Delta}$, but $\Delta = H_1(X,\mathbb{Z})$ and $\hat{\Xi}$ may, of course, be identified with the space of differential forms of the first kind on X. Then, thanks to the Abel–Jacobi theory, the map that assigns to the divisor $p_1+\cdots+p_n-q_1-\cdots-q_n$ the linear form

$$\omega \mapsto \sum_{i=1}^n \int_{q_i}^{p_i} \omega$$

defines an isomorphism — for both the group structure and the holomorphic structure — of $\Delta\backslash\Xi$ with the jacobian of X. The group and $\hat{\Delta}\backslash\hat{\Xi}$ the group (or moduli space) $\mathcal{P}^0$ of line bundles of degree 0, thus the connected component of the Picard group. Each of these line bundles admits a connection and the family of connections is given by a coset of $\hat{\Delta}$ in $\hat{\Xi}$. More precisely, as is explained on page 313 of [GH] in the context of complex tori, but the explanation is also valid here, the exact sequence of sheaves

$$H^1(X,\mathbb{Z}) \to H^1(X,\mathbb{O}) \to H^1(X,\mathbb{O}^\times) \to H^2(X,\mathbb{Z}), \tag{7.7}$$

in which there is a factor 2π in the first arrow that must not be forgotten, leads to an identification

$$\mathcal{P}^0 = H^1(X,\mathbb{Z})\backslash H^1(X,\mathbb{O}), \quad \hat{\Xi} = H^1(X,\mathbb{O}) = H^{0,1}(X), \quad \hat{\Delta} = H^1(X,\mathbb{Z}), \tag{7.8}$$

the notation $H^{0,1}(X)$ being taken from Hodge theory, where the space $H^{1,0}(X)$ is the space of differentials of the first kind. Unfortunately, I have difficulty remembering which is which because of the reversal of the order of the 0 and the 1 in the relations $H^1(X,\mathbb{O}) = H^1(X,\Omega^0) \simeq H^{0,1}(X)$, $H^0(X,\Omega^1) \simeq H^{1,0}(X)$.

It is undoubtedly best that I be as precise as I can because my experience with differentials and Hodge theory, even on curves, is limited. For example, $H^1(X,\mathbb{O}) = H^{0,1}(X)$ is the complex conjugate of $H^{1,0}(X)$, the space of differential forms of the first kind, this identification being given by the Hodge $\star$-operator ([GH, page 82]). There are two isomorphisms of $H^{1,0}(X)$ as a vector space over $\mathbb{R}$ to $\hat{\Xi}$; they are given by the real and imaginary parts of the periods (7.6a). To continue, I return to an enlarged form of the diagram (7.7), suppressing the explicit reference to X from the notation.

$$\begin{array}{ccccccccccc}
\{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{C}) & \longrightarrow & H^1(\mathbb{C}^\times) & \longrightarrow & H^2(\mathbb{Z}) & \longrightarrow & H^2(\mathbb{C}) = \mathbb{C} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{O}) & \longrightarrow & H^1(\mathbb{O}^\times) & \longrightarrow & H^2(\mathbb{Z}) & \longrightarrow & H^2(\mathbb{O}) = \{0\}
\end{array}$$

Diagram I

The central square of the diagram is summarized in [GT, §2], although with reference to a general group G, not just GL(1): *"A flat connection has two components. The* (0, 1) *component, with respect to the complex structure on* X, *defines holomorphic structure, and the* (1, 0) *component defines a holomorphic connection."* According to the Hodge theory an element of $H^1(\mathbb{C}) = H^{1,0}(X) \oplus H^{0,1}(X)$ is realized uniquely as a sum of a holomorphic form and an antiholomorphic form. In the diagram, the third vertical arrow, $H^1(\mathbb{C}) \to H^1(\mathbb{O})$, is the projection on the second factor. Since the last arrow in the first line is an injection, the kernel of $H^1(\mathbb{C}^\times) \to H^1(\mathbb{O}^\times)$ is a complex vector space of dimension g, isomorphic to $H^{1,0}(X)$.

A flat connection is a connection in which there is a local notion of constant section; these are obviously given by $H^1(\mathbb{C}^\times)$, while $H^1(\mathbb{O}^\times)$ parametrizes line bundles. Since the Chern class, given by the degree of a line bundle, is the image of its parameter in $H^2(\mathbb{Z})$, we see that the collection of connections on a given line bundle is parametrized by $H^{1,0}(X)$. Thus the collection of flat connections on a given line bundle, which are parametrized by the inverse image of an element of $H^1(\mathbb{O}^\times)$ form an affine space over $H^{1,0}(X)$. A line bundle of degree 0, thus of Chern class 0, is an element of $H^1(\mathbb{O}^\times)$, thus an element of $H^1(\mathbb{Z})\backslash H^1(\mathbb{O})$ or a coset of $H^1(\mathbb{Z})$ in $H^1(\mathbb{O}) = H^{0,1}(X)$ or, as seems to be demanded by the formalism, by its complex (thus $\star$-)conjugate in $H^{1,0}(X)$. The comment of [GT] cited is the observation that an element of $H^1(\mathbb{C}^\times)$ is the image of an element of $H^1(\mathbb{C}) = H^{0,1} \oplus H^{1,0}$. Mapping this element to $H^1(\mathbb{O})$ amounts to projection on its first component. The horizontal arrow then yields a line bundle. So the bundle determines the first component. (As a test for my orientation: this is $H^1(\Omega^0)$ appearing in the Hodge theory as $H^0(\overline{\Omega})$, thus as an antiholomorphic differential, whose orthogonal complement lies in the holomorphic direction $\partial/\partial z$.) To determine the image of the upper horizontal arrow, we need to know both the first and the second component. So the supplementary information needed to determine the second component of an element of $H^1(\mathbb{C})$ is contained in its image in $H^1(\mathbb{C}^\times)$, thus in the connection.

It will be worthwhile to return to the geometric theory at the end of this section, just to understand better what the Assertion means for $n = 1$ and how it can be proved, but our principal goal is to introduce an analytic form of it that will allow us to introduce a candidate for the abelian quotient $\mathfrak{A}_{\mathrm{ab}}$ of $\mathfrak{A}$. The analytic form has quite a different flavor.

We have already defined $\hat{\Xi}$ as $H^{1,0}(X)$ and defined the periods of an element of $H^{1,0}(X)$, thus of a differential form of the first kind, by (7.6a). Then the conjugate space $\hat{\Xi}^{\mathrm{conj}}$ is $H^{0,1}(X)$ and their sum can be identified with $H^1(X, \mathbb{C})$, a $2g$-dimensional space, represented by holomorphic differential forms with arbitrary periods. We conclude that the span of the periods, either the real periods or the complex periods, for both are not simultaneously necessary,

$$\delta \mapsto \mathrm{Re} \int_\delta \omega, \quad \delta \mapsto \mathrm{Im} \int_\delta \omega, \quad \omega \in H^{1,0}(X),$$

is just Ξ, treated as a real vector space, thus the real dual of $\hat{\Xi}$. If we express the surface X in the usual way as a disc with boundary

$$\delta_1 \delta_{g+1} \delta_1^{-1} \delta_{g+1}^{-1} \cdots \delta_g \delta_{2g} \delta_g^{-1} \delta_{2g}^{-1},$$

the various segments on the boundary being identified as indicated by the subscripts, then any additive mapping of this sort is determined by its values on $\delta_1, \dots, \delta_{2g}$.

Observe that, because there is a multiplication by $2\pi i$, $z \to 2\pi i z$, the image of $H^1(X, \mathbb{Z})$ is characterized by the conditions that the real parts of the periods are 0 and the imaginary parts lie in $2\pi\mathbb{Z}$.

If $\omega \in \hat{\Delta}$ and p_0 is an arbitrary but fixed point on X, then

$$p \mapsto \exp\left(i \operatorname{Im} \int_{p_0}^{p} \omega\right) \tag{7.6b}$$

is a continuous character of the jacobian (or of the Picard variety $\mathcal{P}^0$) and, as ω varies over $\hat{\Delta}$, we obtain in this way a family of $\mathbb{Z}^{2g}$ characters of I_F^0, which can, of course, be extended to I_F, but this is a secondary matter, the choice of a nonzero constant. The character is defined by its differential equation,

$$\chi_R^{-1} d\chi_R = i \operatorname{Re} \omega = i\omega_R \quad \text{or} \quad \chi_I^{-1} d\chi_I = i \operatorname{Im} \omega = i\omega_I, \tag{7.9a}$$

either of which defines in some sense a holonomic system or a *perverse* sheaf, but in a real context. The usual holonomic system would be given by the complex equation

$$\chi^{-1} d\chi = \omega, \tag{7.9b}$$

which may also be treated as two real equations. So it has more boundary conditions, thus conditions of periodicity. If the local coordinate is $z = x + iy$, $\omega = (\mu + i\nu)(dx + i dy)$, and if $\chi = \exp(\alpha(x, y) + i\beta(x, y))$, then (7.9b) amounts to

$$\left(\frac{\partial\alpha}{\partial x}, \frac{\partial\beta}{\partial x}\right) = (\mu, \nu), \qquad \left(\frac{\partial\alpha}{\partial y}, \frac{\partial\beta}{\partial y}\right) = (-\nu, \mu).$$

For the second equation, that in (7.9b), periodic conditions are not appropriate; for one or the other of the first, they are. For the second, boundary conditions would be to combine both conditions of (7.9a). So they are again usually impossible to satisfy. In the analytic or arithmetic theory, it is the second of equations (7.9a) that is pertinent. In the context of perverse sheaves, thus in the context of the Assertion, the issue of a global solution of the differential equation is inappropriate. I was troubled and confused by this difference for some time. Its source has become clearer. One thinks of the exponential function $\exp \lambda z$ on the interval [0, 1], with 0 and 1 identified. If one wants functions, one needs $\lambda \in 2\pi i\mathbb{Z}$; if one accepts sheaves, thus the differential equation

$$\frac{dh}{dz} = \text{constant}$$

is acceptable. One reflects an analytic impulse, my dominant impulse; the other a geometric impulse, by which [CFT] is guided. The notion of a Hecke eigensheaf that appears there is, as we shall see, a clever way of admitting this greater generality. As already observed, it can also be incorporated into the analytic theory mediated by characters of the fundamental group The two possibilities could be examined separately, but as my purpose here was to adumbrate an analytic theory that would not lag behind the geometric theory, I have preferred to incorporate some to-and-fro in the exposition, as well as some redundancy.

To add to it, we reflect just a minute on the equation

$$\frac{1}{h}\frac{dh}{dz} = \lambda$$

on the circle, realized as the real line modulo $2\pi z$. On the line it defines a flat connection on the trivial bundle because the quotient of any two solutions $c_1 \exp(\lambda z)$, $c_2 \exp(\lambda z)$ differ by a multiplicative constant. It also defines a flat connection on the circle because $c_1 \exp(\lambda z)$, $c_1 \exp(\lambda(z + 2\pi))$ also differs by a multiplicative constant. It does not, however, define a section of the trivial bundle as a bundle on $\mathbb{Z}\backslash\mathbb{R}$. Trivial as the difference is, I find it, as the reader will discover, hard to fix in my mind. The integral of a constant function becomes linear and after passing to the exponential even more difficult to recognize. This becomes even worse with a curve and its jacobian. The jacobian is a quotient of a linear space on which a differential of the first kind is just a constant element of the dual; on the curve itself, it is hardly linear. The danger of confusing the intuition is even more severe for differentials with values in a vector bundle or in a Lie algebra. Another feature that leads to confusion is that the equations (7.9a) and (7.9b) describe the development of a complex line, thus of a real plane, or better a local section of a $\mathbb{U}(1) \subset \mathrm{GL}(1)$ bundle over X, even of a local system of X, because for it there is a local notion of constant section. I have only increased the possibility of confusion by referring to boundary conditions; at best, we are dealing with boundary conditions on a rectangle.

In the analytic theory, we are dealing with characters, thus with functions with values in the group $\mathbb{U}(1)$ of complex numbers of absolute value 1. So we are dealing with one or the other of the equations (7.9a), say the first, or, in other words, with $\mathbb{U}(1)$-bundles. The sequence

$$\{1\} \to \mathbb{Z} \to \mathbb{R} \to \mathbb{U}(1) \to \{1\}$$

yields an analogue of Diagram I, in which the vertical arrow $H^1(\mathbb{R}) \to H^1(\mathbb{O})$ is an isomorphism.

$$\begin{array}{ccccccccccc}
\{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{R}) & \longrightarrow & H^1(\mathbb{U}(1)) & \longrightarrow & H^2(\mathbb{Z}) & \longrightarrow & H^2(\mathbb{R}) = \mathbb{R} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{O}) & \longrightarrow & H^1(\mathbb{O}^\times) & \longrightarrow & H^2(\mathbb{Z}) & \longrightarrow & H^2(\mathbb{O}) = \{0\}.
\end{array}$$

Diagram II

The significance of the diagram is that every holomorphic line bundle is realized as a $\mathbb{U}(1)$-bundle, that each of them carries a unique local system in the real sense, thus with constant transition functions in $\mathbb{U}(1)$. All this is simple, but it has taken me some time to appreciate the consequence: the analytic theory is closely related to the holomorphic theory but different from it. In the analytic theory, the local systems have automorphisms: sections of the associated $\mathbb{U}(1)$-bundle. These are automorphic forms, but this may not be pertinent.

It is worthwhile to explain this further. In the theory of algebraic curves, there is a great deal of structure crammed into a very small space and it is difficult to describe it in an orderly fashion. Starting with an element η in $H^1(\mathbb{O}) = H^{0,1}$, we add its image $\star\eta$, which lies in $H^0(\Omega^1) = H^{1,0}$, to it and divide by 2 to obtain a form $\operatorname{Re}\omega$. Then the element of $H^1(\mathbb{U}(1))$, thus a flat sections of the bundle is given by

$$\exp\left(2\pi i \int_{p_0}^{p} \operatorname{Re}\omega\right).$$

Notice that, because of the presence of $H^1(\mathbb{Z})$ at the beginning of each line and because of the factor $2\pi i$ that appears in the passage from $\mathbb{R}$ to $\mathbb{U}(1)$, $\operatorname{Re}\omega$ is determined only up to a form with integral periods.

When the geometric theory is treated as an offshoot of the arithmetic theory, the restriction to unramified representations or forms is unnatural. It is also unnecessary. For the global theory on a nonsingular algebraic curve X, the space of differentials of the first kind is replaced by the space $\Omega = \Omega_X$ of global meromorphic differentials ω with local expressions ω_x. There is one condition on the forms ω considered that one might be tempted to impose: the residue at each point must be integral. I omit it for a brief moment, because it took me sometime to understand the significance of such a condition, but I shall very quickly impose it. Even if it is unnecessary, there are already enough other complications to master. One consequence is that the solutions of the differential equation $df/f = \omega$ are single-valued in a neighborhood of each point, so that no singularities are introduced locally into the sheaf of solutions. So the distinction is — perhaps — between a *sheaf*, thus by the differential equation, that is by ω, which is well defined in all cases, and, up to a constant, a *single-valued* function, its solution, which is not! The analytic impulse, as well as the arithmetic, is to emphasize the function; the geometric impulse is to emphasize the sheaf. The consequences of the two points of view have already revealed themselves. If the condition of integrality is imposed, the periods of ω, thus its integrals over one-cycles, are well-defined modulo $2\pi i\mathbb{Z}$. In particular the real and imaginary parts are defined modulo $2\pi\mathbb{Z}$. The periods as such are not defined because an integral of ω even over a cycle homologous to zero is given by the sum of the residues in the 1-chain that it bounds.

There are two difficulties in the Assertion, apart from understanding how it is verified. First of all, it is an assertion only in an unramified context. Coming from the arithmetic theory of automorphic forms or representations, I find this an unacceptable restriction. We shall remove it. Secondly, there is no immediate link to a theorem of Weil that we shall recall later and that offers a partial solution to the problem of identifying the geometric galoisian group $\mathfrak{A}$, a kind of self-duality similar to that of class field theory.

I recall the structure of the vector space of meromorphic differentials on X. First of all the space of differentials without singularities, thus differentials of the first kind, has dimension equal to g, the genus of the curve. Secondly, the singularities may be assigned almost arbitrarily. There is only one constraint: the sum of the residues must be 0. This is a consequence of, for example, the theorem of Riemann–Roch, which can be cited in the form given in [Sp]. Take a finite set of points $x_1, \dots, x_n$ on the curve X and integers $d_1, \dots, d_n$. Then the space of possible singularities concentrated on this finite set and of degree at most d_i at x_i is of dimension $\sum_i d_i$. If a given singularity can be realized by a meromorphic differential ω then any other realization is of the form $\omega + \omega'$, where ω' lies in the g-dimensional space of holomorphic differentials. So to prove that differentials can be assigned arbitrarily, we need only verify that for all choices of $x_1, \dots, x_n$ and of $d_1, \dots, d_n$ with

$$d_1 + \cdots + d_n > 0, \tag{7.10}$$

the space of differentials with the singularities allowed by these choices is of dimension $g - 1 + \sum_i d_i$. The condition (7.10) takes account of the constraint that the sum of the residues is 0. Take the divisor $\mathfrak{a}$ on page 264 of Springer's book [Sp] to be $-\sum_i d_i x_i$. Then, according to the form of the Riemann–Roch theorem given there, the dimension of the space of possible differentials is

$$i(\mathfrak{a}) = g - 1 + \sum d_i,$$

because, in the notation of [Sp], $d(\mathfrak{a}) = -\sum d_i$ and $r(-\mathfrak{a}) = 0$.

It is convenient to introduce an increasing sequence of differential forms: the forms with no singularities, thus the forms of the first kind; the forms whose only singularities are simple poles; finally, the forms with arbitrary singularities. We can then add the supplementary condition, already introduced, that the residues be integral. If $\delta_1, \delta_2, \dots, \delta_{2g}$ is the base of the integral cycles, then $\omega \to \mathrm{Re} \int_{\delta_i} \omega$ (or $\mathrm{Im} \int_{\delta_i} \omega$) defines $2g$ linear forms linearly independent over $\mathbb{R}$ on the g-dimensional complex space of forms of the first kind. Moreover, as ω varies, the $2g$-dimensional vectors in $\mathbb{R}^{2n}$ or, perhaps better, in $\hat{\Xi}$,

$$\begin{aligned}\varpi_R &= \{\varpi_{R,i}\} = \bigl\{\mathrm{Re}\textstyle\int_{\delta_i}\omega \bigm| i = 1, \ldots, 2g\bigr\},\\ \varpi_I &= \{\varpi_{I,i}\} = \bigl\{\mathrm{Im}\textstyle\int_{\delta_i}\omega \bigm| i = 1, \ldots, 2g\bigr\},\end{aligned} \tag{7.11}$$

are arbitrary, but not independently arbitrary. Without the condition on the residues, they are both path-dependent. Even if the condition on the residues is imposed, the second is only defined modulo $2\pi\mathbb{Z}$, but, as already observed, that is all we need because we use $\exp(i\,\mathrm{Im}\int\omega)$. The condition on the residue prevents us from multiplying all ω by i or $-i$, so that the set of ϖ_I and ϖ_R may be different.

For the moment, we are dealing with line bundles, so that $n = 1$. My impulse was to look for a theorem in which irreducible, thus one-dimensional, automorphic representations of the geometric form of the group of idele classes appear. If there is no ramification — and if we admit as automorphic representations only continuous functions in the parameter $x \in X$, thus only continuous characters of the group of divisors modulo those linearly equivalent to 0, a group whose connected component is the jacobian, thus a complex variety of a dimension g that, as a group, may be identified with $\mathbb{U}^{2g}$, $\mathbb{U} = \{z \in \mathbb{C}^\times \mid |z| = 1\}$ — its group of unitary characters, thus the set of irreducible unramified automorphic representations, is isomorphic to $\mathbb{Z}^{2g}$, or to the group of differentials ω of the first kind for which ϖ_I lies in $(2\pi\mathbb{Z})^{2g}$. It is easy enough to make the isomorphism explicit in terms of $F^\times\backslash\mathbb{I}_F^0$ — rather than in terms of the jacobian — by applying the method of [GH] for proving a theorem of Weil, and we shall do so.

What then is the purpose of the remaining ω, either the remaining ω of the first kind or, more generally, the differentials with singularities? For the comparison with the Assertion, which is implicitly stated in an unramified context, it is the differentials of the first kind that are relevant, but for the description of the global group $\mathfrak{A}$ in the geometric context, it will be necessary to admit differentials with singularities, thus with negative powers in their local Laurent expansions. For this prologue, however, it is best to consider only those with integral residue, since a nonintegral residue introduces ramification in the line bundles themselves — local sections at some points behave like z^α, $\alpha \in \mathbb{C}$. That would, at this stage, be one complication too many.

For the moment, we remain with differentials with no singularities. We count — once again — the parameters available. I refer to Diagram I. It is clear from the lower line of the diagram that line bundles are parametrized by $\mathbb{Z}^{2g}\backslash\mathbb{R}^{2g}$. The upper line then shows that the possible local systems on a given line bundle are parametrized by $H^{1,0}(X) = H^0(X, \Omega^1)$, thus by $\mathbb{R}^{2g}$. So all in all, we need $\mathbb{Z}^{2g}\backslash\mathbb{R}^{4g}$ to specify a local system. On the other hand, in the analytic/arithmetic context the set of unramified automorphic forms is given by the $\mathbb{Z}^{2g}$ characters of the jacobian, or rather by their extensions to characters of $F^\times\backslash\mathbb{I}_F$, parametrized by $\mathbb{U}$ or by characters of $\mathbb{I}_F^0\backslash\mathbb{I}_F$. These extensions are incidental to the central issue. So the puzzling matter is the presence in the geometric theory of supplementary parameters in $\mathbb{Z}^{4g}\backslash\mathbb{R}^{4g}$. We shall introduce them artificially. I was, initially, made more than a little uneasy by the artifice.

We shall return to this point, but only after broaching the question of attaching, in the geometric theory, an idele-class character to a differential ω, perhaps singular but with integral residues. It turns out that this entails an enlargement of the notion of idele class. We take the product of $\mathbb{I}_F$ with $4g$ copies of $\mathbb{Z}$, thus with two copies $\mathfrak{Z}_R$ and $\mathfrak{Z}_I$ of $\mathbb{Z}^{2g}$, so that the dual of the modified group is the group of characters of $\mathbb{I}_F$ multiplied by two copies of the $2g$-fold product of $\mathbb{Z}\backslash\mathbb{R}=\mathbb{U}(1)$ with itself. We introduce an imbedding of $F^\times$ in $\mathfrak{J}_F=\mathfrak{Z}_R\times\mathfrak{Z}_I\times\mathbb{I}_F$ by

(7.12) $$f\mapsto\prod_{i=1}^{2g}\frac{\int_{\delta_i}d\ln f}{2\pi i}\times\prod_{i=1}^{2g}\frac{\int_{\delta_i}d\ln f}{2\pi i}\times f.$$

There will be a finite number of points $q_1, q_2,\dots$ at which ω has a singularity and, for any given idele f, a finite number of points $p_1, p_2,\dots$ at which $f=\prod_x f_x$ has a zero or pole. If the sets $D_\omega=\{q_1,q_2,\dots\}$ and $D_f=\{p_1,p_2,\dots\}$ are disjoint and if $f\in F^\times$ is a principal idele we may introduce λ_R as the difference of

(7.13a) $$\mathrm{Re}\Big\{\sum_j\mathrm{res}_{q_j}(\omega\ln f)-\sum_i\mathrm{ord}_{p_i}(f)\int_p^{p_i}\omega-\gamma\sum_i\mathrm{ord}_{p_i}(f)\Big\}$$

and

(7.13b) $$\frac{1}{2\pi i}\sum_{k=1}^{g}\Big\{\varpi_{R,g+k}\cdot\int_{\delta_k}d\ln f-\varpi_{R,k}\cdot\int_{\delta_{g+k}}d\ln f\Big\},$$

where γ is a supplementary complex parameter, p is a supplementary point, and a choice of path from p to p_i that avoids the singularities of ω is implicit for each i. It modifies the value of λ_R only by an additive constant in $2\pi\mathbb{Z}$. We want to introduce a pairing $(\omega,f)_R=\exp(i\lambda_R)$ defined for all ideles f. The expression (7.13a) is certainly defined; the expression (7.13b) is not, but it is defined if we replace f by an element $\tilde f$ of $\mathfrak{Z}_R\times\mathfrak{Z}_I\times\mathbb{I}_F$ and $\int_{\delta_k}d\ln f$ by 2π times the appropriate coordinate of the $\mathfrak{Z}_R$ component of $\tilde f$. This defines $(\omega,\tilde f)_R$ in general. We define $(\omega,\tilde f)_I$ in the same manner. It is simpler to abbreviate $\tilde f$ to f, and I do so in the following discussion, inserting the tilde if its omission would lead to confusion or as a reminder.

The parameter γ only affects the pairing at those f whose total degree $\sum_i\mathrm{ord}_{p_i}f$ does not vanish and two pairs (ω,γ), (ω',γ') yield the same pairing if $\gamma'-\gamma=\int_p^{p'}\omega$. As we did for the local parameters, we shall have to use both pairings $(\omega,f)_R$

and $(\omega, f)_I$ unless all the residues β_0 of ω are zero. So the parametrization of characters might best be expressed in terms of pairs (ϖ_R, ϖ_I) with an appropriate equivalence relation, but this would be too fastidious for a prologue and, in any case, obvious.

The key to the global definition of $(\omega, f)_R$ or $(\omega, f)_I$, whose properties have yet to be discussed, is a generalization of a theorem attributed in [GH] to Weil. I formulate the generalization as a lemma that implies that, for each ω, $f \mapsto (\omega, f)_R$, $f \mapsto (\omega, f)_I$ define idele-class characters.

Lemma 7.1. *If f is a principal idele and D_ω and D_f are disjoint, then $(\omega, f)_R = (\omega, f)_I = 1$.*

The theorem of Weil affirms that if f and g are meromorphic functions on the compact Riemann surface X such that the set of zeros and poles of f is disjoint from the set of zeros and poles of G then

$$\prod_p f(p)^{\mathrm{ord}_p(g)} = \prod_p g(p)^{\mathrm{ord}_p(f)}. \tag{7.14}$$

For the simplest example, the projective line $\mathbb{P}^1$, the theorem is elementary and easy to prove. Suppose, for example that $f = (x - a_1)/(x - b_1)$, $g = (x - a_2)/(x - b_2)$. Then

$$(f, g) = \frac{g(a_1)}{g(b_1)} \frac{f(b_2)}{f(a_2)} = \frac{a_1 - a_2}{a_1 - b_2} \frac{b_1 - b_2}{b_1 - a_2} \frac{b_2 - a_1}{b_2 - b_1} \frac{a_2 - b_1}{a_2 - a_1} = 1$$

In general, the theorem is a consequence of a relation like that of the lemma, but for $\omega = dg/g$. The idele f is still to be principal. The relation becomes

$$\sum_j \mathrm{res}_{q_j}(\omega \ln f) - \sum_i \mathrm{ord}_{p_i}(f) \int_p^{p_i} \omega = \lambda \in 2\pi i \mathbb{Z}. \tag{7.15}$$

The new relation is stronger, or, rather, more compact, because the periods of ω themselves now lie in $2\pi i \mathbb{Z}$. This is not just a condition on the real or imaginary parts. We recall the proof given on page 229 and on pages 242–243 of [GH], following, so far as possible, the notation of that book. We have already followed it with the usual description of the basic cycles $\delta_1, \dots, \delta_g, \delta_{g+1}, \dots, \delta_{2g}$ that display the surface as a planar polygon Δ with sides identified. We have a function f with poles and zeros at p_i and a form $\omega = dg/g$ with first-order poles at q_j. The sets $\{p_i\}$ and $\{q_j\}$ are taken to be disjoint. The p_i and the q_j are to lie in the interior of the planar region and we join each p_i to a common point p on the boundary by a curve α_i that avoids the q_j, thus introducing incisions that reduce Δ to a region Δ' and add several curves to its boundary, the curve α_i and the curve in the inverse direction.

Since $\sum_i \operatorname{ord}_{p_i}(f) = 0$ 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 affect on the relation (7.15).

As in [GH], we integrate the form $\varphi = \omega \ln f$ over the boundary of Δ'. By the residue theorem, this integral is given by

$$\int_{\partial\Delta'} \varphi = 2\pi i \sum_{q_j} \operatorname{res}_{q_j} \varphi = 2\pi i \sum_{q_j} \operatorname{res}_{q_j}(\omega \ln f). \tag{7.16}$$

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},

$$\ln f(p') = \ln f(p) + \int_{\delta_{g+i}} d \ln f, \tag{7.17}$$

so that

$$\int_{\delta_i + \delta_i^{-1}} \varphi = \Big(\int_{\delta_i} \omega\Big)\Big(-\int_{\delta_{g+i}} d \ln f\Big). \tag{7.18}$$

In the same way,

$$\int_{\delta_{g+i} + \delta_{g+i}^{-1}} \varphi = \Big(\int_{\delta_{g+i}} \omega\Big)\Big(-\int_{\delta_i} d \ln f\Big). \tag{7.19}$$

Moreover for identified points $p \in \alpha_i$ and $p' \in \alpha_i^{-1}$,

$$\ln f(p') - \ln f(p) = -2\pi i \operatorname{ord}_{p_i}(f). \tag{7.20}$$

so that[1]

$$\int_{\alpha_i + \alpha_i^{-1}} \varphi = 2\pi i \operatorname{ord}_{p_i}(f) \int_p^{p_i} \omega.$$

As in [GH], the conclusion is that

$$2\pi i \Big\{ \sum_j \operatorname{res}_{q_j}(\omega \ln f) - \sum_i \operatorname{ord}_{p_i}(f) \int_p^{p_i} \omega \Big\}$$

is equal to

$$\sum_{k=1}^{g} \{ \int_{\delta_k} d \ln f \cdot \int_{\delta_{g+k}} \omega - \int_{\delta_k} \omega \cdot \int_{\delta_{g+k}} d \ln f \} \tag{7.21$'$}$$

[1] 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.

or

$$\sum_{k=1}^{g}\left\{\int_{\delta_k} d\ln f\cdot\int_{\delta_{g+k}} d\ln g-\int_{\delta_k} d\ln g\cdot\int_{\delta_{g+k}} d\ln f\right\}. \tag{7.21$''$}$$

In (7.21″), 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 (7.14) follows.

To prove the lemma itself, we deal with $(\omega,\cdot)_R$ and, implicitly, $(\omega,\cdot)_I$ with essentially the same sequence of formulas. Since f is now a principal idele, the term in (7.13a) that contains γ is 0, and (7.13a) itself is reduced to the real part of

$$\lambda=\sum_j \operatorname{res}_{q_j}\{\omega\ln f(q_j)\}-\sum_i \operatorname{ord}_{p_i}(f)\int_p^{p_i}\omega \tag{7.22}$$

and the assertion is that the difference between the real part of (7.22) and (7.13b) lies in $2\pi\mathbb{Z}$. The proof is the same as before; we deal with (7.22) as we dealt with (7.15), collecting terms in the same way:

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

The conclusion is that

$$2\pi i\sum_j\left(\operatorname{res}_{q_j}(\omega)\ln f(q_j)-\sum_i \operatorname{ord}_{p_i}(f)\int_p^{p_i}\omega\right)$$

is equal to

$$\sum_{k=1}^{g}\left(\int_{\delta_k} d\ln f\cdot\int_{\delta_{g+k}}\omega-\int_{\delta_k}\omega\cdot\int_{\delta_{g+k}} d\ln f\right). \tag{7.23}$$

To calculate λ, we take the imaginary part of this, divide by 2π, and subtract (7.13b). This yields

$$\sum_{k=1}^{g}\left(\int_{\delta_k} d\ln f\cdot\left(\operatorname{Re}\int_{\delta_{g+k}}\omega-\varpi_{R,g+k}\right)-\left(\operatorname{Re}\int_{\delta_k}\omega-\varpi_{R,k}\right)\cdot\int_{\delta_{g+k}} d\ln f\right). \tag{7.24}$$

The periods of $d\ln f$ are all multiples of $2\pi i$ and the numbers $\operatorname{Re}\int_{\delta_k}\omega-\omega_{R,k}$, $k=1,\ldots,2g$, are also all integral multiples of 2π. Indeed they are 0, but that is not the point here. This proves the lemma!

There is a difficulty with the pairings $(\omega, f)_R$ and $(\omega, f)_I$ that is resolved by the lemma. For a given ω, it is not defined for all ideles f, or, to be precise, $\tilde{f}$, only for those for which D_f and D_ω are disjoint. We can extend it to all ideles by setting any given idele f equal to $f_1 f_2$, where f_2 is principal and f_1 is an idele whose set of zeros and poles is disjoint from the set of singularities of ω. Then we set $(\varpi_R, f) = (\varpi_R, f_1)$, $(\varpi_I, f) = (\varpi_I, f_1)$. Thanks to the lemma, the result will be independent of the choice of the factorization of f.

There is a second difficulty, not resolved by the lemma, at least not without closer examination. What do we do if the function f or the differential ω has a singularity at a point x on the boundary of $\triangle$, say in δ_l and thus in δ_l^{-1}. So it can be approached in two ways from within $\triangle$, one through a half-neighborhood of a subinterval of δ_l, the other through a half-neighborhood of δ_l^{-1}. If the limiting results for the differences of (7.13a) and (7.13b) are the same modulo 2π, there is no problem. We just deform δ_l a little around the offending point and the choice of the deformation, whether we deform a little to the left in the sense of δ_l or in the sense of δ_l^{-1} to make the calculation does not matter. Since the singularities of ω are assumed not to fall on the singularities of f, we can treat the two independently.

The contribution of a singularity of ω to the first term of (7.13a) does not depend on the relation of its position to the curve δ_l. On the other hand, the second term is affected as are the factors $\varpi_{R,k}$. The first is affected because the integral, inside $\triangle$, from p to p_i as a point on δ_l^{-1} is replaced by an integral over a path inside $\triangle$ from p to p_i as a point on δ_l. The difference is a multiple of $2\pi i$ and is multiplied by $\mathrm{ord}_{p_i}(f)$. So it causes no problem. The factor $\varpi_{R,l}$ is deformed but the result is an additive modification by $2\pi i$ times the residue of ω, which is assumed to be integral.

The singularities of f appear in both (7.13a) and (7.13b). Since it is easier, we consider first the effect on (7.13b). The path δ_l first passes to the right of the point and then to the left. So the modification in $\int_{\delta_l} d\ln f$ is $2\pi i\, \mathrm{ord}_{p_i}(f)$, and in (7.13b) $\pm\varpi_{R,g+l'}\, \mathrm{ord}_{p_i} f$, where l' is l or $l-g$ according as $l \le g$ or $g < l \le 2g$. It is evident that something similar will happen with (7.13a). The factor $\mathrm{ord}_{p_i} f$ is already in evidence. The modification is therefore given by the negative of the integral over the path from the point p_0 to p_i on δ_l followed by the inverse path from p_i on δ_l^{-1} to p_0. The two together, with sign, yield a closed path within $\triangle$ from p_i on δ_l^{-1} to p_i on δ_l. Since we can deform the path inside the contour at the cost of adding an integral multiple of 2π, we might as well move directly along the boundary. The integrals along δ_l and δ_l^{-1} cancel and we are left with the integral along δ_{l+g} if $l \le g$ and along the inverse of δ_{l-g} if $l > g$. So up to an additive factor that is an integral multiple of 2π, the difference does not change. I apologize to the more skillful reader for the clumsy argument. I hope it is correct!

The conclusion is that we have attached to ω two characters $\tilde{f} \mapsto (\omega, \tilde{f})_R$ and $\tilde{f} \mapsto (\omega, \tilde{f})_I$ of $F^\times \backslash \mathfrak{I}_F$. In order to persuade ourselves that we indeed have, in a useful way, identified all idele class characters, but also to understand what we have in hand, we remind ourselves of the structure of the group of ideles, or rather of $\mathfrak{Z}_F$, and of its character group, and then of the structure of the group of characters constructed from the admissible differentials.

I observe first of all, to make the task easier, that $\mathfrak{Z}_R \times \mathfrak{Z}_I$ is a subgroup of $F^\times\backslash\mathfrak{I}_F$ and that the characters $(\omega, \cdot)_R$ and $(\omega, \cdot)_I$, certainly yield, upon restriction, all characters of this subgroup. The restrictions are trivial if

$$\varpi_{R,i} \equiv 0 \pmod{2\pi}, \quad \varpi_{I,i} \equiv 0 \pmod{2\pi}, \qquad 1 \le i \le 2g. \tag{7.25}$$

So the issue is whether we obtain all characters of $F^\times\backslash I_F$ from forms ω satisfying one or the other of the two conditions.

We are dealing with a great deal of structure in a very small space. We begin with the curve X, an intuitively difficult object. Then we pass to its jacobian jac_X, the quotient of a vector space Ξ, which is a vector space over $\mathbb{C}$ and thus over $\mathbb{R}$ as well by a distinguished lattice Δ. The jacobian carries not only the structure of a complex manifold, but also the structure of an algebraic variety, and of a group. There are also algebraic mappings of $X \times X$, $(x, y) \mapsto x - y$ of $X \times X$ to jac_X. Analytically — and if we exclude all ramification — the functions of immediate interest are functions on a subgroup of the group of idele-classes, namely on the group $I_F^{\mathrm{unr}}\backslash I_F^0 = F^\times \mathbb{I}^{\mathrm{unr}}\backslash\mathbb{I}_F^0$, indeed they are characters of this group. Such characters are determined by their values on the elements represented by $f_{u,v} = \prod_x f_x$, where $f_x = 1$, except for $x = u, v$ and $f_u = z_u^{-1}$, $f_v = z_v$, z_u and z_v being local parameters at u and v. Such functions are obtained by taking characters χ of $\Delta\backslash\Xi$ and pulling them back to functions χ' on $I_F^{\mathrm{unr}}\backslash I_F^0$ by setting $\chi'(fu, v) = \chi(u - v)$, $u - v$ being the image of (u, v) in the jacobian. This does not function in the geometric context because the functions χ are not holomorphic. It does function in the geometric context if we take χ as a holomorphic character of Ξ, thus a function $\exp(\lambda(\cdot))$ where λ lies in the dual of Ξ as a complex space. This appears to be the expedient found by the geometers. It suggests that analysts, too, not demand that χ be a character of $\Delta\backslash\Xi$, only that it be given by a real linear form λ on Ξ, $\chi(\cdot) = \exp(i \operatorname{Re} \lambda(\cdot))$. This is effectively what we have done.

Each element of the parameters that we propose for the characters of I_F is determined by two elements $\operatorname{Re}\omega$, $\operatorname{Im}\omega'$ — the first element satisfying the first set of conditions (7.25), the second the second set — because we allow products of $(\omega, \cdot)_R$ and $(\omega', \cdot)_I$, where ω' may or may not be equal to ω, and by a constant γ that may be taken as real and is only pertinent modulo $2\pi\mathbb{Z}$. It is clear that with the duality proposed, the function of γ is to generate the characters of $I_F^0\backslash I_F$. It is the characters determined by $\operatorname{Re}\omega$ and $\operatorname{Im}\omega'$ that matter. We pass to them, thus implicitly passing to the quotient by the subgroup of characters generated by the γ. It has already been observed that there is a classical filtration: forms of the first kind (with no singularities) are a subset of forms of the third kind (with at most simple poles, where for our purposes the residues must be integral), and these are in turn a subset of the forms with singularities of arbitrarily high order (but always with integral residues.)

If ω is a form of the first kind, the form $i\omega$ satisfies the condition on the residue of integrality because the residues are all 0. It is therefore unnecessary to include the second $\omega(=\omega')$ or, rather, the contribution $(\omega', \cdot)_I$. More precisely, we have to divide by pairs (ω, ω') of differentials of the first kind for which $\omega = i\omega'$, but this is a fastidious point of the parametrization. Since we know that γ accounts for all characters of $I^0\backslash I$, to establish the desired duality we need only examine the restriction of the remaining characters to I^0. The p that appears in (7.22) is a matter of indifference. The differential forms of the first kind can be regarded as complex linear forms on the complex vector space defining the jacobian. For the exponential $\exp(i \operatorname{Re} \int_0^p \omega)$ to define a character of the jacobian, the real parts of the $2g$ periods of ω must lie in $2\pi\mathbb{Z}$. This is the real part of the condition (7.25). The imaginary part is not relevant here. It clearly defines a lattice in the g-dimensional complex dual of the space defining the jacobian. Thus the characters defined by the ω chosen give exactly the continuous characters of $I^{\mathrm{unr}}\backslash I^0$, which by the classical theory may be identified with the jacobian. This is a repetition — and not the first — of previous reflections. I should probably observe as well that with the conditions (7.25), the formula for λ_R given by the difference between (7.13a) and (7.13b) reduces when $\gamma = 0$ to

$$\operatorname{Re}\Big\{\sum_j \operatorname{res}_{q_j}(\omega \ln f) - \sum_i \operatorname{ord}_{p_i}(f) \int_p^{p_i} \omega\Big\} \tag{7.13c}$$

Since the differential forms of the first kind give, what may be regarded as a complete set of characters on the quotient $I^{\mathrm{unr}}\backslash I$, all we have to do is assure ourselves that differentials with arbitrary singularities, but otherwise satisfying our conditions, give a complete set of characters on I^{unr}, where, of course, the characters defined by the differential forms of the first kind give 1. We must now employ both ω and ω'. On the other hand, we need no longer concern ourselves with the behavior outside of I^{tr}. If we can match, at least on I^{tr}, a given continuous character χ with one χ_1 given by a differential, then we can complete the matching by identifying $\chi\chi_1^{-1}$ with a character associated to a differential form of the first kind, perhaps multiplied, in addition, by the character associated to one of the supplementary parameters $\gamma \in \mathbb{R}$. We first consider forms of the third kind, or rather their real and imaginary parts, treating the two separately. They define characters of $I^{\mathrm{tr}}\backslash I^{\mathrm{unr}}$.

It is clear from (8.6) that for a form ω of the third kind and an idele in I^{unr} the value of $(\omega, f)_R$ is

$$\prod_x (f_x \bar{f}_x)^{in_x/2} = \prod_x \exp(in_x a_x)$$

where n_x is the residue of ω at x, the only constraint being $\sum_x n_x = 0$, and where

$$f_x = \exp(a_x + ib_x).$$

For the imaginary part we obtain

$$\prod_x \left(\frac{f_x}{\bar{f}_x}\right)^{n_x/2} = \prod_x \exp(i n_x b_x).$$

These together yield a complete set of characters of $I^{\mathrm{tr}}\backslash I^{\mathrm{unr}}$. The differentials of the first kind yield of course the trivial character.

All that is left to show is that the real and the imaginary parts of all differentials yield all characters of I^{tr}, the differentials of the third kind yielding the trivial character. This is clear from formula (7.4)

One point of view, the analytic, has been explained. Although it is not the immediate issue in this prologue, it is important to explain how the geometric theory and the notion of Hecke eigensheaf accommodate the same — or similar — structures. It seems to me that with some of these matters, whether geometric or analytic, one is walking a fine line between the manipulation of definitions and genuine theorems. So there is reason to be uneasy. One goal, here and in the following section, is to offer, at least conjecturally, a precise description of the group $\mathfrak{A}$ in the global geometrical theory. For its abelian quotient this will be, almost inevitably, a reformulation of classical results for abelian integrals, well understood by specialists and, to some extent, familiar to all. We have just rehearsed those necessary for the analytic theory. I found that there was a kaleidoscopic variability in the way these results presented themselves. I hope I have finally arrived at a stable configuration of the constitutive elements. I now describe briefly the geometric theory, but without attempting to include ramification. In the analytic theory, the parametrization by $\mathcal{P}_X = \mathrm{Bun}_1(X)$ is optional; it seems, on the other hand, to be intrinsic to the Assertion.

The Hecke eigensheaves are supported, according to the definitions of [CFT] on $\mathrm{Bun}\, G$, thus in the context of $G = \mathrm{GL}(1)$ on $\mathrm{Bun} = \mathrm{Bun}_1$. This is also a double coset space of $G(\mathbb{A}_F)$, namely

$$\mathcal{B} = G(F)\backslash G(\mathbb{A}_F)/K,$$

where F is the field of algebraic functions on F, $K = \prod_{x\in X} K_x$, where $K_x = G(\mathcal{O}_x)$ for almost all x but for a finite number of places, thus for $x \in S$, K_x lies between $G(\mathcal{O}_x)$ and a congruence subgroup $\{g \in G(F_x) \mid g \equiv I \pmod{z_x^n}\}$, $n \in \mathbb{N}$. Of course, $G(\mathbb{A}_F) = \prod G(F_x)$. In [CFT] — for $\mathrm{Bun}\, G$ itself — the set S is taken to be empty, but this can scarcely be necessary, and it must be possible, with just a little care, to incorporate the congruence conditions into the discussion. They may even simplify matters, because the introduction of a level structure can remove, I suppose, the vexing complications introduced by stacks.

Hecke eigensheaves accommodate many possibilities because they are sheaves, namely perverse sheaves, but for our purposes here, which is just to make the connection between the geometric theory and the analytic theory, we can take these perverse sheaves to be a local systems of dimension 1, thus line bundles with a connection or locally distinguished constant sections. According to my innocent reading of the notion of perverse sheaf, these are the simplest possibilities. A possibility at a higher level would be the flat structure given not by differentials of the first kind, but by differentials with singularities. Whether they have to be singularities with integral residue, so that the sheaves are single valued locally, I am not yet certain. Perverse sheaves with support are outside my range of experience, as is the extension of a local system over the complement of a proper subvariety to a perverse sheaf over the whole variety. For a first explanation of the notion of a Hecke eigensheaf and its relation with geometric automorphic forms — in the more general form envisaged as functions on $F^\times\backslash\mathfrak{I}_F$ — differentials of the first kind are adequate. The rest the reader can discover on his own. We shall incorporate ramification into the discussion only in so far as necessary to make the ideas clear, perhaps not at all. It is important to understand that the complexities introduced by ramification are an essential feature of the theoretical structure even in the geometric theory, but that the notion of a Hecke eigensheaf as such is of interest in itself and that its extension to the ramified context offers only a very modest addition to one's intuitive understanding.

In the context of line bundles, we consider the Picard group $\mathcal{P}$, which is the moduli space for line bundles. Given a line bundle $\mathcal{L}$ on X, thus a point in $\mathcal{P}$, and a point $x \in X$, we can create a bundle $\mathcal{L}_x$ on X by modifying the notion of a section of L in a neighborhood of x. If the local coordinate near x on X is taken to be z, $z(x) = 0$, then the sections of the modified bundle $\mathcal{L}_x$ are the sections of $\mathcal{L}$ divided by z. As a part of the construction of $\mathcal{P}$ as an algebraic variety, which is, of course, a core element of the theory of algebraic curves, the map h from $X \times \mathcal{P}$ to itself given by $x \times \mathcal{L} \to x \times \mathcal{L}_x$ is algebraic or, if one prefers, holomorphic, with a holomorphic inverse. A perverse sheaf $\mathcal{K}$ on $\mathcal{P}$ can be pulled back to $X \times \mathcal{P}$ and then transferred by h to one on the same space. For our purposes at present, this perverse sheaf need be nothing more than a line bundle provided with a local notion of a constant section, thus a local system, but it is best to be aware of the possibilities. I denote the new sheaf by $h_*\mathcal{K}$. The sheaf K is called a Hecke eigensheaf with respect to a local system E on X if

$$h_*\mathcal{K} = E \otimes \mathcal{K}. \tag{7.26}$$

For those who, like me, are not fully at ease with contemporary mathematics, I recall that a local system is also a perverse sheaf. For $n = 1$, the Assertion is that, given E, we can find a $\mathcal{K}$ that satisfies this equation.

The intuition is clear. Translating within $\mathcal{P}$ by the action of $x \in X$, we modify K, but not in a way that can be detected locally, not even locally over X, although it can be detected globally over X. It is difficult, however, not to become entangled in the various strands of the geometry. The connected component $\mathcal{P}^0$ of the Picard variety parametrizes bundles of degree 0 and differs only slightly from the full variety, but it differs in an important way. The homology and cohomology groups of $\mathcal{P}^0$ over $\mathbb{Z}$ and $\mathbb{C}$ are the same as those of X in degrees 0 and 1. So, in the following form the first part of Diagram I applies to both X and $\mathcal{P}^0$,

$$\begin{array}{ccccccc} \{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{C}) & \longrightarrow & H^1(\mathbb{C}^\times) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \{0\} & \longrightarrow & H^1(\mathbb{Z}) & \longrightarrow & H^1(\mathbb{O}) & \longrightarrow & H^1(\mathbb{O}^\times)_{c_1=0}, \end{array} \tag{7.27}$$

where in the lower right-hand corner only those elements with Chern class equal to 0 are allowed, thus line bundles of degree 0.

Consequently, in the case of $\mathcal{P}^0$, we may continue to consider local systems on $\mathcal{P}^0$ as line bundles together with a differential form of the first kind. Local systems on $\mathcal{P}$ — the only kind of perverse sheaf that I want to consider here — are just pieced together from local systems on its various components. Different components are linked by (7.26), which appears in [CFT] as Equation (3.9). Let $\mathcal{P}^n$ be the elements of $\mathcal{P}$ of degree n, $n \in \mathbb{Z}$. The comparison (7.26) effectively compares a sheaf on the connected component $\mathcal{P}^n$ on the right with the same sheaf but over $\mathcal{P}^{n+1}$ on the left, but on both sides there is an extra parameter, one of which, that on the left, is modifying the sheaf, while the other does not. So if we apply the equality twice, once in one sense, once in the other, and take the varying parameters into account, we see that we are imposing a condition on $\mathcal{K}$, a condition that is described by E. All we need to do is ensure that the condition is satisfied as we pass from 0 to 1 and then, back again, from 1 to 0. That takes care of the necessary equality at the level 0, and then (7.26) routinely takes us through the other integers $n = \pm 1, \pm 2, \ldots$.

From the identity of (7.27) for X and $\mathcal{P}^0$, we may identify a line bundle with Chern class 0 on X and with one on $\mathcal{P}^0$ and a flat connection on the first with one on the second. How does this function? We denote the construction in which rather than admitting a pole of order 1 at x, we add a zero, passing from $\mathcal{L}$ to $\mathcal{L}_{-x}$ and introduce the corresponding map from $X \times \mathcal{P}$ to $\mathcal{P}$ by $h' : x \times \mathcal{L} \to x \times \mathcal{L}_{-x}$. Then $y \times x \times \mathcal{L} \to y \times x \times \mathcal{L}_{x-y}$ takes $X \times X \times \mathcal{P}^0$ to $\mathcal{P}^0$ and (7.26) is replaced by an equation for the restriction $\mathcal{K}^0$ of $\mathcal{K}$ to $\mathcal{P}^0$,

$$h'_* h_* \mathcal{K}^0 = E^{-1} \otimes E \otimes \mathcal{K}^0 \tag{7.28}$$

on $X \times X \times \mathcal{P}^0$. The notation $\mathcal{L}_{x-y}$ is simply a more elegant, and perhaps more suggestive, way of writing $(\mathcal{L}_x)_{-y}$.

To prove (7.28) we need to know:

($\star$) The jacobian, thus the group of divisors of degree 0, is identical with the elements of degree 0 in $\mathcal{P}$, this identification being given by mapping the divisor $\delta = \sum_i \pm x_i$, $x_i \in X$ to the line bundle $\mathcal{L}_\delta$ whose sections are functions f with $\operatorname{div} f + \delta \geq 0$.

($\star\star$) The isomorphism between the various cohomology groups appearing in Diagram I on the one hand and (7.27) on the other can be obtained by pull-back from $x \to \mathcal{L}_{x-y}$ with a fixed y and a fixed $\mathcal{L}$.

So (7.28) is simply the assertion that E is the pull-back of $\mathcal{K}$. It seems to be much ado about nothing, but that would be, I suspect, a view that failed to appreciate the marvels of the theory created by Abel, Jacobi and others.

This discussion suggests that, at least for GL(1), one neither wins nor loses by working with the arithmetic/analytic structures rather than the geometric, but it does not suggest to me a direct equivalence. The space $\mathrm{Bun}_1(X) = \mathcal{P}$ is implicated in an essential way in the statement of the (geometric) Assertion. In the analytic theory $\mathrm{Bun}_1(X)$, or rather its connected component, appears as an optional enlargement of the group of characters. There is one respect in which the analytic theory appears to offer an advantage: the description of the group $\mathfrak{A}$. This description, which shall be formulated and verified for GL(1) in this section, and for general quasisplit G in the following section, but only as a conjecture that will not be entirely precise, has to serve as my apology for an irritatingly lengthy rehearsal of familiar classical material and the modern geometrical viewpoint.

For the local theory, an analytic theory, the group to be parametrized is formed by the characters of $F_x^\times$. Apart from the ambiguities in the extension of diagram (7.5), the parametrization is given by differentials. So, to be as precise as possible, because we are (almost!) dealing with definitions rather than theorems, just as the characters of GL(1, F_x) are identified with homomorphisms of the Weil group into GL(1) in the local arithmetic, so characters of GL(1, F_x) (or, at first, GL(1, $\mathcal{O}_x$)) are associated with differentials ω with values in the Lie algebra of GL(1) over F_x, or rather with their principal parts. These form a group and should be regarded as the abelian form $\mathfrak{A}_x$ of the local Weil group in the geometric context, with multiplication given by addition of differentials, except that the extension of Ω_x to $\tilde{\Omega}_x$ of diagram (7.5) is needed to complete the construction.

Globally, we have introduced a similar relation between differentials and characters, except that there is no longer a question of discarding the regular parts of the differentials. Moreover, the characters are not characters of idele classes I_F but of an enlarged group $F^\times\backslash\mathfrak{I}_F$. Multiplication of characters becomes addition of differentials. It is this group, or rather an extension of it by the group of characters of the group $\mathcal{P}^0\backslash\mathcal{P} \simeq \mathbb{Z}$, that functions as the abelianized form $\mathfrak{A}_{\mathrm{ab}}$ of the group

$\mathfrak{A}_F$. So it is an analogue of the abelianized Weil group with multiplication given by addition of differentials,

$$\frac{df}{f} = \omega_1 + \omega_2. \tag{7.29}$$

I add that class field theory has accustomed us to identify, in the arithmetic theory, the abelianized form of the Weil group with I_F and the Weil group itself with a subgroup of the Galois group. There is a merging of definition and theorems that, if we are not careful, obscures for us the accomplishments of the past.

Before turning to the theory for a general group, I remark that I may have found partial answers to two questions while struggling not with proofs, but just with the formulation of conjectures and assertions in the geometric theory: (i) what are the respective merits of the geometric and analytic standpoint? (ii) what is the interest of the geometric theory in itself, thus what are the principal theorems or conjectures, independently of any relation to quantum field theory? The response to the second question is best left to §9. The response to the first question is tentative, especially as there are a number of clumsy aspects to the analytic theory for a general group and even for GL(1). The difficulty with the geometric theory is that there are so many possibilities that they are never exhausted. In the theory of Fourier transforms there are many possibilities: the spectral theory for square-integrable functions; Paley–Wiener theorems; theorems related to Schwartz distributions of various sorts; the Laplace transform. I am inclined to take the spectral theory as central. For the geometric theory, there is a similar difficulty. What is the core problem? My hope for a spectral theory is that one could formulate a clearly defined spectral problem, thus an L^2-problem — differential operators with boundary conditions — whose solutions on Bun_G could be regarded in at least some respects as a definitive formulation of the existence problem for Hecke eigensheaves: an eigensheaf (or eigenfunction) $\mathcal{K} = \mathrm{Aut}_E$ on Bun_G with eigenvalue a ${}^LG$ local system E on X is a pair characterized by a certain set of conditions on E and by the relation between E and $\mathcal{K}$.

The eigenvalue — in a sense like that of the geometric theory — is $\exp(i\,\mathrm{Re}\,\omega)$. It is $\mathrm{Re}\,\omega$ (or $\mathrm{Im}\,\omega$) that is characterized by a differential equation, as the real part of an analytic function it is harmonic outside of the singularities and with circumscribed behavior at the singularities, for the residue is integral. Notice, in passing, that we can recover $\mathrm{Im}\,\omega$ or ω — up to an unimportant constant from $\mathrm{Re}\,\omega$ — and the Cauchy–Riemann equations. The function $\mathrm{Re}\,\omega$ is moreover implicitly subject to a boundary condition. We have made the boundary condition more flexible, even removed it, by introducing $\mathfrak{Z}_R$ (or $\mathfrak{Z}_I$), but that was necessary only to keep up with the geometers. The boundary condition is a condition not on $\mathrm{Re}\,\omega$ as a function on X but on the function (sheaf for the geometers) associated to it on $\mathcal{K}$. Boundary conditions on ω itself would double their number and yield an overdetermined eigenvalue problem.

When we allow singularities, $\mathcal{H}$ is replaced by a quotient $\mathbb{I}_F/\prod_x K_x$, where $K_x = \mathbb{O}_x^\times$ for almost all x, but for a finite number of x it is the set of $f_x \in \mathbb{O}^\times$ that are congruent to 1 modulo some positive power $z_x^{n_x}$ of z_x. It may not be immediately apparent, but the definitions (7.13a), (7.13b), and the formulas for $(\omega, \cdot)_R$, $(\omega, \cdot)_I$ are an expression of the condition (7.26), although that condition refers more to the characters defined by $\omega = \operatorname{Re}\omega + i \operatorname{Im}\omega$. In the geometric theory it is only the local conditions on ω that matter, not the boundary conditions or conditions of periodicity. As a consequence, or so it seems to me, there is for a general group some difficulty in formulating the problem of existence.

It is difficult to recognize the equation (7.26) in the conditions (7.13a) and (7.13b) for at least three reasons: (i) the condition (7.13b) is a matter of conditions of periodicity; (ii) the final term of (7.13a) is constant on connected components of Bun_1, so that in the geometric or sheaf-theoretic context it has no meaning; (iii) the first term is not present when ω has no singularities. Although the geometers are well aware of the possibility of singularities [FG], they are not studied in [CLG]. Even when singularities are present, (7.26) is likely to remain, in that form, the telling geometric condition.

One of the purposes of the next section is to begin the search, for a general G, for a construction of automorphic representations analogous to that given for GL(1) by $(\omega, \cdot)_R$ or $(\omega, \cdot)_I$. Since these are unitary characters, we have to expect unitary characters (representations) for a general group as well. The middle term of expression 7.12.a controls the unramified contribution. The first term controls the character on each $\mathbb{O}_x^\times$. The condition that the residue n_x of ω at each point x be integral implies that $(\omega, \cdot)_R$ and $(\omega, \cdot)_I$ yield respectively r^{in} or e^{imn}, where $n = \operatorname{res}_{q_j} \omega$, $f(q_j) = re^{im}$.

8.a. The geometric theory for a general group (provisional)

Such a theory is not yet available even in embryonic form, although some reflections are suggested by the previous constructions. As I observed in the previous section, these constructions are perhaps not merely my interpretation of those explained in [CFT], but are implicit in the proof of Weil's theorem. The relation between Lemma 7.1 and the calculation that yields it differ on the face of things from formula (7.28) and its proof. Their interpretations are also informed by a different impulse: sheaves are replaced by differential equations. For regular holonomic systems, there is presumably an equivalence available [HTT]. After the admission of differentials with more general singularities, this may no longer be so, although that is unlikely. My impulse arises, however, from other sources: from a greater familiarity and perhaps even greater ease with differential equations than with sheaves, or perhaps from a suspicion that, important, powerful, and fascinating as they are, in recent decades an excessive, sometimes inappropriate, appeal to sheaves has, inadvertently, had an unhealthy influence on some parts of mathematics or, rather, on some of its practitioners; and from an attachment, already expressed, to representation theory, as introduced, in a remarkable, but little read, sequence of papers by Frobenius, in response to a suggestion of Dedekind and then developed by several major mathematicians of the twentieth century.

For a general group G, even if it is split, as I suppose in this text, there is not only no global geometric theory yet available, there is also no local theory. Moreover, there is an extra question. What is the relation between, on the one hand, the functoriality of the geometric theory, the identification of the group $\mathfrak{A} = \mathfrak{A}_{\text{geom}}$, and a description of its properties and, on the other, the Langlands duality featured in gauge theory? Are they one and the same, or are they different? That they are different, occurred to me on reading a brief, but instructive and suggestive letter, that I received from David Nadler in March of 2011. Nadler writes: "The 6-dimensional theory Z depends not on a group G but only on the combinatorics of G in a way that is unbiased towards G and its dual group $G^\vee$." This is not so for the theory to whose preliminary exploration this section is devoted: G and ${}^LG$ (or $G^\vee$) do not play symmetric roles! Moreover, there is no 6-dimensional field theory in sight. So there is a great deal left for me, and perhaps not for me alone, to understand. It will be best not to broach this question until Section 9, yet to be written. It requires a good deal of supplementary reflection, informed by some knowledge of field theory.

Indeed, even my attempt to broach the purely mathematical questions turned out to be premature. One of the principal mathematical problems of the geometric theory, perhaps *the* principal one, is the identification of the geometric galoisian group $\mathfrak{A}_{\text{geom}}$ in terms of differentials, thus the general form of the identification of its abelian quotient in the previous section. This is by no means a simple matter, for it demands a serious understanding, not merely a formal understanding, of moduli spaces for vector bundles and G-bundles, of the differential geometry of these bundles as in [Si], and of the relation between $\mathfrak{D}$-modules and perverse sheaves. These are all very rich subjects, of which I could not hope to acquire an adequate understanding before the deadline imposed by the editor of this volume, if ever. So I was forced to content myself with some provisional suggestions just to intimate to the reader what I have in mind. As will be almost immediately evident, there are major unresolved difficulties left open.

The identification of $\mathfrak{A}_{\text{geom}}$ entails functoriality for the geometric theory. If there is some form of reciprocity — different as our title implies from functoriality — in the geometric theory, I do not know how to formulate it. The geometric Langlands program as envisioned in [KW] contains, I suspect, much, much more than the identification of the geometric galoisian group $\mathfrak{A}_{\mathrm{geom}}$ in terms of differentials envisaged in this section. It may, indeed, have little relation to it. It does contain a kind of *duality*, but it may be best to distinguish this duality from the *reciprocity* in the arithmetic theory and from any concrete identification of $\mathfrak{A}_{\mathrm{geom}}$, although it is clearly related to this. It had been my intention to begin, in a ninth section, the attempt to understand [KW] and, more generally, the many and various contributions to the geometric theory and its relation to quantum field theory, but that, as I have already confessed, is matter for an even more distant and more uncertain future.

In Section 7, the emphasis was on functions on $\mathrm{Bun}_G = \mathcal{P}_X$, $G = \mathrm{GL}(1)$; sheaves were not emphasized. There is, indeed, a major difference. The forms ω were parametrized by a local system and then by a second parameter in $H^1(\mathbb{Z})$. Two elements of $H^1(\mathbb{R})$ in the first line of Diagram I that differ by an element of $H^1(\mathbb{Z})$, define isomorphic local systems or, viewed from another angle, an automorphism of a given line bundle replacing one local system by another, thus, in terms of one of the local systems multiplication of the flat connection by the character $\exp(i \operatorname{Re} \omega)$, where the exponent is constrained to be the imaginary part of a holomorphic function. So there is a mixing of a real (unitary) theory and a complex (holomorphic) theory. This brings with it advantages but also difficulties. One of the difficulties for me is that — as is clear from [Si] and the works there cited — the mixing for a general group demands very serious differential-geometric preparation, not merely the Cauchy–Riemann equations. One advantage, already explained, is that, in the analytic theory, we can hope to formulate the problems in the context of a spectral theory in an L^2-space.

I expected, on first reflecting on the matter, that, as for GL(1), the group $\mathfrak{A}_{\mathrm{geom}}$ will be given by a kind of inverse limit of differentials with values in ${}^L\mathfrak{g}$, the inverse limit being taken over $\omega \to \omega'$, where ω' is the image of ω under a homomorphism ${}^LG \to {}^LG'$ in the sense of L-groups. So the inverse limit is over the group, the direct limit over the differentials. In order to deal with all automorphic representations, we would have to admit, as for GL(1), differentials with singularities. Whether there should be restrictions on the residues similar to those we described for GL(1) can be left moot. Even for those without singularities, there is a great deal of theory to understand.

As an aside, I mention that, following [CFT], I shall take the group action on G-bundles and on ${}^LG$-bundles to be on the right. So, once we have fixed a local trivialization, the differentials generate along curves a function with values in ${}^LG$ according to the differential equation $dg \cdot g^{-1} = \omega$.

My first expectations were perhaps, in the light of our understanding of the geometric theory, too naive, too influenced by the construction of the geometric $\mathfrak{A}_{\mathrm{ab}}$. If there is homomorphism ${}^LH \to {}^LG$, then the differential with values in $\hat{\mathfrak{h}}$ transfers to $\hat{\mathfrak{g}}$, so that if a parametrization of automorphic representations or forms in the geometric theory is established, functoriality will be an immediate consequence. In the arithmetic theory the definition of the galoisian group $\mathfrak{A}_{\mathrm{arith}}$ is based on functoriality and on the notion of a hadronic representation, itself based on the properties of $L(s, \pi, \rho)$, $\rho : {}^LG \to \mathrm{GL}(n)$. This assumes, in particular, that in establishing functoriality we have also completely understood the nature of the Ramanujan conjecture and the Arthur parameters. Although I have alluded to these in the arithmetic context, I have not attempted any, even conjectural, definition in a geometric context. This would be reckless without more experience with the classification of bundles for specific groups on specific curves, with the parameters, and with the corresponding automorphic representations. We need more concrete assurance that the parametrization proposed here is correct and some insight into its specific consequences. There may be surprises. This is one of the many reasons that this section is provisional. In one way or another, the parameter obtained from ω on transfer to $\mathrm{GL}(n)$ under an irreducible ρ will be a direct sum of irreducible parameters, ω_i, for $\mathrm{GL}(n_i)$, $\sum_i n_i = n$,

$$\omega \to \bigoplus_i \omega_i. \tag{8.1}$$

The initial parameter ω would be hadronic if there were no i for which $n_i = 1$ and ω_i is trivial.

There are also many other many other matters to consider. We have somehow to reconcile the unitary and holomorphic (or meromorphic) forms of this equation. I am not yet in a position to do so and am uneasy about suggesting definitions that I do not understand. For the moment, the definition of $\mathfrak{A}_{\mathrm{geom}}$ remains, at best, imprecise. To make it clear that anyone who, like me, has little or no differential geometrical experience has much to learn, I quote one of the first paragraphs in [Si], which treats $\mathrm{GL}(n)$, which can for us be regarded as typical. Although our concern is with complete nonsingular curves, the statement in [Si] refers more generally to smooth, projective X. “A *harmonic bundle* on X is a C^∞ vector bundle E with differential operators δ and $\bar{\delta}$ and algebraic operators θ and $\bar{\theta}$ (operators from E to one-forms with coefficients in E), such that the following hold. There exists a metric K so that $\partial + \bar{\partial}$ is a unitary connection and $\theta + \bar{\theta}$ is self-adjoint. And if we set $D = \partial + \bar{\partial} + \theta + \bar{\theta}$ and $D'' = \bar{\delta} + \theta$, then $D^2 = 0$ and $(D'')^2 = 0$. With these conditions, (E, D) is a vector bundle with flat connection, and $(E, \bar{\partial})$ is a *Higgs bundle*: a holomorphic vector bundle with holomorphic section θ such that $\theta \wedge \theta = 0$.” Of course, for a surface some of the assertions are superfluous. This statement is followed by a theorem whose first sentence I repeat. “There is a natural equivalence between the categories of harmonic bundles on X and semisimple flat bundles (or representations of $\pi_1(X)$).” I observe that this statement does not take into account a possibility that we encountered for $\mathrm{GL}(1)$: automorphisms of the unitary bundles and associated automorphic forms.

There are so many questions in the geometric theory, both local and global, that have never been touched, that I am more than a little uncertain of the similarities and differences between it and the arithmetic theory. Perhaps we should begin by stating clearly the difference between the objects in the analytic theory and the objects in the geometric theory. They are all constructed from the curve X, a set of points, but also a Riemann surface and the "set" $\mathrm{Bun}_G(X)$ of G-bundles on X, with or without the extra structure that allows the introduction of ramified automorphic forms. The set Bun_G may be identified with $G(F)\backslash G(\mathbb{A}_F)/\prod_{x\in X} G(\mathbb{O}_x)$ (see [CLG]) or, if there is extra structure, with $G(F)\backslash G(\mathbb{A}_F)/\prod_{x\in X} K_x$, where K_x is equal to $G(\mathbb{O}_x)$ for almost all x, say $x \notin S$, S finite, and, for example, equal to $G_x^{n_x}$ for $x \in S$, where n_x is a nonnegative integer and $G_x^{n_x}$ the set of elements in $G(\mathbb{O}_x)$ congruent to 1 modulo $\mathfrak{m}_x^{n_x}$, where $\mathfrak{m}_x$ is the maximal ideal in $\mathbb{O}_x$, but it may also be constructed geometrically.

Although these two descriptions yield the same two sets, they yield functionally dissimilar objects. As identified with $G(F)\backslash G(\mathbb{A}_F)/\prod_{x\in X} G(\mathbb{O}_x)$, $\mathrm{Bun}_G(X)$ is a set in the familiar sense; constructed as a *stack* it has, essentially, the structure of an algebraic variety, which can be recovered from that on the quotients $G(F_x)/G(\mathbb{O}_x)$, a matter to which we shall return but not to examine it in depth. In the one context, the appropriate objects are functions; in the others, sheaves. As we have seen, we can expect — or hope — that for functions more precision is possible because more structure is possible.

It is not clear where it is best to begin, perhaps by reviewing the very little we know about the local theory, or, if one prefers, what we clearly do not know. Like the theory over a number field or one of its completions, the analytic theory over the function field and over its completions F_x is a theory about representations, usually infinite-dimensional. The group $G(F_x)$ has a sequence of decreasing subgroups: $G(F_x) \supset G_x^{\mathrm{unr}} \supset G_x^{\mathrm{tr}}$, where $G_x^{\mathrm{unr}} = G(\mathbb{O}_x)$, $G_x^{\mathrm{tr}} = G_x^1$ is the set of g in G_x^{unr} whose power series expansion begins with the identity. These subgroups are of course not normal, but we can consider, as a first, coarse classification locally, irreducible representations whose restriction to one or the other of the subgroups contains this or that irreducible representation. The most important are those whose restriction to G_x^{unr} contains the trivial representations. Their theory is the theory of spherical functions and characters of the Hecke algebra. We can admit all characters; we can admit "tempered" characters; we can admit those characters that correspond over a local arithmetic field to Arthur parameters, although Arthur parameters in the geometric context are certainly not a topic to broach in this prologue. The theory of spherical functions, a generalized form of the theory of elementary divisors, will be, in many respects, the same for the geometric theory as for the arithmetic theory. So unramified characters will be parametrized by a conjugacy class t in ${}^LG(\mathbb{C})$, or, even, if we assume, as we have done, that G is split, by a semisimple class in $\hat{G}(\mathbb{C})$. We can take this class to be arbitrary and this corresponds to a geometric theory or we can take it to lie in a compact form $\hat{U}(\mathbb{R})$ of $\hat{G}(\mathbb{C})$ and this corresponds to the tempered analytic theory, the adjectives "geometric" and "analytic" having for the moment only the imprecise meaning suggested by various remarks in the previous section.

In contrast to the arithmetic theory, the representations of the group $G_x^{\text{unr}}/G_x^{\text{tr}}$ may be infinite-dimensional. There are two possibilities. We can consider either representations of $G(\mathbb{C})$ or the representations of its compact form $U(\mathbb{R})$. As for GL(1), when dealing with a higher-dimensional G we have to choose between a holomorphic theory and an analytic theory. I choose, because of preferences already acknowledged and for reasons already explained, the analytic theory. It is not entirely clear what this implies even locally. The abelian theory suggests that it is the representations of $G(\mathbb{C})$ that we need. Although this is a noncompact reductive group and the distinctions of §2 — the class of tempered representations, the Arthur class, general representations whether unitary or not — may be pertinent, it may be the finite-dimensional representations, these suggested because the trivial representation certainly appears in the unramified theory, to which we should pay the most attention. They can be holomorphic or antiholomorphic or some mixture of the two. The usual (Langlands) classification parametrizes the tempered representations by conjugacy classes of homomorphisms of $W_{\mathbb{C}} = \mathbb{C}^\times$ into ${}^LG$, thus by $z = re^{i\theta} \to r^\lambda e^{i\mu(\theta)}$ where λ is a real linear combination of weights of G and μ a weight of G, the pair (λ, μ) being given up to conjugation up to the action of the Weyl group. The holomorphic finite-dimensional representations correspond to unitary representations of the unitary form of G and these correspond to homomorphisms of $U(1)$ into ${}^LG$, thus to $\lambda = 0$. The similarity of the parametrizations of the characters of the Hecke algebra (related to unramified representations) and of certain representations of $G_x^{\text{unr}}/G_x^{\text{tr}}$ (related to tamely ramified representations) is curious and gives pause for reflection.

It is suggested by the theory for GL(1) — and confirmed by various reflections, although by no means certain — that a central role will be played by differentials with values in the Lie algebra $\hat{\mathfrak{g}} = {}^L\mathfrak{g}$ and their real parts, taken in an appropriate sense, which I hesitate to attempt to make precise without a better understanding of the differential geometric theory [Si]. They will define the local system E of the Assertion of §7 or of the related Conjecture that follows in this section. What are the restrictions on these differentials and what do we mean by their real parts? The first question to be answered is what the nature of their residues must be, for — as I suppose — the residue at a point controls, when their is no higher order singularity, the representation of $G_x^{\text{unr}}/G_x^{\text{tr}}$. So far as I know, the representations of G_x^{unr}/G_x^n, $n > 1$, have been little studied, nor, of course, have those of $G(F_x)/G_x^n$, $n \geq 1$. The discussion in this section is predicated on the hypothesis that they are controlled by the singular part of a Laurent expansion of a differential ω with values in $\hat{\mathfrak{g}}$. The residue, thus the coefficient of $1/z$, will be an element of $\hat{\mathfrak{g}}$. The demand that it be integral is compatible with our discussion of the tamely ramified parameter μ in the preceding paragraph.

The group G_x^{tr} is an infinite-dimensional Lie group. We shall only be concerned with representations of the finite-dimensional quotients G_x^{tr}/G_x^n. It is a finite-dimensional simply connected nilpotent Lie group and its irreducible unitary representations are classified by the method of coadjoint orbits [VE], thus by conjugacy classes in the dual of the Lie algebra over $\mathbb{C}$ of G_x^1/G_x^n. Thanks to the sequence $G_x^1 \supset G_x^2 \supset \cdots$ these coadjoint orbits form an increasing sequence of sets. This can be interpreted to state that they are parametrized by the singular parts of local differentials ω at x with residue 0 and values in $\hat{\mathfrak{g}}$.

These facts together suggest, but hardly prove, that the local parametrization for a general group is very much like that of the diagram (7.5) for GL(1), although I do not yet know how to define in general the patching of conjugacy classes of GL($\mathbb{C}$) that appears in its upper line. This suggestion will be taken as an hypothesis for the remainder of this section. I have made no attempt to prove it. The local theory is only a part of the unresolved difficulties, and this prologue, even the essay *Functoriality and Reciprocity* that I hope will follow it, is intended to be no more than a first exploration of possibilities.

As with the arithmetic theory, the major issues in the geometric theory will be global. They may not be so difficult as for the arithmetic theory, but the theory of vector bundles or of G-bundles on curves over $\mathbb{C}$ is very rich and for me largely unfamiliar, so that I could very easily overstep the limits of my knowledge, which are severe. It would certainly be presumptuous for me to say too much at this stage, but I do want to sketch the possibilities. Although the problem of describing the global geometrical galoisian $\mathfrak{A}_{\mathrm{geom}}$ may be more accessible than that of describing $\mathfrak{A}_{\mathrm{arith}}$, we can expect it to be difficult and to require a good deal of experience and technical skill. It is my hope that the arguments for GL(1), especially the proof of Lemma 7.1 in which the calculus of residues is applied, will serve as a model.

The principal issue is to understand the unramified theory or, better, the theory at the unramified places. For the unramified theory, the basic object is

$$\begin{gathered} G(F)\Big\backslash \prod_{x\in X} G(F_x)\Big/\prod_x G(\mathbb{O}_x) = \varinjlim_T G_T(F)\Big\backslash \prod_{x\in T} G(F_x) \prod_{x\notin T} G(\mathbb{O}_x)\Big/\prod_x G(\mathbb{O}_x), \\ G_T(F) = G(F)\cap \Big\{\prod_{x\in T} G(F_x) \prod_{x\notin T} G(\mathbb{O}_x)\Big\}, \end{gathered} \tag{8.2a}$$

where T can be taken as large as appropriate. If there is ramification, the first line will be replaced by

$$\text{(8.2b)}\quad G(F)\Big\backslash \prod_{x\in X} G(F_x)\Big/\prod_x K_x = \varinjlim_T G_T(F)\Big\backslash \prod_{x\in T} G(F_x)\prod_{x\notin T} G(\mathbb{O}_x)\Big/\prod_x K_x,$$

where K_x is open in $G(\mathbb{O}_x)$ for all x and $K_x = G(\mathbb{O}_x)$ for $x \notin S \subset T$.

The set (8.2a) is Bun_G and the set (8.2b) is Bun_G with frills. There are explanations to be given because Bun_G is, whether as a variety, as an injective limit of varieties, or as a stack, an algebro-geometric object. For us, however, who want to make the connection with the geometric theory of automorphic forms, it is also simply a set. Whether as a variety or as a set it is a quotient. We begin in the unramified context with the trivial G-bundle, which we modify by an element of $\prod_{x\in T} g_x$, $g_x \in G(F)$. If we take $g_x = 1$, $x \in T' - T$, $T' \supset T$, this can be regarded as an element of $\prod_{x\in T'} G(F_x)/\prod_{x\in T'} G(\mathbb{O}_x)$, so that we have an injective family of sets. Since each g_x is defined in a neigborhood of x and regular in this neighborhood except at x and since we can enlarge T to T', we can suppose these neighborhoods cover X, so that the collection $\{g_x\}$ defines a G-bundle. We can even suppose that $g_x \in G(F)$ because the set $g_x G(\mathbb{O}_x) \cap G(F)$ will not be empty, and we can replace g_x by an element of this set. The choice does not affect the bundle. The conclusion is that any choice of $g \in \prod G(F_x)$ defines a G-bundle on X and that all G-bundles are obtained in this way.

Thus Bun_G is constructed as a limit of the quotient of

$$\prod_{x\in T} G(F_x)/G(\mathbb{O}_x)$$

by $G_T(F)$. Each point of $G(F_x)/G(\mathbb{O}_x)$ represents a modification of Bun_G at the point x; it is an extremely complicated variety, the direct limit of finite-dimensional subvarieties. Thus, starting from single point of $\mathrm{Bun}\, G$, the trivial bundle, and repeatedly modifying the bundle, each modification at perhaps a different point, we can reach any bundle. So there are two sources of complexity in the construction of Bun_G. They are the modifications and the divisions by $G_T(F)$. It is well to give some examples for vector bundles, thus for the groups $\mathrm{SL}(n)$ and $\mathrm{GL}(n)$, to see how the parameters in $G(F_x)/G(\mathbb{O}_x)$ and the parameter x together lead to very complex modifications that may, because of the division by $G_T(F)$, yield a bundle isomorphic to that with which we began. The topological or geometrical structure is a combination of the double parametrization: by the parameter of $x \in X$ and by the coordinates on $G(F_x)/G(\mathbb{O}_x)$. For $\mathrm{GL}(1)$ this double parametrization is simple because $G(F_x)/G(\mathbb{O}_x) \simeq \mathbb{Z}$. Thus, as in the second term on the left of (7.15), the only relevant parameters are p_i, basically a point in a neigborhood of p, and the integer $\mathrm{ord}_{p_i}(f)$. For groups of higher dimension, the parameters are far more complex.

Consider $G = \mathrm{GL}(n)$ and, first of all, the structure of $\mathrm{Bun}_x = G(F_x)/G(\mathbb{O}_x)$ as a space or variety, whose dimension is infinite, on which $G(\mathbb{O}_x)$ acts to the left. If z is the local coordinate at x, the representatives of the double cosets in $G(\mathbb{O}_x)\backslash G(F_x)/G(\mathbb{O}_x)$ are the matrices

$$t(m_1,\dots,m_n)=\begin{pmatrix} z^{m_1} & 0 & 0 & \dots & 0\\ 0 & z^{m_2} & 0 & \dots & 0\\ 0 & 0 & z^{m_3} & \dots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & z^{m_n}\end{pmatrix},$$

where the m_i are integers, implicitly subject to the condition $m_1 \ge m_2 \ge \cdots \ge m_n$. The variety Bun_x is a union of the connected components $\mathrm{Bun}_{x,m}$ defined by the condition $\sum_i m_i = m$. Multiplying by a scalar matrix, we replace $\mathrm{Bun}_{x,m}$ by $\mathrm{Bun}_{x,m+kn}$, $k\in\mathbb{Z}$. We shall consider some examples, taking small values of m and n.

For $m=0$ and all n, there is a distinguished point,

$$q_0 = G(\mathbb{O}_x)t(0,0,\dots,0)G(\mathbb{O}_x)/G(\mathbb{O}_x).$$

If $n=1$, it is the only point in $\mathrm{Bun}_{x,0}$. In general,

(8.3)

$$\mathrm{Bun}_{x,m_1,\dots,m_n} = G(\mathbb{O}_x)t(m_1,\dots,m_n)G(\mathbb{O}_x)/G(\mathbb{O}_x) \simeq G(\mathbb{O}_x)/P(m_1,m_2,\dots,m_n),$$

where

$$\begin{aligned} P(m_1,m_2,\dots,m_n) &= G(\mathbb{O}_x)\cap (tG(\mathbb{O}_x)t^{-1}\cap G(\mathbb{O}_x))\\ &= \{(a_{i,j}) \ni a_{i,j}\equiv 0 \pmod{z^{m_i-m_j}}\},\end{aligned}$$

with $t=t(m_1,\dots,m_n)$.

As in [CFT], we can try to grasp the full space $G(F_x)/G(\mathbb{O}_x)$ by writing the elements of $G(F_x)$ as products nak, where n is unipotent and upper-triangular, thus $n\in N(F_x)$, a is a diagonal matrix $T=t(m_1,\dots,m_n)=\mathrm{diag}(t^{m_1},\dots,t^{m_n})$, and $k\in G(\mathbb{O}_x)$. The connected components are then given as algebraic varieties by

$$N(F_x)/(N(F_x)\cap TG(\mathbb{O}_x)T^{-1})\times G(\mathbb{O}_x), \tag{8.4}$$

which is closed in the full variety. The structure of the first factor has to be explained, but it is intuitively clear. For example, if $n=2$ and $m=m_1-m_2$, then a full set of representatives for the quotient in (8.4) is given by

$$n(p)=\begin{pmatrix}1 & p(t)\\ 0 & 1\end{pmatrix}, \tag{8.5}$$

where p is a finite Laurent series with an indefinite number of nonzero terms of negative degree, $p(t)=\sum_{k<m} a_k t^k$. If $p(t)$ is identically 0, then $n(p)T$ lies in the

double coset with parameter (m_1', m_2'), where $\{m_1', m_2'\} = \{m_1, m_2\}$ and $m_1' \geq m_2'$. Otherwise, let l be the least k for which $a_k \neq 0$. Then

$$n(p)T = \begin{pmatrix} t^{m_1} & t^{m_2+l}\alpha(t) \\ 0 & t^{m_2} \end{pmatrix}, \tag{8.6}$$

where $\alpha(t)$ is a polynomial with nonzero constant term. If $m_2 + l \geq \min\{m_1, m_2\}$, this lies in the same double coset as when $p(t) = 0$, otherwise it lies in the double coset with parameter $\{m_1', m_2'\} = \{m_1 - l, m_2 + l\}$ and $m_1 - l \geq m_2 + l$. Since we can choose l to lie as far to the left as we like and then let all the coefficients of $\alpha(t)$ approach 0, we conclude that one coset can lie in the closure of many others.

On the other hand, some of the double cosets (8.3) are closed. If $m_1 \geq m_2$, the relation between the parameters at the end of the preceding paragraph is $m_1 - l \geq m_1 \geq m_2 \geq m_2 + l$. If $(m_1 - l) - (m_2 + l) = 1$, this is out of the question. If $m_2 \geq m_1$, the relation is $m_1 - l \geq m_2 \geq m_1 \geq m_2 + l$ and there is the same difficulty. So the set $\mathrm{Bun}_{x,1,0}$ is closed.

There are clear relations of containment between the various groups $P(m_1, m_2, \ldots, m_n)$, that yield mappings between the various sets $\mathrm{Bun}_{x,m_1,\ldots,m_n}$ or, more generally, between the analogous varieties for a general G. They are usually referred to as a blowing-up or a blowing-down, or as Hecke correspondences, or as modifications. It is certainly appealing and useful to keep the geometric language and the geometric context in mind, but we shall not always do so. A partial order on the weights of the usual kind and, for example, arguments along the lines of the discussion of the previous paragraph provide a partial order by inclusion on the set of closures $\overline{\mathrm{Bun}}_{x,\tau}$ of the varieties $\mathrm{Bun}_{x,\tau} = G(\mathbb{O}_x)\tau/\,\mathrm{Gal}\, O_x)$ and these closures are complete. The element τ is a matrix $t(m_1, \ldots, m_n)$ for $\mathrm{GL}(n)$ and an element in, say, a split torus for a general (split) G. The varieties themselves are open in their closure. It is possible — it is so already for $\mathrm{GL}(1)$ — that we cannot find a cofinal set of $\overline{\mathrm{Bun}}_{x,\tau}$, but we can find a cofinal family of finite unions $\bigcup_i \overline{\mathrm{Bun}}_{x,\tau_i}$. So we can introduce the union of these varieties to obtain a variety $\mathrm{Bun}_x = G(F_x)/G(\mathbb{O}_x)$, infinite-dimensional but the union of closed, finite-dimensional subvarieties.

This is very likely all familiar. We use it to construct the global Bun_X. Let T be a finite set in X and $\{\tau_x, x \in T\}$ a collection of τ. Consider

$$\prod_{x\in T} \overline{\mathrm{Bun}}_{x,\tau_x} \subset \prod \overline{\mathrm{Bun}}_{x,\tau_x}, \tag{8.7}$$

where for the imbedding of the left side in the right, it is understood that $\tau_x = 1$, $x \notin T$. In principle, we can fix T and take a union to arrive at

$$\prod_{x\in T} G(F_x)/G(\mathbb{O}_x), \tag{8.8a}$$

which we can divide by

$$G_T(F) = G(F) \cap \left\{ \prod_{x \in T} G(F_x) \prod_{x \notin T} G(\mathbb{O}_x) \right\}, \tag{8.8b}$$

but this will lead to a discrete object. There is a second, more important limit implicit in (8.8a). We can first fix the number of elements in $T = \{x_1, \ldots, x_n\} \in X \times \ldots X = X^{(n)}$, so that we introduce n supplementary parameters, as in the theory of the Picard variety, although repetitions are not necessary. They are already at hand in (8.8a). It is presumably better to take the limits in the order opposite to that suggested in (8.2a) with a finite number of double cosets at first but with all possible T. Then, as for GL(1) and the Picard variety, we may reach the limit before exhausting the possibilities offered by (8.8a). This construction poses problems of various kinds, with stability, stacks and with other matters. I am in no position to deal with them at the moment and prefer to pass on to another issue, the central question of this section. So I simply take them as solved or, at least, solvable. In essence, however, we arrive at the algebro-geometric form of the set appearing as a limit in the modified form of (8.2a),

$$\lim_n \bigcup_{|T|=n} G_T(F) \backslash \prod_{x \in T} G(F_x)/G(\mathbb{O}_x). \tag{8.8c}$$

It appears that, in spite of its formidable appearance, the algebro-geometric result is finite-dimensional. Indeed, for some curves X, it is, I find, strangely simple [At; Le].

The central question for us here is whether the method used in the previous section to construct a character from the differential form ω can function for nonabelian groups. There are three issues raised by formula (7.13a) and (7.13b): the periods that appear in (7.13b) and whose existence was accommodated by the introduction of $\mathfrak{z}_R$ and $\mathfrak{z}_I$; the contributions of the singularities of ω; the contributions of the singularities of f. It is clear that ω in both its holomorphic (in general, meromorphic) and unitary form will yield a homomorphism of the fundamental group into $G(\mathbb{C})$ or into its unitary form, thus a nonabelian form of the periods. Any study of this will have to wait until I better understand the issues arising from a study of [Si].

We do not yet understand the local ramified theory. So we have to exclude, at least provisionally, all ramification. One possibility is to assume that ω itself has no singularities. Another possibility to keep in mind is that we can fix a finite set $S \subset X$, which T is always supposed to contain, and, for $x \in S$, fix $g_x \in G(F_x)$ and a nonnegative integer n_x and work not on (8.8a) but on

$$\prod_{x \in S} g_x G^{n_x}(\mathbb{O}_x) \prod_{x \in T/S} G(F_x)/G(\mathbb{O}_x) \tag{8.9}$$

and to divide not by (8.8b) but by the subset of this set consisting of g for which $gg_xG^{n_x} \subseteq g_xG^{n_x}$, for all $x \in S$. If we work on the set (8.9), we are excluding the effect of the singularities of ω, thus effectively imposing the condition of Lemma 7.1 that the singularities of f and ω be disjoint. So we are left with the middle term of (7.13a).

What is the issue? For GL(1), the form ω leads not immediately but in connection with the Abel–Jacobi theory to a Hecke parameter at every place and the identity of Lemma 7.1 obtained from the residue theorem shows that the character constructed as a product of the local parameters is an idele-class character. Apart from the ambiguity already noted, which has to be resolved, the form ω should also give a local parameter everywhere and thus local spherical functions ϕ_x (normalized, say, to be 1 at the identity). The problem is to show that the local parameters together yield an admissible global parameter. In other words, there is a compatible family of functions f_T (associated in some way to a perverse sheaf) on the varieties of (8.2a), each f_T being, first of all, a linear combination (perhaps in some general sense — a direct image of some perverse sheaf) of left-translations of $\prod_x \phi_x$ and, secondly, invariant under the group $G_T(F)$. The function f_T once chosen — in whatever way imagination suggests — integration around the outside of $\triangle'$ and the residue, can with any luck, be used to show that it is invariant under $G_T(F)$. There is, as will be apparent, a gulf here, maybe two, that I make, for the moment, no proposal for bridging.

I had initially hoped that even if I was unable in this, the first part of the prologue, to reach the relation of the geometric theory to quantum field theory, I would be able to make a convincing suggestion about the construction of the mathematical theory, thus about the construction of the group $\mathfrak{A}_{\text{geom}}$. The possibility of constructing it in terms of differentials with values in ${}^L\mathfrak{g}$ is suggested by the abelian theory and I had hoped — and still hope — that one could prove the appropriate theorem with the help of the residue theorem as for Lemma 7.1. There are encouraging signs, but there are, as I have just explained, also obstacles: for example the full determination from the differentials of the parameters of the spherical functions at each unramified place. On the other hand, the example of elliptic curves [At] suggests that the moduli spaces for G-bundles may be simpler, at least in some respects, than one fears. Although I still had a few weeks grace until the deadline for submitting the paper, I concluded in the face of this and other formidable obstacles that it would be best to stop at the point I had reached, where an uncertain optimism was still possible, and to give myself the leisure — more than a few weeks — to understand better not only the spaces Bun_G and their differential-geometric and algebro-geometric properties but also the quotients $G(F_v)/G(\mathbb{O}_v)$ and the geometrical spherical "functions".

Certainly my limited understanding of the construction of Bun_G is a handicap. There are two puzzles that I have already mentioned. The first is the construction of the local parameters $\mu_x \in {}^LG$ for the spherical functions, defined only at the unramified points. I have suggested that for the abelian theory they are to be defined by the integral of the differential form as $\exp \int_{p_0}^{p} \omega = \mu_p \mu_{p_0}^{-1}$, but without being myself sufficiently clear of how μ_{p_0} was defined. In fact, for $G = \mathrm{GL}(1)$, $\mathrm{Bun}_X = \mathrm{Pic}_X = \bigcup_{n\in\mathbb{Z}} \mathcal{P}^n$ and $p_0 \in X$ is to be interpreted as a point in $\mathcal{P}^1$, the bundle attached to it being the trivial bundle modified by permitting a pole at p_0. The value μ_{p_0} is given by γ in formula (7.13a), thus by a supplementary parameter, so that the parameter of the automorphic representations contains, in addition, to the differential ω a complex number $e^{i\gamma}$ of absolute value 1. Until I understand better the nature of differentials on X with values in ${}^LG$ and the structure of Bun_G, it is idle to make suggestions about the form of $\mathfrak{A}_{\mathrm{geom}}$ that are more precise than that already made at the beginning of this section.

Although I prefer to fix my attention on the geometrical theory as a theory of automorphic forms, thus on functions on Bun_G, it is still necessary to reckon with algebro-geometric aspects of the problem. Geometrically, the Hecke algebra has, it appears, to be defined geometrically. The double-coset space

$$G(\mathbb{O}_v)\backslash G(F_v)/G(\mathbb{O}_v)$$

may be discrete, but $G(\mathbb{O}_x)$ and the spaces $G(F_v)/G(\mathbb{O}_v)$ are algebraic varieties, so that convolution of two elements in the Hecke algebra, formally $\int f_1(gh^{-1}) f_2(h)dh$, has to be defined — so far as I can seen — not as an integral but in terms of direct images of (perverse) sheaves under the mapping of $(g_1, g_2) \mapsto g = g_1 g_2$. I would suppose that, if convolution is to be defined, the spherical functions will also have to be interpreted as sheaves, again perverse, and as sheaves will not have support in a compact subvariety of $G(F_v)$ or $G(F_v)/G(\mathbb{O}_v)$. I am not absolutely certain that we will need a formula for these sheaves, but suspect we will; I am also not certain what form it will take. I take the existence of the formula, in some form, for granted below. The necessary formula could very well be discovered and proved by taking the theory over p-adic fields as a model. So far as I know this has not yet been done. What we have to do, at least for the unramified theory, is first establish, or at least surmise, what the local parameters are. Each of them is supposed to be a conjugacy class in ${}^LG(\mathbb{C})$ or even, if a form of the Ramanujan conjecture is valid, in the unitary form ${}^LU$ of this group, although my interpretation of the results of [At] suggests that the possibility of Arthur parameters intervening has to be kept in mind. It may be that some authors have reflected on the presence of the distinctions familiar from the arithmetic context — tempered or of Arthur type — in the geometric context.

The potential local parameters for a given differential form on hand, one has to show that among the “linear combinations” of left translates of the associated spherical function on (8.8a), a product — over T, a set that has to be allowed to grow larger and larger — of the spherical functions on the individual factors, there is one invariant under the group (8.8b). I should think that to establish this it would be a help to have an explicit formula for the spherical functions.

With the reader's permission, I introduce an intuitive fashion of thinking, based on our experience with spherical functions over archimedean and nonarchimedean local fields. In the representation of $G(F_v)$ (resp. $\prod_T G(F_x)$) on $G(F_x)/G(\mathbb{O}_x)$ (resp. $\prod_T G(F_v)/G(\mathbb{O}_v)$) each representation that occurs, occurs with multiplicity one. We consider the representation with parameter μ_x or rather, at this stage, with parameter $\prod_T \mu_x$. We need to establish that it contains a vector fixed by all elements of $G_T(F)$. For GL(1), this vector was essentially unique and could be determined by integrating the differential. For a general group, we can expect, because of endoscopy and multiplicity, both familiar from the arithmetic theory, the unicity to fail.

We have, at the same time, to contend with something more serious. Functions are not sheaves; sheaves are not functions. Rather they are not, even with the Riemann–Hilbert correspondence, uniquely functions. As I have already made clear, it seems to me that completeness theorems, to assure that we have in hand all pertinent objects of some given sort, require something less ethereal than sheaves, even than perverse sheaves. Nevertheless, for the sake of the argument I confound briefly, in the following observations, functions and sheaves.

The object $\mathrm{Bun}_G(X)$ as defined in terms of (8.8a) and (8.8b) appears far too large, far too coarse, to admit any analysis, but the goal of the theory of moduli spaces as expounded in [Le] is to show that they are, in essence, algebraic varieties, on which differential equations of various sorts can be introduced and studied, so that the exclusive use of sheaves, as in [CFT] is not obligatory, not, in my view, even to be recommended! The general theory does, however, differ from the abelian theory in that a given parameter does not correspond to a single function — up to a scalar factor — but to an infinite-dimensional space of functions, but many of us are already familiar with this from the arithmetic theory.

So what are we to do, keeping in mind that we are working — initially — with the group G_T? The function/sheaf for which we are searching will be a product of spherical functions $\prod_{x \notin T} \phi_x$ times some linear combination of left-translates of $\prod_{x \in T} \phi_x$, presumably by elements in $G_F(T)$. We cannot at first take an average because $G_F(T)$ is infinite, but also because, so far as I know, we cannot take the average of sheaves. On the other hand, we might be able to calculate the change in the function imposed by the translations by $g \in G_T(F)$ by adding up the local modifications as an integral around the boundary of $\triangle'$ finding either that the total change is 0 or that it permits an averaging. This is, at the moment, where the problem stands. Nothing is certain, but there is a great deal on which to reflect!

At this point, I cannot be very much clearer about this proposal for constructing $\mathfrak{A}_{\mathrm{geom}}$. A few observations are, however, in order. It is useful, first of all, to compare it with a conjecture in §6.1 of [CFT], although this conjecture is formulated only in an unramified context.

Conjecture. *Let E be an irreducible ${}^L G$-local system on X. Then there exists a nonzero Hecke eigensheaf* Aut_E *on* Bun_G *with the eigenvalue E whose restriction to each connected component of* Bun_G *is an irreducible perverse sheaf.*

The earlier Assertion is this conjecture for $\mathrm{GL}(n)$. Let me try to explain the general form and its relation with our tentative proposal.

As will be obvious, I have been strongly influenced when composing this section by the geometric and sheaf-theoretic formalism for the Hecke theory with which this conjecture is expressed. This formalism is very elegant, but I did not understand the intuition that informs it. Perhaps I still do not. Nonetheless, if I had not struggled to interpret it in a perhaps more mundane but also more concrete analytic context, it may never have meant anything to me at all. Implicit in the conjecture there are conventions and conceptions — familiar in some circles, less so in those to which I belong — of which we remind ourselves before explaining the relation between it and our goal. Locally the trivial bundle is $G(\mathbb{O}_x)$, a set on which $G(\mathbb{O}_x)$ acts from the right. If $\gamma \in G(F_x)$ then the action of $G(\mathbb{O}_x)$ to the right on $G(\mathbb{O}_x)\gamma G(\mathbb{O}_x)$ defines a G-bundle on $G(\mathbb{O}_x)\gamma G(\mathbb{O}_x)/G(\mathbb{O}_x)$, an operation of blowing-up or modification that we can, if desired, repeat, passing to $G(\mathbb{O}_x)\gamma' G(\mathbb{O}_x)\gamma G(\mathbb{O}_x)/G(\mathbb{O}_x)$, and so on, or just blowing up a given point of $G(\mathbb{O}_x)\gamma G(\mathbb{O}_x)/G(\mathbb{O}_x)$. At all events, this operation allows us to introduce a G-bundle structure on

$$\bigcup_T \prod_{x \in T} G(F_x)/G(\mathbb{O}_x)$$

and then, passing to the limit over $T = \{x_1, \ldots, x_n\}$ as before, first over all possibilities for the set T with a given n and then over n, we arrive, at Bun_G and the G-bundle over it.

It is bundles and modifications that are the preferred form of expression in [CFT]. A central diagram is found in §6.1 of those lectures.

(8.10)
$$\begin{tikzcd} & \mathit{Hecke} \arrow[dl, "\mathfrak{h}^{\leftarrow}"'] \arrow[dr, "\mathfrak{h}^{\rightarrow}"] & \\ \mathrm{Bun}_G & & X \times \mathrm{Bun}_G \end{tikzcd}$$

I do not find the definition of *Hecke* in [CFT, §6.1] perfectly transparent, but I think it safe to take it to be the union over increasing T of the union (or sum) over $x \in T$

of the quotient of[2]

$$\{x \times G(F_x) \times G(F_x)/G(\mathbb{O}_x)\} \times \prod_{\substack{y\in T\\ y\neq x}} G(F_y)/G(\mathbb{O}_y) \tag{8.11}$$

by $G_T(F)$, whose action on $G(F_x)\times G(F_x)/G(\mathbb{O}_x)$ is through the first factor alone, a definition compatible — I hope and, indeed, believe — with that of [CFT]. The arrow on the left of (8.10) takes $x\times h_x\times g_x\times\prod_{y\in T,y\neq x} g_y$ to $x\times h_x\times\prod_{y\in T,y\neq x} g_y$; the arrow on the right takes it to $x\times h_x g_x\times\prod_{y\in T,y\neq x} g_y$. This seems equivalent to the assertion in [CFT], which I have difficulty understanding, and I, myself, see no reasonable alternative to (8.11). Informally, the object *Hecke* consists of quadruples $(\mathcal{M}, \mathcal{M}', x, \beta)$, where $\mathcal{M}$ is a G-bundle on X, $\mathcal{M}'$ is a modification at a single point x, and β is an expression of the identity of $\mathcal{M}$ and $\mathcal{M}'$ outside of x. Of course T grows to include more and more points. Observe that it defines a correspondence that commutes with the action of $G_T(F)$.

In the Conjecture, the initial object is the local system. Our initial object is more, it is the differential. The difference is somewhat difficult to describe, but its source is clear. It is the difference between a local system and a local system with isomorphism. For example, in the analytic context, there is, on the curve X or on its jacobian, the trivial bundle itself, but there is also the trivial bundle plus a section, $\exp(i\,\mathrm{Re}\,\omega)$, where ω is a holomorphic differential with real periods in $2\pi\mathbb{Z}$, a set parametrized by $\mathbb{Z}^{2g}$. Here we distinguish between them. In the Conjecture and in the earlier Assertion, both taken from [CFT], they are confounded. The advantage of the local systems with isomorphism is that it refers to the set of solutions of a precise analytic problem, an eigenvalue problem for the Laplacian, so that we can treat the set without having to exhibit its individual elements. This is what, I hope, the differentials — with whatever supplementary data are necessary — will offer in general.

Having affirmed, for the second time, that there is a difference between the local system and the differential, I now retract and explain that, when trying to understand the meaning of the conjecture, I discovered that this supposed difference was the result either of my careless reading of [CFT] or of the author's careless writing. The author speaks of local systems, local systems for vector bundles and "local systems" for ${}^LG$-bundles — the latter seem to be no more than ${}^LG$-bundles — for they are what allow the definition of the vector bundles V_λ^E, which are local systems, defined

[2]The pertinent phrase from [CFT] is, "Note that the fiber of *Hecke* is the moduli space of pairs (M,β), where $\mathcal{M}$ is a G-bundle on X and $\beta:\mathcal{M}'|_{X\backslash x}$. It is known that this moduli space is isomorphic to a twist of $\mathrm{Gr}_x = G(F_x)/G(\mathbb{O}_x)$ by the $G(\mathbb{O}_x)$-torsor $\mathcal{M}'(\mathbb{O}_x)$ of sections of $\mathcal{M}'$ over $\mathrm{Spec}\,\mathbb{O}_x$:

$$(h^{\to})^{-1}(x,\mathcal{M}') = \mathcal{M}'(\mathbb{O}_x)\times_{G(\mathbb{O}_x)}\mathrm{Gr}_x\text{."}$$

I hope it means what I suggest.

by the constant sections of an ${}^LG$-bundle. For groups with no center, the distinction between the two notions of local system is barely perceptible, and this may be the source of the confusion. The emphasis in [CFT] is often on semisimple groups. The group GL(1) that we were examining in §7 is, however, all center!

This confusion, all being well, clarified, let us try to understand the conjecture. Certainly, whatever we manage to establish, we want it to imply the conjecture! What does it mean for the sheaf $\mathcal{F}$ on Bun_G to be an eigensheaf with eigenvalue E? The condition is formulated sheaf-theoretically as equations (6.1) and (6.2) of [CFT].

$$\text{(CFT-6.1)} \qquad H_\lambda(\mathcal{F}) = \mathfrak{h}_*^{\rightarrow}(\mathfrak{h}^{*\leftarrow}\mathcal{F} \otimes \mathrm{IC}_\lambda);$$

$$\text{(CFT-6.2)} \qquad \iota_\lambda : H_\lambda(\mathcal{F}) \simeq V_\lambda^E \boxtimes \mathcal{F}, \qquad \lambda \in P_+.$$

The first line is a definition. In the second line λ is a dominant weight of ${}^LG$ or a double coset $G(\mathbb{O}_x)\tau_\lambda G(\mathbb{O}_x)$, ι_λ is an isomorphism and V_λ^E is the vector bundle $E \times_{{}^LG} V_\lambda$. The sheaf IC_λ is a perverse sheaf associated to the subvariety $G(\mathbb{O}_x)\tau_\lambda/G(\mathbb{O}_F)$ of $\mathrm{Bun}_{x,\tau_\lambda}$, the Goresky–MacPherson or intersection cohomology sheaf described in [CFT] and many other places. It appears to be the cohomological representative of this subvariety in the context of perverse sheaves, thus the cohomological representative of a spherical function, the characteristic function of a double coset. In any case, the second line is the condition that $\mathcal{F}$ has eigenvalue E. The almost imperceptible mixing of G-bundles and ${}^LG$-bundles is striking!

We replace the ${}^LG$ local system by the differential ω or rather by the set of parameters $\{\mu_x \mid x \in X\}$ associated to it, without troubling ourselves by the imprecisions that this entails at this stage. Let us try to understand the situation in the context of group representations, but only in a grossly informal manner. At all but a finite set S of points in X, we have a representation $\pi_x = \pi(\mu_x)$ of $G(F_x)$, a representation that contains a nonzero vector fixed by $G(\mathbb{O}_F)$. It occurs in the space of functions on $G(F_x)/G(\mathbb{O}_x)$ and, as we infer — for the sake of the argument — from the usual theories of spherical functions, with multiplicity one. So we have a clearly defined space of functions on $\prod_{x\in T} G(F_x)/G(\mathbb{O}_x)$. On each $G(F_x)$, we take a left translate of the spherical function with parameter μ_x. Then we take a tensor product of such functions over $x \in T$ and then linear combinations, perhaps in a topological sense — for example, by convolution with a function, a measure, or a distribution. The group $\prod_{x\in T} G(F_x)$ acts on this space and we assume that there is a nonzero vector Φ in it invariant under $G_T(F)$. That would be our solution of the problem. Does it offer a sheaf $\mathcal{F} = \mathrm{Aut}_E$ satisfying the conditions of the conjecture? The question, at the moment, is not in what sense it might be a sheaf, or in what sense a function, but whether and why we can expect the equation (CFT-6.2) to be valid.

The appropriate construction works entirely from the right, so that the invariance under $G_T(F)$ on the left plays no role in the arguments. It only assures, because the constructions are undertaken from the right, that the result continues to be invariant under $G_T(F)$, so that it can be transferred to Bun_G. In other words, we replace the diagram (8.10) with

$$
\begin{tikzcd}
 & H \arrow[dl, "\mathfrak{h}^{\leftarrow}"'] \arrow[dr, "\mathfrak{h}^{\rightarrow}"] & \\
\prod_{x\in T} G(F_x)/G(\mathbb{O}_x) & & T\times\prod_{y\in T} G(F_y)/G(\mathbb{O}_y)
\end{tikzcd}
\tag{8.12}
$$

The diagram defines H; it is given by (8.11), but there is now no division by $G_T(F)$, neither of H nor of $\prod G(F_y)/G(\mathbb{O}_y)$

If $\mathfrak{h}^{\leftarrow}$ and $\mathfrak{h}^{\rightarrow}$ can be interpreted as actions on the right, then (8.12) may be interpreted as a covering of (8.10). A typical element of H is $(x, h_x, g_x, \prod_{y\neq x} g_y)$. The maps $\mathfrak{h}^{\leftarrow}$ and $\mathfrak{h}^{\rightarrow}$ are defined independently on the various summands and on the various factors, in particular:

$$\mathfrak{h}_x^{\leftarrow} : (h_x, g_x) \mapsto h_x G(\mathbb{O}_x); \qquad \mathfrak{h}_x^{\rightarrow} : x\times(h_x, g_x) \mapsto x\times h_x g_x G(\mathbb{O}_x). \tag{8.13}$$

All these morphisms commute with the action of $G_T(F)$. The fiber of $\mathfrak{h}^{\rightarrow}$ over $(x, g_x G(\mathbb{O}_x))$ is, if I am not mistaken, the set $\{(x, h_x, h_x^{-1} g_x)\}$, thus $G(F_x)$. It is perhaps important to stress as well that the local factors of IC_λ are sheaves on $G(F_x)/G(\mathbb{O}_x)$, so that the global product is a sheaf

The intersection cohomology sheaves IC_λ are defined in [CFT] locally, one at each point of X. We have agreed that a provisional section of a prologue is not the place to describe them precisely. They are, as suggested, the intersection-cohomological representatives of the subvarieties

$$G(\mathbb{O}_x)\tau_\lambda G(\mathbb{O}_x)/G(\mathbb{O}_x)$$

of $G(F_x)/G(\mathbb{O}_x)$. It is plausible that, whatever the precise definition is, we can, as in [CFT] extend it from a local construct to a global construct. Indeed, from the point of view adopted in this prologue, we just define it on (8.11) by pulling back the local IC_λ through the projection on $g_x \in G(F_x)/G(\mathbb{O}_x)$, the third coordinate in (8.11). The result is not invariant under $G_T(F)$. Moreover, there are implicit parameters with an algebraic or function-theoretical significance that are being kept in reserve, the points x in T. One might want to verify that the constructions were compatible with this aspect of the construction — but not now.

I am a tyro here and have by no means understood in any genuine sense intersection cohomology. So I am reduced to guessing what the relations (CFT-6.1) and (CFT-6.2) mean not only in that context, but in the context of functions, if they have an interpretation there. In (CFT-6.1) the sheaf $\mathcal{F}$ or our function Φ depends on the first coordinates h_x alone; the sheaf $\Psi = \mathrm{IC}_\lambda$ depends on the coordinate g_x

alone. The direct image $H_*^{\rightarrow}$ is an integral, in this case,

$$\int \Phi(hg)\Psi(g^{-1})dg = \prod_x \int \Phi_x(h_x g_x)\Psi_x(g_x^{-1})\, dg_x \tag{8.14}$$

thus convolution on the right by Ψ, which does not destroy the invariance under $G_T(F)$.

The function Φ has moreover been obtained as a limit of linear combination of left translates of a product of spherical functions $\otimes_x \phi_x$, whose eigenvalues we know. If we have a formula for these functions, we can calculate

$$\prod_x \int \Phi_x(h_x g_x)\Psi_x(g_x^{-1})\, dg_x,$$

explicitly. This achieved, it should not be too difficult to deduce the relation CFT-6.2)!

On closer examination, there are several troubling aspects to these reflections. It appears to me, as already explained, that the theory of spherical functions in the geometric context necessarily entails the use of sheaves because there is no G-invariant measure with respect to which convolutions of spherical functions on $G(F_x)$ can be defined. The integral in (8.14) is fictional. A graver flaw is that we have not succeeded in introducing into our discussion the essential ingredient of the proof of Weil's identity and Lemma 7.1, namely the residue theorem. To a large extent, although not entirely, this is because I am working around my ignorance of the theory of Bun_G. The conviction that we can deal, even in a geometric context, with automorphic forms on Bun_G as functions is because the spaces defined by (8.2a), or at least large pieces of them, are finite-dimensional algebraic varieties. Their construction as such is difficult and technical and it is fatuous to attempt, as I have been doing, to discuss the geometric theory without having understood it and its results.

The form ω defines a ${}^LG$-bundle with singularities, the bundle E of (CFT - 6.2) Then each irreducible representation $\rho = \rho_\lambda$ of ${}^LG$, λ being the highest weight of ρ, defines a local system, the local system V_λ^E of that formula. If, as before, μ_x are the parameters defined by integrating ω, $\alpha \in V_\lambda$, $\alpha^\star \in V_\lambda^*$, its dual space, then $\alpha^*(\mu_x\alpha)$ is a function on Bun_G. The question is how to combine it with a rational function f on X with values in G so that the result can be integrated over the boundary of $\triangle'$ as in the proof of Lemma 7.1. Although there is a duality between G and ${}^LG$ or between $\mathfrak{g}$ and ${}^L\mathfrak{g}$, it is coarse and I cannot see, at the moment, how it can be used. I am handicapped not only by an aging brain but also by a lack of facility with all the pertinent notions. In addition, the meddle of promising clues and doubtful juxtapositions is daunting to all but the very determined. Nevertheless, although

I have less confidence in the suggestions of this section than in those of the first six, I think there is something to be done.

For example, it is troubling that, as I have observed, the pairing between G and ${}^LG$ or, perhaps better, $\mathfrak{g}$ and ${}^L\mathfrak{g}$ is very coarse, apparently at most a pairing at the level of conjugacy classes, but that may be just as well, because what we sum are residues or products of residues with factors defined by ω with values in ${}^L\mathfrak{g}$. The residues themselves are logarithmic derivatives $df \cdot f^{-1}$. The following relations are clear.

(i) If $f_1 = uf$,

$$df_1 \cdot f_1^{-1} = du \cdot u^{-1} + udf \cdot f^{-1}u^{-1},$$

and the first term has no residue at x if $u \in G(\mathbb{O}_x)$. The conjugacy class of the second term is that of $df \cdot f^{-1}$

(ii) There is a similar relation for $f_1 = fu$,

$$df_1 \cdot f_1^{-1} = df \cdot f^{-1} + fdu \cdot u^{-1}f^{-1}.$$

If $u \in G(\mathbb{O}_x)$, the conjugacy class of the second term is regular at x.

(iii) If $f_1 = ufu^{-1}$,

$$df_1 \cdot f_1^{-1} = du \cdot u^{-1} + udf \cdot f^{-1}u^{-1} - uf^{-1}(u^{-1}du)fu^{-1}.$$

Thus any linear function of the residue at x of $df \cdot f^{-1}$ that is invariant under conjugation does not change on passing from f to f_1. This linear function should be the substitute for the right-hand side of (7.20). It will have to be matched with a substitute for $\int \omega$ as in the proof of the lemma.

Consider for example the group $\mathrm{GL}(n)$, then if z is a local parameter at x, the matrix-valued function f may be written locally as u_1Tu_2, where u_1, u_2 lie in $G(\mathbb{O}_x)$ and

$$T = t(-m_1, -m_2, \ldots, -m_n), \quad m_1 \geq m_2 \cdots \geq m_n,\ m_i \in \mathbb{Z},$$

so that the residue of $f^{-1}df$ is the matrix

$$-\begin{pmatrix} m_1 & 0 & \ldots & 0 \\ 0 & m_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & m_n \end{pmatrix},$$

which can be considered a parameter for the double cosets $G(\mathbb{O}_x)\backslash G(F_x)/G(\mathbb{O}_x)$, or as a highest weight for $\hat{G}$ or ${}^LG$. We of course have a pairing of it with the Lie algebra of ${}^LG$ or with the group ${}^LG$ itself, through the trace of the corresponding representation, thus with ω or $\int \omega$.

So there are a good number of clues that could lead to a nonabelian theory similar to the abelian theory of §7. I do not have a clear notion of how to follow them. I hesitate moreover to search for the theory so long as I have not mastered the techniques for constructing moduli spaces described in [Le]. The moduli space as described in (8.8a) and (8.8b) is convenient in some respects, but it is analytically awkward and, as we found when discussing the conjecture, it encourages us to work not with functions, thus not with solutions of partial differential equations, but with sheaves, for which convolution is possible, at least in a topological sense. I tried in the essay to pass from one to the other by sleight-of-hand, but was not, as even a casual examination reveals, successful. The usual convolution is not defined because $G(F_x)/G(\mathbb{O}_x)$ does not carry an invariant measure. On the other hand, the moduli spaces — or at least large parts of them — are finite-dimensional, even compact, as with the jacobians, algebraic varieties and we might expect to define the eigensheaves as functions satisfying differential equations. I do not yet know what these might be. Moreover, whatever form the final theory takes, I certainly hope it embraces all possibilities: sheaf-theoretic, analytic, and geometric.

It seems best to leave all these questions aside until I acquire a more intimate understanding not only of the nature, both algebro-geometric and differential-geometric, of the moduli spaces, but also of the contributions of the mathematical physicists to what they refer to as the geometric Langlands program.

Contrary to my hopes, which were, in part, unreasonable, the last two sections of this first half of the prologue have turned out to differ sharply from the first six. Although the first six are speculative, they are informed by years of reflection, which has sufficiently matured that I have considerable confidence not only in the correctness of the theory suggested but also in the soundness of the methods proposed for arriving at it. This is not so for the last section, for which the penultimate section was preparation. The last section is only provisional. I hope that on returning to the material in §8.b I can do better!

References

[A1] J. Arthur, "On some problems suggested by the trace formula", pp. 1–49 in *Lie group representations, II* (College Park, MD, 1982–1983)), edited by R. Herb et al., Lecture Notes in Math. **1041**, Springer, Berlin, 1984. MR 85k:11025 Zbl 0541.22011

[At] M. F. Atiyah, "Vector bundles over an elliptic curve", *Proc. London Math. Soc.* (3) **7** (1957), 414–452. MR 24 #A1274 Zbl 0084.17305

[BC] A. Borel and W. Casselman (editors), *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), Proc. Symp. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR 80g:10002a Zbl 0477.00007

[BE] R. Langlands, "Beyond endoscopy", 2004, http://publications.ias.edu/rpl/paper/91.

[CFT] E. Frenkel, "Lectures on the Langlands program and conformal field theory", pp. 387–533 in *Frontiers in number theory, physics, and geometry. II*, edited by P. Cartier et al., Springer, Berlin, 2007. MR 2007k:11102

[CLG] E. Frenkel, *Langlands correspondence for loop groups*, Cambridge Studies in Advanced Mathematics **103**, Cambridge University Press, 2007. MR 2008h:22017 Zbl 1133.22009

[D] R. Dedekind, "Konstruktion von Quaternionkörpern", in *Gesammelte mathematische Werke*, vol. 2, Chelsea Publishing Co., New York, 1968. MR 38 #5572

[DL] P. Deligne and G. Lusztig, "Representations of reductive groups over finite fields", *Ann. of Math.* (2) **103**:1 (1976), 103–161. MR 52 #14076 Zbl 0336.20029

[FG] E. Frenkel and B. Gross, "A rigid irregular connection on the projective line", *Ann. of Math.* (2) **170**:3 (2009), 1469–1512. MR 2012e:14020 Zbl 1209.14017

[FLN] E. Frenkel, R. Langlands, and B. C. Ngô, "Formule des traces et fonctorialité: le début d'un programme", *Ann. Sci. Math. Québec* **34**:2 (2010), 199–243. MR 2012c:11240 Zbl 05998868

[GH] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978. MR 80b:14001 Zbl 0408.14001

[GT] E. Frenkel, "Gauge theory and Langlands duality", pp. 369–403 (exp. 1010) in *Séminaire Bourbaki* 2008/2009, Astérisque **332**, Société Mathématique de France, Paris, 2010. MR 2011m:22033 Zbl 1209.22009

[Ha] G. Harder, "Wittvektoren", *Jahresber. Deutsch. Math.-Verein.* **99**:1 (1997), 18–48. MR 98e:13025 Zbl 0883.13022

[HC] Harish-Chandra, *Collected papers*, edited by V. S. Varadarajan, Springer, New York, 1984. MR 85e:01061d

[He] E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*, Chelsea, New York, 1970. MR 50 #4524 Zbl 0208.06101

[HTT] R. Hotta, K. Takeuchi, and T. Tanisaki, *D-modules, perverse sheaves, and representation theory*, Progress in Mathematics **236**, Birkhäuser, Boston, MA, 2008. MR 2008k:32022 Zbl 1136.14009

[JY] C. U. Jensen and N. Yui, "Quaternion extensions", pp. 155–182 in *Algebraic geometry and commutative algebra*, vol. 1, edited by H. Hijikata et al., Kinokuniya, Tokyo, 1988. MR 90a:12007 Zbl 0691.12011

[K] A. W. Knapp, *Lie groups: Beyond an introduction*, 2nd ed., Progress in Mathematics **140**, Birkhäuser, Boston, MA, 2002. MR 2003c:22001 Zbl 1075.22501

[KW] A. Kapustin and E. Witten, "Electric-magnetic duality and the geometric Langlands program", *Commun. Number Theory Phys.* **1**:1 (2007), 1–236. MR 2008g:14018 Zbl 1128.22013

[Le] J. Le Potier, *Lectures on vector bundles*, Cambridge Stud. Adv. Math. **54**, Cambridge University Press, 1997. MR 98a:14019 Zbl 0872.14003

[LP] M. Larsen and R. Pink, "Determining representations from invariant dimensions", *Invent. Math.* **102**:2 (1990), 377–398. MR 92c:22026 Zbl 0687.22004

[LSP] R. P. Langlands, "Reflexions on receiving the Shaw Prize", pp. 297–308 in *On certain L-functions*, edited by J. Arthur et al., Clay Math. Proc. **13**, Amer. Math. Soc., Providence, RI, 2011. MR 2012d:22001 Zbl 1242.22001

[LW] R. Langlands, Letter to André Weil, 1967, http://publications.ias.edu/rpl/paper/43.

[M] J. S. Milne, "Motives — Grothendieck's dream", expository notes, 2012, http://www.jmilne.org/math/xnotes/MOT.pdf.

[ND] R. Langlands, "Is there beauty in mathematical theories?", in *The many faces of beauty*, edited by D. Stelluto and V. Hösle, University of Notre Dame Press, 2012.

[Re] H. Reichardt, "Über Normalkörper mit Quaternionengruppe", *Math. Z.* **41** (1936), 218–221. MR 1545614 Zbl 0014.00603

[Ri] H. Richter, "Über die Lösbarkeit einiger nicht-Abelscher Einbettungsprobleme", *Math. Ann.* **112**:1 (1936), 69–84. MR 1513036 Zbl 0013.14704

[Si] C. T. Simpson, "Nonabelian Hodge theory", pp. 747–756 in *Proc. ICM* (Kyoto, 1990), vol. 1, edited by I. Satake, Math. Soc. Japan, Tokyo, 1991. MR 93c:14010 Zbl 0765.14005

[Sp] G. Springer, *Introduction to Riemann surfaces*, Addison-Wesley, Reading, MA, 1957. MR 19, 1169g Zbl 0078.06602

[ST] R. Langlands, "Stabilité et transfert", To appear in *Ann. Sci. Math. Québec*.

[T] R. Taylor, "Galois representations", *Ann. Fac. Sci. Toulouse Math.* (6) **13**:1 (2004), 73–119. MR 2005a:11071 Zbl 1074.11030

[VE] J. van Ekeren, "The orbit method for nilpotent Lie groups", lecture notes, http://math.mit.edu/~jethro/kirillov.pdf.

[Wa] J.-L. Waldspurger, "Endoscopie et changement de caractéristique", *Journ. Inst. Math. Jussieu* **5**:3 (2006), 423–525. MR 2007h:22007 Zbl 1102.22010

Received June 30, 2012.

ROBERT LANGLANDS
SCHOOL OF MATHEMATICS
INSTITUTE FOR ADVANCED STUDY
1 EINSTEIN DRIVE
PRINCETON, NJ 08540
UNITED STATES
rpl@ias.edu

PACIFIC JOURNAL OF MATHEMATICS

http://pacificmath.org

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

See inside back cover or pacificmath.org for submission instructions.

The subscription price for 2012 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
http://msp.org/
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS

Volume 260 No. 2 December 2012

Special issue
devoted to the memory of Jonathan Rogawski

0030-8730(201212)260:2;1-A