

THE PRINCIPLE OF FUNCTORIALITY

JAMES ARTHUR

PREFACE

Following the explicit instructions of the organizers, I have tried to write an article that is suitable for a general mathematical audience. It contains some analogies and metaphors that might even be put to nonmathematicians. I hope that experts will be tolerant of the inevitable simplifications.

The principle of functoriality is one of the central questions of present day mathematics. It is a far reaching, but quite precise, conjecture of Langlands that relates fundamental arithmetic information with equally fundamental analytic information. The arithmetic information arises from the solutions of algebraic equations. It includes data that classify algebraic number fields, and more general algebraic varieties. The analytic information arises from spectra of differential equations and group representations. It includes data that classify irreducible representations of reductive groups.

1. SPECTRA

The expected relationships between arithmetic and analytic objects are through properties of spectra. In this sense, the theory bears a resemblance to some key notions from classical physics.

One of the great discoveries of nineteenth century physics was the recognition of absorption spectra in extraterrestrial sources of light. The spectral decomposition of starlight revealed dark bands at characteristic wavelengths across the spectrum. This suggested that light of certain discrete frequencies was being absorbed before it reached telescopes on earth. Physicists also observed that the missing wavelengths matched wavelengths of light emitted by chemical elements in laboratory experiments. They were thereby able to deduce that stars in our galaxy contain the very elements on which the chemistry on earth is based. Later observations of light from other galaxies revealed a downward red shift in characteristic absorption spectra. This indicated that the galaxies were moving apart at enormous velocities, and suggested that the universe might have had a cataclysmic origin.

It was not until the advent of quantum mechanics in the twentieth century that absorption spectra were given a satisfactory theoretical explanation. They were

Received by the editors October 10, 2000, and, in revised form, December 1, 2000, and February 21, 2002.

2000 *Mathematics Subject Classification*. Primary 11R39; Secondary 22E55.

Key words and phrases. Spectra, automorphic representations, Galois group, functoriality, Langlands group, motives.

The author was supported in part by a Guggenheim Fellowship, the Institute for Advanced Study, and an NSERC Operating Grant.

shown to correspond with eigenvalues of appropriate Schrödinger operators. A given atom could absorb or emit light only at certain frequencies, corresponding to the energy levels of bound states represented by different eigenvalues. The mathematical spectra of differential operators thus carried fundamental information about the physical world, which even now seems almost magical.

The analogy with number theory is through spectra of other differential operators. These are Laplace-Beltrami operators (and variants of higher degree) attached to certain Riemannian manifolds. The spectra of these and other operators are expected to carry fundamental information about the arithmetic world, a possibility that also seems quite magical.

2. AUTOMORPHIC REPRESENTATIONS

The spectral data that is believed to be related to number theory is framed in the language of automorphic forms. Suppose that H is a locally compact, unimodular topological group, and that Γ is a discrete subgroup of H . Let R be the regular representation of H on the Hilbert space $L^2(\Gamma \backslash H)$, taken with respect to a Haar measure on H . Then R is the unitary representation of H defined by

$$(R(y)\phi)(x) = \phi(xy), \quad \phi \in L^2(\Gamma \backslash H), \quad x, y \in H .$$

One is interested in the decomposition of R into irreducible unitary representations of H . For example, $\Gamma = \mathbb{Z}$ is a discrete subgroup of the additive group $H = \mathbb{R}$. The decomposition of R in this case amounts to the theory of Fourier series.

The theory of automorphic forms concerns a much richer family of examples. For Γ , one takes a group of rational points $G(\mathbb{Q})$, where G is a (connected) reductive algebraic group defined over \mathbb{Q} . For example, one could take G to be the general linear group $\mathrm{GL}(n)$ of rank n . Then $\mathrm{GL}(n, \mathbb{Q})$ is just the multiplicative group of invertible, rational, $(n \times n)$ -matrices. In general, it seems pretty clear that $\Gamma = G(\mathbb{Q})$ has an interesting arithmetic structure. What may not be obvious at first glance is that there should be an associated locally compact group H .

The rational field comes with the usual absolute value

$$v_\infty(t) = |t|_\infty, \quad t \in \mathbb{Q} ,$$

and its corresponding completion $\mathbb{Q}_\infty = \mathbb{R}$. For each prime number p , there is also a p -adic absolute value

$$v_p(t) = |t|_p, \quad t \in \mathbb{Q} ,$$

on \mathbb{Q} . Its completion is the field $\mathbb{Q}_{v_p} = \mathbb{Q}_p$ of p -adic numbers. We recall that if

$$t = (ab^{-1})p^r, \quad (a, p) = (b, p) = 1, \quad r \in \mathbb{Z} ,$$

then

$$|t|_p = p^{-r} .$$

In particular, t is of small p -adic absolute value if it is highly divisible by p . The p -adic absolute value satisfies the strong form

$$|t_1 + t_2|_p \leq \max \{|t_1|_p, |t_2|_p\}, \quad t_1, t_2 \in \mathbb{Q} ,$$

of the triangle inequality. It follows that the compact subset

$$\mathbb{Z}_p = \{t_p \in \mathbb{Q}_p : |t_p|_p \leq 1\}$$

of \mathbb{Q}_p is actually a subring.

The fields

$$\mathbb{Q}_v = \begin{cases} \mathbb{R}, & \text{if } v = v_\infty, \\ \mathbb{Q}_p, & \text{if } v = v_p, \end{cases}$$

are all locally compact. They can be put together into a locally compact ring. Since the fields are actually noncompact, their direct product will not in fact be locally compact. However, one can form the restricted direct product

$$\begin{aligned} \mathbb{A} &= \prod_v^{\text{rest}} \mathbb{Q}_v = \mathbb{R} \times \prod_p^{\text{rest}} \mathbb{Q}_p \\ &= \{t = (t_v): t_p = t_{v_p} \in \mathbb{Z}_p \text{ for almost all } p\}. \end{aligned}$$

Endowed with the obvious direct limit topology, $\mathbb{A} = \mathbb{A}_{\mathbb{Q}}$ becomes a locally compact ring, called the adèle ring of \mathbb{Q} . The diagonal image of \mathbb{Q} in \mathbb{A} is easily seen to be discrete. One deduces from this that $H = G(\mathbb{A})$ is a locally compact group, in which $\Gamma = G(\mathbb{Q})$ embeds as a discrete subgroup.

The regular representation R of $G(\mathbb{A})$ on $L^2(G(\mathbb{Q})) \backslash G(\mathbb{A})$ is not generally a direct sum of irreducible subrepresentations. It does often have a part that decomposes discretely, like the theory of Fourier series, but it can also have a part that decomposes continuously, like the theory of Fourier transforms. An *automorphic representation* of G can be described informally as an irreducible representation π of $G(\mathbb{A})$ that is a constituent of R . (The formal definition in [L4] includes representations obtained by analytic continuation in the parameters of the continuous spectrum.) Suppose that π is an arbitrary irreducible representation of $G(\mathbb{A})$. Under a weak continuity condition, which is always satisfied by automorphic representations, one can show that π is a restricted tensor product

$$\pi = \bigotimes_v \pi_v$$

of irreducible representations of the groups $G(\mathbb{Q}_v)$, in which almost every π_v is in a natural sense unramified. Conversely, one could obviously construct an irreducible π from any such choice of representations $\{\pi_v\}$. However, the condition that π be automorphic is very rigid. It imposes deep and interesting relationships among the components $\{\pi_v\}$ of π .

3. DUAL GROUPS AND CONJUGACY CLASSES

As an example, let us describe the unramified representations of general linear groups. Suppose that $G = \text{GL}(n)$. An irreducible representation π_v of $G(\mathbb{Q}_v) = \text{GL}(n, \mathbb{Q}_v)$ is *unramified* if $v = v_p$ is p -adic, and if the restriction of $\pi_{v_p} = \pi_p$ to the maximal compact subgroup $G(\mathbb{Z}_p)$ of $G(\mathbb{Q}_p)$ contains the trivial representation. The unramified representations can be classified as follows.

Let

$$B(\mathbb{Q}_p) = \left\{ b = \begin{pmatrix} b_{11} & & * \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix} \right\}$$

be the Borel subgroup of upper triangular matrices in $G(\mathbb{Q}_p)$. Suppose that $u = (u_1, \dots, u_n)$ is an n -tuple of complex numbers. Then

$$\chi_u(b) = |b_{11}|^{u_1 + \frac{n-1}{2}} |b_{22}|^{u_2 + \frac{n-3}{2}} \dots |b_{nn}|^{u_n - \frac{n-1}{2}}, \quad b \in B(\mathbb{Q}_p),$$

is a 1-dimensional representation of $B(\mathbb{Q}_p)$. (The exponents $(\frac{n-1}{2}, \dots, -(\frac{n-1}{2}))$ represent a normalization that compensates for the failure of the group $B(\mathbb{Q}_p)$ to be unimodular.) Let $\tilde{V}_{p,u}$ be the space of continuous functions ϕ from $G(\mathbb{Q}_p)$ to \mathbb{C} such that

$$\phi(bx) = \chi_u(b)\phi(x), \quad b \in B(\mathbb{Q}_p), \quad x \in G(\mathbb{Q}_p).$$

There is then an induced representation $\tilde{\pi}_{p,u}$ of $G(\mathbb{Q}_p)$ on $\tilde{V}_{p,u}$ by right translation:

$$(\tilde{\pi}_{p,u}(y)\phi)(x) = \phi(xy), \quad x, y \in G(\mathbb{Q}_p).$$

The representation $\tilde{\pi}_{p,u}$ is usually irreducible. It always has a unique irreducible constituent $\pi_{p,u}$ that is unramified. If $u' = (u'_1, \dots, u'_n)$ is any other n -tuple, one can show that the representation $\pi_{p,u'}$ is equivalent to $\pi_{p,u}$ if and only if

$$(u'_1, \dots, u'_n) \equiv (u_{\sigma(1)}, \dots, u_{\sigma(n)}) \left(\text{mod} \left(\frac{2\pi i}{\log p} \right) \mathbb{Z}^n \right),$$

for some permutation σ of $(1, \dots, n)$. Conversely, any unramified irreducible representation π_p of $G(\mathbb{Q}_p)$ is equivalent to $\pi_{p,u}$, for some u .

Set \hat{G} equal to the complex general linear group $\text{GL}(n, \mathbb{C})$. An unramified representation $\pi_p \simeq \pi_{p,u}$ of $G(\mathbb{Q}_p)$ can then be represented uniquely as a semisimple conjugacy class

$$c(\pi_p) = \begin{pmatrix} p^{-u_1} & & 0 \\ & \ddots & \\ 0 & & p^{-u_n} \end{pmatrix}$$

in \hat{G} . (The powers p^{-u_i} eliminate the ambiguity caused by the congruence condition above, while the identification of $c(\pi_p)$ with a conjugacy class takes care of the ambiguity caused by permutations.)

Suppose now that π is an automorphic representation of $G = \text{GL}(n)$. Then π determines a family

$$c(\pi) = \{c_p(\pi) = c(\pi_p) : p \notin S = S(\pi)\}$$

of conjugacy classes in \hat{G} , one class of almost every prime number p . An automorphic representation thus provides us with an explicit set of elementary data.

The construction we have described for $\text{GL}(n)$ was carried out by Langlands for a general reductive group G . His first step was to define a dual group \hat{G} . In general, \hat{G} is a complex reductive group that is related to G in two ways. First, any maximal torus \hat{T} in \hat{G} is isomorphic to the complex dual torus

$$X^*(T) \otimes \mathbb{C}^* = \text{Hom}(X_*(T), \mathbb{C}^*)$$

of any maximal torus T in G . Secondly, the Coxeter-Dynkin diagram for \hat{G} is the dual of the associated diagram for G . The dual group provides a natural parameter space for representations, but it has to be supplemented by a second piece of information.

The general situation is complicated by the fact that G need not be *split* over \mathbb{Q} , in the sense that it contain a maximal torus T that is isomorphic over $\overline{\mathbb{Q}}$ to a product of groups $\text{GL}(1)$. One can always choose an isomorphism ψ over $\overline{\mathbb{Q}}$ from G to such a group. The map

$$\sigma \longrightarrow \sigma(\psi^{-1})\psi, \quad \sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}),$$

then represents the obstruction to ψ being defined over \mathbb{Q} . This map determines a homomorphism

$$\Gamma_{\mathbb{Q}} = \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \text{Out}(G) = \text{Aut}(G)/\text{Int}(G)$$

from the absolute Galois group of \mathbb{Q} to the group of outer automorphisms of G . The last homomorphism does not determine the isomorphism class of G over \mathbb{Q} uniquely. To every such homomorphism, there does correspond a unique G over \mathbb{Q} that is *quasisplit*, in the sense that it contains a Borel subgroup (i.e. a maximal solvable subgroup) that is defined over \mathbb{Q} . For this reason, quasisplit groups play a central role in the theory. In general, the group $\text{Out}(G)$ is canonically isomorphic to $\text{Out}(\widehat{G})$. The group G therefore determines both a complex dual group \widehat{G} and a homomorphism from $\Gamma_{\mathbb{Q}}$ to $\text{Out}(\widehat{G})$.

Given G , Langlands introduced what is now called the *L-group* of G . It is defined as a semidirect product

$${}^L G = \widehat{G} \rtimes \Gamma_{\mathbb{Q}},$$

relative to the homomorphism from $\Gamma_{\mathbb{Q}}$ to $\text{Out}(\widehat{G})$, and a suitable section from $\text{Out}(\widehat{G})$ to $\text{Aut}(\widehat{G})$. The action of $\Gamma_{\mathbb{Q}}$ on \widehat{G} factors through a quotient $\text{Gal}(E/\mathbb{Q})$ of $\Gamma_{\mathbb{Q}}$, for a finite Galois extension E of \mathbb{Q} . For many purposes, one can replace the profinite group $\Gamma_{\mathbb{Q}}$ by the finite quotient $\text{Gal}(E/\mathbb{Q})$. In particular, if G is a group such as $\text{GL}(n)$ that splits over \mathbb{Q} , one can often work with the dual group \widehat{G} instead of the full *L-group* ${}^L G$.

The notion of an unramified representation of $G(\mathbb{Q}_p)$ can be defined whenever G is quasisplit over \mathbb{Q}_p and is split over an unramified extension of \mathbb{Q}_p . Under these conditions, which hold for almost all p , the construction for $\text{GL}(n)$ above is easily extended to $G(\mathbb{Q}_p)$. It provides a bijection $\pi_p \longleftrightarrow c(\pi_p)$ between the unramified irreducible representations π_p of $G(\mathbb{Q}_p)$, and the conjugacy classes $c(\pi_p)$ in

$$\widehat{G} \rtimes \text{Gal}(E/\mathbb{Q})$$

whose image in \widehat{G} is semisimple, and whose image in $\text{Gal}(E/\mathbb{Q})$ equals Frob_p , the Frobenius class of p .

At this point, we ought to recall that Frob_p is a well defined conjugacy class in $\text{Gal}(E/\mathbb{Q})$, for every p outside a finite set $S(E)$. For example, if E is represented as a splitting field of an irreducible monic polynomial with integral coefficients, the primes p with $\text{Frob}_p = 1$ are the ones for which the polynomial breaks into distinct linear factors modulo p . A rather deep theorem from algebraic number theory asserts that these primes completely determine E . In other words, the map

$$E \rightarrow \text{Spl}_E = \{p: \text{Frob}_p = 1\},$$

from finite Galois extensions E to families of prime numbers, is injective. An independent characterization of the image of this map could be regarded as a classification of the finite Galois extensions E .

Given the parametrization of unramified representations, we see that an automorphic representation π of G determines a family of conjugacy classes

$$c(\pi) = \{c_p(\pi) = c(\pi_p): p \notin S = S(\pi)\}$$

in $\widehat{G} \rtimes \text{Gal}(E/\mathbb{Q})$. The dependence of this family on $S(\pi)$ and E is, incidentally, quite inessential. One can remove it by declaring two families of conjugacy classes $c = \{c_p\}$ and $c' = \{c'_p\}$ to be equivalent if $c_p = c'_p$ for almost all p . An automorphic

representation thus provides a well defined equivalence class $c(\pi) = \{c_p(\pi)\}$. In any case, our discussion to this point leads us to draw the following conclusion. The rigidity inherent in the condition that π be automorphic will be reflected in rather concrete objects. We should expect there to be interesting relationships among the different conjugacy classes in the family $c(\pi)$ attached to π .

The conjugacy classes $\{c_p(\pi)\}$ attached to an automorphic representation

$$\pi = \bigotimes_v \pi_v = \pi_{\mathbb{R}} \otimes \left(\bigotimes_p \pi_p \right)$$

represent analytic data. The language of group representations is really what is most natural, but the constructions could have been carried out in terms of the spectral theory of differential operators. Any irreducible representation $\pi_{\mathbb{R}}$ of $G(\mathbb{R})$ determines a set of eigenvalues of an algebra of Laplace-type differential operators. These operators act simultaneously on a tower of Riemannian locally symmetric spaces attached to G . The p -adic representations $\{\pi_p\}$ have their own spectral interpretation. In particular, any family $\{c_p(\pi) : p \notin S\}$ determines a homomorphism

$$h^S = \bigotimes_{p \notin S} h_p$$

from a certain commutative \mathbb{C} -algebra

$$\mathcal{H}^S = \bigotimes_{p \notin S} \mathcal{H}_p$$

into \mathbb{C} . Elements in \mathcal{H}^S are called *Hecke operators*. They act on the underlying tower of locally symmetric spaces, and commute with the Laplace operators. They are in fact combinatorial analogues of Laplace operators, in the precise sense of graph theory (and its higher dimensional analogues). The conjugacy classes $\{c_p(\pi)\}$ thus provide eigenvalues of Hecke operators. In the language of quantum mechanics, they represent energy levels of momentum observables.

4. FUNCTORIALITY

The principle of functoriality describes deep relationships among automorphic representations on different groups. In its most fundamental form, it is a remarkably simple conjecture that relates the associated families $\{c_p(\pi)\}$ of semisimple conjugacy classes.

Suppose that G' is a second reductive group over \mathbb{Q} , and that ρ is an algebraic homomorphism from ${}^L G'$ to ${}^L G$ such that the diagram

$$\begin{array}{ccc} {}^L G' & \xrightarrow{\rho} & {}^L G \\ & \searrow & \swarrow \\ & \Gamma_{\mathbb{Q}} & \end{array}$$

is commutative. Suppose also that π' is an automorphic representation of G' . Then $\{c_p(\pi')\}$ is a family of semisimple conjugacy classes in ${}^L G'$. Composing with ρ , we obtain a family of semisimple conjugacy classes $\{\rho(c_p(\pi'))\}$ in ${}^L G$.

Conjecture (Langlands [L1]). *Suppose that G , G' and ρ are given, and that G is quasisplit. Then for any automorphic representation π' of G' , there is an automorphic representation π of G such that*

$$c_p(\pi) = \rho(c_p(\pi')), \quad p \notin S(\pi) \cup S(\pi') .$$

The principle of functoriality, in the basic form we have just stated, is really quite concrete. Semisimple conjugacy classes in complex groups can certainly be described in explicit terms. Homomorphisms ρ between L -groups are also pretty concrete objects. For example, if $G = \mathrm{GL}(n)$ and G' is split, ρ can be decomposed in terms of irreducible representations of the finite group $\mathrm{Gal}(E/\mathbb{Q})$ and the complex reductive group \widehat{G}' . The former objects go back to Frobenius, while the latter are classified by the highest weight theory of Weyl. The general principle of functoriality asserts that the relationships imposed on the classes $c_p(\pi')$ by the condition that π' be automorphic, whatever form they might take, are supplemented by a completely separate set of relationships imposed by the condition that π be automorphic. Any case of functoriality that can be established thus represents a reciprocity law between the arithmetic data implicit in the automorphic representations π and π' .

To get a sense of the depth of the conjecture, we consider what might at first glance seem to be an elementary special case. Suppose that $G' = \{1\}$ and $G = \mathrm{GL}(n)$. Then π' is of course trivial. The choice of ρ amounts to that of a homomorphism

$$r: \mathrm{Gal}(E/\mathbb{Q}) \rightarrow \mathrm{GL}(n, \mathbb{C}) ,$$

for a finite Galois extension E of \mathbb{Q} . Functoriality in this case therefore reduces to the following assertion.

Conjecture (Langlands [L1]). *For any r , there is an automorphic representation π of $\mathrm{GL}(n)$ such that*

$$c_p(\pi) = r(\mathrm{Frob}_p), \quad p \notin S(E) .$$

We can assume that r is faithful. The conjecture then asserts that

$$\mathrm{Spl}_E = \{p \notin S(E) : c_p(\pi) = 1\} .$$

This amounts to a characterization of the set Spl_E in analytic terms and, in this sense, represents a classification of the finite Galois extensions E of \mathbb{Q} . The conjecture remains for the most part unproved. The cases that have been established are summarized in the following examples.

(a) $n = 1$. The conjecture in this case is the Kronecker-Weber theorem. An elementary analysis of the cyclotomic fields

$$\mathbb{Q} \left(e^{\frac{2\pi i}{N}} \right), \quad N \geq 1 ,$$

and of the automorphic representations of $\mathrm{GL}(1)$ converts this case of the conjecture to the assertion that any extension E with $\mathrm{Gal}(E/\mathbb{Q})$ abelian is contained in some $\mathbb{Q} \left(e^{\frac{2\pi i}{N}} \right)$. Langlands actually formulated the principle of functoriality for a general global field F in place of \mathbb{Q} . In this more general context, the case of $n = 1$ is the Artin reciprocity law, which is the essence of class field theory. (See [H].)

(b) $n = 2$, $\mathrm{Gal}(E/\mathbb{Q})$ solvable. This is the theorem of Langlands [L3] and Tunnell [T] that was an important ingredient of Wiles's proof of Fermat's Last Theorem. It is a consequence of cyclic base change for $\mathrm{GL}(2)$, among other things, and is valid if \mathbb{Q} is replaced by any number field F .

(c) n arbitrary, $\mathrm{Gal}(E/\mathbb{Q})$ nilpotent. The conjecture was established in this case as an application of cyclic base change for $\mathrm{GL}(n)$ [AC]. The result is again valid if \mathbb{Q} is replaced by an arbitrary number field F .

(d) Further results. Partial results for $n = 2$ have recently been established in the remaining case that the image of $\mathrm{Gal}(E/\mathbb{Q})$ in $\mathrm{PGL}(2, \mathbb{C})$ is the simple group

A_5 [BDST]. These exploit the p -adic properties of holomorphic modular forms. In the case that $n = 4$ and $\text{Gal}(E/\mathbb{Q})$ is solvable, the conjecture has recently been established for representations ρ that factor through the group $\text{GO}(4, \mathbb{C})$ of orthogonal similitudes [Ra]. The result is valid for a general ground field F (of characteristic 0).

5. THE AUTOMORPHIC LANGLANDS GROUP

Suppose for the moment that $G = \text{GL}(n)$. The special case of functionality we discussed in the last section asserts that there is an automorphic representation of G attached to any continuous, n -dimensional representation of the compact, totally disconnected group $\Gamma_{\mathbb{Q}} = \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. It would be natural to inquire about the converse. Given any automorphic representation π of G , is there a continuous homomorphism

$$r: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \text{GL}(n, \mathbb{C})$$

such that

$$c_p(\pi) = r(\text{Frob}_p), \quad p \notin S(E) \cup S(\pi) ?$$

The answer is no! For it is known that there are rather stringent necessary conditions on the local component $\pi_{\mathbb{R}}$ of any such π . To show that these are also sufficient conditions is an open problem, on which there has been progress in the case $n = 2$. However, the point we are making is that there are more automorphic representations π than there are Galois representations r . To what might they correspond?

In [L5], Langlands suggested that there ought to be a universal group $L_{\mathbb{Q}}$ whose n -dimensional representations parametrize automorphic representations of $\text{GL}(n)$, or rather, the set of families $c = \{c_p\}$ attached to automorphic representations. Formulated in terms of a later suggestion of Kottwitz, $L_{\mathbb{Q}}$ would be a locally compact group, equipped with a surjective map $L_{\mathbb{Q}} \rightarrow \Gamma_{\mathbb{Q}}$. It should have many of the formal properties of the absolute Galois group $\Gamma_{\mathbb{Q}}$. In particular, it should have a local analogue $L_{\mathbb{Q}_v}$, for each v , that fits into a commutative diagram

$$\begin{array}{ccc} L_{\mathbb{Q}_v} & \longrightarrow & \Gamma_{\mathbb{Q}_v} \\ \downarrow & & \downarrow \\ L_{\mathbb{Q}} & \longrightarrow & \Gamma_{\mathbb{Q}} \end{array}$$

of continuous homomorphisms. The vertical embedding on the left would be determined only up to conjugacy, and would extend the familiar conjugacy class of embeddings $\Gamma_{\mathbb{Q}_v} \hookrightarrow \Gamma_{\mathbb{Q}}$ of the local Galois group $\Gamma_{\mathbb{Q}_v} = \text{Gal}(\overline{\mathbb{Q}_v}/\mathbb{Q}_v)$.

The main property of $L_{\mathbb{Q}}$ should be its relationship to automorphic representations. In the special case of $\text{GL}(n)$, one would ask for a bijection

$$r \longleftrightarrow c$$

between equivalence classes of continuous, completely reducible, n -dimensional representations r of $L_{\mathbb{Q}}$ and equivalence classes of automorphic families c of conjugacy classes for $\text{GL}(n)$. This would be supplemented by local bijections

$$r_v \longleftrightarrow \pi_v$$

between equivalence classes of continuous, completely reducible, n -dimensional representations r_v of $L_{\mathbb{Q}_v}$ and equivalence classes of continuous irreducible representations π_v of $\mathrm{GL}(n, \mathbb{Q}_v)$. The local and global bijections should be compatible in the natural sense that for any r , there is an automorphic representation π with $c(\pi) = c$, such that for each v , the restriction r_v of r to $L_{\mathbb{Q}_v}$ corresponds to the local component π_v of π .

The general correspondence would be slightly weaker. If G is quasisplit, one would expect surjective maps $\phi \rightarrow c$ and $\phi_v \leftarrow \pi_v$, the latter at least with finite fibers, in which ϕ and ϕ_v are conjugacy classes of homomorphisms $L_{\mathbb{Q}} \rightarrow {}^L G$ and $L_{\mathbb{Q}_v} \rightarrow {}^L G_v$ that commute with the maps to the associated Galois groups. If G is arbitrary, the expectation is similar, except that the local correspondence $\phi_v \leftarrow \pi_v$ need not be surjective, while the global correspondence $\phi \rightarrow c$ would have to be restricted to those maps ϕ such that each ϕ_v lies in the image of the local correspondence. In general, the local and global correspondences would again have to be compatible. For any ϕ , there should be an automorphic representation π of G with $c(\pi) = c$, such that for each v , the local component π_v of π maps to the localization ϕ_v of ϕ .

Our choice of \mathbb{Q} for all this discussion has been purely for simplicity. The discussion is essentially the same if \mathbb{Q} is replaced by any global field F . In particular, there ought to be a locally compact group L_F , equipped with an embedding $L_{F_v} \rightarrow L_F$ for each completion F_v of F . This group should classify automorphic families of conjugacy classes for any reductive group G over F , in the manner outlined above.

The local Langlands groups are quite elementary. They are given by

$$L_{F_v} = \begin{cases} W_{F_v}, & \text{if } v \text{ is archimedean,} \\ W_{F_v} \times \mathrm{SU}(2, \mathbb{R}), & \text{if } v \text{ is nonarchimedean,} \end{cases}$$

where W_{F_v} is the Weil group of the local field F_v . In general, Weil groups are locally compact groups attached to local or global fields. They are defined in a relatively concrete way from the corresponding Galois groups, and fit together into a commutative diagram

$$\begin{array}{ccc} W_{F_v} & \longrightarrow & \Gamma_{F_v} \\ \downarrow & & \downarrow \\ W_F & \longrightarrow & \Gamma_F . \end{array}$$

The local Langlands group L_{F_v} is thus a (split) extension of W_{F_v} by a compact, simply connected Lie group (either $\{1\}$ or $\mathrm{SU}(2, \mathbb{R})$). If v is archimedean, Langlands has established the local correspondence $\phi_v \leftarrow \pi_v$ in general. His proof [L2] uses fundamental results of Harish-Chandra on harmonic analysis on $G(F_v)$. For nonarchimedean v there has been significant progress, but a proof of the general local correspondence in this case still seems a long way off.

The global Langlands group L_F will be much larger. Using the theory of automorphic L -functions, and analytic properties of such functions implied by functoriality, one can construct a reasonably explicit candidate for L_F . It is an infinite fiber product of (nonsplit) extensions

$$1 \rightarrow K_c \rightarrow L_c \rightarrow W_F \rightarrow 1$$

of W_F by compact, semisimple, simply connected Lie groups K_c . However, one would have to establish something considerably stronger than functoriality in order to show that L_F has all the desired properties.

6. THE MOTIVIC GALOIS GROUP

It would be appropriate to say something at this point about the conjectural theory of motives. I do not feel confident in doing so. However, the notion of a motive is surely one of the great ideas of twentieth century mathematics. It ought to have at least some mention at this meeting.

The theory is due to Grothendieck. It could be regarded in naïve terms as an attempt to classify the arithmetic data in nonsingular, projective algebraic varieties over \mathbb{Q} . In this sense, it is a kind of class field theory for algebraic varieties of higher dimension. Varieties, however, turn out not to be the fundamental objects. Grothendieck discovered a universe within algebraic geometry, so to speak, whose “elementary particles” he called irreducible motives. The theory is quite precise. However, it depends on the so-called “standard conjectures”, which describe properties of algebraic cycles and their images in the cohomology of varieties.

Consider the case of class field theory. At its most elementary level, class field theory is an attempt to study solutions of polynomial equations with rational coefficients in one variable. Such objects of course define algebraic varieties over \mathbb{Q} of dimension zero. Galois theory introduces a more intrinsic object, the absolute Galois group $\Gamma_{\mathbb{Q}}$. It is a profinite group that captures much of the essential information from the original family of objects. To organize the information into manageable form, one takes continuous representations

$$r: \Gamma_{\mathbb{Q}} \rightarrow \mathrm{GL}(n, \mathbb{C}) .$$

As we have seen, the conjecture of Langlands relates the arithmetic information in any r with analytic information from some automorphic representation of $\mathrm{GL}(n)$.

In general, one can take the larger family of nonsingular, projective algebraic varieties over \mathbb{Q} of arbitrary dimension. Grothendieck’s conjectural theory of motives introduces a more intrinsic object, the motivic Galois group $\mathcal{G}_{\mathbb{Q}}$. In its most down to earth form, $\mathcal{G}_{\mathbb{Q}}$ is a reductive proalgebraic group over \mathbb{C} , which again captures much of the essential information in the original family of objects. A *motive* of rank n is a proalgebraic representation

$$m: \mathcal{G}_{\mathbb{Q}} \rightarrow \mathrm{GL}(n, \mathbb{C}) .$$

It is conjectured that the arithmetic information in any m is again directly related to analytic information from some automorphic representation of $\mathrm{GL}(n)$. The group $\mathcal{G}_{\mathbb{Q}}$ comes with a proalgebraic projection $\mathcal{G}_{\mathbb{Q}} \rightarrow \Gamma_{\mathbb{Q}}$. Any continuous representation r of $\Gamma_{\mathbb{Q}}$ therefore pulls back to a proalgebraic representation $\mathcal{G}_{\mathbb{Q}}$, and can in this way be regarded as a motive.

The conjectural theory of motives also applies to any completion \mathbb{Q}_v of \mathbb{Q} . It yields a proalgebraic group $\mathcal{G}_{\mathbb{Q}_v}$ over \mathbb{C} that fits into a commutative diagram

$$\begin{array}{ccc} \mathcal{G}_{\mathbb{Q}_v} & \longrightarrow & \Gamma_{\mathbb{Q}_v} \\ \downarrow & & \downarrow \\ \mathcal{G}_{\mathbb{Q}} & \longrightarrow & \Gamma_{\mathbb{Q}} \end{array}$$

of proalgebraic homomorphisms. By the 1970s, Deligne and Langlands were prepared to conjecture very general relations between motivic Galois groups and the automorphic representations of any G . Expressed in terms of the automorphic Langlands groups of §5, they are as follows.

Conjecture (Langlands [L5, §2]). *There is a commutative diagram*

$$\begin{array}{ccc} L_{\mathbb{Q}} & \xrightarrow{\Phi} & \mathcal{G}_{\mathbb{Q}} \\ & \searrow & \swarrow \\ & \Gamma_{\mathbb{Q}} & \end{array}$$

together with a compatible commutative diagram

$$\begin{array}{ccc} L_{\mathbb{Q}_v} & \xrightarrow{\Phi_v} & \mathcal{G}_{\mathbb{Q}_v} \\ & \searrow & \swarrow \\ & \Gamma_{\mathbb{Q}_v} & \end{array}$$

for each completion \mathbb{Q}_v of \mathbb{Q} , in which Φ and Φ_v are continuous homomorphisms.

The conjecture attaches to any proalgebraic homomorphism μ from $\mathcal{G}_{\mathbb{Q}}$ to ${}^L G$ over Γ_F , an associated automorphic representation of G . In particular, it implies that if m is any motive of rank n , there is an automorphic representation π of $\mathrm{GL}(n)$ with the following property. The family of conjugacy classes $c(\pi) = \{c_p(\pi)\}$ in $\mathrm{GL}(n, \mathbb{C})$ associated to π is equal to the family of conjugacy classes $c(m) = \{c_p(m)\}$ obtained from m , and the local homomorphisms $\mathcal{G}_{\mathbb{Q}_p} \rightarrow \mathcal{G}_{\mathbb{Q}}$ at primes p that are unramified for m . The importance of the arithmetic information in $c(m)$ will unfortunately not be clear from this brief discussion. We can only remark that $c_p(m)$ is the image of the Frobenius class Frob_p under a different kind of representation of $\Gamma_{\mathbb{Q}}$, namely a compatible family

$$\Gamma_{\mathbb{Q}} \rightarrow \prod_{\ell \notin S(m) \cup \{p\}} \mathrm{GL}(n, \mathbb{Q}_{\ell})$$

of ℓ -adic representations attached to m . This object comes from the piece of the cohomology of some variety that is the ultimate origin of m .

Suppose that E is an elliptic curve defined over \mathbb{Q} . Then E is a direct sum of a motive m_E of rank 2 and two simpler motives of rank 1. The family of conjugacy classes

$$c(E) = c(m_E) = \{c_p(m_E)\}$$

has an elementary description in terms of the Tate module for E . The conjecture above asserts in this case that there is an automorphic representation π of $\mathrm{GL}(2)$ such that the family $c(\pi) = \{c_p(\pi)\}$ equals $c(E)$. This is the celebrated conjecture of Shimura, Taniyama, and Weil. It has now been proved in complete generality, in a paper [BCDT] that extends the special case established by Wiles in his proof of Fermat's Last Theorem.

The general discussion of this section makes sense if \mathbb{Q} is replaced by any global field F . The conjectural theory of motives over F yields a reductive, proalgebraic group \mathcal{G}_F over \mathbb{C} and a map $\mathcal{G}_F \rightarrow \Gamma_F$. It is natural to inquire about the structure of this group [S2]. As in the case of the group L_F , one ought to be able to construct a fairly explicit candidate for \mathcal{G}_F , at least up to conjugation by elements in the connected component \mathcal{G}_F^0 . The conjectural properties of automorphic

L -functions, and the conjectural characterization of those automorphic representations that come from motives, suggest taking \mathcal{G}_F to be a proalgebraic fiber product of certain extensions

$$1 \rightarrow \mathcal{D}_c \rightarrow \mathcal{G}_c \rightarrow \mathcal{T}_F \rightarrow 1$$

of a fixed group \mathcal{T}_F by complex, semisimple, simply connected algebraic groups \mathcal{D}_c . The group \mathcal{T}_F is an extension

$$1 \rightarrow \mathcal{S}_F \rightarrow \mathcal{T}_F \rightarrow \Gamma_F \rightarrow 1$$

of Γ_F by a complex, proalgebraic torus \mathcal{S}_F . (See [S1, chapter II], [L5, §5], [S2, §7].) The contribution to \mathcal{G}_F of any \mathcal{G}_c is required to match the contribution to L_F of a corresponding group L_c , in which K_c is a compact real form of \mathcal{D}_c . This construction of \mathcal{G}_F , along with that of L_F at the end of §5, comes with a continuous homomorphism

$$L_F \rightarrow \mathcal{G}_F ,$$

up to conjugation by \mathcal{G}_F^0 . However, it would not solve any of the problems. One would still have to show that \mathcal{G}_F was indeed isomorphic to the motivic Galois group of Grothendieck.

7. RECENT PROGRESS

We shall conclude with a very brief description of four recent advances on the problem of functoriality, and the related questions discussed in §5 and §6. These results have come in the past couple of years. For the most part, they were quite unexpected.

(a) Suppose that F_p is a finite extension of \mathbb{Q}_p , and that $G = \mathrm{GL}(n)$. This is the setting for Langlands' nonabelian generalization of local class field theory. There are now two separate constructions of the conjectured local correspondence $r_p \longleftrightarrow \pi_p$. One is by M. Harris and R. Taylor [HT], and uses global methods. It generalizes results by Langlands, Deligne and Carayol in the case $n = 2$. The other, by G. Henniart [He], is of a more local character, but still uses global information. It generalizes results by Kutzko for $n = 2$. In each case, the authors characterize the correspondence by showing that r_p and π_p satisfy certain natural conditions of compatibility. (See [Ro] and [C].)

(b) Suppose that F is a finite extension of the global function field $\mathbb{F}_p(t)$, and that $G = \mathrm{GL}(n)$. It is in this setting that L. Lafforgue [Laf] has established a form of the global correspondence $r \longleftrightarrow c$. The groups L_F and \mathcal{G}_F are of course not known to exist. However, in the function field case, one can replace either L_F or \mathcal{G}_F by the Galois group Γ_F , provided that one takes ℓ -adic representations $\lambda: \Gamma_F \rightarrow \mathrm{GL}(n, \mathbb{Q}_\ell)$ (for some prime number $\ell \neq p$) instead of complex representations $r: L_F \rightarrow \mathrm{GL}(n, \mathbb{C})$, or proalgebraic representations $m: \mathcal{G}_F \rightarrow \mathrm{GL}(n, \mathbb{C})$. Generalizing results of Drinfeld for $n = 2$, Lafforgue establishes a bijection $\lambda \longleftrightarrow c$ between the set of equivalence classes of continuous ℓ -adic representations λ and a set of automorphic families of conjugacy classes c whose complement is easy to characterize. The results include a bijection $m \longleftrightarrow c$ between motives of rank n over F and this set of automorphic families. They also yield Langlands' non-abelian generalization of global class field theory for F , namely, the function field version of the special case of functoriality that applies to complex n -dimensional representations of Γ_F .

(c) Suppose that F is a number field, that G' is the split form of the special orthogonal group $\mathrm{SO}(2n+1)$, and that $G = \mathrm{GL}(2n)$. Then \widehat{G}' is the symplectic group $\mathrm{Sp}(2n, \mathbb{C})$, which comes with a standard representation

$$\rho: \widehat{G}' = \mathrm{Sp}(2n, \mathbb{C}) \hookrightarrow \mathrm{GL}(2n, \mathbb{C}) = \widehat{G}.$$

J. Cogdell, H. Kim, I. Piatetskii-Shapiro and F. Shahidi [CKPS] have established the functoriality conjecture in this situation for generic representations π' . Generic representations are expected to comprise a significant subset of the families of all cuspidal automorphic representations of $\mathrm{SO}(2n+1)$. The proof is based on the converse theorems of Cogdell and Piatetskii-Shapiro, which incidentally also play a role in the results (b) of Lafforgue.

(d) Suppose that F is a number field, that $G' = \mathrm{GL}(2) \times \mathrm{GL}(3)$ and that $G = \mathrm{GL}(6)$. H. Kim and F. Shahidi [KS] have established functoriality for the tensor product representation

$$\rho: \widehat{G}' = \mathrm{GL}(2, \mathbb{C}) \times \mathrm{GL}(3, \mathbb{C}) \rightarrow \mathrm{GL}(6, \mathbb{C}) = \widehat{G}.$$

They have also used this result to establish a second new case of functoriality, in which $G' = \mathrm{GL}(2)$, $G = \mathrm{GL}(4)$, and

$$\rho: \widehat{G}' = \mathrm{GL}(2, \mathbb{C}) \rightarrow \mathrm{GL}(4, \mathbb{C}) = \widehat{G}$$

is irreducible. In addition, the arguments have recently been extended by Kim [K] to the case of 5-dimensional representations of $\mathrm{GL}(2)$. These results lead to a significant improvement of estimates relating to a classical spectral theory problem, as well as providing new reciprocity laws between automorphic representations. They yield new bounds for the eigenvalues of both the Laplacian and associated Hecke operators for arithmetic quotients of the upper half plane.

POSTSCRIPT

The results we have mentioned, in §6 and earlier, represent a good sample of what is known. When measured against the general statement of functoriality, they must seem rather fragmentary. Indeed, they are. However, they have been hard won. Each has been the end product of much thought, sometimes by a succession of mathematicians over a period of years. Moreover, future progress is as likely to build on existing results as it is to subsume them in more general methods. New ideas will certainly be needed. But to judge from the past, a deciding factor will be the mathematical power to bring ideas to fruition. In this light, the principle of functoriality seems all the more remarkable for its ultimate simplicity. Its secrets will challenge the imagination and energy of mathematicians for a long time to come.

REFERENCES

- [AC] J. Arthur and L. Clozel, *Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula*, Ann. of Math. Stud. **120**, Princeton Univ. Press, Princeton, 1989. MR **90m**:22041
- [BCDT] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over \mathbb{Q} : Wild 3-adic exercises*, J.A.M.S. **14** (2001), 843–939. MR **2002d**:11058
- [BDST] K. Buzzard, M. Dickinson, N. Shepard-Barron, and R. Taylor, *On icosahedral Artin representations*, Duke Math. J. **109** (2001), 283–318.

- [C] H. Carayol, *Preuve de la conjecture de Langlands locale pour GL_n : Travaux de Harris-Taylor et Henniart*, Seminaire Bourbaki, 1998-99, Exposé 857, Astérisque No. 266 (2000), 191–243. MR **2001i**:11136
- [CKPS] J. Cogdell, H. Kim, I. Piatetski-Shapiro, and F. Shahidi, *On lifting from classical groups to GL_N* , Publ. Math. Inst. Hautes Études Sci. **93** (2001), 5–30. MR **2002i**:11048
- [H] H. Hasse, *History of class field theory*, in *Algebraic Number Theory*, edited by J. Cassels and A. Fröhlich, Academic Press, 1967, pp. 266–279. MR **36**:1417
- [HT] M. Harris and R. Taylor, *On the geometry and cohomology of some simple Shimura varieties*, Ann. of Math. Stud. **151**, Princeton University Press, Princeton, 2001.
- [He] G. Henniart, *Une preuve simple des conjectures de Langlands de $GL(n)$ sur un corps p -adique*, Invent. Math. **139** (2000), 439–455. MR **2001e**:11052
- [K] H. Kim, *Functoriality for the exterior square of GL_4 and symmetric fourth power of GL_2* , preprint.
- [KS] H. Kim and F. Shahidi, *Functorial products for $GL_2 \times GL_3$ and symmetric cube for GL_2* , Ann. of Math. **155** (2002), 837–893.
- [L1] R. Langlands, *Problems in the theory of automorphic forms*, Lecture Notes in Math. **170**, Springer-Verlag, 1970, pp. 18–61. MR **46**:1758
- [L2] ———, *On the classification of irreducible representations of real algebraic groups*, in *Representation Theory and Harmonic Analysis on Semisimple Lie Groups*, AMS Mathematical Surveys and Monographs, **31**, 1989, pp. 101–170. MR **91e**:22017
- [L3] ———, *Base Change for $GL(2)$* , Ann. of Math. Stud. **96**, Princeton Univ. Press, Princeton, 1980. MR **82a**:10032
- [L4] ———, *On the notion of an automorphic representation*, in *Automorphic Forms, Representations and L-functions*, Proc. Sympos. Pure Math., Vol. 33, part 1, Amer. Math. Soc., 1979, pp. 203–207. MR **81m**:10055
- [L5] ———, *Automorphic representations, Shimura varieties and motives*, in *Automorphic Forms, Representations and L-functions*, Proc. Sympos. Pure Math., Vol. 33, part 2, Amer. Math. Soc., 1979, pp. 205–246. MR **83f**:12010
- [Laf] L. Lafforgue, *Chtoucas de Drinfeld et correspondance de Langlands*, Invent. Math. **147** (2002), 1–241.
- [Ra] D. Ramakrishnan, *Modularity of solvable Artin representations of $GO(4)$ -type*, Int. Math. Research Notices, no. 1 (2002), 1–54.
- [Ro] J. Rogawski, *The nonabelian reciprocity law for local fields*, Notices Amer. Math. Soc. **47** (2000), 35–41. MR **2001a**:11192
- [S1] J.-P. Serre, *Abelian ℓ -adic Representations and Elliptic Curves*, Benjamin, New York, 1968. MR **41**:8422
- [S2] ———, *Propriétés conjecturales des groupes de Galois motiviques et des représentations ℓ -adiques*, in *Motives*, Proc. Sympos. Pure Math., Vol. 55, part 1, Amer. Math. Soc., 1994, pp. 377–400. MR **95m**:11059
- [T] J. Tunnell, *Artin’s conjecture for representations of octahedral type*, Bull. Amer. Math. Soc. **5** (1981), 173–175. MR **82j**:12015

The references in the article have been updated, but the original text remains unchanged otherwise.

Given the general nature of this article, I have tried to keep the list of formal references to a minimum. Elementary introductions to various aspects of the Langlands program can be found in [Ro] and the following two articles:

J. Arthur, *Harmonic analysis and group representations*, Notices Amer. Math. Soc. **47** (2000), 26–34.

S. Gelbart, *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. **10** (1984), 177–219. MR **85e**:11094

Other introductory articles on representation theory and automorphic forms are contained in the proceedings of the Edinburgh instructional conference:

Representation Theory and Automorphic Forms (T. N. Bailey and A. W. Knap, eds.), Proc. Sympos. Pure Math., **61**, Amer. Math. Soc., 1996. MR **98i**:22001.

For more advanced articles that are still of a general nature, the reader can turn to the Corvallis proceedings of [L4] and [L5], and the Washington proceedings of [S2].

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO, M5S 3G3, CANADA
E-mail address: `arthur@math.toronto.edu`