Eisenstein Series, the Trace Formula, and the Modern Theory of Automorphic Forms*

1. Eisenstein Series and Automorphic L-functions

The modern theory of automorphic forms is a response to many different impulses and influences, above all the work of Hecke, but also class-field theory and the study of quadratic forms, the theory of representations of reductive groups, and of complex multiplication, but so far many of the most powerful techniques are the issue, direct or indirect, of the introduction by Maass and then Selberg of spectral theory into the subject.

The spectral theory has two aspects: (i) the spectral decomposition of the spaces $L^2(\Gamma\backslash G)$ by means of Eisenstein series; (ii) the trace formula, which can be viewed as a striking extension of the Frobenius reciprocity law to pairs $(\Gamma, G), G$ a continuous group and Γ a discrete subgroup.

The attempt to discover a class of Euler products attached to automorphic forms that would include the Dirichlet series with Grössencharakter attached to real quadratic fields led Maass in 1946, under difficult circumstances in the chaos of immediate postwar Germany, to the study of eigenfunctions of the Laplacian

$$\Delta = y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)$$

on the upper half-plane, eigenfunctions that transform simply, or are even invariant, under discrete groups, Γ, especially subgroups of the modular group [27]. Apparently he was influenced to some extent by the work of Fueter on quaternionic function theory.

It is simplest to consider functions actually invariant under the discrete group. If it is Fuchsian and the fundamental domain is not compact, then Δ has a continuous spectrum, and the corresponding eigenfunctions are given by analytic continuation of functions defined by infinite series, the Eisenstein series. They are attached to cusps. If, for example, the cusp is at infinity so that the group

$$\Gamma_\infty = \left\{\gamma = \begin{pmatrix} 1 & x \\ 0 & 1\end{pmatrix} \mid \gamma\in\Gamma\right\}$$

* Appeared in Number Theory, Trace Formulas and Discrete Groups, Academic Press (1989)

is infinite, then the attached series is

$$\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} y^{(s+1)/2} = \sum \left(\frac{y}{|cz+d|^2}\right)^{(s+1)/2}. \tag{1.1}$$

This series is easily seen to converge for $\mathrm{Re}\ s > 1$ and to yield eigenfunctions of Δ, but not the ones needed for the spectral decomposition, for they correspond to parameters satisfying $\mathrm{Re}\ s = 0$. Thus an analytic continuation is required. It is often carried out in two steps as in [22], the functions being first continued to the region $\mathrm{Re}\ s > 0$, either with the help of the resolvent or the Green's function, the method used by Roelcke, or by a truncation process as employed by Selberg. The continuation across $\mathrm{Re}\ s = 0$ was first effected by Selberg by means of a further truncation and the reflection principle.

To recapitulate briefly, Maass's work on Euler products and automorphic forms drew his attention to a problem in spectral theory that, in turn, led to a problem in analytic continuation. The purpose of the first part of this lecture is to recall how a much larger class of Euler products, the *automorphic L-functions*, one of the central notions of the modern theory of automorphic forms, arose, a little by accident, from the solution of the problem in analytic continuation.

Recall first that in the fifties there was a tremendous surge of interest in automorphic forms on groups of higher dimension – due largely, I suppose, to the papers of Siegel on orthogonal groups and the symplectic groups; so it is hardly surprising that Selberg and others attempted to extend to them his techniques and ideas, the analytic continuation of the Eisenstein series and the trace formula. Decisive progress, however, had to await the introduction by Gelfand in 1962 [15] of the general notion of cusp forms that lies at the center of the spectral theory in higher dimensions.

Although it is not necessary, it is extremely convenient, if only to avoid elaborate notational complications, to work with adelic groups. In addition, the spectral decomposition is then made with respect to the largest possible family of operators, including the Hecke operators and the differential operators. It entails working on the group rather than the symmetric space, but then the obvious symmetries are recognized and the same problem not solved repeatedly.

For example, if $G = GL(n)$ then an automorphic form is a function ϕ on $G(\mathbf{Q}\backslash G(\mathbf{A})$. It is a cusp form if, for every block decomposition of the $n\times n$ matrices and every $g\in G(\mathbf{A})$, we have

$$\int \phi\left(\begin{pmatrix} I & X \\ 0 & I\end{pmatrix} g\right) dx = 0.$$

Here X is an $n_1\times n_2$ matrix, $n_1+n_2=n$, with adelic entries.

The notion of a cusp form clearly isolated, the analytic continuation of the general Eisenstein series is effected in three steps.

A. Series in One Variable Attached to Cusp Forms. If, for example, G is $SL(n)$ and $n = n_1 + n_2$, then such a series is associated to a cusp form ϕ on

$$M(A) = \left\{ m = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \in G(A) \right\}$$

and to a parameter s. Here m_i is an $n_i \times n_i$ matrix. For simplicity, I am confining myself to functions invariant on the right under a maximal compact subgroup K.

If

$$N = \left\{ \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \right\}$$

$$P = \left\{ \begin{pmatrix} A & X \\ 0 & B \end{pmatrix} \right\}$$

then

$$P = MN = NM.$$

Set

$$F_s(nmk) = \phi(m) \left(\frac{|\det\, m_1|^{n_2}}{|\det\, m_2|^{n_1}} \right)^{s+1/2}. \tag{1.2}$$

Then the Eisenstein series is

$$E_s(g) = \sum_{\gamma \in P(\mathbf{Q})\backslash G(\mathbf{Q})} F_s(\gamma g). \tag{1.3}$$

Provided that ϕ is a cusp form, the methods used for the Eisenstein series on the upper half-plane will deal with those that like (1.3) are associated to maximal parabolic subgroups P and thus involve only a single parameter, although the extension should perhaps not be thought of as entirely routine.

B. Series in Several Variables Associated to Cusp Forms. They are attached to non maximal parabolic subgroups, thus for $SL(n)$, to partitions $n = n_1 + n_2 + \cdots + n_r$ with more than two elements. The group M is given by

$$M = \{ \begin{pmatrix} m_1 & & & \\ & m_2 & & \\ & & \ddots & \\ & & & m_r \end{pmatrix} \mid \prod \det\, m_i = 1 \},$$

and P and N are defined accordingly. The functions F_s and the series E_s now depend on several parameters $s = (s_1, \dots, s_r)$, $\sum n_i s_i = 0$.

$$F_s(nmk) = \phi(m) \prod_{i=1}^{r} |\det\, m_i|^{s_i + \rho_i} \tag{1.4}$$

where the ρ_i are real numbers chosen to simplify the formulas for the functional equations. To deal with these Eisenstein series, one combines the results from A with forms of Hartog's lemma. I observe that Hartog's lemma was introduced into the subject quite early, and by several mathematicians independently (cf. Appendix I of [22]).

C. The General Eisenstein Series. Apart from a more complicated dependence on $k \in K$, these are defined by functions like (1.4), with ϕ being any square-integrable automorphic form on $M(\mathbf{A})$. That ϕ is no longer necessarily a cusp form entails altogether new difficulties. Even for classical groups of low dimension, the analytic continuation of these series and the spectral decomposition involve quite different ideas than those that suffice for A or B [22].

One need not look far for fatuous and misleading comments on the techniques involved in the three steps. They are regrettable, but fortunately need not concern us here, for it is the series of step A, or rather the calculation of their constant term, that led to the introduction of the general *automorphic L-functions* and the *principle of functoriality*. Thus these ideas could have appeared before the general theory of Eisenstein series and independently of it, but the psychological inhibitions may have been too great.

A perhaps too simple formulation of the principle of functoriality is that certain natural operations on the L-series attached to automorphic forms reflect possible operations on the forms themselves. The usual Hecke theory, for example, associates to an automorphic form an Euler product

$$\prod_{\rho \notin S} \frac{1}{(1 - \alpha_p p^{-s})(1 - \beta_p p^{-s})}. \tag{1.5}$$

Since it is only the unordered pair $\{\alpha_p, \beta_p\}$ that matters, we may think of it as a conjugacy class $\{t_p\}$ of complex 2×2 matrices with these eigenvalues. A natural operation on elements of $GL(2)$ is to let them act on symmetric tensors of a given degree n. This transforms conjugacy classes in $GL(2)$ into conjugacy classes in $GL(n+1)$. So we pass from

$$\{t_p\} = \left\{ \begin{pmatrix} \alpha_p & 0 \\ 0 & \beta_p \end{pmatrix} \right\}$$

to

$$\left\{\begin{pmatrix} \alpha_p^n & & & & \\ & \alpha_p^{n-1}\beta_p & & & \\ & & \ddots & & \\ & & & \alpha_p\beta_p^{n-1} & \\ & & & & \beta_p^n \end{pmatrix}\right\}$$

This leads us from the Euler product (1.5) to

$$\prod_{p\notin S} \frac{1}{(1-\alpha_p^n p^{-s})(1-\alpha_p^{n-1}\beta_p p^{-s})(1-\beta_p^n p^{-s})},$$

and the principle of functoriality predicts that there is an automorphic form on $GL(n+1)$ to which this series is attached.

Before indicating how the calculation of the constant term of Eisenstein series suggested the introduction of automorphic L-functions, I recall, for it is easy to forget, that 20 years ago it was by no means clear how, or even whether, the Hecke theory could be extended to groups other than $GL(2)$. Ideas of varying quality were proposed, and it is surprising that this calculation, carried out more to pass the time than with any precise aim, should yield not only specific series whose analytic continuation and functional equation could be proved, but also a class of series with a natural completeness. Repeated efforts to rework the Hecke theory had led nowhere, except perhaps to a clearer understanding of how it functioned, which could only later be turned to profit.

There is a classical paradigm for the calculation, that giving the constant term of the series (1.1) or, more precisely, since we have passed to the adèle group, of

$$E_s(g) = \sum_{P(\mathbf{Q})\backslash G(\mathbf{Q})} F_s(\gamma g),$$

with

$$F_s(g) = \left|\frac{a}{b}\right|^{s+1/2}, \;\; g = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} k.$$

Here P is the group of upper-triangular matrices in $G = GL(2)$, and K a maximal compact subgroup of $G(\mathbf{A})$.

If N is the subgroup

$$N = \left\{\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\right\}$$

of P, then the constant term of E_s is the function

$$\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} E_s(ng)dn = \int_{\mathbf{Q}\backslash\mathbf{A}} E_s\left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g\right) dx.$$

To calculate it, we substitute the series expansion for E_s, and combine terms appropriately to obtain

$$\sum_{\gamma\in P(\mathbf{Q})\backslash G(\mathbf{Q})/N(\mathbf{Q})} \int_{P(\mathbf{Q})\cap\gamma^{-1}N(\mathbf{Q})\gamma\backslash N(\mathbf{A})} F_s(\gamma ng)dn.$$

The double coset space $P(\mathbf{Q})\backslash G(\mathbf{Q})/N(\mathbf{Q})$ appearing here has a simiple structure, for it consists of only two elements with representatives

$$\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1, \quad \gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \omega.$$

This is special case of the Bruhat decomposition that will appear later. For $\gamma = 1$ the integral is simply $F_s(g)$, because

$$\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} dn = 1.$$

For $\gamma = \omega$ the integral is a product

$$\prod_v \int_{N(\mathbf{Q}_v)} F_s(\omega n_v g_v)dn_v, \tag{1.6}$$

because

$$P(\mathbf{Q}) \cap \omega^{-1}N(\mathbf{Q})\omega = 1.$$

The integrals appearing in (1.6) can be calculated place by place. Take g to be 1, so that each g_v is 1. For a nonarchimedean place, we calculate $F_s(wn_v)$ easily. First of all,

$$\omega n_v = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & x \end{pmatrix}.$$

If x is integral this matrix belongs to K and

$$F_s(\omega n_v) = 1.$$

Otherwise

$$\begin{pmatrix} 0 & 1 \\ 1 & x \end{pmatrix} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x^{-1} & 0 \\ 0 & x \end{pmatrix}\begin{pmatrix} * & * \\ x^{-1} & 1 \end{pmatrix}, \quad |x| > 1.$$

Thus

$$\int_{N(\mathbf{Q}_v)} F(\omega n_v)dn_v = 1 + \left(1 - \frac{1}{p}\right)\sum_{n=1}^{\infty}\frac{1}{p^{ns}} = \frac{1 - \frac{1}{p^{s+1}}}{1 - \frac{1}{p^s}}.$$

Taking the product over the finite places, we obtain

$$\frac{\zeta(s)}{\zeta(s+1)}.$$

The infinite place yields the usual supplementary Γ-factor.

In general one carries out the calculation in a similar fashion and, at the end, a result obtained, one looks for a transparent way to express it. It is at this stage that the automorphic L-functions suggest themselves. What are the ingredients of the calculation?

i) The Eisenstein series associated to a cusp form ϕ on a Levi subgroup M of a maximal parabolic subgroup of G. Because ϕ is a cusp form the constant term will be expressed in terms of cusp forms for the same group M. (I observe, in passing, that square-integrable forms are not closed in the same way. This is one of the reasons that C is more difficult than A and B.) The constant term is itself a sum of one or two terms, depending upon the nature of the Bruhat decomposition. (I observe, again in passing, that for the classical groups the concepts of a parabolic subgroup or the Bruhat decomposition are elementary notions of linear algebra that could profitably be included in the education of all pure mathematicians.)

ii) Since we are working adelically, the calculation is ultimately local, and we are able to draw on our experience with local harmonic analysis, especially Harish-Chandra's theory of spherical functions on real groups, and the explicit formulas that Bhanu-Murty and Gindikin-Karpelevich contributed to it.

The calculation itself was carried out for a large number of illustrative cases in [21]. One begins with P and writes an arbitrary element g of $G(\mathbf{A})$ as $g = nmk, n \in N(\mathbf{A}), m \in M(\mathbf{A}), k \in K$. The function F_s has the form

$$F_s(g) = \chi_s(m)\phi(m)\psi(k),$$

where χ_s is a character of $M(\mathbf{A})$ that depends on the complex parameter s. The Eisenstein series is

$$E_s(g) = \sum_{P(\mathbf{Q})\backslash G(\mathbf{Q})} F_s(\gamma g),$$

with constant term

$$\int_{N'(\mathbf{Q})\backslash N'(\mathbf{A})} E_s(n'g)dn'. \tag{1.7}$$

The group N' belongs to a parabolic subgroup P', that may or may not be P itself.

Since the Eisenstein series can be continued to the whole complex plane as a meromorphic function of s, so can (1.7). The Bruhat decomposition allows us to write (1.7) as a sum of one or two terms. If there are two, one is simply $F_s(g)$, which is entire, so that the remaining term, the one in which we are interested, is as well behaved as (1.7) itself. It is

$$\int_{N'(\mathbf{A})} F_s(\omega n'g)dn', \tag{1.8}$$

with a suitable ω.

To calculate (1.8) as a product it is necessary to be somewhat careful in the choice of ϕ and ψ and, in addition, to recall that for each g the function $\phi_g : m \to \phi(mg)$ belongs to a space V of functions on $M(\mathbf{Q})\backslash M(\mathbf{A})$ transforming according to an irreducible representation σ of $M(\mathbf{A})$. Thus ϕ may be regarded as a function ϕ_g on $G(\mathbf{A})$ with values in V and, more precisely, as an element of the space of the induced representation

$$\mathrm{Ind}_P^G \sigma_s,$$

σ_s being the tensor product of σ with χ_s. It is extended to $P(\mathbf{A})$ by making it trivial on $N(\mathbf{A})$.

It turns out – as a result of formal considerations – that the operation (1.8) depends only on the class of σ and not on its realization on a subspace of $L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$. The only information we need from the realization, but this is of course decisive, is that (1.8) can be analytically continued. To calculate (1.8) as a product, we work abstractly, realizing σ as a product $\otimes\sigma_v$, and at the places where there is no ramification we can take a simple model for σ_v to perform the calculations.

The calculations reduce finally to simple summations like those for $GL(2)$, but this requires some understanding of the unramified representations of the local groups $G(\mathbf{Q}_p)$ or, what amounts to the same thing, of the structure of the local Hecke algebras. This is an elementary but somewhat elaborate topic.

Almost everywhere the reductive group G with which we began is split, or at worst quasisplit, and split over an unramified extension. It is only such groups that have representations that we can call unramified. Their unramified representations can be parametrized, and an elegant form of the parametrization is forced upon us by the need to express the results of the calculation in transparent form. The parameter attached to an unramified representation is a semisimple conjugacy class $\{t(\pi_p)\}$ in a *complex* reductive group ${}^LG$. (For $G = GL(n)$ the group ${}^LG$ is again $GL(n)$, as was implicit in our earlier remarks; for other groups the relation between G and ${}^LG$ is less than direct.)

At almost all places the local factors of the constant term can be expressed in terms of the classes $\{t(\pi_p)\}$. There are finite-dimensional complex analytic representations $r_1, \dots, r_n$ of ${}^LM$ and constants $a_1, \dots, a_n$ such that the local factor is

$$\prod_{i=1}^{n} \frac{\det(I - r_i(t(\pi_p))p^{-a_i(s+1)})}{\det(I - r_i(t(\pi_p))p^{-a_i s})}.$$

This suggests the introduction, for any complex analytic representation r of any ${}^LG$, of the Euler product

$$L_S(s, \pi, r) = \prod_{p \notin S} \frac{1}{\det(I - r(t(\pi_p))p^{-s})}. \tag{1.9}$$

Here S is a finite set of places that a finer treatment would remove. Thus the constant term (1.8) is in essence a product

$$\prod_{i=1}^{n} \frac{L_S(a_i s, \pi, r_i)}{L_S(a_i s + 1, \pi, r_i)}.$$

The analytic continuation of the series (1.9) can then be obtained in sufficiently many cases to justify their further study by choosing some group M, which becomes the group of primary interest, and then searching for a goup G that contains a parabolic subgroup of which it is a Levi factor. A large number of examples were given in [21].

2. The Structure of Trace Formulas and their Comparison

The L-functions attached by Hecke to modular forms in the upper half-plane are of great arithmetical importance, and so are the general automorphic L-functions. Of course, as usual for objects attached to reductive groups, the more familiar the groups the more important the object, so that the functions attached to the general linear group or the symplectic group will appear more frequently and in more critical circumstances than the others. That does not make them any easier to treat, and methods must be found to establish in general the basic analytic properties: analytic continuation to the whole complex plane and the functional equation.

There are methods available that, in addition, have led to substantial progress with outstanding arithmetical problems and have suggested new concepts and theorems in the study of harmonic analysis or theta series. They are only partially explored, and further development promises a deeper understanding of the theory of automorphic forms, and not a few surprises. However, they have limits of which we are becoming ever more keenly aware and that temper our pleasure at the success of those who have pursued them – but not our admiration. However, not every stone has been turned.

The methods fall into three classes: a more profound exploration of the expansion of Eisenstein series at the cusps, looking beyond the constant term; the multitude of zeta integrals described by Gelbart-Shahidi [14] that include, in particular, the Rankin-Selberg technique and that sometimes involve theta series; and the trace formula.

The general problem of analytic continuation of automorphic L-function leads (with a little imagination) quickly to the circle of questions referred to by the convenient catch phrase *principle of functoriality*, which implies the possibility of transporting automorphic representations from one group to another. The first two methods effect the analytic continuation directly, but the use of the trace formulas proceeds through the principle of functoriality.

In contrast to its initial purpose, which was apparently to analyze the spectrum of the Laplace-Beltrami operator on the quotient of the upper half-plane by a Fuchsian group, the arithmetical applications of the trace formula usually involve a comparison of two or more trace formulas or of a trace formula with a Lefschetz formula, and this leads to a multitude of questions in arithmetic and harmonic analysis, appealing in themselves, in whose solution the trace formula, in turn, can often be put to good effect.

One such problem, which belongs to the elements of the general theory of the trace formula and of Eisenstein series, but which viewed historically is a deep arithmetical statement, the result of successive efforts by several of the best mathematicians, is the conjecture of Weil that the Tamagawa number of a simply-connected semisimple group over $\mathbf{Q}$ is 1.

Recall that if G is a semisimple group over $\mathbf{Q}$, and $G(\mathbf{A})$ the group of adelic-valued points on G, so that $G(\mathbf{Q})$ is a discrete subgroup of $G(\mathbf{A})$, then the Tamagawa number is the volume of the quotient $G(\mathbf{Q})\backslash G(\mathbf{A})$ with respect to a canonically defined measure. An explicit volume for

$$\tau(G) = \mathrm{vol}(G(\mathbf{Q})\backslash G(\mathbf{A}))$$

is of immediate appeal when $G(\mathbf{R})$ is compact (for example, if it is the orthogonal group of a definite form), for it yields directly a class-number formula. This is the case to which Siegel confined himself in his Oslo lecture of 50 years ago, when he discussed his extensions of classical formulas of Eisenstein and Minkowski [31]. He also dealt with groups for which $G(\mathbf{R})$ is not compact and was concerned, as should be stressed, although they are not pertinent here, with more general formulas than those for volumes of fundamental domains.

Tamagawa's realization much later that Siegel's formula for volumes had a strikingly simple adelic form led Weil to his general conjecture [32], which was verified in many cases, but by no means all. The problem was to find a general, uniform method. That we now have, and as part of the general trace formula, although for reasons that will be explained later factors of type E_8 that are not quasisplit will have to be excluded.

Connected reductive algebraic groups over **Q** are broken up into families defined cohomologically, a family consisting of all groups that can be obtained one from the other by an inner twist. For example, all special orthogonal groups attached to forms of the same dimension and with discriminants differing by a square lie in the same family, for if one of the groups, G, is defined by a symmetric matrix X and the other, G', by X' then there is a matrix U with coefficients from $\bar{\mathbf{Q}}$ and determinant from **Q** such that $X' = UX^tU$. Then

$$\psi : A \to UAU^{-1}$$

defines an isomorphism from G to G' over $\bar{\mathbf{Q}}$ and the isomorphisms $\psi_\sigma = \psi^{-1}\sigma(\psi), \sigma \in Ga\ell(\bar{\mathbf{Q}}/\mathbf{Q})$ are given by

$$A \to V_\sigma A V_\sigma^{-1}, \quad V_\sigma = U^{-1}\sigma(U),$$

and

$$\det\, V_\sigma = (\det\, U)^{-1}\sigma(\det\, U) = 1,$$

so that $\sigma \to V_\sigma$ defines a Galois cocycle $\{\psi^{-1}\sigma(\psi)\}$ with values in the adjoint group of G. Thus ψ defines G' as an inner twisting of G.

Each inner family contains a distinguished element, determined up to isomorphism, and this is the quasisplit element of the group. It is the group for which $G(\mathbf{Q})\backslash G(\mathbf{A})$ is most noncompact and for a family containing orthogonal groups is characterized as the orthogonal group associated to a form with isotropic subspace of largest dimension. Abstractly it is characterized by the property that it contains a Borel subgroup defined over $\mathbf{Q}$. Since the Levi factors of a Borel subgroup are tori, the Eisenstein series associated with it have functional equations expressed in terms of L-functions attached to $Grössencharakter$ and are thus particularly easy to handle. On the other hand, a Borel subgroup over $\mathbf{Q}$ is necessarily a minimal parabolic over $\mathbf{Q}$, and the constant function is always a residue of Eisenstein series associated with the minimal parabolic. Combining these two facts with the elements of spectral theory for self-adjoint operators, one readily calculates $\tau(G) = \mathrm{vol}(G(\mathbf{Q}))\backslash G(\mathbf{A})$ [20], [19]. It is found to be 1 for simply connected semisimple groups.

The second, more difficult step is to show that if G^* is quasisplit and G an inner form of it then $\tau(G) = \tau(G^*)$. It has been carried out by Kottwitz [17], who compares the trace formula for G, in which $\tau(G)$ appears, with that for G^*, in which $\tau(G^*)$ appears. The L-groups of G and G^* are the same, and one of the simpler manifestations of the principle of functoriality will be an assertion that the L-functions associated with G are all found among the L-functions associated with G^*. Thus for each automorphic π of G there will be an automorphic representation π^* of G^* such that

$$L(s,\pi,r) = L(S,\pi^*,r)$$

for all representation of ${}^LG$. One expects to prove this by a comparison of trace formulas for G and G^*, and methods are being developed for this purpose. What Kottwitz has in effect done is to anticipate their complete development by a judicious choice of the functions to be substituted in the formulas, so that a great many difficult terms are removed but not the volumes of the fundamental domains.

The proof of Weil's conjecture by Kottwitz, and of cyclic base-change by Arthur-Clozel [12], apart from any other applications, contemplated or already accomplished, are of sufficient importance to justify a description of the gross structural features of Arthur's general trace formula and of the scheme for comparison, although this entails some overlap with [11]. Moreover, a clear warning is necessary that the relative trace formulas of Jacquet have not yet been fitted into this scheme, that a great deal remains to be done, and that many of the techniques have been most highly developed for comparisons that involve disconnected groups, the so-called twisted trace formulas, and these have been omitted from this discussion.

In contrast to the trace formula of Selberg, which was, as its name implies, a formula for a trace, but which could only be proven for groups of rank one, the formula developed by Arthur begins with an equality between two functions that are then integrated separately over $G(\mathbf{Q})\backslash G^1(\mathbf{A})$ and the integrals calculated in completely different manners. The resultant identity is the first form of the trace formula [1], [2]. (For semisimple groups $G^1(\mathbf{A}) = G(\mathbf{A})$; for reductive groups it is the kernel of a family of homomorphisms from $G(\mathbf{A})$ to $\mathbf{R}^+$.)

Recall that the spectral analysis of the action of $G(\mathbf{A})$ by right translation on $\mathfrak{H} = L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$ is best accomplished with the operators

$$K(f) : \phi \to \int_{G(\mathbf{A})} \phi(gh)f(h)dh,$$

where f is a smooth function of compact support on $G(\mathbf{A})$, and $K(f)$ is an integral operator with kernel

$$K(g,h) = \sum_{G(\mathbf{Q})} f(g^{-1}\gamma h).$$

In addition, there is available a direct sum decomposition of $\mathfrak{H}$ defined by the theory of Eisenstein series. It is parametrized by pairs (M,ρ), M being a Levi factor over $\mathbf{Q}$ of a parabolic subgroup of G defined over $\mathbf{Q}$ and ρ a cuspidal representation of $M(\mathbf{A})$. The pairs are subject to a suitable equivalence relation. Denote the set of parameters by $\mathfrak{X}$. The summand $\mathfrak{H}_\chi$ labelled by $\chi \in \mathfrak{X}$ has itself a direct-integral structure, but for the moment only the coarse decomposition matters.

The basic equality, or identity, is between two truncations of the kernel of K. The first, a geometrically defined truncation of the restriction of K to the diagonal, is integrated over $G(\mathbf{Q}\backslash G^1(\mathbf{A}))$ to yield the coarse $\mathcal{O}$-expansion

$$J^T_{\mathrm{geom}}(f) = \sum_{\mathcal{O}} J^T_{\mathcal{O}}(f).$$

The sum runs over conjugacy classes of semisimple elements in $G(\mathbf{Q})$, and T is a multidimensional parameter determining the location of the truncation. The sum converges absolutely, and each term is a polynomial in T. The distributions $J^T_{\mathcal{O}}$ have at least two disadvantages. They are not invariant under conjugation because the truncation demands a choice of a maximal compact subgroup of $G(\mathbf{A})$, and they are not expressible as integrals over a single conjugacy class, or even a single type of conjugacy class. For example, the term corresponding to $\mathcal{O} = \{1\}$ involves all unipotent classes, and the term

$$\mathrm{vol}(G(\mathbf{Q})\backslash G^1(\mathbf{A}))f(1), \tag{2.1}$$

that must be isolated if the trace formula is to be applied to Weil's conjecture, does not appear explicitly.

The second truncation is defined by first composing K with a truncation operator ΛT on functions on $G(\mathbf{Q})\backslash G(\mathbf{A})$ that converts slowly increasing functions into rapidly decreasing functions and that is an idempotent, and then restricting the kernel of the composition $K \circ \Lambda^T$ to the diagonal. The operator $K \circ \Lambda^T$ is again an integral operator, and its kernel can be expressed as an integral of truncated Eisenstein series. Integrating over the diagonal we obtain a sum over χ, referred to as the course χ-expansion,

$$J^T_{\mathrm{spec}}(f) = \sum_{\chi} J^T_{\chi}(f).$$

Once again the distributions J^T_χ are not invariants. Moreover, although the J^T_χ themselves are polynomials in T, they are calculated as an integral over the parameters (partly continuous, partly discrete) of the direct integral decomposition of $\mathfrak{H}_\chi$ of inner products of truncated Eisenstein series, and these are not polynomials.

Thus the first form of the trace formula

$$\sum_{\mathcal{O}} J^T_{\mathcal{O}}(f) = \sum_{\chi} J^T_{\chi}(f)$$

is not too useful. What Arthur does next is to transform the left and right sides of the formula into forms that, although not final, at least appear more applicable. The appearance of the variable T on these formulas is of no significance since the differences

$$J^{T'}_{\mathrm{geom}}(f) - J^T_{\mathrm{geom}}(f)$$

and

$$J^{T'}_{\mathrm{spec}}(f) = J^T_{\mathrm{spec}}(f)$$

can be expressed in terms of the geometrical and spectral sides of the trace formula for the Levi factors of proper parabolic subgroups of G. For example, if G is anisotropic, the polynomials are in fact of degree 0 and the dependence on T is fictitious. Arthur himself prefers to work with a carefully chosen and fixed truncation parameter that he suppresses from the notation, and it is best to follow him so that the coarse trace formula becomes

$$\sum J_{\mathcal{O}}(f) = \sum J_{\chi}(f).$$

Both transformations are difficult, but the geometric side is perhaps easier. Here the transformation introduces some disagreeable new features, but it appears that there is no choice but to accommodate ourselves to them. They are aesthetically displeasing but do not obstruct the arguments. We have to choose some compact neighborhood Δ of the identity in $G(\mathbf{A})$, to which are then attached finite sets S of places, and S appears explicitly in the new form of the geometric side, the fine $\mathcal{O}$-expansion, which is then only valid for functions that depend only on the coordinates in S and have support in Δ, being outside of S the product of characteristic functions of the standard maximal compact subgroups.

The fine $\mathcal{O}$-expansion [7] is a sum over the Levi subgroups containing a fixed Levi subgroup of some fixed parabolic subgroup over $\mathbf{Q}$.

$$J_{\text{geom}}(f) = \sum_M \frac{|\Omega_0^M|}{|\Omega_0^G|} J_M(f),$$

and $J_M(f)$ is a sum over conjugacy classes in M of terms

$$a^M(s,\gamma) J_M(\gamma, f).$$

The two groups, Ω_0^M, Ω_0^G, whose order appears in the formula are Weyl groups, and $a^M(S,\gamma)$ is a constant that is not determined explicitly, and for many purposes need not be, except for certain γ. It implicitly involves a measure and, for some groups of low dimension and presumably in general, values of apparently unmanageable Dirichlet series. The notion of conjugacy employed is somewhat unusual. For the semisimple part of γ it is conjugacy in $M(\mathbf{Q})$; for the unipotent part it is conjugacy in $\prod_{v \in S} M(\mathbf{Q}_v) = M(\mathbf{Q}_S)$. The function f is a product of a function f_S on $\prod_{v\in S} G(\mathbf{Q}_v) = G(\mathbf{Q}_S)$ with a standard characteristic function outside of S; and

$$J_M(\gamma, f) = J_M(\gamma, f_S)$$

is a weighted orbital integral over

$$\{g^{-1}\gamma n g \mid g \in G(\mathbf{Q}_S), n \in N(\mathbf{Q}_S)\}$$

if $P = MN$ is a parabolic subgroup with Levi factor M. Thus it is a finite sum of weighted orbital integrals.

The distributions J_M are still not all invariant, but the larger M is the more invariant they are. In particular, J_G is invariant, and a sum of invariant orbital integrals, because

$$J_G(\gamma, f_S) = \int_{G_\gamma(\mathbf{Q}_S)\backslash G(\mathbf{Q}_S)} f_S(g^{-1}\gamma g)\, dg,$$

where G_γ is the connected centralizer of γ in G. If $\gamma = 1$ this is $f_S(1) = f(1)$. Since

$$a^G(S,1) = \tau(G),$$

the fine $\mathcal{O}$-expansion contains a large invariant contribution in which the term (2.1) appears explicitly. I observe in passing that $a^G(S,\gamma)$ is 0 if the semisimple part of γ is not elliptic.

The distributions $J_M(\gamma), M \neq G$, are not so simply expressed, and in advanced applications of the trace formula [12] they must be treated without the help of explicit formulas, to which in earlier, simpler applications [23] recourse could be had. It is best, nonetheless, to work them out in simple cases in order to develop a feel for them.

For $G = SL(2)$ the only pertinent M is the group of diagonal matrices. If

$$g = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} k,$$

with $k \in K_S, x = (x_v) \in \mathbf{Q}_S$, set

$$v_M(g) = -\sum_{v \in S} \lambda(x_v)$$

with

$$\lambda(x_v) = \begin{cases} \frac{1}{2}\ln(1 + x_v^2), & v \text{ infinite}, \\ \ln(\max\{1, |x|\}), & v \text{ finite}. \end{cases}$$

If $\gamma \in M(F)$ has distinct eigenvalues a, b then $J_M(\gamma, f_S)$ is equal to

$$\left(\prod_{v \in S} \left|\frac{a}{b}\right|_v^{1/2} \left|1 - \frac{b}{a}\right|_v\right) \int_{M(\mathbf{Q}_S)\backslash G(\mathbf{Q}_S)} f_S(g^{-1}\gamma g) v_M(g) dg.$$

A change of variables in the integral turns this into the product of

$$\prod_{v \in S} \left|\frac{a}{b}\right|_v^{1/2}$$

and

$$-\sum_{v \in S} \int_{K_S} \int_{Q_S} f_S(k^{-1}\gamma n(x) k) \lambda((1 - b/a)^{-1} x_v) dx\, dk, \tag{2.2}$$

where

$$n(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Since $a^M(S,\gamma) = 1$ for such γ, the singular behavior of this integral as γ approaches ± 1 is brought into the trace formula. If $h_S : \gamma \to J_m(\gamma, f_S)$ were smooth on $M(\mathbf{Q}_S)$, the contribution $J_M(f)$ to the fine

$\mathcal{O}$-expansion would, apart from whatever the terms associated to $\gamma = \pm 1$ yield, be the fine $\mathcal{O}$-expansion for a function on M obtained by multiplying h_S with the product of the characteristic functions of the maximal compact subgroups outside of S. This would facilitate comparisons enormously, but h_S is not smooth. Fortunately, there is a way to circumvent the difficulties this causes, Arthur's principle of the *cancellation of singularities*, to which we shall come later.

The expression (2.2) is a sum of two other expressions,

$$\sum_{v \in S} \ln \Big|1 - \frac{b}{a}\Big|_v \int_{K_S} \int_{\mathbf{Q}_S} f_S(k^{-1}\gamma n(x)k)dx\ dk, \tag{2.3}$$

all of whose singular behavior at ± 1 arises from the factor in front of the integral, and a second expression, which although also singular at ± 1, does have limiting values at these points. They are

$$-\sum_{v \in S} \int_{K_S} \int_{\mathbf{Q}_S} f(k^{-1}\gamma n(x)k) \ln |x|_v dx\ dk, \quad \gamma = \pm 1. \tag{2.4}$$

The integral appearing in (2.3) has the limit

$$\int_{K_S} \int_{S} f_S(k^{-1}\gamma n(x)h)dx\ dk, \quad \gamma \pm 1, \tag{2.5}$$

which is the orbital integral of (2.4) without the weight, and is, in fact, the sum of the integrals over the unipotent conjugacy classes of $G(\mathbf{Q}_S)$. The terms $J_M(\gamma, f_S), \gamma \pm 1$, are linear combinations of (2.4) and (2.5). The coefficient of (2.5) can be chosen freely, for the coefficients $a^G(S, \gamma n(x)), x \in \mathbf{Q}$, can then be modified accordingly. Arthur introduces a procedure for arriving at a definite choice that seems to be as good as any other, but it is by no means sacred.

Some arbitrariness notwithstanding, the fine $\mathcal{O}$-expansion is a definite advance over the coarse, but is still neither invariant nor stable and thus not yet useful for comparisons. Before describing how Arthur achieves an invariant form, we must consider the fine χ-expansion. The conversion of the coarse χ-expansion to the fine requires a Paley-Wiener theorem for reductive groups [4] and a remarkably astute and skillful use of ordinary Fourier analysis [5], [6].

In the fine χ-expansion the inner products of truncated Eisenstein series disappear and the factors describing the functional equations appear. These are matrix-valued transcendental functions, but they factor as a product of a scalar-valued transcendental function, expressible in terms of automorphic L-functions as in Section I, and a matrix-valued rational function (in arguments s and p^s).

Taking residues of the Eisenstein series associated to the parameter χ can yield square-integrable eigenfunctions on $G(\mathbf{Q})\backslash G^1(\mathbf{A})$ or on intermediate Levi subgroups. This leads to a fine direct-integral decomposition of the space $\mathfrak{H}_\chi$.

$$\mathfrak{H}_\chi = \int_{\Pi(G,\chi)} d\pi,$$

where $\prod(G,\chi)$ is a well determined set of unitary representations. There are analogous sets $\prod(M,\chi)$ for all Levi subgroups, that are empty if M does not contain the Levi factor in any pair defining χ.

The fine χ-expression is a double sum [6]

$$J_{\mathrm{spec}}(f) = \sum_\chi \left\{ \sum_M \frac{|\Omega_0^M|}{|\Omega_0^G|} \int_{\Pi(M,\chi)} a^M(\pi) J_M(\pi, f) d\pi \right\}$$

that is not yet known to be absolutely convergent in general. The functions $a^M(\pi)$ contain a local contribution that is defined by intertwining operators and is quite subtle and a global contribution that is expressible in terms of logarithmic derivatives of automorphic L-functions.

Such derivatives result from formal operations with collections of functions called (G, M)-families by Arthur [3], operations that pervade his work, appearing already in the proof of the fine $\mathcal{O}$-expansion. They can be described briefly.

If M is a Levi factor of a parabolic, let $\mathcal{P}(J)$ be the finite set of parabolic subgroups containing M. (The case to think of is the group M of diagonal matrices in $G = GL(n)$, then $\mathcal{P}(M)$ is parametrized by the permutations of $\{1, \dots, n\}$.) Suppose that for each $P \in \mathcal{P}(M)$ we are given a smooth function $c_P(\lambda)$ on $\mathfrak{a}_M^*$ (the coordinate space $\mathbf{R}^n$ in the example). The collection $\{c_P\}$ is said to form a (G, M)-family if whenever P and P' are adjacent (the corresponding permutations σ, σ' differ by right multiplication by the interchange of adjacent integers $i, i+1$), then c_P and $c_{P'}$ agree on the hyperplane passing through the wall common to the chambers associated to P and P' (the hyperplane $\lambda_{\sigma(i)} = \lambda_{\sigma(i+1)}$). To each $P \in \mathcal{P}(M)$ there is an associated set of simple roots Δ_P and their coroots. A simple but important lemma states that the formula

$$c_M(\lambda) = \sum_{P \in \mathcal{P}(M)} c_P(\lambda) \theta_P(\lambda)^{-1}$$

with

$$\theta_P(\lambda) = \prod_{\alpha \in \Delta_P} \lambda(\check{\alpha})$$

defines a smooth function on $\mathfrak{a}_M^*$.

If, for example, we take $G = GL(2)$ and M as above, then $\mathcal{P}(M)$ consists of two elements P and P', and if we take $X_P, X_{P'}$ in $\mathfrak{a}_M$ with

$$X_P - X_{P'} = (x, -x),$$

then

$$c_P(\lambda) = e^{\lambda(X_P)}, \quad c'_P(\lambda) = e^{\lambda(X_{P'})}$$

is a (G, M)-family and

$$c_M(0) = x$$

is essentially the length of the integral joining X_P and $X_{P'}$. The analogous construction in higher dimensions leads to the volumes of convex sets that are ubiquitous in Arthur's papers.

A second construction of a (G, M)-family starts from a set, $\{r_\alpha | \alpha \text{ a root}\}$, of nonzero complex-valued functions on $\mathbf{R}$. If Σ_P is the set of roots in P, and $\bar{P}$ the parabolic in $\mathcal{P}(M)$ opposite to P so that $P \cap \bar{P} = M, \Sigma_P \cap \Sigma_{\bar{P}} = \emptyset$, set

$$r_{P|P'}(\lambda) = \prod_{\alpha \in \Sigma_P \cap \Sigma_{P'}} r_\alpha(\lambda(\check{\alpha})). \tag{2.6}$$

Then for each $P' \in \mathcal{P}(M), \nu \in \mathfrak{a}_M^*$,

$$r_P(\lambda; v, P') = r_{P|P'}(\nu)^{-1} r_{P|P'}(\nu + \lambda)$$

defines a (G, M)-family, and $r_M(0; \nu, P')$ can be expressed in terms of logarithmic derivatives.

The product $d_P = c_P d_P$ of two (G, M)-families is again a (G, M)-family, and there are formulas for calculating d_M that can be regarded as variants either of Leibniz's rule or of partitions of convex sets.

It is principally these formulas that lead to the produce $a^M(\pi) J_M(\pi, f)$ appearing in the fine χ-expansion. The factors that appear in the functional equation of the Eisenstein series are intertwining operators $J_{P'|P}(\pi_\lambda)$ for which formulas like (2.6) are available. They now involve operator-valued functions, but the formalism of (G, M)-families is still applicable and a factorization

$$J_{P'|P}(\pi_\lambda) = r_{P'|P}(\pi_\lambda) R_{P'|P}(\pi_\lambda)$$

in which $r_{P'|P}(\pi_\lambda)$ is a complex-valued function given by automorphic L-functions and $R_{P'|P}(\pi_\lambda)$ is an intertwining operator given by elementary functions, yields the factorization $a^M(\pi)J_M(\pi, f)$, in which all lack of invariance is in $J_M(\pi, f)$. (The notation is unfortunate. $a^M(\pi)$ is defined by $r_{P'|P}(\pi_\lambda)$ and $J_M(\pi, f)$ with the help of $R_{P'|P}(\pi_\lambda)$.)

Arthur's final step in his development of the trace formula is to render it invariant by a brutal transposition of terms from the spectral side to the geometric side, a process that has little to recommend it but its successes, but these are overwhelming, for with his invariant formula, cancellation of singularities becomes a very supple tool and it becomes possible to substitute functions for which many disagreeable terms vanish while the essential ones remain.

The source of all lack of invariance in the trace formula lies in the initial truncation; so the lack of invariance, the difference between the values of the distribution on f and on f^x, where $f^x(g) = f(x^{-1}gx)$, or between its value on the two convolutions

$$h * f : g \to \int h(gg_1^{-1})f(g_1)dg_1$$

and $f * h$, can be expressed in a universal way for all the distributions $J_M(\gamma)$ or $J_M(\pi)$ occurring in the trace formula.

Denote any one of these distributions by J_M. Then the simplest form of the universal formula is

$$J_M(f^x) = \sum_{Q \supseteq M} J_M^{M_Q}(f_{Q,x}). \tag{2.7}$$

The sum is over all parabolic subgroups over $\mathbf{Q}$ containing M; $J_M^{M_Q}$ is the analogue of J_M appearing in the trace formula for the Levi factor M_Q of Q; and $f_{Q,x}$ is a function on $M_Q(\mathbf{A})$ given in terms of f by a simple integral formula of the form ([3])

$$f_{Q,x}(m) = \delta_Q(m)^{1/2} \int_K \int_{N_Q(\mathbf{A})} f(k^{-1}mnk)u'_Q(k,x)dk\, dn.$$

In particular,

$$f_{G,x} = f.$$

Strictly speaking, the trace formula may not apply to f^x, because it may not be an element of the Hecke algebra, so that (2.7) has to be rewritten in terms of convolutions to obtain a correct formula.

When working with the trace formula, one fixes a finite set of places containing the infinite place, and all freedom is in f_S so that the achievement of invariance is really a local problem on

$$\prod_{v \in S} G(\mathbf{Q}_v) = G(Q_S).$$

Keeping this in mind, we now simplify the notation, often replacing f_S by f.

To attain invariance, Arthur introduces a map $f \to \phi_M(f)$ from functions on $G(Q_S)$ to functions on $M(Q_S)$, or rather from functions on $G(Q_S)$ to a class of functions on $M(Q_S)$, for only the orbital integrals of $\phi_M(f)$ on regular semisimple elements are specified. (In view of theorems of Harish-Chandra and Kazhdan, this is equivalent to specifying the values on $\phi_M(f)$ of distributions supported on characters in the sense of [8], Section 1.) The function $\phi_M(f)$ is defined by the equation

$$\mathrm{tr}\ \pi(\phi_M(f)) = J_M(\pi, f),$$

π running over the irreducible tempered representations of $M(Q_S)$, thus exactly those necessary for the local harmonic analysis. Then all distributions J_M, both those from the spectral and those from the geometrical side of the trace formula, are converted into invariant distributions I_M by the inductive formula

$$J_M(f) = \sum_{L \supseteq M} I_M^L(\phi_L(f)),$$

the sum running over all Levi factors over **Q** containing M.

Observe that to carry out the inductive definition implicit in this formula it must be shown that the distributions I_M^L are supported on characters. Notice also that the construction simplifies distributions from the spectral side. For example, if π is tempered then

$$I_M(\pi, f) = \begin{cases} 0 & M \neq G, \\ \mathrm{tr}\ \pi(f) & M = G. \end{cases}$$

However, it will tend to complicate distributions from the geometric side. Of course, for $M = G$ the procedure effects no change, and

$$I_G(\gamma) = J_G(\gamma), \quad I_G(\pi) = J_G(\pi).$$

Since the distributions $I_M(\gamma), I_M(\pi)$ are obtained from $J_M(\gamma), J_M(\pi)$ by a uniform procedure, they can be substituted for $J_M(\gamma), J_M(\pi)$ in the trace formula to obtain an *invariant trace formula.*

Moreover, the distributions appearing in the invariant trace formula being supported on characters, it makes sense to ask whether they simplify, or even vanish, on functions whose orbital integrals are subject to suitable conditions. This is indeed so and yields simpler terms of the trace formula that are very effective for specific purposes [13, 18].

It has been observed that arithmetical applications of the trace formula usually involve comparisons. In essence one shows that some linear combination of the geometric sides is zero, infers that the same linear combination of the spectral sides is zero, and then from this deduces relations between automorphic representations of the groups involved.

Since the very purpose and natue of the trace formula is to allow a term-by-term analysis of the geometric side, the comparison can proceed only if there is some correspondence between conjugacy classes in the groups involved. This is simplest for base change for $G(n)$ under cyclic extensions of the ground field, for then the norm map is a well-defined function from the conjugacy classes of one group to those of another. For this application the concomitant problems in local harmonic analysis (transfer of orbital integrals and fundamental lemmas) are consequently easier, and the methods for global comparison, especially cancellation of singularities, much more advanced [12].

For other problems the correspondence is not given by a function, and the notions of stabilization and endoscopy necessary to circumvent the attendant difficulties are perhaps the most startling that the trace formula has suggested to harmonic analysis, although not yet nearly so well understood as some others, such as Harish-Chandra's Selberg principle, Arthur's Paley-Wiener theorem, or formulas for characters as weighted orbital integrals.

Suppose G^* is quasisplit and $\psi : G \to G^*$ is an inner twist so that $\psi^{-1}\sigma(\psi)$ is an inner automorphism of G for all $\sigma \in Ga\ell(\bar{\mathbf{Q}} \mid \mathbf{Q})$. Thus if γ is an element of $G(\mathbf{Q}), \psi(\gamma)$ is an element of $G^*(\bar{\mathbf{Q}})$. It is a theorem of Steinberg and Kottwitz that the conjugacy class of $\psi(\gamma)$ in $G^*(\bar{\mathbf{Q}})$ always meets $G^*(\mathbf{Q})$, and this allows us to define a correspondance between conjugacy classes in $G(\mathbf{Q})$ and $G^*(\mathbf{Q})$; but to obtain a function one has to introduce the notion of $stable$ conjugacy classes in $G(\mathbf{Q})$, which is for simply-connected, semisimple groups just conjugacy in $G(\bar{\mathbf{Q}})$, but which is slightly more delicate for other groups [16]. The notion is also defined over local fields. Thus for a term-by-term comparison of the trace formulas of G and G^*, one needs a trace formula in which the geometric side is expressed as a sum over stable conjugacy classes and in which the distributions are stably invariant. Primitive forms of such a stable trace formula are available in general. In a very few special cases it is completely developed [25], [29].

It is the primitive form that leads to the proof of Weil's conjecture, but in order to illustrate cancellation of singularities we proceed with more general considerations. Stabilization and stably-invariant harmonic analysis lead to the introduction of endoscopic groups of G. These are quasisplit reductive groups H such that the L-group ${}^LH$ is imbedded in ${}^LG$. Among them is the group G^* with ${}^LG^* = {}^LG$. There is a map from regular semisimple conjugacy classes of $G(\mathbf{Q})$ to stable semisimple conjugacy classes of $H(\mathbf{Q})$ and a notion of transfer $f \to f^H$ from functions on $G(A)$ to functions on $H(A)$ [26]. The function f^H is not well defined; only its stably invariant orbital integrals are. In addition to each elliptic endoscopic group there is attached [17] a cohomological invariant $\imath(G,H)$.

If

$$I_M^{\mathrm{geom}}(f) = \sum_\gamma I_M(\gamma, f)$$

and if $I_M^{\mathrm{spec}}(f)$ is defined in a similar fashion, then the invariant trace formula is an equality

$$\sum_M I_M^{\mathrm{geom}}(f) = \sum_M I_M^{\mathrm{spec}}(f).$$

Recall that the sum is over the Levi factors containing a fixed minimal one. The stably-invariant trace formula will be a similar identity, but it is best to take the sum over the Levi factors M^* of G^* containing a fixed minimal one, noting that to every M there is associated a unique M^*. Thus the identity will take the form

$$\sum_{M^*} SI_{M^*}^{\mathrm{geom}}(f) = \sum_{M^*} SI_{M^*}^{\mathrm{spec}}(f), \tag{2.8}$$

in which all distributions appearing are stably invariant so that their value on a function f is determined by the stable orbital integrals of f on semisimple elements.

One can associate to each Levi factor M_H of H a Levi factor M^* of G^*. We write $M_H \to M^*$. When f^H can be defined for all H, we define $SI_{M^*}^{\mathrm{geom}}$

$$SI_{M^*}^{\mathrm{geom}} = I_{M^*}^{\mathrm{geom}}(f) - {\sum_M}' \imath(G,H) \sum_{M_H \to M^*} SI_{M_H}(f^H) \tag{2.9}$$

and, in a similar fashion, $SI_{M^*}^{\mathrm{spec}}(f)$. The prime indicates that the sum is over all elliptic endoscopic groups, with the exception of G^*. Such a definition at least guarantees the validity of (2.8).

It is the distribution $SI^{\mathrm{spec}}_{G^*}$ that carries the interesting information about automorphic forms, and one would like to show that

$$SI^{\mathrm{spec}}_{G^*}(f) = SI^{\mathrm{spec}}_{G^*}(f^*), \quad f^* = f^{G^*}, \tag{2.10}$$

from which it follows, in particular, that $SI^{\mathrm{spec}}_{G^*}$ is stable, for f^* is specified by the stable orbital integrals of f. Observe that SI_{G^*} denotes two distributions, one on G and one on G^*, distinguished only by their arguments.

As a first application of these ideas, one can, following Kottwitz, prove for G simply-connected and semisimple the existence of pairs f, f^* such that $f^* = f^{G^*}$ and such that both f^H and f^{*H} can be taken to be zero for $H \neq G^*$. In addition

$$I^{\mathrm{geom}}_{M^*}(f) = I^{\mathrm{spec}}_{M^*}(f) = 0, \quad M^* \neq G^*,$$
$$I^{\mathrm{geom}}_{M^*}(f^*) = I^{\mathrm{spec}}_{M^*}(f^*) = 0, \quad M^* \neq G^*.$$

For this pair, we thus obtain the identity

$$SI^{\mathrm{geom}}_{G^*}(f) - SI^{\mathrm{geom}}_{G^*}(f^*) = SI^{\mathrm{spec}}_{G^*}(f) - SI^{\mathrm{spec}}_{G^*}(f^*). \tag{2.11}$$

The right side is an infinite linear combination of irreducible traces. That f^H and f^{*H} can be taken to be zero for $H \neq G^*$ implies strong relations among the orbital integrals, and if Hasse's principle is valid for G as well as Weil's conjecture for groups of lower dimension, then the left side of (2.11) reduces to

$$\tau(G)f(1) - \tau(G^*)f^*(1) = (\tau(G) - \tau(G^*)).$$

This can be equal to a sum of traces only if it is zero, so Weil's conjecture follows inductively.

Hasse's principle intervenes because it is, more generally, necessary to the calculations that allow one to show that when f^H exists for all H then the contribution of the regular semisimple conjugacy classes to the right side of (2.9) is stable. One might hope that if the Hasse principle is so strongly enmeshed in these calculations it could be possible to use the trace formula to prove it. So far as I know, this has not been attempted.

Further general development of the stable trace formula awaits the proof of the existence, both locally and globally, of the transferred functions f^H. One can hope that once this is done the methods developed by Arthur will overcome the other obstacles.

The results of Arthur-Clozel suggest that not only will (2.10) be valid but, more generally

$$SI_{M^*}^{\mathrm{geom}}(f) = SI_{M^*}^{\mathrm{geom}}(f^*) \tag{2.12}$$

$$SI_{M^*}^{\mathrm{spec}}(f) = SI_{M^*}^{\mathrm{spec}}(f^*) \tag{2.13}$$

and that, moreover, it will be possible to express the two sides of (2.12) and (2.13) as sums of more primitive stably invariant distributions. On the geometric side will appear, in particular, stably invariant orbital integrals and on the spectral side, among others, characters of tempered L-packets. The identities (2.12) and (2.13), of which (2.10) is the most interesting, realize a final form of the trace formula, the stably invariant trace formula, that can presumably be applied without further transformation.

A proof along the lines of [12] would of course involve an elaborate induction, upwards on the dimension of G and downwards on that of M. The definition of f^* and the manipulation of [17] and [24] then show that the difference

$$SI_{G^*}^{\mathrm{geom}}(f) - SI_{G^*}^{\mathrm{geom}}(f^*) \tag{2.14}$$

is the sum of very few terms, those corresponding to unipotent elements. Thus if it could be shown that it was a discrete sum of characters, then a little harmonic analysis on one of the noncompact factors $G(\mathbf{Q}_v)$ should show that it is zero.

According to (2.8) the difference (2.14) can be expressed in terms of the differences

$$SI_{M^*}^{\mathrm{spec}}(f) - SI_{M^*}^{\mathrm{spec}}(f^*) \tag{2.15}$$

and

$$SI_{M^*}^{\mathrm{geom}}(f) - SI_{M^*}^{\mathrm{geom}}(f^*), \quad M^* \neq G^*. \tag{2.16}$$

Now $I_G(\gamma, f)$ is equal to $J_G(\gamma, f)$ by its very definition. However, for $M \neq G$ there is an inexplicit element introduced in the passage from $J_M(\gamma, f)$ to $I_M(\gamma, f)$, namely, $\phi_L(f)$, so that the only handle on the distributions $I_M(\gamma)$ is their formal properties of splitting and descent.

In the case of base change, which one will take as a model for the development of the stable trace formula, one uses these properties and a progressively less restrictive choice of functions f and f^* to show that all differences (2.14), (2.15), and (2.16) are zero.

One begins with a class of functions for which (2.14) and all but one of (2.16), that corresponding to the M^* at which we have arrived by the downward induction, are zero. More precisely, it is analogues of these differences that appear, but rather than continuously qualifying my remarks with references to base change, I prefer to speak of the deduction of the stably invariant trace formula as being already achieved, with no intention of slighting the difficulties that are still to be overcome. Now $SI_{M^*}^{\mathrm{geom}}(f)$ or $SI_{M^*}^{\mathrm{geom}}(f^*)$ when expanded as a sum over stable conjugacy class of M^* presumably resemble very closely the expansion of the stable trace formula for some function on $M^*(\mathbf{A})$, but because of singularities like those we have seen in (2.2) cannot be equal to such an expansion. However, one can expect that the singularities cancel when the difference is taken, so that there is a function ϵ on $M^*(\mathbf{A})$ to which the trace formula can be applied such that

$$SI_{M^*}^{\mathrm{geom}}(f) - SI_{M^*}^{\mathrm{geom}}(f^*) = SI_{M^*}^{\mathrm{geom}}(\epsilon).$$

Finally, one uses (2.8) for M^* rather than G and for ϵ rather than f to replace $SI_{M^*}^{\mathrm{geom}}(\epsilon)$ by a spectral-theoretic expansion, probably just $SI_{M^*}^{\mathrm{spec}}(\epsilon)$, the other terms very likely vanishing. This gives

$$\sum_{M^*} SI_{M^*}^{\mathrm{spec}}(f) - \sum_{M^*} SI_{M^*}^{\mathrm{spec}}(f^*) = SI_{M^*}^{\mathrm{spec}}(\epsilon).$$

This spectral equality involves measures of Lebesque type of various dimensions, and they must vanish separately. In particular, the term corresponding to the measure of lowest dimension, the split rank of the center of G, is just

$$SI_{G^*}^{\mathrm{geom}}(f) - SI_{G^*}^{\mathrm{spec}}(f^*);$$

so this difference must be zero.

Once this difference is shown to be zero for the restricted class of functions, it can be shown that it is zero for an even larger class, for which one deduces the vanishing of (2.14) and (2.16). To deal with completely general f and f^* it is necessary to work one's way down through the M^* to the minimal Levi factor, verifying at each stage that the vanishing of (2.16) for that M^* and the restricted class of functions implies its vanishing in general. The vanishing of (2.14), in general, is only obtained at the last stage, at which one also has the vanishing of (2.15) for $M^* = G^*$ and general f and f^*. For $M^* \neq G^*$, the vanishing of (2.15) is proved with the help of local results linking the distributions $I_{M^*}(\gamma)$ and $I_{M^*}(\pi)$ and a supplementary induction.

All of this looks to be elaborate and extremely difficult, as indeed it is; so it is very helpful to understand it for simple examples, where the convoluted inductive arguments and much of the technique are unnecessary and the constructions more explicit. The group $U(3)$ of [29] is perhaps the best choice, but even $G = SL(2)$ yields considerable insight.

Here $G = G^*$, and the only Levi factor besides G itself that need be considered is the group M of diagonal matrices. It is Abelian so that $\Phi = \Phi_M(f)$ can be defined more directly than usual as the Fourier transform of

$$\pi \to J_M(\pi, f),$$

π a character of $M(\mathbf{Q}_S)$. However, this is not too direct, and a real understanding of Φ_M requires an examination of the normalized intertwining operators $R_{P'|P}(\pi_\lambda)$, but that would entail an elementary but extended digression.

Since

$$I_M(\gamma, f) = J_M(\gamma, f) - \Phi(\gamma),$$

the singularities of $I_M(f)$ are those of $J_M(\gamma, f)$. Since $G = G^*$ we may write M rather than M^* in the definitions of the stable trace formula. Moreover, the elliptic endoscopic groups other than G itself are all anisotropic tori, so that (2.9) reduces to

$$SI_M^{\mathrm{geom}}(f) = I_M^{\mathrm{geom}}(f)$$

and

$$SI_M^{\mathrm{geom}}(f) = \sum_{\gamma \in M(\mathbf{Q})} I_M(\gamma, f).$$

Since

$$\gamma \to I_M(\gamma, f)$$

is not smooth, the right side is not the trace of a smooth function on $M(\mathbf{A})$, but the difference

$$\sum_{\gamma \in M(\mathbf{Q})} I_M(\gamma, f) - \sum_{\gamma \in M(\mathbf{Q})} I_M(\gamma, f^*)$$

is, because

$$\gamma \to I_M(\gamma, f) - I_M(\gamma, f^*)$$

is smooth. Of course this is trivial if we take $f^* = f$, but the point is that we need not so so.

The first sign that it is smooth is that the difference of (2.3) for f and (2.3) for f^* is 0 because the integrals appearing there are stable, and thus equal, so that the difference is zero, and the singular factor in front innocuous. The difference between (2.2) and (2.3) involves the factor $\prod_{v\in S}|a/b|_v^{1/2}$, which we may ignore, and a sum over v, each term of the sum being smooth away from ± 1 but not necessarily at ± 1. For example, for v nonarchimedean the corresponding term is the sum of

$$-\int_{K_S}\int_{\mathbf{Q}_S} f_S(k^{-1}\gamma n(x)k)\ln|x|_v dx\ dk$$

and

$$-\int_{K_S}\int_{|x|_v\le|1-b/a|_v} f_S(k^{-1}\gamma n(x)k)\ln|x|_v dx\ dk.$$

The first expression is smooth, and the singularity of the second is that of the product of

$$-\int_{|x_v|\le|1-b/a|_v}(\ln\Big|\frac{1-b}{a}\Big|_v-\ln|x_v|)dx_v$$

and

$$\int_{K_S}\int_{\mathbf{Q}_S} f(k^{-1}\gamma n(x)k)dx\ dk,\quad S'=S-\{v\}. \tag{2.17}$$

Since (2.17) is a stable orbital integral, this singularity is cancelled by the corresponding one for f^*. The contributions to the singularity from the Archimedean place also cancel, but the proof is more elaborate.

References

[1] Arthur, J. *A trace formula for reductive groups I: Terms associated to classes in $G(\mathbf{Q})$*, Duke Math. Jour. **45** (1978).

[2] Arthur, J. *A trace formula for reductive groups II: Applications of a truncation operator,* Comp. Math. **40** (1980).

[3] Arthur, J. *The trace formula in invariant form*, Ann. of Math. **114** (1981).

[4] Arthur, J. *A Paley Wiener theorem for real reductive groups*, Acta Math. **150** (1983).

[5] Arthur, J. *On a family of distributions obtained from Eisenstein series I: Applications of the Paley-Wiener theorem*, Amer. Jour. Math. **104** (1982).

[6] Arthur, J. *On a family of distributions obtained from Eisenstein series II: Explicit formulas*, Amer. Jour. Math. **104** (1982).

[7] Arthur, J. *On a family of distributions obtained from orbits,* Can. Jour. Math. **37** (1986).

[8] Arthur, J. *The invariant trace formula I: Local theory,* J. Amer. Math. Soc., to appear.

[9] Arthur, J. *The invariant trace formula II: Global theory,* J. Amer. Math. Soc., to appear.

[10] Arthur, J. *Intertwining operators and residues II: Invariant distributions*, Comp. Math., to appear.

[11] Arthur, J. *The trace formula and Hecke operators*, this volume.

[12] Arthur, J. and Clozel, L. *Simple algebras, base change, and the advanced theory of the trace formula*, Ann. of Math. Studies, to appear.

[13] Bernstein, J. et al. *Représentations des groupes réductifs sur un corps local*, Hermann, Paris (1984).

[14] Gelbart, S. and Shahidi, F. *Analytic properties of automorphic L-functions*, Academic Press (1988).

[15] Gelfand, I.M. *Automorphic functions and the theory of representations.* Proc. of the ICM, Stockholm, (1962).

[16] Kottwitz, R. *Rational conjugacy classes in reductive groups*, Duke Math. Jour. **49** (1982).

[17] Kottwitz, R. *Stable trace formula: elliptic singular terms*, Math. Ann. **275** (1986).

[18] Kottwitz, R. *Tamagawa numbers*, Ann. of Math., 127 (1988).

[19] Lai, K.F. *Tamagawa number of reductive algebraic groups*, Comp. Math. **41** (1980).

[20] Langlands, R.P. *The volume of the fundamental domain for some arithmetical subgroups of Chevalley groups.* In: Algebraic Groups and Discontinuous Subgroups, Proc. Symp. Pure Math. **9** (1966).

[21] Langlands, R.P. *Euler Products.* Yale Mathematical Monographs (1971).

[22] Langlands, R.P. *On the functional equations satisfied by Eisenstein series,* Springer Lecture Notes **544** (1976).

[23] Langlands, R.P. *Base change for $GL(2)$*, Ann. of Math. Studies **96** (1980).

[24] Langlands, R.P. *Les débuts d'une formule des traces stable,* Publ. Math. de l'Univ. de Paris VII. **13** (1983).

[25] Langlands, R.P. and Labesse, J.-P. *L-indistinguishability for $SL(2)$*, Can. Jour. Math. **31** (1979).

[26] Langlands, R.P. and Shelstad, D. *On the definition of transfer factors,* Math. Ann., 278 (1987).

[27] Maass, H. *Uber eine neue Art von nichtanalytischen automorphen Formen,* Math. Ann. **121** (1949).

[28] Roelcke, W. *Analytische Fortsetzung der Eisensteinreihen zu den parabolischen Spitzen von Grenzkreisgruppen erster Art,* Math. Ann. **132** (1956).

[29] Rogawski, J. *Automorphic representations of unitary groups in three variables,* Ann. of Math. Studies, to appear.

[30] Selberg, A. *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series,* J. Indian Math. Soc. (N.S.) **20** (1956).

[31] Siegel, C.L. *Analytische Theorie der quadratischen Formen,* Com. Rend. du Cong. Inter. des Math., Oslo (1937).

[32] Weil, A. *Adeles and Algebraic Groups.* Birkhäuser (1982).