The Trace Formula and its Applications

An introduction to the work of James Arthur*

I balanced all, brought all to mind,
An Irish airman foresees his death
W. B. Yeats

In May, 1999 James Greig Arthur, University Professor at the University of Toronto was awarded the Canada Gold Medal by the National Science and Engineering Research Council. This is a high honour for a Canadian scientist, instituted in 1991 and awarded annually, but not previously to a mathematician, and the choice of Arthur, although certainly a recognition of his great merits, is also a recognition of the vigour of contemporary Canadian mathematics.

Although Arthur's name is familiar to most mathematicians, especially to those with Canadian connections, the nature of his contributions undoubtedly remains obscure to many, for they are all tied in one way or another to the grand project of developing the trace formula as an effective tool for the study, both analytic and arithmetic, of automorphic forms. From the time of its introduction by Selberg, the importance of the trace formula was generally accepted, but it was perhaps not until Wiles used results obtained with the aid of the trace formula in the demonstration of Fermat's theorem that it had any claims on the attention of mathematicians – or even number theorists – at large. Now it does, and now most mathematicians are willing to accept it in the context in which it has been developed by Arthur, namely as part of representation theory.

The relation between representation theory and automorphic forms, or more generally the theory of numbers, has been largely misunderstood by number theorists, who are often of a strongly conservative bent. Representation theory appears, in number theory and elsewhere, because symmetries create redundance, and this disorderly redundance has to be reduced to an orderly, perceptible uniqueness. Physicists and chemists, who are often bolder than mathematicians, accepted this and incorporated unhesitatingly representation theory into their analysis. In spite of Gauss, Dirichlet, Dedekind and Frobenius, number theorists balked and sought alternatives, but in the hope that such obscurantism belongs to the past, I am going to begin immediately with representations, and with the language of adèles, adding only – as a corrective – that there are circumstances, even with the trace formula, when it

* During the preparation of this introduction I was able to profit from frequent conversations with Arthur himself. I am grateful for the patience with which he responded to my many questions.

Appeared in Canad. Math. Bull., vol. XX (2001).

is useful to calculate in as concrete a manner as possible, and that then, in spite of the great conceptual ease they offer, representations and adèles have to be replaced by bases and by congruences. But these circumstances are not so frequent as many would like and do not arise at all in the work of Arthur.

Arthur, like Harish-Chandra, is inclined to work not with particular groups but with a general reductive group. The generality sometimes veils the difficult technical or conceptual achievements that inhere in various of his results. Nevertheless, any attempt to present an outline of his theory in terms of specific groups would become entangled in a welter of largely irrelevant detail, so that I have chosen to assume that the reader is familiar, at least at a formal level, with the basic concepts from the theory of algebraic groups and from other pertinent domains.

Although I have tried here to expound Arthur's papers in a coherent, logical sequence and to avoid all fatuous metaphor, I expect – and encourage – most readers to skim the report. Some passages, unfortunately far too many, are laborious, and of interest only to those who are contemplating a serious study of the domain, but others can, I hope, be read quickly and easily and will give even those with no time and no inclination to trouble themselves with details some notion of Arthur's accomplishments.[1]

Selberg first introduced the trace formula to study the asymptotic properties of the spectrum of automorphic forms for the group $\mathrm{SL}(2)$, and then went on to apply it to other analytic problems, in particular to the eponymic zeta-function and to some automorphic L-functions, but Arthur has been much more concerned with questions associated with functoriality, thus the comparison of automorphic forms or representations for different groups. Although the trace formula for groups with compact quotient is the result of a simple formal analysis, the formula for groups with noncompact quotient, and thus with continuous spectrum, is fraught with difficult analytic problems, some appearing already for groups of rank one, in particular for $\mathrm{SL}(2)$, others, of a different nature, appearing first for groups of higher rank. Although not the very first, one of the earliest signs of Arthur's mathematical powers was his derivation in [5], [7], [9] of an identity, valid for a general reductive group, that was a genuine trace formula, thus an identity with a sum over conjugacy classes on one side, and a sum over spectral terms on the other. These two sides are usually referred to as the geometric and the spectral sides of the trace formula.

1. The first general trace formula. Recall that the theory of automorphic forms, in its purely analytic formulation, refers to the study of invariant functions on $\Gamma\backslash G$, where for the present purposes – with

[1] It may be manifest from this report, but I nonetheless warn the innocent reader that my own knowledge and understanding of Arthur's papers is far less than it should be. Much of the following that is obscure would otherwise be clear. The trace formula cries out for competent exposition, as a whole or in part, and for smoother, more direct treatments of a number of issues that Arthur overcame only by circuitous means.

a notation that will trouble the purist – $\Gamma = G(\mathbb{Q})$ and $G = G(\mathbb{A})$, the group G on the right being a connected reductive algebraic group over $\mathbb{Q}$. (Unless the contrary is explicitly indicated, the symbol G standing alone without argument will denote, at least in the first part of this report, the topological group $G(\mathbb{A})$.) One could replace the field $\mathbb{Q}$ by any finite extension F of it, thus by an algebraic number field, but because of the possibility of reducing the scalar field back to $\mathbb{Q}$, this is not a more general situation. On the other hand, F could also be a function field, a possibility not covered by Arthur's theory and not yet, so far as I know, treated.

The structure of G as a topological group is, like that of $\mathbb{A}$ itself, far from simple, but rather than discuss it, I simply take various attendant notions as self-evident. If f is a (smooth) function (with compact support) on G it defines an operator

$$\phi \to \phi', \qquad \phi'(x) = \int \phi(xy)f(y)dh$$

on the functions on $\Gamma\backslash G$ with kernel

$$K(x,y) = \sum_{\gamma\in\Gamma} f(x^{-1}\gamma y) \tag{1}$$

We can think of this as an operator on $L^2(\Gamma\backslash G)$. It will not be of trace class unless $\Gamma\backslash G$ is compact, but that is seldom the case. There are two sources of noncompactness. The centre of the algebraic group G may contain a split torus or there may be a split torus in its derived group. If it were not that the principal applications of the theory were to groups with centre, it would be best to avoid the first possibility simply by assuming that the algebraic group is semisimple. I do this for now, observing that otherwise we can replace G by a subgroup defined by some conditions modulo the derived group, or we can form, rather than $L^2(\Gamma\backslash G)$, an L^2-space of functions transforming according to a character of a subgroup of the centre, or we can employ, as Arthur does to gain flexibility, some combination of the two devices.

The existence of a split torus in the derived group is a more serious matter and entails a truncation of the kernel K that is dictated by the general reduction theory developed by Siegel and completed by Borel and Harish-Chandra and that (for groups of rank greater than one) is one of the key earlier inventions of Arthur. The central result of reduction theory is that *grosso modo* the geometry of $\Gamma\backslash G$ is that of a cone in the Lie algebra $\mathfrak{a}$ over $\mathbb{R}$ of a maximal $\mathbb{Q}$-split torus A of the semisimple part of the algebraic group G. This cone is defined by the set Δ of simple roots of the torus, thus as

$$\{H \mid \alpha(H) > 0, \quad \alpha\in\Delta\}. \tag{2}$$

In addition to the set of roots, there is also the set $\hat{\Delta} = \{\varpi_\alpha\}$ of dual roots defined by

$$\alpha(\varpi_\beta) = \delta_{\alpha,\beta}.$$

It is common in the theory to fix a minimal parabolic subgroup over $\mathbb{Q}$ that contains A and to consider only parabolic subgroups P that are defined over $\mathbb{Q}$ and that contain this fixed minimal one. The subsets of Δ are then identified with these parabolic subgroups. So we denote them by Δ^P. Since P contains A, it has a distinguished Levi factor M. Let N be its unipotent radical.

It is usual to distinguish a maximal compact subgroup K of G. Then any element of G can be written as $x = nmk$ with n in the unipotent radical of a given P, m in its Levi factor and k in K. The vector space $\mathfrak{a}$ can be identified with the set of real-valued functions on Δ and those functions that vanish on Δ^P form then the Lie algebra $\mathfrak{a}_P$ over $\mathbb{R}$ of the centre of M and, in addition, can be thought of as the space of homomorphisms of the group of characters of M defined over $\mathbb{Q}$ into R. Thus if m lies in M, we can define $H_M(m) = H(m)$ in $\mathfrak{a}_P$ by

$$e^{\langle H(m),\chi\rangle} = |\chi(m)|.$$

The map

$$x \to H_P(x) = H_M(m)$$

allows us to pull functions on $\mathfrak{a}_P$ back to functions on G.

If the trace of the operator defined by f or the kernel K were defined, it would be the integral of $K(x, x)$ over a fundamental domain. In general it is not, because the spectrum has continuous components in all dimensions up to $dim(A)$. An analysis of the asymptotic behaviour of $K(x, y)$ on a fundamental domain, thus essentially on a cone as in (2), suggests a truncation of $K(x, y)$ that is defined with the help of the simple roots and their duals and the algebraic groups P. Both the kernel and its truncation can be expressed in two different ways, as a sum over classes in Γ that remain to be introduced or as a spectral sum. Integrating one or the other expression of the truncated kernel over the diagonal yields two results that must be equal. This equality or identity is the basis for all further development. So we begin by reviewing the two expressions it, the geometric side given by a sum over classes in Γ and the spectral side.

Two elements in Γ are said to be in the same class $\mathfrak{o}$ if their semisimple components are conjugate in Γ. Thus for example all unipotent elements have as semisimple component the unit element and therefore form a single class. The truncation is defined by an element T deep in the cone (2), thus with $\alpha(T) \gg 0$ for all $\alpha \in \Delta$. If P is a parabolic subgroup we set

$$K_P(x,y) = \sum_{M(\mathbb{Q})} \int_{N(\mathbb{A})} f(x^{-1}\gamma n y)dn.$$

If Δ contains a single element, thus if the dimension of A is one so that the fundamental domain is essentially a half-line, then the truncation is straightforward. We take P to be the minimal parabolic and subtract from the original kernel, or rather its restriction to the diagonal, the expression

$$\sum_{P(\mathbb{Q})\backslash G(\mathbb{Q})} K_P(\delta x, \delta x)\hat{\tau}_P(H(\delta x) - T), \tag{3}$$

in which $\hat{\tau}_P$ is the characteristic function of a half-line. In general, such expressions are formed for each P and alternately added and subtracted.

If $\mathfrak{o}$ is a class, we can introduce the partial sum

$$K_{\mathfrak{o}}(x,y) = \sum_{\gamma\in\mathfrak{o}} f(x^{-1}\gamma y)$$

as well as the truncations of the partial sums. Thus the truncation itself $k^T(x,f)$, a function on the diagonal and thus of x alone, becomes a sum over the set $\mathcal{O}$ of classes $\mathfrak{o}$,

$$k^T(x,f) = \sum_{\mathfrak{o}} k^T_{\mathfrak{o}}(x,f).$$

The integration over x can proceed term by term and yields the first form of the geometric side of the trace formula.

Thus the basic objects on the geometric side, at least in its first form, are the integrals

$$J^T_{\mathfrak{o}}(f) = \int_{\Gamma\backslash G} k^T_{\mathfrak{o}}(x,f)dx. \tag{4}$$

Expressing these integrals in a usable form was one of the main tasks undertaken by Arthur in the years following the appearance of the first papers. In some cases, however, a suggestive form lies close at hand.

Suppose that $\mathfrak{o}$ is an equivalence class that consists entirely of semisimple elements. We can choose $\gamma \in \mathfrak{o}$ and a P with Levi factor M such that γ is contained in no proper parabolic subgroup of M over

$\mathbb{Q}$. If the centralizer of γ is contained in M the class is called unramified. Then, apart from a factor given by a volume, the expression (4) is equal to a *weighted orbital integral*,

$$(5) \qquad \int_{G_\gamma(\mathbb{A})\backslash G(\mathbb{A})} f(x^{-1}\gamma x)v(x,T)dx.$$

The weights $v(x,T)$ can be made explicit, and we shall return to this in the course of our discussion, which will be lengthy because there are a large number of defects in the geometric side as it now stands and in (5). First of all, we need explicit expressions for (4) for all classes, not just the unramified ones. Secondly, even for the unramified ones, the term (5) is not yet amenable to local harmonic analysis and almost all applications of the trace formula reduce ultimately to local comparisons. Moreover, the parameter T has been introduced simply to effect a truncation and has no intrinsic significance, so that it should not appear in a final result. In addition, in order to truncate, we have introduced a fixed maximal compact subgroup of G, so that neither the geometric nor the spectral side of the trace formula as it at first appears is invariant. Thus they certainly cannot be traces. Since it is traces that are to be compared, we have to see about passing from the noninvariant form to an invariant form. Before undertaking all these modifications, we examine the spectral side.

The spectral side, as its name implies, entails a knowledge of the spectral decomposition of $L^2(\Gamma\backslash G)$, in other words of its decomposition into a direct integral of irreducible representations of G. The basic structure of the space is defined in terms of the parabolic subgroups P. In particular, the space of cusp forms is defined by the vanishing of the integrals over their unipotent radicals,

$$\int_{N(\mathbb{Q})\backslash N(\mathbb{A})} \phi(ng)dn.$$

The space of cusp forms is invariant and irreducible under the action of G and the representation on it is a discrete direct sum of irreducible representations. Many questions can be asked about the components, but that is an arithmetical problem not part of the spectral theory. The spectral theory takes the structure of the space of cusp forms for G and for the Levi factors $M = M_P$ of the parabolic subgroups P as given and constructs in terms of these unknown objects, whose deeper properties it is now the goal of the trace formula to investigate, the full decomposition of $L^2(\Gamma\backslash G)$. Although we assumed, for simplicity, that G was semisimple, this will no longer be so for the groups M. The centre of M will contain a split torus A_P whose dimension is equal to that of $\mathfrak{a}_P$. As a result, if ρ is an automorphic representation of M and λ an element in the dual of $\mathfrak{a}_P$ then

$$\rho_\lambda = \rho \otimes e^{i\lambda(H(m))}$$

is also an automorphic representation of M. According to the theory of Eisenstein series, the representations induced from the ρ_λ, ρ cuspidal, all appear in the decomposition of $L^2(\Gamma\backslash G)$, so that there is a continuous spectrum of dimension $dim\, \mathfrak{a}_P$ associated to the pair (M, ρ). There is a simple notion of equivalence on the pairs (M, ρ) with ρ cuspidal. An equivalence class will be denoted χ and the set of all classes by $\mathfrak{X}$. To each such class is associated an invariant subspace $L^2_\chi(\Gamma\backslash G)$ of $L^2(\Gamma\backslash G)$ and

$$L^2(\Gamma\backslash G) = \oplus L^2_\chi(\Gamma\backslash G).$$

The space $L^2_\chi(\Gamma\backslash G)$ may contain not only continuous spectrum of dimension $dim\, \mathfrak{a}_P$ but possibly spectra of any dimension between 0 and $dim\, \mathfrak{a}_P$. The exact nature of this spectrum is unknown, but there is a conjecture of Arthur that was suggested by his reflections on the trace formula and that is related to Ramanujan's conjecture, or rather to its failure for general groups in the first, crude form in which it was proposed. This conjecture, to which we shall return, allows a calculation of the precise spectrum of the space $L^2_\chi(\Gamma\backslash G)$ from a knowledge of certain automorphic L-functions. Some cases have been established, so that there is place for confidence in it, but a general proof will probably demand a much deeper understanding of trace formulas than we have at present. The projections on the spaces $L^2_\chi(\Gamma\backslash G)$ entail a decomposition of the kernel $K(x, y)$ as

$$K(x, y) = \sum_{\chi \in \mathfrak{X}} K_\chi(x, y),$$

and the truncation can be performed on the individual K_χ, so that

$$k^T(x, f) = \sum k^T_\chi(x, f).$$

Arthur establishes, but his proof is not direct, that this decomposition can be integrated term by term, so that

$$\sum_{\mathfrak{o} \in \mathcal{O}} J^T_{\mathfrak{o}}(f) = \sum_{\chi \in \mathfrak{X}} J^T_\chi(f), \tag{6}$$

where $\mathcal{O}$ is the collection of classes $\mathfrak{o}$ and

$$J^T_\chi(f) = \int_{\Gamma\backslash G} k^T_\chi(x, f)\, dx. \tag{7}$$

The formula (6) is the first form of the trace formula, but is useless without a better understanding of the terms (4) and (7).

Just as it will be useful and instructive to have the evaluation (5) of (4) in the simplest case, so will it be useful to have an evaluation of (7) in the simplest case, for χ unramified – in a sense that need not be defined more precisely here. Then the space $L^2_\chi(\Gamma\backslash G)$ does not contain the additional spectrum of dimension less than $dim\,\mathfrak{a}_P$. It may, and usually does, happen that there are several P and ρ such that (M,ρ) is in the class χ. Then

$$J^T_\chi(f) = \sum_P \frac{1}{n(A_P)} \int_{\Pi(M)} tr(\Omega^T_P(\pi)_\chi \cdot I_P(\pi, f)_\chi) d\pi. \tag{8}$$

The formula is from [9] but the notation is that of the paper [13]. Although I have modified it a little in an attempt to make it more transparent, it remains less than ideal. The sum is over the pertinent P and $n(A_P)$ is an integer, the number of chambers in $\mathfrak{a}_P$. The integration appears at first glance to be over the unitary dual of $M(\mathbb{A})$. If $I_P(\pi)$ is the representation of G unitarily induced from π then

$$I_P(\pi, f) = \int_G f(x) I_P(\pi, x) dx.$$

The expression $I_P(\pi, f)_\chi$ is defined similarly in terms of $I_P(\pi)_\chi$, which is 0 if (P,π) does not lie in χ and is $I_P(\pi)$ with an appropriate multiplicity if it does.

The central objects in (8) are the operators $\Omega^T_P(\pi)_\chi$. They are in essence logarithmic derivatives of intertwining operators. Because of the presence of a derivative, the expression (8) is not invariant. Thus it has the same defects as (5). Not only does it depend on the ancillary and ultimately irrelevant parameter T but also it is not invariant.

Since Arthur's bibliography is long, it may appear that we are spending far too much time with only two of the very earliest papers and that we are never going to get off the ground. Nonetheless, if we are to understand at all the nature of the objects with which Arthur had to deal, we have to look at (5) and (8) more closely.

2. The geometry of convex sets defined by roots and coroots. The sum in (6) – of either the left or the right side – turns out to be a polynomial in the variable T. The term of principal interest in this polynomial is the constant term. The higher-order terms in T can be interpreted as the constant term of a trace formula for functions on the Levi factors that are obtained as simple integrals of f. The individual terms in the sums of (6) are also polynomials in T, and the desired trace formula, in which T does not appear, is obtained by observing that the constant term of the sum is the sum of the constant terms. To be more precise, the constant term of a polynomial is its value at the origin, so that to specify it an origin has to be chosen. Arthur defines his preferred origin T_0 in [10], in which as well as in [4] some geometrical principles basic to many of his later arguments are introduced.

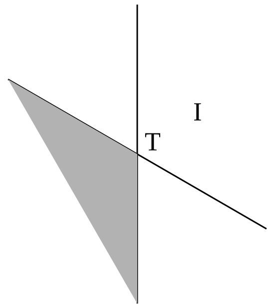

Figure 1

The truncation is particularly simple when $\mathfrak{a}$ is of dimension one. Then T varies on part of an affine line and to establish that a function $j(T)$ of T is a polynomial we need only verify that $j(T') - j(T)$ is of the form $\lambda(T' - T)$, where λ is linear. The main term, thus the kernel itself, does not survive when the difference is taken; only the truncating terms do. If $T' - T$ lies in the positive chamber, then according to (3) the difference between the two truncated kernels is

$$\sum_{P(\mathbb{Q})\backslash G(\mathbb{Q})} K_P(\delta x, \delta x)\sigma(H(\delta x)) \tag{9}$$

if σ is the characteristic function of the interval $[T, T')$.

The integral over the diagonal of (9) is

$$\Big\{ \int_{M(\mathbb{Q})\backslash M(\mathbb{A})^1} \sum_{M(\mathbb{Q})} f_M(m^{-1}\gamma m) dm \Big\}\Big\{ \int_T^{T'} dH \Big\} \tag{10}$$

if

$$f_M = \int_K \int_{N(\mathbb{A})} f(k^{-1}mnk) dn\, dk.$$

Although we have been betrayed by our assumption that the algebraic group G was semisimple, the formula (10) shows that the difference between the two truncations is a linear term in $T' - T$, namely

$$\int_T^{T'} dH,$$

in essence just the length of the interval from T' to T, times a factor that does not depend on T or T'. Since M is not semisimple and indeed contains a split torus the quotient $M(\mathbb{Q})\backslash M(\mathbb{A})$ is not of finite volume. To have a group of finite volume, we can replace $M(\mathbb{A})$ by $M(\mathbb{A})^1$, the kernel in $M(\mathbb{A})$ of the map H_M. Since the quotient is then even compact, the spectrum is discrete and the first factor is the trace of f_M.

In addition to the simple roots Δ^P of $\mathfrak{a}$ in the Levi factor of the parabolic P, there is also the set Δ_P of simple roots of $\mathfrak{a}_P$ in the algebraic group G, and more generally, whenever $P' \subset P$, sets $\Delta_{P'}^P$ of the roots of $P' \cap M$, $M = M_P$, in M. There are natural decompositions

$$\mathfrak{a}_{P'} = \mathfrak{a}_{P'}^P \oplus \mathfrak{a}_P,$$

so that we can regard the elements of the basis $\{\varpi_\alpha\}$ (of the dual of $\mathfrak{a}_{P'}^P$) that is dual to $\Delta_{P'}^P$ as functions on $\mathfrak{a}_{P'}$. Then the truncations are defined by the characteristic functions $\hat\tau_{P'}^P$ of

$$\Big\{H \mid \varpi(H) > 0 \,\forall\, \varpi_\alpha \in \Delta_{P'}^P\Big\}. \tag{11}$$

Indeed, for G itself only $P = G$ is used, and then it is dropped from the notation. The other possibilities are included only because they appear in arguments by induction, of which there are many. They will certainly not be presented here. All we want is some feeling for the geometry.

If the dimension of $\mathfrak{a}$ is two, there are four of these functions. For $P' = G$, $\hat\tau_G$ is identically one, as there are no conditions in (11). If $\mathfrak{a}_{P'}$ is of dimension one, then $\hat\tau_{P'}$ is the characteristic function of a half-plane and if P' is the minimal parabolic then $\hat\tau_{P'}$ is the characteristic function of a cone, the cone (I) of Figure 1.

Although the various terms of the truncation are defined by different parabolics and therefore different kernels, these kernels have similar asymptotic behaviour so that what in effect is happening is that terms are being subtracted that correspond to the two half-planes that intersect in I and that then a term supported on I is added, so that we are left with a term concentrated on the shaded region, a cone extending to infinity. Although itself unbounded, the shaded region defines a bounded region in the fundamental domain because its intersection with the cone (2) is bounded.

To understand what is obtained from the difference of two truncations, we translate the diagram of Figure 1 from T to T'. The principal term, that corresponding to $P' = G$, is removed. The term corresponding to an intermediate P', with $dim\,\mathfrak{a}_{P'} = 1$, remains, but has been truncated by multiplication by -1 times the characteristic function of an infinite band of finite width, whose two sides are determined by the projections of T and T' on $\mathfrak{a}_{P'}$. For the minimal P', the difference corresponds to a truncation by the full strip of Figure 2, itself the union of the three pieces, I, II and

III. The truncation by I combines with the band for one of the intermediate parabolics P' and yields an expression that is defined by a kernel for M', and that therefore can be treated inductively, the width of the band appearing as an additional factor that is linear in T and T'. The truncation by III corresponds to the second intermediate parabolic. Finally, the truncation by II yields the trace for a kernel on the Levi factor of the minimal parabolic multiplied by the area of II, which is a polynomial in T and T'.

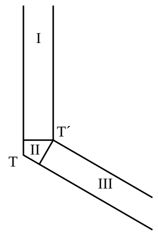

Figure 2

The argument is less transparent when $T' - T$ is not in the cone (2) but the conclusion is the same, and applies not only to the complete kernel but to the partial kernels defining the geometric and the spectral terms. Until now, we have only needed the parabolic subgroups over $\mathbb{Q}$ containing a fixed minimal one, but for many purposes it is necessary to introduce the finite set $\mathcal{P}(M)$ of those for which a given M is a Levi factor. All those M that contain the originally chosen A need be considered, so that we are enlarging the set of P to be considered and the set of M. The $\mathbb{Q}$-parabolic groups containing M can be parametrized by the chambers in $\mathfrak{a}_M$ (or in its dual a^*_M), where $\mathfrak{a}_M = \mathfrak{a}_P$ for any $P \in \mathcal{P}(M)$. A simple, but important, notion appears in (10), that of a (G, M) family. It is a collection $\{c_P(\lambda)\}$ of smooth functions on $i\mathfrak{a}^*_M$ with the property that $c_P(\lambda) = c_{P'}(\lambda)$ whenever P and P' are associated to adjacent chambers and λ lies in the hyperplane defining the common wall of these two chambers. The simplest example of such a family is a collection $\{e^{\lambda(X_P)}\}$, where $\{X_P\}$ is an A_M-orthogonal family. This means that $X_P - X_{P'}$ is orthogonal to the wall separating P and P' whenever P and P' are adjacent. If $X_P - X_{P'}$ points from the chamber associated to P to that associated to P', then the family is said to be positive. A typical example of such a family is given in Figure 3, where the points of an A_M-orthogonal family appear as well as the convex set that they span.

There is a basic principle that appears in [10]. If $\{c_P(\lambda)\}$ is a (G, M)-family, then

$$c_M(\lambda) = \sum_{P \in \mathcal{P}(M)} c_P(\lambda)\theta_P(\lambda)^{-1} \tag{12}$$

can be extended to a smooth function on $i\mathfrak{a}^*_M$. The function θ_P is, apart from a constant, the product of linear functions given by the simple coroots. For example, if the family $\{c_P(\lambda)\}$ is given by an

A_M-orthogonal family, then $c_M(\lambda)$ is the Fourier transform of the characteristic function of their hull, so that the value c_M of $c_M(\lambda)$ at the origin is its volume.

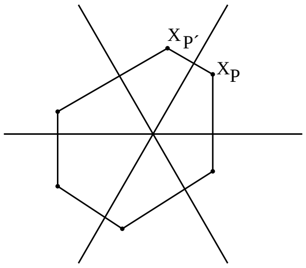

Figure 3

The sum of two A_M-orthogonal families is again an A_M-orthogonal family. More generally, the product of two (G, M)-families $\{c_P\}$ and $\{d_P\}$ is a (G, M)-family $\{e_P\}$. It is very important that e_M, or even $e_M(\lambda)$, can be expressed in terms of functions associated to $\{c_P\}$ and $\{d_P\}$. This is the source of a *splitting principle* that, among other things, reduces the global contributions to the trace formula to a sum of products of local contributions. If, for example, the two families are given by positive A_M-orthogonal families $\{X_P\}$ and $\{Y_P\}$, so that $\{e_P\}$ is given by $\{X_P + Y_P\}$, then e_M is a volume. In Figure 4 the area of the darkly shaded region is c_M, the areas of the six lightly shaded regions add up to d_M. The regions themselves are parametrized by the parabolic subgroups with M as Levi factor, thus by the elements of $\mathcal{P}(M)$, and the area is given by a number d'_P, which is the value at $\lambda = 0$ of a smooth function $d'_P(\lambda)$ given by a formula very similar to (12). Such a function is defined for each parabolic Q. Each of the remaining regions has area equal to $c_M^Q d'_Q$ where Q is an intermediate parabolic, c_M^Q is defined relative to a Levi factor of Q (rather than relative to G itself). The general formula reads

$$e_M(\lambda) = \sum_Q c_M^Q(\lambda) d'_Q(\lambda), \tag{13}$$

where the sum is over the set of parabolic subgroups containing M.

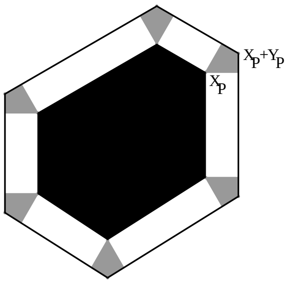

Figure 4

3. In pursuit of a trace formula. As was emphasized, the formula (6) is not invariant and contains, in addition, an undesirable dependence on the variable T. The dependence on T has been discovered to be weak and has been eliminated by setting $T = T_0$, a simple, but persuasive, device as all terms on both sides are polynomials in T whose higher order terms are of no intrinsic interest because they are associated to Levi factors of proper parabolic subgroups of G.

Although there are circumstances (cf. [K1]) in which the trace formula simplifies on its own, in general it is best to work with an invariant form. The invariant form has an appealing structural simplicity, but the apparent disadvantage that the geometric sides contain hidden spectral information. It is

$$\sum_M \frac{W_0^M}{W_0^G} \sum_{\gamma\in M(\mathbb{Q})} a^M(\gamma) I_M(\gamma, f) = \sum_M \frac{W_0^M}{W_0^G} \int_{\Pi(M)} a^M(\pi) I_M(\pi, f) d\pi. \tag{14}$$

The sum, taken from [26], is over the Levi factors containing the fixed A. The factors W_0^M are orders of Weyl groups. So it is the other terms that need to be understood. The expression on the right anticipates the resolution of a problem of convergence, partially dealt with by W. Müller in [Mü]. In the formula, as it stands to date, some attention has to paid to the order in which the sum and the integral on the right are taken ([27]). The left side too is simpler in [26] than in [27]. The factors $a^M(\gamma)$ are in reality factors $a^M(S, \gamma)$ that, for reasons that will appear later, depend in addition on the choice S of a large finite set of places.

Although the basic technique for creating invariance was introduced in [10], the final form does not appear until later, in [26], [27]. In particular, an important local principle, a Paley-Wiener theorem, was needed but not proved until later. It is a theorem for real groups of considerable interest in its own right, but also essential for rendering the spectral side of the trace formula at all explicit as in [13] and [15], and this was apparently the original purpose of the theorem of [15]. The Paley-Wiener theorem on the real line is well known. It characterizes the Fourier transforms of functions (or of smooth functions) with compact support as entire functions with supplementary properties. Arthur's theorem characterizes the Fourier transforms of smooth, compactly supported functions f on a reductive or semisimple real group whose translates, on the left and right, under a maximal compact subgroup generate a finite-dimensional space in terms of the behaviour of the operators

$$\pi(f) = \int f(x)\pi(x)dx,$$

as π runs over the representations $I(\sigma, \lambda)$ induced from representations $\sigma \otimes e^{\lambda(H)}$ of the Levi factor of a minimal parabolic subgroup of the real group. Since the induced representations have a structure that is, at best, difficult, the characterization is not at first glance very useful, but a simple easily applicable consequence is deduced from it in [15] in the form of a multiplier theorem, that provides a large family of functions $\Gamma(\sigma, \lambda)$ such that when $I(f, \sigma, \lambda)$ is defined by an element of the Paley-Wiener space, then so is $\Gamma(\sigma, \lambda)I(f, \sigma, \lambda)$.

Arthur's characterization is not, however, adequate to the arguments of [10]. For that another characterization, not of the family of operators $\pi(f)$ but of their traces and for tempered representations π alone, that is due to Clozel and Delorme [CD] is necessary.

The truncation used to obtain a trace formula was a truncation on the diagonal tailored to the geometric side. Although it was applied to the partial kernels $K_\chi(x, y)$, it is not adapted to them, as they are defined by a decomposition of $L^2(\Gamma\backslash G)$. An important step at the very beginning of the development of the spectral side is to show that there is a geometrically defined projection operator $\Lambda = \Lambda^T$ on $L^2(\Gamma\backslash G)$ such that the integrated truncation is also the integration (over the diagonal) of the kernel of ΛK and the individual terms, therefore the integration of the kernel of ΛK_χ. Of course ΛK can be replaced by $K\Lambda$ or, as Λ is a projection, by $\Lambda K \Lambda$. The operator Λ does not commute with the action of the group G.

In order to understand the implications of this replacement, we recall – for the first time – the definition and the structure of the spaces $L^2_\chi(\Gamma\backslash G)$. The class χ is defined by a pair (M, ρ), but to avoid an excess of subscripts and superscripts we take it as given and free M for other uses.

The spectral decomposition of $L^2(\Gamma\backslash G)$ as a whole and of $L^2_\chi(\Gamma\backslash G)$ in particular is given, as is any spectral decomposition, by functions depending on the spectral parameters. Since it is the spectral decomposition of a group representation and not that of a single operator, these parameters include not only π but a second parameter ϕ, an element of the vector space on which the group acts. In the case under consideration, the functions $E(x,\pi,\phi)$ are called Eisenstein series. The first step in their description is to use the collection χ to classify, for each possible Levi factor M, the discrete spectrum of $M(\mathbb{A})$ in $L^2(M(\mathbb{Q})\backslash M(\mathbb{A})^1)$. We have to pass to $M(\mathbb{A})^1$ and not remain with $M(\mathbb{A})$ because the inevitable presence of the centre A_M of M when $M \neq G$ prevents M from having a true discrete spectrum.

An irreducible unitary representation π of $M(\mathbb{A})$ determines by restriction one of $M(\mathbb{A})^1$, two representations π and π' determining the same restriction if and only if

$$\pi'(m) = e^{i\lambda(H_M(m))}\pi(m),$$

for some element λ of the real dual $\mathfrak{a}_M^*$ of $\mathfrak{a}_M$. Conversely, if π is a representation of $M(\mathbb{A})^1$ and $\lambda \in \mathfrak{a}_M^*$ then π can be extended to a representation π_λ of $M(\mathbb{A})$ by writing m as the product of e^H, $H \in \mathfrak{a}_M$, and m' in $M(\mathbb{A})^1$ and then setting

$$\pi_\lambda(m) = e^{i\lambda(H)}\pi(m'). \tag{15}$$

A Lebesgue measure on the parameter λ together with the uniform atomic measure on the unitary representations of $M(\mathbb{A})^1$ provides a measure $d\pi$ on the set of unitary representations of $M(\mathbb{A})$. To each χ is associated a subspace

$$L^2_\chi(M(\mathbb{Q})\backslash M(\mathbb{A})^1) = \sum_\pi \mathfrak{H}_{M,\chi}(\pi)$$

of $L^2(M(\mathbb{Q})\backslash M(\mathbb{A})^1)$, the sum running over the irreducible unitary representations of $M(\mathbb{A})^1$. The space $\mathfrak{H}_{M,\chi}(\pi)$ is a space that decomposes as a finite direct sum of irreducible representations all equivalent to π. Moreover, it will be 0 except for a countable number, even in general a finite number, of π. (In order not to encumber the language, we shall frequently refer to the representation on this space as π, thus suppose, for simplicity, that the multiplicity is one, as is often the case.) Spaces $\mathfrak{H}_{M,\chi}(\pi_\lambda)$ are then defined by the construction (15). The Eisenstein series $E(x,\phi)$ are attached to elements ϕ in $\mathfrak{H}_{M,\chi}(\pi_\lambda)$ and to a parabolic subgroup P with M as Levi factor. Then the family of functions

$$x \to \int_{\Pi(M)} E_P(x,\phi(\pi))d\pi$$

(where of course $\pi \to \phi(\pi)$ is subject to constraints on its integrability, on its smoothness or on its support) generate a subspace $L^2_{P,\chi}(\Gamma\backslash G)$ of $L^2_\chi(\Gamma\backslash G)$. A simple equivalence relation, that of being associate, is introduced on parabolics and

$$L^2_{P,\chi}(\Gamma\backslash G) = L^2_{\mathfrak{P},\chi}(\Gamma\backslash G)$$

depends only on the class $\mathfrak{P}$ of P. The result is that

$$L^2_\chi(\Gamma\backslash G) = \sum_{\mathfrak{P}} L^2_{\mathfrak{P},\chi}(\Gamma\backslash G). \tag{16}$$

The spectral dimension of the component $L^2_{\mathfrak{P},\chi}(\Gamma\backslash G)$ is $dim\, \mathfrak{a}_M$ if M is the Levi factor of $P \in \mathfrak{P}$.

As a consequence of the formula (16) the kernel K_χ can be expressed as an integral over the parameters P and $\Pi(M)$ of Eisenstein series

$$\sum_{\mathfrak{P}} \sum_{P\in\mathfrak{P}} \frac{1}{n(A)} \int_{\Pi(M)} E(x,\phi_j(\pi))\bar{E}(x,\phi_k(\pi))d\pi,$$

and the truncation $\Lambda K_\chi \Lambda$ as

$$\sum_{\mathfrak{P}} \sum_{P\in\mathfrak{P}} \frac{1}{n(A)} \int_{\Pi(M)} \Lambda^T E(x,\phi_j(\pi))\Lambda^T \bar{E}(y,\phi_k(\pi))d\pi.$$

The family $\{\phi_j(\pi)\}$ is an orthonormal basis of the space of π, usually chosen so as to vary in a coherent manner. Thus the integral of the truncated kernel $\Lambda K_\chi \Lambda$ is

$$J^T_\chi(f) = \sum_{\mathfrak{P}} \sum_{P\in\mathfrak{P}} \frac{1}{n(A)} \int_{\Pi(M)} tr(\Omega^T_\chi(P,\pi)\rho_\chi(P,\pi,f))d\pi, \tag{17}$$

where the operator $\Omega^T_\chi(P,\pi)$ is so defined that

$$\int_{G(\mathbb{Q})\backslash G(\mathbb{A})} \Lambda^T E(x,\phi_j(\pi))\Lambda^T \bar{E}(x,\phi_k(\pi))dx = tr(\Omega^T_\chi(P,\pi)\rho_\chi(P,\pi,f)). \tag{18}$$

I recall that the integral over $\Pi(M)$ in (17) is a Lebesgue integral over the dual of $A_M = A_P$ followed by a discrete sum. The notation is not exactly that of Arthur, for his emphasis is different than ours, but the comparison with papers [12], [13], [14] should not be difficult.

Such abstract considerations can hardly be very meaningful. If it is noticed that the class χ has a rank, the dimension of the centre of the Levi factor to which it is attached, it is seen that the spectral dimensions of the components of (16), thus the dimension of $A = A_P$, vary from 0 to this rank. It may happen, as in the unramified case appearing in formula (8), that only the largest of these dimensions appear. That is the favorable, or at least the simplest, situation, for one of the difficulties that plague Arthur's development of the trace formula, a residual effect of the methods used to establish the theory of Eisenstein series, is that the spectral side is only a sum over χ so that there is no clean separation of the individual terms into spectra of different dimension and no guarantee that such a separation is possible without destroying convergence.

In order to reduce the level of abstraction and to introduce some concrete formulas, I shall examine the case that the rank of χ is one, adequately represented, both structurally and analytically, by the group $\mathrm{SL}(2)$. With some understanding of this case and the standard analytic techniques on the line used to analyse it in hand, we shall then try to appreciate the new type of analysis that Arthur substituted for them in higher rank.

Before beginning, we recall from [9] a formula not for (18) but for a perturbation of it. It is valid only when the Eisenstein series are associated to cusp forms, thus when the the dimension of A_P is equal to the rank of χ. We fix a given π and consider all the other defined by

$$\pi_\lambda(m) = e^{\lambda(H(m))}\pi(m). \tag{19}$$

This notation differs slightly from that of (15). We replace (18) by

$$\int_{G(\mathbb{Q})\backslash G(\mathbb{A})} \Lambda^T E(x,\phi,\lambda)\Lambda^T \bar{E}(x,\psi,-\bar{\mu})dx, \tag{20}$$

in which $E(x,\phi,\lambda)$ is just $E(x,\phi_j(\pi_\lambda))$, but in which, for the next formula, it is best not to assume that $\mu = -\bar{\lambda}$. By the definition, the space of π_λ is the same as that of the given π used to define it. So the Eisenstein series are parametrized by elements of a common space and by λ. If we let λ approach $-\bar{\mu}$ in (20) we recover (18).

For an arbitrary π, the theory of Eisenstein series associates intertwining operators $M(t,\pi)$ to appropriate elements of various Weyl groups. Since we are concerned with the behavior of $M(t,\pi_\lambda)$ as a function of λ, we write $M(t,\lambda)$ rather than $M(t,\pi_\lambda)$. The formula for (19) is

$$\int_{G(\mathbb{Q})\backslash G(\mathbb{A})} \Lambda^T E(x,\phi,\lambda)\Lambda^T \bar{E}(x,\psi,-\bar{\mu})dx = \sum_{\mathfrak{P}}\sum_{s,t} \frac{(M(s,\lambda)\phi, M(t,-\bar{\mu})\phi')}{\theta_P(s\lambda + t\bar{\mu})}. \tag{21}$$

The function θ_P has already appeared in (12). The sums over s and t are finite sums, and s and t may, without much loss of precision, be thought of as elements of a Weyl group.

Although it does not simplify, the formula (21) becomes more transparent and the limit $\lambda \to -\bar{\mu}$ easier to take when the rank of χ is one and, therefore, the cuspidal Eisenstein series functions of a single complex variable. The dimensions of the spectrum appearing in (12) are then one and zero, the zero-dimensional spectrum being of course the discrete spectrum. Moreover, the variable λ can be identified – not canonically – with a complex number $i\lambda = \sigma + i\tau$. In (19) and (21), we are at first preoccupied with purely imaginary λ, for they yield the one-dimensional spectrum. Set

$$\Theta(\tau) = e^{-2i\tau T} M(i\tau). \tag{22}$$

Then, as in [L], the formula (21) becomes in the limit $\lambda = \mu$, $\sigma = 0$

$$-\Big\{\frac{1}{i}\Theta^{-1}(\tau)\Theta'(\tau)\phi, \psi) + \frac{2}{i\tau}(\Theta(\tau)\phi, \psi) - (\phi, \Theta(\tau)\psi)\Big\}. \tag{23}$$

It can happen – this is basically the unramified case – that there are very few s and t in the sum (21) so that the second term does not appear. Then the following considerations simplify.

We have noted how important it is for the further development of the trace formula that the left side of (17), and thus also the right side, is a polynomial in T, indeed, when the rank of χ is one, a polynomial of degree one. This ceases to be so if we separate the terms of various rank. The terms of rank one are obtained by taking ϕ and ψ in (23) to be functions of τ and then integrating over $(-\infty, \infty)$. The first term, a logarithmic derivative, is what appears in the unramified case and clearly yields a linear function of T. The second term, in its dependence on T is essentially an integral with kernel $\sin(\tau T)/\tau$, so that as T grows larger and larger it will just pick up the value of whatever it multiplies at $\tau = 0$. Thus it will approach a constant, but there will be a remainder which is not a polynomial in T, but is very small, at least for T large. Since the complete sum (17) is a polynomial, this small remainder will be cancelled by a small remainder from the discrete spectrum.

The discrete spectrum is also obtained from the Eisenstein series, as functions of $x \in G$ obtained as the residue of $E(x, \phi, \lambda)$ at certain real points $\lambda > 0$. We can readily obtain the inner product of two truncated residues by taking residues in the formula (21). As in [L], it will have, apart from a constant factor, the form

$$(m(\lambda)\phi, \psi) - \frac{1}{\lambda + \mu} e^{-(\lambda+\mu)T}(m(\lambda)\phi, m(\mu)\psi)). \tag{24}$$

The first term is a constant, thus a linear polynomial in T; the second is small for large T. So the result that (16) is a polynomial in T is compatible with these formulas. Moreover, we now have a little more information for χ of rank one. Up to remainders that become small as T grows large, individual terms in (16) are polynomials in T for which there are explicit expressions in terms of logarithmic derivatives of the operators $M(t, \pi)$.

These results are established in general in [13], [14]. In rank one, the constant term is the sum of a one-dimensional logarithmic integral, the δ-function generated by the $\sin(\tau T)/\tau$ kernel, and the contribution from the first term in (24), which is nothing but the inner product of two nontruncated square-integrable automorphic forms in the discrete spectrum. In general, each term of (17), except for those of rank zero and of the highest rank, will combine all of these features. So I refer to Theorem 8.2 of [14] for their precise description. On the other hand, I would like to comment briefly on the proof, since it illustrates Arthur's ability to overcome difficult analytic obstacles with unexpected extensions of traditional analytic techniques. I fear that, imbedded as they are in developments that also carry a heavy algebraic and even arithmetic burden, traditionally oriented analysts have not given these contributions the attention they deserve.

His first step ([12]) was a form of (21) valid for Eisenstein series associated to all elements of the discrete spectrum and not just the cusp forms. It is suggested by (24), in which only the first term counts. It turns out that in general the inner product of two truncated Eisenstein series is the sum of a term for which there is an analogue of (21) and a term that grows exponentially small as T grows large. A simple case appears in the formula (24).

This formula yields an asymptotic expression for the functions appearing in (17)

$$\Psi_\pi^T(\lambda, f) = \frac{1}{n(A)}\, tr(\Omega_{\chi,\pi}^T(P,\lambda)\rho_{\chi,\pi}(P,\lambda,f)),$$

where, once again, as in (19) the notation is adapted to a fixed π and a varying π_λ. It is tempting to insert this asymptotic formula directly into (17) and to ignore the small terms. The difficulty is that there is no uniformity in λ that permits an integration or the discard of the small terms.

So a cut-off function in λ with compact support on which it is largely one or close to one has to be introduced in the integrals of (17) and then dilated so that the result approaches the original integral. As a result of this modification, the expression is no longer a polynomial in T, because the cut-off is not a result of modifying f. (Cutting off the Fourier transform of a function with compact support does not yield the Fourier transform of a function with compact support!) What Arthur does in [13] is to exploit the multiplier theorem resulting from his Paley-Wiener theorem to show that the cut-off is still approximately a polynomial as T grows and that, in addition, as the cut-off function approaches in a suitable sense the constant 1, this polynomial approaches (17). The possible failure of the Ramanujan conjecture creates an additional difficulty in the analysis, but once it is accomplished he is free to use his approximate formula.

He has then to extend the calculation leading to (24) from dimension one to arbitrary dimension. The functions that result from the general form of (21) are, like (22), products of exponentials and intertwining operators and appear in the form of (G, M)-families. The $\sin(\tau T)/\tau$ kernel is essentially the Fourier transform of the characteristic function of the interval $(-T, T)$ and its properties can readily be deduced by passing to Fourier transforms. This interval is replaced in general by the convex sets described in §2, so that the $\sin(\tau T)/\tau$ kernel is replaced by a function of the form (12). The formula (13) allows him to treat the exponential factors separately from the intertwining operators.

This is by no means the only application of (13). An automorphic representation π can be expressed as the tensor product $\otimes\pi_v$ over the places finite and infinite of the field F in question (for us F has been $\mathbb{Q}$) of representations of the local groups $G(F_v)$. There is a corresponding factorization of the intertwining operators $M(s, \pi)$ or, if π is replaced by π_λ, $M(s, \lambda)$. The local operators are themselves products of a scalar that depends on λ and a second simple factor that is almost everywhere 1, so that the tensor product of the second factors is of more algebraic than analytic interest. The product of the scalar factors is of analytic interest and turns out to be an L-function, classical and equal to the ζ-function or to a Dirichlet L-function in the simplest cases but a more general Euler product otherwise. Once again, it is formula (13) that enables Arthur to separate these two factors. Although this separation is of major importance when applying the trace formula and is used on the way to (14), its immediate relevance is that it yields an estimate that allows a passage to the limit in the mollifier after the approximate formula for $\Psi_\pi^T(\lambda, f)$ have been inserted. The result is a decisive step on the way to the right side of the compact formula (14) in which that of Theorem 8.2 of [14] is implicit. This theorem is the precise general form of formula (8). We have yet to consider the precise general form of (5).

4. Weighted orbital integrals: reduction to local integrals. If we consider the truncations only for $T = T_0$, it is only the weight $v(x, T_0)$ that matters in (5). It can be expressed as $v_M = v_M(x)$, where $v_M(x)$ is defined by a (G, M)-family $\{v_P(\lambda, x)\}$ associated to an A_M-orthogonal family that is introduced by observing that the functions $H_P(x)$ described earlier for the $\mathbb{Q}$-parabolic subgroups containing a given minimal one can be defined more generally and that for each x the family $\{H_P(x)\}$ is A_M-orthogonal.

Although we have not yet undertaken the comparison of trace formulas, the purpose of such comparisons is usually an equality between part or all of the spectral sides of the trace formula for appropriate functions f and f' on two different groups, an equality that results from an equality on the geometric side, in turn an equality, at least in part, between the *local* orbital integrals of f and

f', usually taken as products $f(x) = \prod f_v(x_v)$, $f'(x') = \prod f'_v(x'_v)$. If this strategy is to be used, the geometric side of the trace formula has to be expressed, perhaps with the help of the splitting principle, in local terms. The expression (5) is not local, but it has a local form. The one preferred by Arthur is not an integral over [2] $G_\gamma(F_v)\backslash G(F_v)$, but one over $G_\gamma(F_S)\backslash G(F_S)$, where S is a finite set of places, usually large and usually containing all infinite places. In addition he multiplies it by a simply defined factor

$$|D(\gamma)|^{1/2} = \prod_{v \in S} |\det_{\mathfrak{g}/\mathfrak{g}_\sigma}(1 - \mathrm{Ad}(\sigma))|_v^{1/2},$$

where σ is the semisimple part of γ. Thus when γ itself is semisimple, as in (5), $\sigma = \gamma$. The result is that (5) is replaced by

$$J_M(\gamma, f) = |D(\gamma)|^{1/2} \int_{G_\gamma(F_S)\backslash G(F_S)} f(x^{-1}\gamma x) v_M(x) dx, \tag{25}$$

an expression that is defined for all γ in $M(F_S)$ whose centralizer G_γ in G is equal to its centralizer M_γ in M. The element γ need no longer be semisimple and $G_\gamma(F_S)$ is really a product $\prod G_{\gamma_v}(F_v)$. Because of the splitting principles deduced from (13), it will also be appropriate to consider (25) relative to a Levi factor intermediate between M and G, but that is a formal matter. Although the weighted orbital integrals (25) appear first in connection with the unramified classes on the geometric side, the next step in the development of the trace formula is to define them for all classes and to express each summand on the geometric side in terms of them.

Arthur's treatment of the geometric terms appears in [20], [21], [24]. One part, carried out in [21], is simply to analyse the combinatorics of the truncation to reduce the geometric term associated to a general orbit $\mathfrak{o}$ to the term associated to the unipotent orbit $\mathfrak{o} = unip$. Beyond that, there are two things to do: to define the distributions $J_M(\gamma, f)$ in general, thus for arbitrary $\gamma \in M(F_S)$ and to express J_{unip} in terms of them. These are the tasks of [24] and [20] respectively. The result is a formula ([21], Theorem 8.1) that looks just like a sum of some of the terms on the left of (14), except that the terms are not yet invariant. Rather than try to understand immediately how this procedure is carried out in general, we examine the simplest case, that of $\mathrm{PGL}(2)$, observing that it suffices to define $J_M(\gamma, f)$ when S consists of a single place.

The matrix

$$\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

[2] At this point, it becomes more convenient to work explicitly with a general number field and not to reduce all questions to $\mathbb{Q}$.

represents a typical unipotent element and $D(\gamma) = 1$. Any element of $G(F_v)$ can be written as $x = nak$ with

$$n = n(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \quad a = a(\alpha) = \begin{matrix} \alpha & 0 \\ 0 & 1 \end{matrix}$$

and with k in the standard maximal compact subgroup. Then

$$dx = dz \frac{d\alpha}{|\alpha|_v^2} dk.$$

Thus, apart from some simple constants that are overlooked in the following sequence of equalities,

$$\begin{aligned} J_G(\gamma, f) &= \int_{G_\gamma(F_v)\backslash G(F_v)} f(x^{-1}\gamma x)dx \\ &= \int dk \int \frac{d\alpha}{|\alpha|_v^2} \{f(k^{-1}n(1/\alpha)k)\} \\ &= \int dk \int dz \{f(k^{-1}n(z)k)\}. \end{aligned} \tag{26}$$

The final integral is clearly convergent if, for example, the support of f is compact. If, on the other hand, $\gamma = 1$, then

$$J_G(\gamma, f) = f(1).$$

There is only one more unipotent weighted orbital integral in addition to this.

Suppose P is the group of upper triangular matrices and P' the group of lower triangular matrices. Their common Levi factor M is the group of diagonal matrices. The remaining weighted unipotent orbital integral is $J_M(1, f)$. According to the procedure introduced in [24], it is introduced as a limit. The matrix $n = n(z)$ is in P, so that $H_P(n) = I$, and if

$$n(z) = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \begin{matrix} 1 & 0 \\ y & 1 \end{matrix} k,$$

then, as we are in the projective group, $H_{P'}(n) = a(\alpha^2)$, α being easily calculated. As a result – a factor depending on the normalization of measure aside – the weight $v_M(n) = v_M(nk)$ is $\ln(1 + z^2)/2$ if v is real and if v is nonarchimedean it is 0 for z integral and $\ln|z|$ if $|z| > 1$.

Consider $J_M(a(\alpha), f)$ for α close to 1. It is

$$\frac{|\alpha - 1|}{|\alpha|^{1/2}} \int dk \int dz \{f(k^{-1}n(-z)a(\alpha)n(z)k)v_M(n(z)k)\}$$

or

$$\frac{|\alpha - 1|}{|\alpha|^{1/2}} \int dk \int dz \{f(k^{-1}n((\alpha - 1)z)a(\alpha)k)v_M(n(z))\}.$$

The integral over k is easily dealt with as $v_M(n(z)k) = v_M(n(z))$. We replace the variable z by $z/(\alpha-1)$ and let α approach 1. The denominator $|\alpha|^{1/2}$ of $|D(\gamma)|^{1/2}$ becomes 1 and the numerator is cancelled by the change in measure. If F is the real field, the term $v_M(n)(z/(\alpha-1))$ is

$$\ln((\alpha-1)^2+z^2)/2 - \ln|\alpha-1|.$$

The second term leads to a contribution that blows up as α approaches 1, although in a simple way, for the contribution is the product of a term in α that is independent of f with (26). The first term approaches, except at $z=0$, the function $\ln|z|$. We set

$$J_M(1,f) = \int dk \int dz \{f(k^{-1}n(z)k)\ln|z|\}. \tag{27}$$

A similar calculation leads to the same result for nonarchimedean fields.

Thus, for $\gamma = 1$, we have defined $J_M(\gamma, f)$ as a limit

$$\lim_{a\to 1} \sum_{L\supset M} r_M^L(\gamma, a) J_L(a\gamma, f),$$

where a lies in $A_M(F)$, provided we define (constants aside)

$$r_M^M(\gamma, a) = 1, \qquad r_M^G(\gamma, a) = \ln|\alpha - 1|.$$

The sum runs over those Levi factors of parabolic subgroups that contain M, in the present case just G and M itself. The purpose of [24] is to do this for all groups G, all M and all $\gamma \in M(F)$.

If we pass to F_S, then the factor $\ln|z|$ in (27) is replaced by

$$\sum_{v\in S} \ln|z_v|_v.$$

In the sequence of modifications (26), an integral over the multiplicative group $\{a(\alpha)\,|\,\alpha \in F^\times\}$ is replaced by an integral over the additive group $\{n(z)\,|\,z\in F\}$, $z = 1/\alpha$. If we had naively attempted to compute the integral of the kernel K_1 without truncating, then in addition to the contribution from

$$\int_{\Gamma\backslash G} f(1)dx, \tag{28}$$

we would have, with $\gamma = n(1)$

$$\mathrm{meas}(G_\gamma(\mathbb{Q})\backslash G_\gamma(\mathbb{A})) \int_{G_\gamma(\mathbb{A})\backslash G(\mathbb{A})} f(x^{-1}\gamma x). \tag{29}$$

Since $G_\gamma = N$, we can replace this integral by an integral over $A(\mathbb{A}) \times K$, where A is the group of diagonal matrices in G and K is the maximal compact subgroup of the global group $G(\mathbb{A})$. We might attempt to apply the sequence of transformations (26), but we would run into a difficulty, and this difficulty is the principal reason for the truncation. If v is a nonarchimedean prime, let $A_0(F_v) = \{a \in A(F_v) \,|\, |a|_v = 1\}$. Since $A(\mathbb{A})$ is the limit of the sets

$$U_S = \prod_{v \in S} A(F_v) \prod_{v \notin S} A_0(F_v),$$

where S is finite and contains all infinite places, we can first try to calculate the integral over $A(\mathbb{A})$ as an integral over U_S. If S is so large that outside of S the function f_v is the characteristic function of the group K_v of integral matrices with integral inverse, then the integral is a product of an integral over $A(F_S)$ with the integrals

$$\int_{A_0(F_v)} 1 = \int_{|\alpha|_v=1} \frac{d\alpha}{|\alpha|^2}.$$

The global integral can and must be defined by normalising the multiplicative measure so that each of these integrals is 1. Suppose $F = \mathbb{Q}$ and v is the prime p. If we tried to pass to the additive measure, taking $z = 1/\alpha$, then we would have

$$\int_{|z_v|=1} dz = 1 - \frac{1}{p},$$

the equality resulting from the necessary normalization of the additive measure. So the transition from the penultimate line of (26) to the final line would involve not a harmless finite constant but a meaningless infinite product

$$\prod_{p \notin S} \left(1 - \frac{1}{p}\right).$$

If, however, we had truncated, then the integral would not be over $A(\mathbb{A})$, but only over those elements $a(\alpha)$ in this group for which $|\alpha| < c$, $c = c_T$ being a constant that depends on the truncation parameter. Fix once again S and suppose once again that f_v is the characteristic function of K_v outside of S. Then for each $\{\alpha_v \,|\, v \in S\}$, we would only have to integrate over the collection of

$$\Big\{\prod_{v \notin S} \alpha_v \,|\, 1 \le |\alpha_v|_v \quad \text{and} \quad \prod_{v \notin S} |\alpha_v|_v \le c/\prod_{v \in S} |\alpha_v|_v\Big\}.$$

When $F = \mathbb{Q}$ so that each of these places is a prime p, this integral is

$$\sum_{n \le C} \frac{1}{n}, \tag{30}$$

the sum being taken only over those n that are not divisible by any prime in S. The number C is $c/\prod_{v\in S}|\alpha_v|_v$. It is well known that the sum (30) is asymptotic to

$$a + b\ln C = a' + b'\ln\Big(\prod_{v\in S}|\alpha_v|_v\Big) = a' - b'\Big(\sum \ln|z_v|_v\Big), \tag{31}$$

if $z_v = 1/\alpha_v$. We conclude that (29) can be expressed as a sum of multiples of $J_G(\gamma, f)$ and $J_M(1, f)$, the coefficients depending on the choice of S, which must be taken sufficiently large to accomodate the given function f.

The constants a' and b' clearly depend on S. (The constant a' also depends linearly on T, but b' is independent of it.) When describing the invariant form of the trace formula, I cautioned the reader that the coefficients $a_M(\gamma)$ depended on the choice of a set S. Since the invariant geometric side with the distributions $I_M(\gamma, f)$ is obtained from the geometric side with the distributions $J_M(\gamma, f)$, we see clearly, in the simple case just treated, that the source of the dependence on S is the asymptotic behaviour of (30).

The argument in general is similar, but the estimates more difficult, and the measures not directly identifiable. The conclusion is a formula,

$$J_{unip}(f) = \sum_M \frac{|W_0^M|}{|W_0^G|}\sum_u a^M(S,u)J_M(u,f), \tag{32}$$

in which the outer sum is over the Levi factors of the parabolic subgroups. The inner sum is over unipotent u with rational representatives. A decisive observation of [10], a paper to which we shall return, is that the distributions $J_{\mathfrak{o}}^T(f)$ and $J_\chi^T(f)$, although not invariant, have similar formal behaviours under the similarity transformations, $f \to f^y$, $f^y(x) = f(yxy^{-1})$. Many other distributions that appear in the analysis of the trace formula behave in the same fashion. In particular, the weighted orbital integrals $J_M(\gamma, f)$ do. As a result, it is possible to define inductively coefficients $a^M(S,u)$ such that

$$J_{unip}(f) - \sum_{M\neq G}\frac{|W_0^M|}{|W_0^G|}a^M(S,u)J_M^I(u,f)$$

is invariant. It is also supported on the unipotent classes, so that once it is shown that it is a measure, it is clear that it is the sum of multiples of the distributions $J_G(u, f)$, u unipotent. If these u can be shown to have rational representatives, the formula (32) follows. What is required is that sums similar to (30) but much more complicated be estimated with conclusions much coarser than (31).

5. The invariant trace formula: introduction. The paper [10] was Arthur's first attempt at an invariant trace formula. It was not complete as it invoked some assumptions that were only verified later, but in addition to the geometric ideas already described in §2 it contains several ideas that reappear as a rogue's yarn throughout the later work.

I have already alluded to the formal structure of the noninvariance of the distributions $J_{\mathfrak{o}}^T$ and J_χ^T and thus of $J_{\mathfrak{o}}$ and J_χ. We have already agreed that the only parabolic subgroups to be considered are those that are defined over the ground field F, usually $\mathbb{Q}$, and that contain a fixed Levi factor M_0 of a fixed minimal parabolic P_0 over F. The only Levi factors to be considered are Levi factors of this collection of parabolic subgroups that contain the fixed Levi factor M_0. The symbol $\mathcal{L}^L(M)$ denotes the set of Levi factors between two Levi factors $M \subset L$; the symbol $\mathcal{F}^L(M)$ the parabolic subgroups over F between M and L; and $\mathcal{P}^L(M)$ those with M as a Levi factor. If $L = G$, the superscript is often omitted. If Q lies in $\mathcal{F}^L(M)$, then as in [10] we associate to the function f a function on $M_Q(\mathbb{A})$ by

$$f_{Q,y}(m) = \delta_Q(m) \int_K \int_{N_Q(\mathbb{A})} f(k^{-1}mnk)u'_Q(ky)dn\, dk, \tag{33}$$

where the function u'_Q is defined by the geometric constructions of §2. It is a volume! If $f^y(x) = f(yxy^{-1})$, then

$$J_{\mathfrak{o}}^T(f^y) = \sum_Q J_{\mathfrak{o}}^{M_Q,T}(f_{Q,y}), \tag{34$'$}$$

the sum running over all allowed F-parabolic subgroups and the double superscript indicating clearly that the distributions on the right are on the group M_Q. Moreover

$$J_\chi^T(f^y) = \sum_Q J_\chi^{M_Q,T}(f_{Q,y}). \tag{34$''$}$$

Thus the formal structure is the same for both the geometric and the spectral distributions, and is independent of T. The definition (33) admits a local form, and the weighted orbital integrals $J_M(\gamma, f)$, which are local objects, satisfy a relation just like (34′) or (34″), in which, however, the sum is over those Q that contain M and are defined over the local field. There are Q that are defined over a local field F_v but not over the global field F. Another manifestation of the difference between the local field and the global field is that the space $\mathfrak{a}_M$ may grow when one passes to the local field so that the geometry used to define (G, M)-pairs changes.

The method used to convert a distribution satisfying the relations (34) to an invariant distribution is at first far from promising, as it amounts to little more than transferring the offending part of one side of the formula, usually of the spectral side, to the opposite side. It is a construction intermediate between local and global, thus a construction over some product $F_S = \prod_S F_v$. So it is understood that $f=\prod f_v$, where f_v is the unit element of the Hecke algebra outside of S. Thus all that matters is $\prod_S f_v$, which we also denote f and which we regard as an element of an appropriate space $U(G)$. The argument proceeding by induction, we introduce similar spaces $U(M)$ on all pertinent Levi factors. There is a second space $V(M)$ and a map

$$\phi_M^M\colon U(M)\to V(M)$$

through which all invariant distributions uniquely factor. More precisely, the map is surjective and pull-back on the dual space to $V(M)$ has as image the set of invariant distributions on U(M). More generally, the maps $\phi_M^L : U(L)\to V(M)$ are to be defined for every pair $M\subset L$ of pertinent Levi factors and the relations[3]

$$\phi_M^L(f^y)=\sum_{Q\in\mathcal{F}^L(M)}\phi_M^{M_Q}(f_{Q,y}) \tag{35}$$

are to be satisfied.

These objects available, we can pass from a family of distributions J^L, one for each Levi factor L, that satisfy [4]

$$J^L(f^y)=\sum_{Q\in\mathcal{F}^L(M)}\frac{c(M_Q)}{c(L)}J^{M_Q}(f_{Q,y})$$

to a family of invariant distributions I^L by demanding that

$$J^L(f)=\sum_{M\in\mathcal{L}^L(M_0)}\frac{c(M)}{c(L)}\hat{I}^M(\phi_M^L(f)). \tag{36}$$

The distribution I^M is the image of $\hat{I}^M$. It is defined inductively by this formula.

Arthur passes quickly – and rather glibly – in [10], returning to the matter in [26], [27], [28], from F_S to F_v, on the grounds that if these assumptions are satisfied for the maps ϕ_M^L that he constructs when S consists of a single place v alone then a product over v in S yields them in general. At all events, the simpler case is the decisive case. The definition of ϕ_M^L is based on the local analogue of the integrals appearing in (17).

6. Weighted characters and the invariant trace formula. In this section, we at first fix the place v and let F denote the local field. The question as to which space to take for $U(G)$ is not fully resolved

[3] For technical reasons, this relation must sometimes take a different, integrated form. The path to the invariant trace formula is, in contrast to others more heavily trodden, not yet broad.

[4] I have included in the formula constants $c(L)$ attached to Levi factors because Arthur does, but they are often identically 1.

until the papers [27], [28]. Although there is an important technical complication – to be explained when we are in a position to appreciate it – that imposes a modification to the definition, it is essentially the Hecke algebra of smooth, compactly supported functions that are left and right K-finite, K being now a local maximal compact subgroup. The space $V(G)$ is defined by the Plancherel formula. As is known from the work of Harish-Chandra, the representations needed for the local harmonic analysis of L^2-functions or of Schwartz functions are the irreducible tempered representations. For any G, the space [5] $V(G)$ is a space of functions on the set $\Pi_{temp}(G)$ of tempered representations. The local and global harmonic analysis are structurally similar. The spectrum of $L^2(G(F))$ is described by a sum over Levi factors (with the same conventions as above, but with respect to a minimal parabolic over F and a distinguished Levi factor of it), each Levi factor contributing spectra of dimension equal to that of A_M, the maximal split torus in its centre. The components of the contributions of M, apart from some identifications arising from the action of the Weyl group, are indexed by the representations of $M(F)$ square-integrable modulo its centre, referred to here, by abuse of terminology, as representations of the discrete series. The definition (19) may be used locally, although I now prefer to add a factor i as in (15),

$$\pi_\lambda(m) = e^{i\lambda(H_M(m))}\pi(m).$$

If σ is a tempered representation or more generally any irreducible representation of $M(F)$ and P a parabolic subgroup with M as Levi factor, there is associated to σ and P an *induced* representation $I_P(\sigma)$. The space $V(G)$ is essentially the space of functions on $\Pi_{temp}(G)$ such that for any M and any tempered representation σ of $M(F)$, the function $\phi(I_P(\sigma_\lambda))$ is of Paley-Wiener type. This means in particular that it extends to an entire function of λ.

Since ϕ^L_M is just ϕ^G_M for $G = L$, it suffices to define ϕ^G_M. It is defined by specifying, for each f, the function $\phi(\pi)$, $\phi = \phi^G_M(f)$. As this will be defined by an analogue of the weighted orbital integral, we denote it $J_M(\pi, f)$ and refer to it as a *weighted character*. Then

$$\phi_M(f)(\pi) = J_M(\pi, f). \tag{37}$$

Because it is easy to forget, when attempting to understand the definitions, that for $M = G$, which is the critical case for pulling back distributions, there is nothing to them, I observe in particular that

5 Arthur's notation develops with time, so that the notation of [10] is not always that of the later papers. The reader who consults [26], [27], [28] will have to accommodate himself to the newer notation. In this report, I sometimes adhere to a fixed convention and sometimes, to make consultation of the original papers easier, use two symbols for what is essentially the same object.

$\phi_G(f)(\pi) = J_G(\pi, f) = tr(\pi(f))$. The relation (35) becomes the local analogue of the relations (34′) and (34″) for weighted orbital integrals,

$$J_M(\pi, f^y) = \sum_{Q \in \mathcal{F}^L(M)} J_M^{M_Q}(\pi, f_{Q,y}). \tag{38}$$

With (38) in mind, Arthur defines [6] $J_M(\pi, f)$ by a local analogue of the integrand appearing in the formula (8) for $J_\chi^T(f)$.

The formula (8) is deduced from (21) in which the global intertwining operators $M(s, \lambda)$ of the theory of Eisenstein series appear. These operators have local analogues, denoted $J_{P'|P}(\pi_\lambda)$ in [28] that arise because a representation $I_P(\pi)$ induced from the extension of π to the parabolic P with Levi factor M will usually be equivalent to other induced representations $I_{P'}(\pi')$ induced from a sometimes different π' with a usually different P' but with the same M. The global operators are tensor products of the local operators, defined directly in a large open cone by an integral and then, in general, by analytic continuation. It is convenient to represent both as the product of a scalar factor and a normalized intertwining operator.

The scalar factor is expressed as the quotient of products of automorphic L-functions, local or global, thus in particular as an Euler product that admits an analytic continuation. Although there is no ambiguity about the local scalar factor at those places where no ramification is present, nor in principle at the other places, at the moment the information available about representations of the groups $M(F)$ is inadequate and some not entirely satisfactory expedients have to be invoked ([28]).

At unramified places, the local normalized operator $R_{P'|P}(\pi_\lambda)$ fixes the unramified vector. In general it is a rational function. The weighted characters are defined by the collection of operators, one for each P in $\mathcal{P}(M)$,

$$\mathcal{R}_P(\nu, \pi_\lambda, P_0) = R_{P|P_0}(\pi_\lambda)^{-1} R_{P|P_0}(\pi_{\lambda+\nu}), \tag{39}$$

which turns out to be a (G, M)-family to which the definitions of §2 can be applied[7] The group $P_0 \in \mathcal{P}(M)$ does not affect the final result. The operator $\mathcal{R}_M(\pi_\lambda, P_0)$ is defined by (12), with $c_P(\lambda)$ given by (39), the variable λ of (12) being the variable ν of (39) and the λ of (39) being a variable that

[6] He later, in the paper [47], slightly modifies the original definition, which turns out not to be entirely suitable to his ultimate aims. The original definition is nevertheless not without its merits. The definition affects the values of the factors $a^M(\pi)$ in (14).

[7] The conceptual elegance of the definitions of §2 is not always matched by a practical simplicity, so that turning them into explicit formulas does not appear to be always feasible. For the families under consideration, it is instructive to turn to §7 of [14].

affects the family c_P and thus the value c_M. It is well to recall that when the dimension of $\mathfrak{a}_M$ is one, $\mathcal{R}_M(\pi_\lambda, P_0)$ is going to be a logarithmic derivative. The weighted characters are finally defined as

$$J_M(\pi_\lambda) = tr(\mathcal{R}_M(\pi_\lambda, P_0) I_{P_0}(\pi_\lambda, f))$$

and do not depend on the choice of P_0. The representation $\pi = \pi_\lambda$ is unitary but need not be tempered.

Although we did not describe them explicitly either in formula (8) or when mentioning Theorem 8.2 of [14], the global operators that appear are obtained by a global normalization that is a product of the local normalizations. The global factorization together with the global geometry leads to a global analogue of the weighted characters. Despite the difference between the local and global $\mathfrak{a}_M$ and their geometries, the properties of the global weighted characters are deduced from those of the local. Once they are available, Arthur can apply (36), with $L = G$, to the distributions $J_{\mathfrak{o}}$ and J_χ to obtain $I_{\mathfrak{o}}$ and I_χ. Then, as a result of easily justified formal manipulations, he obtains ([10]) an *invariant* trace formula

$$\sum I_{\mathfrak{o}}(f) = \sum I_\chi(f). \tag{40}$$

It has yet to be explained why this is the formula (14). Moreover, it has to be decided exactly what the spaces $U(L)$ and $V(L)$ are to be and whether the maps ϕ_L^L are indeed open and surjective, thus as Arthur later phrases it, in for example [38], whether all invariant distributions are supported on characters.

As Arthur explains in [38], there are at least two forms of the problem, for all invariant distributions or for all invariant tempered distributions, neither of which is completely solved, and neither of which is appropriate or necessary for the question at hand. Thus, in [10] a false start was made with Assumption 5.1, an assumption that may be valid but has not yet been established in general. There are, so far as I can see, two important ways in which his final arguments differ from those foreseen in [10], which remains none the less a basic reference. Indeed, it appears that to someone with a fundamental understanding of all that is contained in [10], the later modifications are minor, but for those who are attempting to reconstruct the arguments for themselves in the belief that there is a linear progression from [10] through [26], [27], [28] there are pitfalls, from which I had to be rescued by Arthur himself.

Although we mentioned in passing that it was best to consider the trace formula as applying to functions on $G^1 = G(\mathbb{A})^1$, the set of elements x in $G(\mathbb{A})$ such that $|\chi(x)| = 1$ for every rational character χ of G over the base field, we were content to deal implicitly with the case that there were no such rational characters, so that $G(\mathbb{A}) = G(\mathbb{A})^1$. We recognized, however, that this assumption was not valid for the Levi factors, so that it was quite inappropriate as almost all arguments in Arthur proceed by induction. Since we have now to explain an important induction argument, the assumption, made only for notational purposes, is abandoned. The trace formula becomes a distribution on $G(\mathbb{A})^1$, but its form is not at all changed. This is at first surprising, but then one observes that all terms on the geometric side are concentrated on $G(\mathbb{A})^1$, because $|\chi(\gamma)| = 1$ for every $\gamma \in G(F)$, and that all terms on the spectral side will continue to be concentrated on $G(\mathbb{A})^1$ because they will include an integral over $i\mathfrak{a}_G^*$, the dual of $G(\mathbb{A})^1\backslash G(\mathbb{A})$. (I made no attempt to describe exactly the terms on the spectral side that result from Theorem 8.2 of [14]. They are contributed by Levi factors L and contain an integral over $i\mathfrak{a}_L^*/i\mathfrak{a}_G^*$, which is replaced by an integral over $i\mathfrak{a}_L^*$ when an integration over $i\mathfrak{a}_G^*$ whose effect is a contraction to $G(\mathbb{A})^1$ is added.)

The critical feature of the passage from a family of noninvariant distributions $\{J^L\}$ to a family $\{I^L\}$ was the existence of the distributions I^L appearing in formula (36). In lieu of a general theorem, it would be enough to show that the distributions appearing in the trace formula are supported on characters. This is accomplished by an elaborate induction in [26], [27] that is one of the new features of the argument.

The notion of a distribution being supported on characters depends of course on what are to be considered the pertinent distributions, those on functions with compact support, those on rapidly decreasing functions, or those on some other space. The initial choice, in [10], seems to have been tempered distributions, but for lack of the necessary results in this case that paper remains incomplete. So Arthur prefers general distributions, but he then encounters another difficulty. The images $\phi_M(f)$ of the map ϕ_M is not necessarily a function of compact support even when f is. It is not easy to find one's bearings on this constantly shifting terrain. For the purposes of his arguments, Arthur introduces spaces of functions that are neither slowly decreasing nor compactly supported and that, as we already observed, are in addition taken to be K-finite.

Recall that $U(G)$ and $V(G)$, or more generally $U(M)$ and $V(M)$, are to be defined with respect to a finite set S of places. The functions in $U(G)$ are not necessarily of compact support on $G(F_S)$, but their restrictions to every slice

$$G^X = \{x \in G(F_S) \,|\, H_G(x) = X\}$$

are. The space $V(G)$ is defined in a similar way. Arthur is forced to some such definition for he needs to show that if $\phi_M(f)$ is defined by (37), then there is a function f' in $U(M)$ such that $tr(f'(\pi)) = \phi_M(\pi)$ for all π.

Even more care and even more elaborate notation is necessary as

$$f(\pi) = \int_{G(F_S)} \pi(x) f(x) dx$$

may not be defined. The integral over a slice is, and one can work with integrals over slices, so that a function f in $U(G)$ defines not a function on $\Pi_{temp}(G(F_S))$ but a function on $\Pi_{temp}(G(F_S)) \times \mathfrak{a}_G$ (Arthur's notation in [28] is more complicated, because he has to envisage a more general possibility.) If π is a tempered representation of $G(F_S)$ induced from a tempered representation σ of a Levi factor over F_S, and this has to be taken to be a collection of perhaps quite disparate Levi factors, one for each $v \in S$, then the trace on a slice $tr(\pi(f^X))$ can be calculated as usual by attaching to f a function f_M on $M(F_S)$ that will be compactly supported on each slice *with respect to* $\mathfrak{a}_G$. Each $\mathfrak{a}_{M_v}$ is likely to be larger than $\mathfrak{a}_G$ and we can multiply σ with a character $e^{\Lambda(H_{\mathcal{M}})}$ of $\mathfrak{a}_{\mathcal{M}} = \sum_v \mathfrak{a}_{M_v}$ to obtain σ_Λ. Suppose π_Λ is induced from σ_Λ. Since we can calculate $tr(\pi_\Lambda(f^X))$ as $tr(f_M^X(\sigma_\Lambda))$, it will behave modulo $\mathfrak{a}_G$ like the Fourier transform of a function with compact support. This has to be taken into account when defining the space $V(G)$ and made part of the definition. Otherwise the map from $U(G)$ to $V(G)$ will certainly not be surjective. It is, moreover, no longer obvious that ϕ_M maps $U(G)$ to $V(M)$. This is proved in [28] and looks to be a second new feature of the argument, although the difficulties had already been broached in [10] because Arthur treats there the case that $U(G)$ is a space of compactly supported functions and $V(G)$ is defined accordingly. He points out to me that there is a difficulty that was overlooked in [10] (but not in [28]!). The split component $\mathfrak{a}_M$ of a Levi factor may grow when we pass from a global field to a local field because its dimension is that of the maximal split subgroup of the centre of M. This means, as we already observed, that the geometry that subtends the constructions of §2 generally changes and this was not noticed in [10]. The passage from the set S to a single place v and from the global geometry, or rather the intermediate geometry over the finite set S, to the local geometry is effected by the splitting and descent formulas of §7, §8 and §9 of [26] that are deduced from properties of the convex sets appearing in §2 of this report.

The formula (36), in which we now suppose that all the $c(M)$ are 1, can be applied to both $J^L(f) = J_M^L(\gamma, f)$ and $J^L(f) = J_M^L(\pi, f)$ (M ceasing for a moment to be the variable Levi factor on the right of (36) and becoming a given Levi factor). If we want to define I^L and $\hat{I}^L$ inductively, then at the stage G we first define I^G by

$$I^G(f) = J^G(f) - \sum_{\substack{M \in \mathcal{L}(M_0) \\ M \neq G}} I^M(\phi_M^G(f)).$$

Provided that all the I^M on the right are defined, the distribution I^G is well defined and easily shown to be invariant. The problem is then to show that this particular distribution is *supported on characters* so that it can be represented as a distribution $\hat{I}^G$ on $V(G)$.

Here as elsewhere, Arthur, in the style of Harish-Chandra, argues relentlessly in general, without examples. In contrast to those of Harish-Chandra, which are usually algebraic or analytic, Arthur's arguments are often implicitly geometric, so that they cry out for illustration. This is so for those based in one way or another on (G, M)-pairs, and is perhaps even more urgent for the arguments with which he establishes that the distributions I^G are supported on characters. There are two parts to the proof. The second uses the trace formula itself and an idea of Kazhdan to complete the final step which is to show by induction that the distributions $f \to I_M^G(\gamma, f)$ are supported on characters.

I do not pretend to understand the first part of the proof, but Arthur was kind enough to offer some insights that will certainly be useful to anyone studying [26] and [27]. In principle, it is to show that if the distributions $I_M^G(\gamma)$ are supported on characters then so are the distributions $I_M^G(\pi)$. In fact, the distributions $f \to I_M^G(\pi, f)$ do not appear alone in the trace formula; they appear only as in (14) in integrated form. It is an integrated form of $I_M^G(\pi)$ that is shown to be supported on characters. It is a direct formal consequence of the formula (36) that implicitly defines $I_M^G(\pi)$ in terms of the various $J_M^L(\pi)$ that $I_M^G(\pi)$ is 0 if π is a tempered representation of $M(\mathbb{A})$. There are, however, nontempered representations that will occur in the trace formula. For these, $I_M^G(\pi)$ may not be 0, but it is only an integrated form that appears and the distribution does vanish for π tempered. If the rank of M in G is one, these would be the representations on the imaginary axis, so that – in the simplest case that π is induced from σ_Λ with Λ not purely imaginary – the integral can be deformed leaving only the residues. These residues appear in the Fourier transform of $I_M^G(\pi)$ as exponentially decaying terms that are not cancelled by anything in $J_M^G(\gamma)$. Thus they must appear in $I_M^G(\gamma)$, so that the integrated $I_M^G(\pi)$ is supported on characters if the distributions $I_M^G(\gamma)$ are.

Thanks to the treatment of the geometric side in [21], the left side of (40) can be readily converted into the form (14) ([27] §3). For the spectral side, the basic formula is that of Theorem 8.2 of [14] for which I have referred to the original paper. The factor $a_M(\pi)$ that appears in (14) is defined by the scalar normalising factor in the global intertwining operators and thus can be expressed in terms of logarithmic derivatives of automorphic L-functions. Although in the formula (14), taken from [26], there appear to be no problems of convergence, this is not so. This is emphasized by the less elegant form of the formula that is given in [27]. It is to be hoped that the results of [M] will some day be improved to yield unconditional convergence.

7. First applications. An application of the trace formula that shows its importance as a primary analytic tool in the theory of automorphic forms is to the proof that the Tamagawa number of a simply-connected semisimple group is 1. This is a modern, general formulation of results of Eisenstein, Smith and Minkowski on class numbers of quadratic forms. There are two steps in the proof. For quasi-split groups it is an immediate consequence of the elements of the theory of Eisenstein series. For a general simply-connected semisimple group G, it is obtained by comparing the trace formula for G with that for its quasi-split form ([K1]). Although some of the results from [10], [21] and [24] are required, the proof, like that for quasi-split groups, is to be regarded as belonging to the very elements of the analytic theory of automorphic forms. There were other difficulties, arising from instability or endoscopy, to overcome, but Kottwitz was able to finesse these.

Although this is not clear from the publication dates, the first application of the general trace formula was, however, to base change for $\mathrm{GL}(n)$. As in the proof that the Tamagawa number is 1, it is a question of comparing two trace formulas, but now one is a twisted formula, for the group $\mathrm{GL}(n)$ over a cyclic extension E of the global base field F and its outer automorphism $A \to \sigma(A)$, σ a generator of the Galois group of E/F. The value of twisted trace formulas was first recognized in a particular case by H. Saito ([S]). The truncations required to imitate [7] and [9] were, I believe, first found, but not published, by Kottwitz. Because twisted trace formulas are so important for applications, in particular to the applications that appear to be his goal, Arthur has developed most of his later arguments in this broader context. Having enough to explain already, I preferred not to introduce this extension in the previous discussion and shall, by and large, leave it to the reader to consult either Arthur or his own imagination for a formulation of the necessary results and techniques.

When E/F is of prime degree, the trace formula compares the representation of $G' = G(\mathbb{A}_F)$ on $L^2(G(F)\backslash G(\mathbb{A}_F))$ with the natural action of the semidirect product $G = G(\mathbb{A}_E) \rtimes \sigma$, a subset of $G(\mathbb{A}_E) \rtimes \mathrm{Gal}(E/F)$. There is a norm map from conjugacy classes in G to conjugacy classes in G' that, as recognized by Shintani, links the harmonic analysis on the two groups, or more precisely the invariant harmonic analysis on G' and the twisted invariant harmonic analysis on the set G. Associated to the norm map is a transfer[8] from functions f on G to functions f' on G'. It is first defined locally, on the corresponding objects over local fields, and then, by taking products, globally. Given f the function f' is well defined for the purposes of invariant harmonic analysis, but not uniquely defined. Basically, orbital integrals of f' are equal to twisted orbital integrals of f. This permits a comparison of

[8] The role of primed and unprimed objects is at first glance disturbing. The guiding principal of [30], from which the notation is taken, is that the group over E is the primary object and the group over F only one of its *endoscopic* groups.

the twisted trace formula for f with the ordinary trace formula for f', from which the results on base change of [30] are obtained.[9]

A second, related comparison is between the multiplicative group G of a simple algebra over F and its quasi-split form $G' = \mathrm{GL}(n)$ that in principle yields extensions of the results obtained in [JL] for quaternion algebras. Since the two groups G and G' are now isomorphic at almost every place, the construction of f' from f is easier.

In both applications, as in many others, the starting point is the observation that the nonvanishing principal elliptic terms (those given by regular semisimple classes $\mathfrak{o}$ on which the orbital integrals of f' or f do not vanish) for the two trace formulas to be compared are in bijective correspondence and, because of the construction of f' from f, pairwise equal. It has then to be deduced from this, from the two trace formulas, and from general features of spectral analysis that the discrete parts of the two spectral sides are then equal. The conclusion is either base change from F to E for the group $\mathrm{GL}(n)$ or transfer of automorphic representations in the sense of functoriality, as in [56], from the multiplicative group of a simple algebra to its inner form $\mathrm{GL}(n)$.

The obstruction to an immediate inference are the remaining terms, on both the geometric and the spectral sides. There are four expressions appearing in the formula (14), or its twisted analogue, whose nature and meaning are obscure: $a^M(S,\gamma)$; $a^M(\pi)$; $I_M(\gamma, f)$; and $I_M(\pi, f)$. There will be similar expressions occuring in the trace formula for G'. It is, in fact, best to focus on the twisted case, thus on base change, in which more complications appear. Thanks to Hilbert's Theorem 90, which characterizes elements with the same norm, the geometric terms in the two trace formulas can be put in bijective correspondence. This can also be done for the spectral terms if some of those for G' are first collected together in a way that is straightforward in principle, but subtle in fact. It is then proved, using induction and the trace formulas, that each of the four expressions for G, and for γ or π as the case may be, is equal to the corresponding expression for G' and γ or π.

To collect terms, the local base change or lift[10] is first defined for tempered representations using the striking fact observed by Shintani for the group $\mathrm{GL}(2)$ (and for some groups over finite fields as

[9] There is a very difficult, and in large part unresolved, local problem implicit in the passage from f' to f, the *fundamental lemma*. It is a problem to whose partial solution many mathematicians have contributed, and the trace formula can usually only be applied after the pertinent form of it is available. For the purposes of [30], a special but important case treated by Kottwitz sufficed.

[10] I observe, for the sake of those to whom this use of the word *lift* suggests a retreat on my part, that, on the contrary, this is the sole context in which it seems an appropriate substitute for *transfer*. Geometrically a Galois extension corresponds to a covering and the transfer to a lift from the base space (the scheme defined by F) to the covering space (the scheme defined by E). Since the word *transfer* is now used in quite a different sense in endoscopy, there is good reason to search for a new term for functorial transfer, not otherwise taken, that, in contrast to *lift*, does violence neither to our mathematical intuition nor to our linguistic sensibilities (as in the pleonasm *base-change lift*) and that, in contrast to *transfer*, has a chance of gaining general acceptance.

well) that, apart perhaps from a sign, the twisted character of a tempered representation at an element γ in $G(E_v)$ is equal to the value of the character of the representation from which it is obtained by base change at the norm γ'. The argument is however global: the local existence is deduced from the trace formula itself, applied as suggested by Deligne and Kazhdan to a class of functions carefully chosen at two places – but otherwise arbitrary – so that, thanks to splitting formulas, all the difficult terms become zero. Once local base change is available, global base change is also defined although it is not assured, without further argument, that the base change of an automorphic representation is again automorphic. In principle, the collection is effected simply by putting together all representations with the same lift. It is, however, more difficult than that because Shintani's principle is not valid for many nontempered representations.

In passing, I observe that the continuing inability to establish the unconditional convergence of the right side of (14) entails complications in the proofs of [30].

There are four different types of equalities to establish, one for each of the four expressions. Two, those for $a^M(S,\gamma)$ and $a^M(\pi)$, can be considered as global in nature. In so far as the first of these are related to Tamagawa numbers, the desired equality is easier, but, as the discussion prior to (32) suggests, the factor $a^M(S,\gamma)$ will also contain a unipotent contribution to which there is no direct access. The other two are local and are in part amenable to the technique of Deligne-Kazhdan.

Many of the equalities follow from an induction assumption and splitting principles. (I confess that I have yet to understand this part of the argument in any serious way.) Thus when the two geometrical sides are subtracted, one from the other, there is a good deal of cancellation almost immediately. There is also a good deal of cancellation on the spectral side. Indeed, all that is left upon subtraction is the contribution of the discrete spectra. On the geometric side, what appears is a sum over Levi factors of the sum over γ' of

$$I_M(\gamma,f) - I_M(\gamma',f'), \tag{41}$$

the first arising from the ordinary trace formula, the second from the twisted formula.

If it were possible to remove the sum over M, replacing it by a given $M \neq G$ and if the differences in (41) were the orbital integrals of a function h on $M(\mathbb{A})$ of the type exploited by Deligne-Kazhdan, thus for which all terms of the geometric side of the trace formula except those corresponding to elliptic semisimple classes are zero, then we could apply the trace formula to M and h. Comparison of the spectral side of the new trace formula for M and h with the difference of the spectral sides of the trace formulas for G and G' leads to an equality between expansions for, on the one hand, a discrete spectrum and, on the other, a continuous spectrum. Both sides of such an equality have to be zero. This is, of course, a standard device in applications of the trace formula, but its successful use here requires that a large number of difficulties be overcome. In particular, it has to be shown that the singularities that may occur in the two terms of (41) individually but that may not appear in the difference if it is to be an orbital integral cancel each other. The upshot is, finally, the other three types of equality having been established along the way, that the remaining one, that for $a^M(\pi)$, is valid as well.

Both base change and the calculation of the Tamagawa number are the result of comparing the trace formula for two different groups. There is a second type of application, still in an embryonic stage, as only the simplest[11] of examples have been treated in order to avoid all difficulties arising either from the existence of any endoscopic groups but the principal one or from noncompactness ([K2]). Here the trace formula, applied to a special choice of function, is to be compared with a formula of Lefschetz type, usually for a correspondence on an incomplete algebraic variety over a finite field, the reduction of a Shimura variety. What can be concluded from [32] (see also [25], [31]) is that the trace formula yields a result that, with the help of endoscopy, will be comparable[12] with the results from the Lefschetz formula – when they become available.

8. Local harmonic analysis. The local harmonic analysis to which the heading refers is all a consequence of the local trace formula, but this is not immediately apparent from the earlier results.

The two papers [7], [9] on the first trace formula for a general group appeared in 1978 and 1980, but they were preceded by the paper [1] that dealt with groups of rank one. It was easy to anticipate on the basis of the formulas in this simpler case some of the weighted orbital integrals that would appear in general. They are essentially local objects and their analysis, at least at the infinite place where the analytic techniques developed by Harish-Chandra were available, could be taken up immediately. So it is not surprising that the earliest results on them ([2], [4]) appear before the general trace formula. For p-adic groups, the principal theorem of [4] does not appear until much later, in [23], but the proofs of [4] and [23] have much in common, Since the theorem turns out to be an illustration of the local trace formula, it is useful to look more carefully at it.[13]

For a compact group, the integral of a matrix coefficient of an irreducible representation over a conjugacy class is, easily identifiable constants aside, equal to the character of the representation

[11] Simple but difficult!
[12] With the help of the fundamental lemma!
[13] Arthur has observed to me that the results of [23] are exploited in the proof that singularities cancel in the comparison of the trace formula for $\mathrm{GL}(n)$ with that for the multiplicative group of a simple algebra (see p. 119 of [30]).

evaluated on the conjugacy class. It was discovered in the fifties that a similar formula is valid for square-integrable representations of any reductive group over $\mathbb{R}$, but only for elliptic conjugacy classes, thus the classes of those regular elements that are contained in a Cartan subgroup compact modulo the centre. The orbital integrals over the other classes is zero (the Selberg principle). It is shown in [4] that for each of the other Cartan subgroups T there is a similar formula, but the orbital integral must be replaced by the weighted orbital integral defined by the Levi factor M for which A_M is the maximal split torus in T. Moreover there is a sign that appears. Although the character of a square-integrable representation is not given by a simple formula off the elliptic elements, there are clear combinatorial procedures for computing it, so that, as a result of the formula of [4], the weighted orbital integrals of the matrix coefficients of square-integrable representations can be, in principle, calculated. It is the formula of [4] that allows Arthur to deduce from the trace formula a formula for the trace of the Hecke operators that it will be possible to compare with the Lefschetz formula.

However difficult it is to understand all the details of [4], the basic ideas are readily accessible to an older generation familiar with the techniques of Harish-Chandra. There is an induction on the dimension of A_M, starting from A_G which is of dimension zero and for which the theorem is that of Harish-Chandra for ordinary orbital integrals of matrix coefficients. It is a matter of using the action of the centre of the universal enveloping algebra to establish the appropriate differential equations and the appropriate boundary conditions at the points where two tori, one of which has one compact dimension more than the other, meet.

As observed, these formulas can be deduced from a general formula, the *local* trace formula, which resembles the trace formula and is proved in a similar way using the truncation of a kernel associated now to a function f on $G(F) \times G(F)$, F being a local field. The formula, whose possibility was suggested by Kazhdan, is proved in [39] and is used in a number of ways to overcome difficulties that arise in applications of the trace formula. Suppose that

$$f(y_1, y_2) = f_1(y_1)f_2(y_2)$$

is a smooth, compactly supported function on $G(F) \times G(F)$. The left and the right actions or $G(F)$ on itself yield an action of the product on $L^2(G(F))$ and thus an operator associated with f, $\varphi \to \varphi'$,

$$\varphi'(x) = \int_{G(F)} \int_{G(F)} f_1(u)f_2(y)\varphi(u^{-1}xy)dudy,$$

whose kernel is

$$K(x, y) = \int_{G(F)} f_1(xu)f_2(uy)du.$$

The local trace formula is obtained by restricting this kernel to the diagonal, truncating, and integrating. In appearance it resembles (14). On both sides there is a sum over Levi factors containing a given minimal one.

$$\sum_M \epsilon_{M/G} \frac{W_0^M}{W_0^G} \int_{\Gamma_{\mathrm{ell}}(M)} J_M(\gamma, f) d\gamma = \sum_M \epsilon_{M/G} \frac{W_0^M}{W_0^G} \int_{\Pi_{disc}(M)} a_{disc}^M(\pi) J_M(\pi, f) d\pi,$$

where for the sake of a compact formula I have set $(-1)^{dim(A_M/A_G)} = \epsilon_{M/G}$. More explicitly, the geometric terms are given by

$$J_M(\gamma, f) = D(\gamma) \int_{A_M(F)\backslash G(F)} \int_{A_M(F)\backslash G(F)} f_1(x_1^{-1}\gamma x_1) f_1(x_2^{-1}\gamma x_2) v_M(x) dx_1 dx_2,$$

where the weight $v_M(x) = v_M(x_1, x_2)$, $x = (x_1, x_2)$, is again defined by a (G, M)-family,

$$v_P(\Lambda, x) = e^{-\Lambda(H_P(x_2)) + \Lambda(H_{\bar{P}}(x_1))}.$$

The set $\Gamma_{\mathrm{ell}}(M)$ is the set of conjugacy classes (γ, γ), where γ is elliptic in M. Thus the maximal split torus in the centre of its centralizer is A_M.

The spectral side of the local trace formula, like the spectral side of the trace formula itself, is deceptive because there is considerable complexity veiled by the simple notation $\Pi_{disc}(M)$ and $J_M(\pi, f)$. The representation π is a representation of the product $G(F) \times G(F)$. To free the symbol M, we define $\Pi_{disc}(G)$ for an arbitrary G, which can then be one of the Levi factors M. The collection $\Pi_{disc}(G)$ is formed from constituents of induced representations

$$I_P(\check{\sigma} \otimes \sigma) = I_P(\check{\sigma}) \otimes I_P(\sigma),$$

in which σ, with contragredient $\check{\sigma}$ is a square-integrable representation of $M(F)$, M the Levi factor of P, invariant under an element t of the normalizer of A_M such that $\mathrm{ad}t - 1$ is invertible on $\mathfrak{a}_M/\mathfrak{a}_G$. The parabolic subgroup P is, as usual, taken to be defined over F and to contain a fixed minimal one. The factor $a_{disc}(M)$ is simpler than the corresponding global factor and involves no transcendental elements, but there is little to be gained by describing it in full. The terms $J_M(\pi, f)$ are, like the global factors, associated to (G, M)-families defined by intertwining operators, now local and unnormalized.

Although the local trace formula, as it first appears and as it is presented above, is not invariant, it, like the global formula, has an invariant form that is given in [36] and, in more detail although still at a somewhat breathtaking pace, in [40]. The formal appearance of the geometric side does not change much on passing to the invariant formula, although there is an important splitting formula[14] for the invariant distributions $I_M(\gamma, f)$ that permits their separation into those $I_M^L(\gamma, f_{i,Q})$, $i = 1, 2$, defined by weighted orbital integrals for the group G itself. The spectral side of the invariant formula has an elegant form,

$$\sum \epsilon_{M/G} \frac{W_0^M}{W_0^G} \int_{T_{disc}} i^M(\tau) r_M(\tau, f_1 \times f_2) d\tau. \tag{42}$$

The integration is over an elegantly chosen collection of *virtual* representations and $i^M(\tau)$ is an elementary factor. What is striking, although on reflection perhaps not surprising, is that in the term

$$r_M(\tau, f_1 \times f_2) = r_M(\tau, P)\Theta(\check{\tau}, f_{1,P})\Theta(\tau, f_{2,P}), \tag{43}$$

the contributions from f_1 and f_2 are separated into invariant distributions defined once again by virtual characters. I omit their explicit description; it is related to the decomposition of tempered induced representations into irreducible representations. The factor $r_M(\tau, P)$ is defined as usual by a (G, M)-pair, one attached to local scalar normalising factors. So the intertwining operators themselves appear on the spectral side of the invariant local trace formula only implicitly, through the virtual characters τ.

The earlier formulas ([4], [23]) for weighted orbital integrals as characters can be quickly deduced from the invariant local trace formula. Indeed they play, curiously enough, a role in its inductive proof. Although, as reported above, Arthur had earlier established, in the course of proving the existence of an invariant global trace formula, that the distributions $I_M(\gamma)$ are supported on characters, the local trace formula allows him to prove this by purely local methods that are, as before, inductive.

The new proof of the earlier formulas is sketched in [36]. A more detailed proof of a more general formula appears in [40]. Suppose M is a given Levi factor. Take f_1 to be a pseudocoefficient (almost the same thing as a matrix coefficient) of a discrete-series representation π and take f_2 to be such that its orbital integral is an approximation to a δ-function at a given element of $\Gamma_{\text{ell}}(M)$. The splitting formula for $I_M(\gamma, f)$ reduces the geometric side to the term involving an integral of $I_M(\gamma, f_1)I_G(\gamma, f_2)$. Taking the limit in f_2 leaves $I_M(\gamma, f_1)$ at the given element γ which is the same as $J_M(\gamma, f_1)$ because f_1 is a pseudocoefficient. On the other hand, $\Theta(\check{\tau}, f_1)$ will be zero unless τ is π and $\Theta(\pi, f_2)$ will be given by the value of the character of π at the given γ.

[14] It could perhaps just as well be called a descent formula. This report is, in spite of its many pages, far too brief and too hastily written. There is much material that, with more time and more space, I would have liked to treat at greater length. There are many things, among them splitting and descent, over which we might have lingered with both pleasure and profit. Arthur himself does not always find time to do his own ideas justice.

It has already been observed that the formula of [23] was used in [30] in the course of establishing the cancellation of singularities. A number of other consequences of the local trace formula are, or will be, invoked in later papers in which the trace formula is applied to functoriality. Orthogonality relations for characters are familiar from the theory of finite or compact groups, and even, following Harish-Chandra, for the discrete-series representations of reductive groups. There is an extension of the orthogonality relations of Harish-Chandra to the class of elliptic representations, which are in essence, the tempered representations whose character does not vanish on the elliptic elements. It is indispensable in comparisons of trace formulas. In abbreviated form, as it appears in [40], it is

$$\int_{\Gamma_{\mathrm{ell}}(G(F))} \Phi(\gamma)\overline{\Phi'(\gamma)} = \int_{T_{\mathrm{ell}}(G)} \phi(\tau)\overline{\phi'(\tau)}d\tau. \tag{44}$$

The left side is in essence the integral of the product of two characters over the elliptic set, thus exactly what appears in the formula of Harish-Chandra; the right side is a sum not over elliptic representations themselves, but over a more convenient set of virtual representations the span of whose characters is also the span of the characters of the elliptic representations.

9. Unipotent representations. Although functoriality (see, for example, [56]) as originally enunciated suggested many new problems in the theory of automorphic forms and provided a coherent way of viewing otherwise disparate phenomena, there was behaviour that did not obviously fall within its scope until Arthur began to reflect on the prerequisites to developing the stable trace formula beyond its embryonic stages ([KL]) and to comparing trace formulas for different groups. He was led to a series of elegant conjectures ([18], [34], [35]) that are rich in consequences, some of them proved ([ABV], [MW,Mö]), and that throw a great deal of light on outstanding problems in both the global theory and the local theory.

Recall that basic to the notions of functoriality is the possibility of attaching to a connected reductive group G over a global or a local field F a complex group, its L-group ${}^LG$, defined in essence by passing to the complex group[15] whose Cartan matrix is the transpose of that of G. When F is global, it is then possible to attach, by means of the Hecke operators, to any automorphic representation π a sequence $\{A_\mathfrak{p}\}$ of conjugacy classes in ${}^LG$, aptly called Hecke classes or with the case that $G = \{1\}$ in mind Frobenius-Hecke classes, and defined for almost all places $\mathfrak{p}$ of F.

If $G = \mathrm{GL}(2)$ then ${}^LG$ can be taken to be $\mathrm{GL}(2,\mathbb{C})$. Then the classes are described by the two eigenvalues $\alpha_\mathfrak{p}$ and $\beta_\mathfrak{p}$ of $A_\mathfrak{p}$. The Ramanujan conjecture, its generalization, the Ramanujan-Petersson

[15] It is an extension of this connected group by a Galois group or by a Weil group of a sufficiently large extension K of F, the choice between the two possibilities and the choice of extension being dictated by circumstances.

conjecture, and the generalization of that to Maaß forms and to arbitrary number fields assert that $|\alpha_{\mathfrak{p}}| = |\beta_{\mathfrak{p}}| = 1$ for all $\mathfrak{p}$ (at least for all $\mathfrak{p}$ for which the class is defined) if π occurs in the space of cusp forms. The complete conjecture asserts that the local components π_v are tempered at all places, including those at infinity. The generalization of this to arbitrary groups would, at the first, careless glance, suggest that the class $\{A_{\mathfrak{p}}\}$ met the maximal compact subgroup of ${}^LG$ if π occurs in the space of cusp forms or, when the archimedean places are included, that π_v is then tempered for all v. In this form, the conjecture was shown very early on to be invalid ([HPS]).

Arthur's conjectures suggest that it is invalid because the notion of cusp form, useful as it is for the spectral theory, has no structural significance. There is no reason, beyond ingrained scepticism, to doubt that the generalized Ramanujan conjecture is valid for $\mathrm{GL}(n)$ and Arthur's global conjectures and their meaning are best discussed in the firm conviction of its truth in this case. The noncuspidal discrete spectrum for $\mathrm{GL}(n)$ does not satisfy Ramanujan's conjecture and was never expected to. It was described by Moeglin-Waldspurger, who verified thereby a conjecture of Jacquet. Rather than presenting the explicit descriptions of the representations that occur, I describe the *Arthur* parameters that arise naturally from the constructions of the representations and provide the right optic for the general conjectures.

Among specialists there is a fairly widespread belief that the (unitary) irreducible cuspidal automorphic representations of $\mathrm{GL}(n)$ will be classified by the (bounded) n-dimensional irreducible representations of a group L_F. This group, if it exists, will be unconscionably large, the inverse limit of Lie groups that are almost compact (or sometimes of their complexifications), and presumably in no real sense explicitly describable. The mere fact of its existence would, however, be equivalent to fundamental properties of automorphic forms, so that it is well to keep this hypothetical group in mind and to make every effort to render it – or at first the consequences of its existence – concrete. As a result of the conjecture of Jacquet proved by Moeglin-Waldspurger, the discrete automorphic spectrum of $\mathrm{GL}(n)$ is parametrized by pairs (σ, m), where m divides n and σ is a cuspidal automorphic representation of the group $\mathrm{GL}(m)$. Thus σ is attached to an m-dimensional representation ϕ of L_F. Arthur suggests that we complete L_F by multiplying it by the Lie group $SU(2)$, take $l = n/m$ and take as the parameter of the representation π defined by (σ, m) the tensor product

$$\psi = \psi_{ss} \otimes \psi_{unip} = \phi \otimes \psi_{unip}\,,$$

where ψ_{unip} is the unique irreducible l-dimensional representation of $SU(2)$ and $\psi_{ss} = \phi$. The parameter ψ is referred to as the Arthur parameter of π. This formulation is to be justified by what it predicts for automorphic representations on groups other than $\mathrm{GL}(n)$ and, also, for what it predicts about the local theory. It turns out that it suggests a great deal, some of which can be proved.[16]

If π is attached to the parameter ψ then the sequence $\{A_{\mathfrak{p}}(\pi)\}$ is given by

$$A_{\mathfrak{p}}(\pi) = A_{\mathfrak{p}}(\sigma) \otimes \psi_{unip} \begin{pmatrix} \mathrm{Nm}\mathfrak{p}^{1/2} & 0 \\ 0 & \mathrm{Nm}\mathfrak{p}^{-1/2} \end{pmatrix}$$

So it is clear that π does not satisfy the Ramanujan conjecture when $l > 1$. The simplest examples are for $m = 1$. Then π is a one-dimensional representation.

For local fields F, especially but not alone the real and complex fields, the group L_F, whose irreducible representations of dimension n now classify the tempered discrete-series representations of $\mathrm{GL}(n, F)$, is simply described. It is the Weil group W_F of F if F is real or complex and the product $W_F \times SU(2)$, a variant of the Weil-Deligne group, if F is nonarchimedean. The tempered, unitary representations of $\mathrm{GL}(n, F)$ are classified by the bounded representations of L_F, irreducible or not, of dimension n and arbitrary irreducible representations of $\mathrm{GL}(n, F)$, unitary or not, by arbitrary n-dimensional representations of L_F.

A similar classification has been established for all reductive groups over archimedean fields, representations of dimension n being replaced by homomorphisms

$$\phi : \, L_F \to {}^L G.$$

There is every reason to expect that such a classification is valid over all local fields, but the questions elucidated by Arthur's conjectures arise already for the real field. It is not a single irreducible representation of $G(F)$ that is attached to ϕ, but a finite set, an L-packet, of inequivalent representations that have the same L-functions, so that as arithmetic objects they are hardly distinguishable. If the L-packet consists of tempered representations its internal structure is described by the theory of endoscopy, established for real groups by Shelstad. Otherwise this theory, which amounts to a collection of identities between linear combinations of characters of $G(F)$ and characters of lower dimensional groups, may fail. The local form of Arthur's conjectures reestablish the theory not for all representations but for a very important class that includes presumably all local factors of all unitary automorphic representations.

[16] We shall be concerned in the rest of the report with the researches of Arthur himself, which are, above all, consecrated to the consequences of this point of view for the trace formula and for functoriality. For consequences pertaining to the very difficult problem of classifying unitary representations of reductive groups over the real field and other local fields, see, for example, [ABV]; for those for the classification of the discrete automorphic spectrum, see [Mö].

These representations are to be parametrized by maps,

$$\psi = \phi \times \psi_{unip} \, : \, A_F = L_F \times SU(2) \to {}^L G, \tag{45}$$

so that the classification of the tempered packet is recovered from the ψ for which ψ_{unip} is trivial. The existence, even for the real field, of the unitary representations needed for the Arthur packet Π_ψ is by no means evident and, in general, not yet known.

The factor $SU(2)$ that appears in (45) is for nonarchimedean fields a second factor of this type. For both archimedean and nonarchimedean fields, and globally, the $SU(2)$ of (45) seems to have algebro-geometric content, a part of which, its relation to multiplication by the fundamental class in Hodge-Lefschetz theory, is described in §9 of [34]. For the trace formula, and in particular for global endoscopy, it is the character relations for Π_ψ that are critical. The pertinent object here is the group $\mathcal{S}_\psi$ of connected components of the centralizer S_ψ in the connected component of the identity in ${}^L G$ of the image $\psi(L_F)$. The character identities are defined by a pairing between the L-packet Π_ψ and $\mathcal{S}_\psi$. Each s in S_ψ defines an endoscopic group H, also reductive and of dimension no larger than that of G itself. There is (almost) an imbedding of ${}^L H$ in ${}^L G$ and ψ factors through ${}^L H$, so that it defines a packet for H. If χ_π is the character of π and $s \in S_\psi$ then

$$\sum_{\pi \in \Pi_\psi} d(s, \pi) \chi_\pi$$

is simply expressible in terms of the *stable* character associated to the Arthur packet for H associated to ψ. Moreover, for each π the function $s \to d(s, \pi)$ is essentially a character of $\mathcal{S}_\psi$

There is, of course, a troubling vagueness to these statements that is not present in [34], which, however, not only assumed a familiarity with the aims of endoscopy and its earlier results but also some acquaintance with the available unipotent representations (those associated to parameters ψ for which ψ_{ss} is trivial). Since endoscopy for tempered representations, and all the more for Arthur packets, is attempting to express concisely a great deal of information of which only fragments are presently available, fragments that are generated by difficult theories that draw on quite diverse sources for their methods, the choice is between vagueness or a careful treatment of many examples. Rather than describing local examples, I pass, however, to the global conjecture and to its implications.

In contrast to the local conjecture, which does not provide a parametrization of all unitary representations, the global conjecture is thought to provide a description, in terms of L_F, and thus in terms of representations for which the Ramanujan conjecture is satisfied, of the complete spectrum of

$L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))$. Unfortunately, however, and again in contrast to the local theory, the group L_F is not given externally to the theory, but has to be constructed within it, perhaps with the help of trace formulas that build on those of Arthur. It is not yet available, so that Arthur in his attempt to establish some quite general cases of functoriality ([45]) is forced to employ constructions that are strongly influenced by the possible existence of L_F but in which L_F is not allowed to appear explicitly.

This said, the global conjecture still envisages a parametrization by the maps (45), now global. The packets are usually infinite, being defined by the local packets and the relation between the local and global L_F. The global $\mathcal{S}_\psi$ has an even more subtle definition than the local and has a slightly different purpose. The global Arthur packets are products of the local. Each element of each local packet defines a function on $\mathcal{S}_\psi$. Multiplying them together, as is likely to be possible, we can attach to each element π of the global packet a function on $\mathcal{S}_\psi$ that determines (conjecturally) whether π occurs globally and, more precisely, whether it occurs in the discrete spectrum globally and with what multiplicity. There are subtle aspects to the pairing between π and $s \in \mathcal{S}_\psi$ that determines this multiplicity, $m_\psi(\pi)$. In particular there is an important sign factor that was introduced to respond to the exigencies of comparing trace formulas and that is easy, but dangerous, to overlook ([R]). Arthur does not suggest that $m_\psi(\pi) > 0$ for at most one ψ, on the contrary.

In the trace formula, the geometric side is to be regarded as the known side, the spectral side as the unknown. More precisely, it is usually the discrete part $I_{disc}(f)$ that is of primary concern in any comparison.[17] Recall that the spectral side is a sum of multidimensional integrals over automorphic representations. The discrete part of the spectral side is the partial sum over those terms for which the integrals are of dimension zero. It contains not alone the discrete spectrum in $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))$ but also what Arthur refers to as surviving remnants of Eisenstein series, resulting from passing to the limit in the $\sin(\tau T)/\tau$-integrals of §3.

We have not and shall not describe the notion of endoscopy, which has to be introduced for comparison of geometric sides, thus basically for circumventing the discrepancy between, on the one hand, conjugacy classes in $G(F)$ over a given ground field F, local or global, which is the pertinent notion for harmonic analysis, and, on the other, conjugacy classes in $G(F)$ but with respect to $G(\bar{F})$, which are all that can be compared when dealing, for example, with the same group over two different fields, one an extension of the other, or two groups that differ by an inner twisting.

[17] I suppress any complications in the notation arising from any possible, although unlikely, failure of absolute convergence on the spectral side. Moreover, I avoid treating twisted trace formulas explicitly although they are indispensable for Arthur's purposes. For the basic notions, the reader should consult [KS].

The effective use of this notion requires the fundamental lemma, or rather in any given case, a collection of fundamental lemmas, one for each endoscopic group. The lemma available, or assumed, it is possible to begin to *stabilize* the geometric side of the trace formula. It is the stabilized geometric sides of trace formulas that are compared; it is the stabilized spectral sides that yield information about the automorphic spectrum. According to [35], the stabilization of the discrete part of the trace formula should have the form,

$$I_{disc}(f) = \sum_H \iota(G,H) S^H_{disc}(f^H). \tag{46}$$

The sum is over the endoscopic groups, of which all have dimension less than that of G itself in the twisted trace formula and for the usual formula all but one. The function f^H is a transfer of f from G to H, similar to the transfer that appears in base change, a particular case of endoscopy. The global distribution S^H_{disc}, for which my notation, sacrificing precision to simplicity, departs from that of Arthur (itself uncertain), is defined intrinsically on H (or at worst on a closely related group), with no reference to G.

The validity of (46) is by no means evident, even when the geometric side can be fully treated; the purpose of [35] is to demonstrate that it is a consequence, neither easy nor immediate, of the conjectures[18] of [34], or, to put matters in a different light, that the global conjectures of [34], in particular the role of the sign character in the multiplicity formula, is a consequence, neither easy nor immediate, of the validity of (46). For someone with a stake in the outcome and a familiarity with examples, these arguments are persuasive, but, riddled with conjecture as they are, they lead, at first glance, to nothing definitive, and may even provoke some scepticism in outsiders. In the concluding pages of [35], Arthur addresses the problem of deducing definitive, fully established results from the intuitively and logically appealing, although incomplete, notions of [34] and [35]. The bulk of his investigations of the succeeding decade – the papers on the local trace formula and those on endoscopy, of which many have appeared ([44], [45], [46], [47], [48], [49], [50], [51], [53], [54], [55]) while others are still in preparation – are devoted to realising the program for the classical groups sketched there.[20]

[18] It is striking that conjectures for twisted multiplicities are used even to demonstrate (46) for the ordinary trace formula.

[20] He continues, however, to assume the fundamental lemma, a combinatorial problem that has turned out to be surprisingly difficult. Partial results on the fundamental lemma are available for groups of low rank, so that we can expect a complete form Arthur's results to be available soon for the classical groups $SO(4)$ and $SO(5)$ a case of considerable interest, especially in regard to the multiplicity formula, because $SO(5)$ and $Sp(4)$ are isogenous and $SO(5)$ defines some of the first Shimura varieties after the classical quotients of the upper-half plane. No-one yet knows for certain how the lemma is to be attacked in general, although a combination of methods from topology and algebraic geometry seems the most likely possibility. There are three forms of the fundamental lemma: for orbital integrals on the group; for orbital integrals on the Lie algebra; and for weighted orbital integrals. Arthur assumes them all!

10. Classical groups and endoscopy. For the group $\mathrm{GL}(n)$ neither stability nor multiplicity is a problem: stability because conjugation in $\mathrm{GL}(n,F)$ itself and conjugation in $\mathrm{GL}(n,F)$ with respect to $\mathrm{GL}(n,\bar{F})$ are the same; multiplicity because Shalika has shown many years ago that cuspidal automorphic representations occur in $L^2(\mathrm{GL}(n,F)\backslash\mathrm{GL}(n,\mathbb{A}))$ with multiplicity one. On the other hand functoriality predicts a close relation between automorphic representations of the classical groups. For many purposes the L-group of $\mathrm{GL}(n)$ can be taken to be $\mathrm{GL}(n,\mathbb{C})$; the L group of the symplectic group $Sp(n)$ to be $SO(n+1,\mathbb{C})$, that of $SO(n+1)$, n even, to be $Sp(n,\mathbb{C})$ and that of the split orthogonal group $SO(n)$, n even, to be $SO(n,\mathbb{C})$. A nonsplit orthogonal group in an even number of variables n splits over a quadratic extension E of the base field and its L-group is $O(n,\mathbb{C})$, to be considered a semidirect product of $SO(n,\mathbb{C})$ with $\mathrm{Gal}(E/F)$. Thus there is a canonical imbedding ϕ of the L-group of each and every classical group into that of a general linear group.

Functoriality predicts that if there is a homomorphism of L-groups

$$\phi:\ {}^LH \to {}^LG, \tag{47}$$

then to each automorphic representation π of H there is associated an automorphic representation Π of G such that the sequence $\{A_{\mathfrak{p}}(\Pi)\}$ is, at least for almost all $\mathfrak{p}$, equal to $\phi(A_{\mathfrak{p}}(\pi))\}$. Thus, in particular, it predicts that to each automorphic representation of a classical group, and in particular to each one occuring discretely in the L^2-spectrum of the group, there is associated an automorphic representation of a general linear group.

There is a simple and important outer automorphism of $\mathrm{GL}(n)$,

$$\theta:\ A \to J^{-1}{}^tA^{-1}J, \tag{48}$$

with

$$J = \begin{pmatrix} 0 & & & 1 \\ & & \cdot & \\ & \cdot & & \\ 1 & & & 0 \end{pmatrix}.$$

The final key observation is that the imbedded classical groups are all among the endoscopic groups for the corresponding twisted trace formula and that, in addition, all twisted endoscopic groups are either these imbedded classical groups or products of lower-dimensional classical groups. So, Arthur argues in [35], the possibility exists of using the twisted trace formula for θ to establish functoriality for the maps (47) and, in addition, to establish the multiplicity conjectures of [34] for the classical groups. The project that this suggests will, when completed, establish, once and for all, the overwhelming importance of the trace formula for the theory of automorphic forms, but it will not be, I hope, the end. Endoscopy and the trace formula will not be exhausted until, for example, the results of [30] are available for all groups. In addition, as I have maintained on more than occasion, but not yet in print, it is unlikely that the deepest and most fundamental of the problems posed by functoriality will be settled without extending the trace formula far beyond its present limits; and that will require, no doubt, if not a mastery then certainly a thorough understanding of Arthur's methods and, in my present estimation, a willingness to combine them with methods taken from a more traditional analytic number theory.

As observed, endoscopy and stabilization are entailed by local considerations. Endoscopic groups were first introduced to understand in detail how the characters of irreducible representations of the local groups $G(F)$, F a local field, failed to be functions on stable conjugacy classes. (In essence a stable class is the intersection of a conjugacy class in $G(\bar{F})$ with $G(F)$.) The discrepancy is described by stable functions or distributions (functions or distributions constant on stable classes) on endoscopic groups. In principle, a given endoscopic group appears in the description for a given representation π if the putative parameter of ψ factors through an imbedding, or something near an imbedding, of ${}^LH$ in ${}^LG$. All endoscopic groups are, by their very definition, quasi-split!

A fully developed global theory of endoscopy requires, first, that it be possible to define a *stable* trace $S = S_H$ for quasi-split groups and, secondly, that it be possible to express the ordinary *trace*, thus one side or the other of the trace formula, as a sum[20] over endoscopic groups of stable traces,

$$I(f) = I^G(f) = \sum_H \iota(G,H) S^H(f^H), \tag{49}$$

f^H being the transfer of f on $G(\mathbb{A})$ to $H(\mathbb{A})$. The ordinary trace is a trace only in an approximate sense; it is the value of either side of the equation (14). The stable trace is defined intrinsically on H. For the nontwisted trace formula, there is a unique principal endoscopic group, the quasi-split inner form of G. For the twisted trace formula, there may be more than one endoscopic group that is in some sense principal, even an infinite number of them.

The primitive calculations ([KL]) on which the notion of endoscopy was based were for the simplest part of the geometric side of (14), for the terms for which $M = G$ and γ is semisimple. There are no other terms if G is anisotropic! These calculations almost lead to an equality of the form (49), but there

[20] According to Arthur ([53]), only elliptic endoscopic groups are needed, thus those for which the maximal split component of the centre is that of G itself. This is surprising because the spectral side of [14] contains what Arthur refers to as surviving remnants of Eisenstein series, but a converzation with him suggests that what might be treated as a contribution from a nonelliptic endoscopic group can also be incorporated as a part of the contribution from a larger elliptic endoscopic group and that there are sound reasons for doing so.

are supplementary terms that have to be included for a complete matching. The group H may, for example, be abelian, so that there is no difference between I^H and S^H. This occurs for the group $\mathrm{SL}(2)$ as in [LL], where the geometric expansion of S^H for an abelian H, a one-dimensional torus, contains not only terms $f^H(\gamma)$, γ not central in G, which are matched by combinations of terms of the geometric side of (14) for which $M = G$ and γ is semisimple but also terms $f^H(\gamma)$, γ central in G, which are matched by combinations of terms of the geometric side of (14) for which $M = G$ but for which γ is unipotent and not semisimple.

So the early calculations are far from adequate; a great deal remains to be done. If G itself is a quasi-split group and the trace is the nontwisted trace, then G appears also on the right side of (49), which can then be regarded as defining inductively S^G, because $f^G = f$ and $\iota(G,G) = 1$. Otherwise, whether or not the trace is twisted, all terms in the relation (49) are defined and it becomes a relation to be proved.

The proof that the stable distributions S^G exist and that (49) is satisfied is awe inspiring. There is no question of doing justice to it here. Since (49) entails a comparison between the trace formulas for quite different groups, more attention has now to be paid to the exact definition of the normalising factors that were used to define the distributions $I_M(\gamma)$ and $I_M(\pi)$ and the coefficients $a^M(\pi)$ that appear in (14). Arthur prefers a new choice whose relation to the old, from which it does not differ greatly, is described in [47]. He also, before undertaking the proof of (49), modifies in various minor ways at the beginning of [53] the formulation of the trace formula. The result looks quite like (14), but the meaning of the terms has changed and the inner summations are over different sets. There are significant consequences of these apparently minor changes. For example, the suppression of the finite set S of places from the notation is shown, all changes made, to be legitimate.

Like I^G the stable distribution S^H will have two expansions, a geometric expansion and a spectral expansion. Moreover, the relation (49) will be a consequence of term-by-term equalities either on the geometric sides or on the spectral sides. Arthur is obliged to treat them while labouring under a tremendous handicap: neither the local parametrization nor an adequate local theory of endoscopy are yet available. Some substitute, factitious or in appearance factitious, has to be found for the first, and the second has to be established in the course of the treatment, with the help of elaborate inductions.[21]

Although the transfer of functions f on G to functions f^H on its endoscopic groups has been mentioned several times, I have not had the courage to describe it in any detail. The function f^H

[21] Fortunately, Waldspurger had already established in [W] that the existence of the transfer over nonarchimedean fields is a consequence of the fundamental lemma for the Lie algebra. Archimedean fields had been treated by Shelstad.

is not uniquely defined. Only its stable orbital integrals are and they are to be expressed as linear combinations of those of f with coefficients, the transfer factors, in whose definition ([LS]), which is not well understood, a large number of ingredients from the structure theory of semisimple groups and from Galois cohomology enter. One of the keys to the success of Arthur's treatment of endoscopy seems to be the adjoint relations of [44] and [49] which establish just the right amount of linear independence[22] to permit him to express all orbital integrals uniquely as linear combinations of stable orbital integrals of the functions f^H.

On the spectral side, an adequate formulation of the results requires the basis of the stable distributions provided by summing over the characters in an L-packet. In the absence of a local parametrization, except at the archimedean places, the packets are not available. So Arthur is forced ([44]) to define a weak substitute for this basis. The elements of the new basis have no particular significance, except that they are compatible with a natural filtration on the space of orbital integrals of functions in the Hecke algebra, whose first term is given by the space of functions whose orbital integrals are zero on all Cartan subgroups but the elliptic ones. To construct the basis he begins with a basis $\Phi_2(G)$, essentially arbitrary, of the stable elements in the bottom of the filtration, and not just for G but for each of its Levi factors M, and then obtains the full basis by induction of virtual characters from M to G.

The *elliptic* characters $\tau \in T_{ell}(G)$ that appear in (44) form a basis of the orbital integrals at the bottom of the filtration, or better, thanks to the pairing of [44], of the bottom of the filtration and of its dual. If H is an endoscopic group of G and ϕ an element of $\Phi_2(H)$, then $f \to \phi(f^H)$ lies in this dual, so that, for each element ϕ of $\Phi_2(H)$, there is an expansion,

$$\phi(f^H) = \sum_{T_{ell}(G)} \Delta(\phi,\tau) f_G(\tau), \tag{50}$$

where $f_G(\tau) = \tau(f)$ is simply the value of f, supposed to lie in the first term of the filtration.

Arthur proves in [44] that the right side of (50) is a stable distribution ϕ^H when G is quasi-split and $H = G$. Thus for any G and any of its endoscopic groups H the value $\phi^H(f^H)$ is well defined.[23] He shows that, in addition, $\phi(f^H) = \phi^H(f^H)$. The argument is similar to those in the proof (49). Distributions are defined for quasi-split groups and are shown to be stable. They then appear as the stable distributions on one side of a formula that is to be proved and that links the value of a distribution on a group G at a function f with the value of these stable distributions at the transfers f^H.

[22] If the local field is archimedean, the cohomological properties of G and, consequently, the basic notions appearing in endoscopy are not so simple as those for a nonarchimedean field. As a result the necessary independence fails, but it is reestablished in an elegant way, suggested by Vogan and Kottwitz, by considering several twisted forms of the group simultaneously ([49]).

[23] The analogous result for archimedean fields can be deduced from the papers of Shelstad.

There are four different kinds of identity that are proved in order to establish (49); there are *local* and *global* identities for the geometric side and for the spectral side. They correspond to the four different expressions that appear in (14) – in the original or in the modified form. The terms $I_M(\gamma, f)$ and $I_M(\pi, f)$ are local because the information they contain refers to only finitely many places;[24] the terms $a^M(\gamma)$ and $a^M(\pi)$ are on the other hand global. To formulate the four identities, the construction of the distributions $f \to \phi^H(f^H)$ has to be extended to a product of places. In addition, the construction has to be applied to the Levi factors M of G. So it is convenient to follow Arthur, as we did in the discussion of base change, and to denote an endoscopic group of G by G' and one of M by M'. Since we can apply the distributions $\phi^{M'}$ to the function f_M obtained by integrating the function f over the unipotent radical of a parabolic subgroup with M as Levi factor, it can be regarded as a distribution on G.

The endoscopic group M' will be a Levi factor of one or more endoscopic groups G' of G. For a given M', Arthur introduces a collection of these G', in which there may be repetitions, that he denotes $\mathcal{E}_{M'}(G)$.

The local geometric theorem is an expression for $I_M(\gamma, f)$ in the spirit of (49). Although $I_M(\gamma, f)$ is not an orbital integral in the usual sense, we can combine it with the transfer factors associated to an endoscopic subgroup M' of M to form

$$I_M(\delta, f) = \sum_\gamma \Delta_M(\delta, \gamma) I_M(\gamma, f). \tag{51}$$

If $I_M(\gamma, f)$ were the orbital integral of a function on M, $I_M(\delta, f)$ would be the stable orbital integral over the stable class δ of its transfer to M'. With the adjoint relations for the transfer factors on M the number $I_M(\gamma, f)$ can be recovered from the collection of $I_M(\delta, f)$. The analogue of the relation (49) for $I_M(\delta, f)$ is[25]

$$I_M(\delta, f) = \sum_{G' \in \mathcal{E}_{M'}(G)} \iota_{M'}(G, G') S^{G'}_{M'}(\delta, f^{G'}). \tag{52}$$

The factor $\iota_{M'}(G, G')$ is simple and vanishes if G' is not elliptic; $S^{G'}_{M'}(\delta, \cdot)$ is a stable distribution on G'. Since the same endoscopic group G' may appear more than once in the sum, which is over M' and G',

[24] As with the invariant trace formula, the adjective *local* might be better replaced by *semilocal*. Indeed there are two local theorem, one truly local and one semilocal. I shall not discuss the necessary splitting and descent theorems for the stable terms needed to mediate between them. My capacities are already overtaxed by the attempt to understand the statement of the four identities.

[25] I simplify the notation in order to pass over in silence a number of issues that would otherwise have to be treated explicitly. A particularly important point is that at the archimedean places, endoscopy does not function within the context of linear combinations of orbital integrals. The transfer of a stable orbital integral from an endoscopic group to G may introduce multiple layers, thus derivatives transverse to orbits.

it would be possible to collect terms and to express the right side as a sum over stable distributions on elliptic endoscopic groups evaluated at $f^{G'}$.

In order to deduce the expansion (49) from (52), it is necessary to relate the factors $a_M(\gamma)$ for G to those for its endoscopic groups. In the modified trace formula, the sums are no longer over γ in $M(\mathbb{Q})$ or, more generally, in $M(F)$ if the global field is taken to be arbitrary. Rather a fixed, finite set of places V is introduced and the sum is over the projection γ on $\prod_{v\in V} M(F_v)$ of elements $\dot\gamma$ in $M(F)$. There may be several $\dot\gamma$ that project on a given γ and the new $a^G(\gamma)$ are expressed in terms of the old $a^G(S,\dot\gamma)$, the set $S\supset V$ being taken sufficiently large, but otherwise arbitrary. The second ingredient in the stabilization of the geometric side of the trace formula is an identity,

$$a^G(\gamma)=\sum_{\mathcal{E}_{\mathrm{ell}}(G)}\sum_{\delta'}\iota(G,G')b^{G'}(\delta')\Delta_G(\delta',\gamma). \tag{53}$$

The outer sum is over the elliptic endoscopic groups of G. The inner sum should be ignored; δ' is, essentially, the image of γ in the endoscopic group G'. The factor $\Delta_G(\delta',\gamma)$ is the transfer factor. What is new in the formula are the factors $b^{G'}(\delta')$. Defined by the quasi-split group G' and the stable class δ' in it alone, they are the stable analogues of $a^G(\gamma)$. The formula (53) is, of course, proved not for just for G but for all its Levi factors.

The geometric expansion of the stable trace formulas appearing in (49) is then, for any quasi-split G,

$$S^G(f)=\sum_M\frac{W_0^M}{W_0^G}\sum_\delta b^M(\delta)S_M^G(\delta,f), \tag{54}$$

the inner sum on the right being over appropriate stable classes (defined, as in the modified trace formula itself, by semilocal objects). As Arthur observed to me, if G is not quasi-split, (49) is not entirely a straightforward consequence of substition of the local and global geometric theorems, thus of formulas (52) and (53), into the geometric side of the trace formula. In (49) the stable trace $S^{G'}$ (new notation) is applied to a function $f^{G'}$, defined by f on G, and in order to prove (49), it has to be shown that for such $f^{G'}$ and for δ that are not images of classes in G, we have $S_{M'}^{G'}(\delta,f^{G'})=0$.

Arthur does not, however, even for quasi-split groups, simply deduce (54) from (52) and (53). Rather he proves everything at once, by an inductive procedure, and the consequence of dealing with so many provisional objects simultaneously is a notational and conceptual thicket in which it is very easy to lose the way.

The terms on the spectral side of the modified trace formula, like those on the geometric side, are more cleanly broken into the product of a local term and a global term than those in the original formula. In particular, the π that appear are unramified outside of the fixed set V, so that outside of V they are characterized by the associated sequence of Hecke classes

$$\{A_{\mathfrak{p}}(\pi)\,|\,\mathfrak{p}\notin V\}. \tag{55}$$

The local (or better semilocal) term, $I_M(\pi, f)$, in the modified trace formula is associated to a representation

$$\pi = \prod_{v\in V} \pi_v.$$

It is understood, moreover that f is now a function on $\prod_{v\in V} G(F_v)$, extended when passing to the original trace formula to a function on $G(\mathbb{A}_F)$ by multiplying by a product of units in the local Hecke algebra. As a result, the global coefficient $a^G(\pi)$ that occurs in the modified formula is a sum over those automorphic representations, $\dot\pi$, unramified outside of V with the same local components at the places in V as π, but not merely of the original global coefficients $a^G(\dot\pi)$. Rather there is an additional sum over M and, because of the modification of the normalization of the intertwining operators, an additional factor defined by the sequence (55). A similar observation applies of course to the global $a^M(\pi)$.

The construction of the local distributions $f \to \phi^H(f^H)$ from the equation (50) yields, simply by taking products, semilocal distributions as well and provides a basis of the span of the characters on $G(F_V) = \prod_{v\in V} G(F_v)$ given by transfers of stable distributions on endoscopic groups.[26] The analogue of (51) is[27]

$$I_M(\phi, X, f) = \sum_{\pi} \Delta(\phi,\pi) I_M(\pi, X, f), \tag{56}$$

a semilocal definition. The analogue of (52) is[28]

$$I_M(\phi, X, f) = \sum_{G'\in\mathcal{E}_{M'}(G)} \iota_{M'}(G, G') S^{G'}_{M'}(\phi, X, f^{G'}). \tag{57}$$

[26] For the complete theory Arthur refers to preprints that are not yet available. It is necessary to guess at their contents. In particular, it appears that he replaces the sum of (50) by a sum over π rather than over the virtual characters τ. Since this is a formal matter, for one is every bit as much of a basis as the other provided we take care to enlarge the discussion immediately to all characters and not to confine it to elliptic characters, I admit with no compunction factors $\Delta(\phi, \pi)$.

[27] We have already seen X in §6. It is necessary for technical reasons because $I_M(\pi, f)$ appears in an integrated form in the trace formula.

[28] Once again, there are equivalence relations on which I do not insist.

The analogue of (53) is

$$a^G(\pi) = \sum_{\mathcal{E}_{\mathrm{ell}}(G)} \sum_{\phi'} \iota(G, G') b^{G'}(\phi') \Delta_G(\phi', \pi). \tag{58}$$

Finally, the analogue of (54) is

$$S^G(f) = \sum_M \frac{W_0^M}{W_0^G} \sum_\phi b^M(\phi) S_M^G(\phi, f), \tag{59}$$

Thus (54) and (59) lead to two equal expansions for $S^G(f)$ and therefore to a stable trace formula.

The proofs of the four identities are completed in the two papers [54] and [55]. They are similar in many respects to those for base change. In particular, the most difficult case of the global geometric theorem will be that for γ equal to the identity in the group. There is more to be said about it, but at this point my courage fails me. So I take farewell of the reader, wishing him, like an eminent predecessor on another occasion, God-speed but leaving him to continue the journey on his own.

Bibliography of James G. Arthur

[1] *The Selberg trace formula for groups of F-rank one*, Ann. of Math. 100 (1974), 326–385.

[2] *Some tempered distributions on groups of real rank one*, Ann. of Math. 100 (1974), 553–584.

[3] *A theorem on the Schwartz space of a reductive Lie group*, Proc. Nat. Acad. Sci. 72 (1975) 4718–4719.

[4] *The characters of discrete series as orbital integrals*, Inv. Math. 32 (1976), 205–261.

[5] *A truncation process for reductive groups*, Bull. Amer. Math. Soc. 83 (1977), 748–750.

[6] *Eisenstein series and the trace formula*, Proc. Sympos. Pure Math. 33, Part 1 (1979), 253–274.

[7] *A trace formula for reductive groups I: Terms associated to classes in $G(Q)$*, Duke Math. J. 245 (1978), 911–952.

[8] *Harmonic analysis of invariant distributions*, Proc. Conf. on Lie Theories, (Queen's University, 1977), Queen's Papers in Pure and Appl. Math 48 (1978), 384–393.

[9] *A trace formula for reductive groups II: Applications of a truncation operator*, Compositio Math. 40 (1980) 87–121.

[10] *The trace formula in invariant form*, Ann. of Math. 114 (1981), 1–74.

[11] *Automorphic representations and number theory*, C.M.S. Conf. Proc. I (1980), 3–51.

[12] *On the inner product of truncated Eisenstein series*, Duke Math. J. 49 (1982), 35–70.

[13] *On a family of distributions obtained from Eisenstein series I: Application of the Paley-Wiener theorem*, Amer. J. Math. 104 (1982), 1243–1288.

[14] *On a family of distributions obtained from Eisenstein series II: Explicit formulas*, Amer. J. Math. 104 (1982), 1289–1336.

[15] *A Paley-Wiener theorem for real reductive groups*, Acta Math. 150 (l983), 1–89.

[16] *The trace formula for reductive groups*, ournées Automorphes, Publications Mathematiques de L'Universite Paris VII 1983, 1–41.

[17] *Multipliers and a Paley-Wiener theorem for real reductive groups*, Representation Theory of Reductive Groups, Birkhäuser 1983, 1–19.

[18] *On some problems suggested by the trace formula*, Lie Group Representations II, vol. 1041 Springer Lecture Notes in Math. (1984), 1–49.

[19] *The trace formula for noncompact quotient*, Proceedings of International Congress of Mathematicians (Warsaw, 1983) PWN, Warsaw, 1984, 849–859.

[20] *A measure on the unipotent variety*, Canad. J. Math. 37 (1985), 1237–1274.

[21] *On a family of distributions obtained from orbits*, Canad. J. Math. 38 (1986), 179–214.

[22] *The Fourier transform of weighted orbital integrals on* $\mathrm{SL}(2, R)$ (with R. Herb and P. Sally), Contemporary Math. 53 (1986), 17–37.

[23] *The characters of supercuspidal representations as weighted orbital integrals*, Proc. Indian Acad. Sci. 97 (1987), 3–19.

[24] *The local behaviour of weighted orbital integrals*, Duke Math. J. 56 (1988), 223–293.

[25] *Characters, harmonic analysis, and an* L^2*-Lefschetz formula*, Proc. Sympos. Pure Math. 48 (1988), 167–179.

[26] *The invariant trace formula. I. Local theory*, J. Amer. Math. Soc. 1 (1988), 323–383.

[27] *The invariant trace formula. II. Global theory* J. Amer. Math. Soc. 1 (1988), 501–554.

[28] *Intertwining operators and residues. I. Weighted characters* J. Funct. Anal. 84 (1989), 19–84.

[29] *Intertwining operators and residues II. Invariant distributions*, Compositio Math. 70 (1989), 51–99.

[30] *Simple Algebras, Base Change and the Advanced Theory of the Trace Formula* (with L. Clozel), Annals of Math. Studies 120, Princeton University Press, 1989.

[31] *The trace formula and Hecke operators*, Number Theory, Trace Formulas and Discrete Groups, Academic Press, 1989, 11–27.

[32] *The* L^2*-Lefschetz numbers of Hecke operators*, Invent. Math. 97 (1989), 257–290.

[33] *Harmonic analysis of tempered distributions on semisimple Lie groups of real rank one*, in: Representation Theory and Harmonic Analysis on Semisimple Lie Groups (eds. P. Sally and D. Vogan), Math. Surveys Monographs 31, Amer. Math. Soc., 1989, 13–100.

[34] *Unipotent automorphic representations: Conjectures*, Asterisque 171–172 (1989), 13–71.

[35] *Unipotent automorphic representations: Global Motivation*, in: Automorphic Forms, Shimura Varieties and L-functions, vol. I (Ann Arbor, MI, 1988), Perspect. Math. 10 (1989), 1–75.

[36] *Towards a local trace formula*, Proceedings of First JAMI Conference, Johns Hopkins University Press, 1990, 1–23.

[37] *Lectures on automorphic L-functions* (with S. Gelbart), L-functions and Arithmetic, (Durham, 1990), London Math. Soc. Lecture Note Ser. 153 (1991), 1–59.

[38] *Some problems in local harmonic analysis*, Harmonic Analysis on Reductive Groups (Brunswick, ME, 1989), Progr. Math. 101 (1991), 57–78.

[39] *A local trace formula*, Pub. Math. I.H.E.S. 73, (1991), 5–96.

[40] *On elliptic tempered characters*, Acta Math. 171 (1993), 73–138.

[41] *On the Fourier transforms of weighted orbital integrals*, J. Reine Angew. Math. 452 (1994), 163–217.

[42] *The trace Paley-Wiener theorem for Schwartz functions*, Contemp. Math. 177 (1994), 171–180.

[43] *L^2-cohomology and automorphic representations*, in: Canadian Mathematical Society, 1945-1995, vol. 3, 1996, 1–17.

[44] *On local character relations*, Selecta Math. 2 (1996), 501–579.

[45] *The problem of classifying automorphic representations of classical groups*, CRM Proceedings and Lecture Notes 11 (1997), 1–12.

[46] *Stability and endoscopy: informal motivation*, Proc. Sympos. Pure Math. 61 (1997), 433–442.

[47] *Canonical normalization of weighted characters and a transfer conjecture*, C. R. Math. Rep. Acad. Sci., Canada 20 (1998), 35–52.

[48] *Towards a stable trace formula*, in: Proceedings of the International Congress of Mathematicians, vol. II (Berlin, 1998), Doc. Math. 1998, Extra vol. II, 507–517.

[49] *On the transfer of distributions: weighted orbital integrals*, Duke Math. J. 99 (1999), 209–283.

[50] *Endoscopic L-functions and a combinatorial identity*, Canad. J. Math. 51 (1999), 1135–1148.

[51] *Stabilization of a family of differential equations*, Mathematical Legacy of Harish-Chandra (Baltimore, MD, 1998), Proc. Sympos. Pure Math. 68 (2000), 77–95.

[52] *Harmonic analysis and group representations*, Notices of AMS 47 (2000), 26–34.

[53] *A stable trace formula I. General expansions*, preprint.

[54] *A stable trace formula II. Global descent*, to appear in Inv. Math. .

[55] *A stable trace formula III. Proof of the main theorems*, preprint.

[56] *The principle of functoriality*, submitted to Mathematical Challenges of the 21st Century.

Supplementary references

[ABV] J. Adams, D. Barbasch and D. Vogan, *The Langlands Classification and Irreducible Representations for Real reductive Groups*, Progr. Math 104, Birkhäuser, 1992.

[CD] L. Clozel and P. Delorme, *Le théorème de Paley-Wiener invariant pour les groupes de Lie réductifs*, Inv. Math. 77 (1984), 427–453; *Sur le théorème de Paley-Wiener invariant pour les groupes de Lie réductifs réels*, C. R. Acad. Sci. Paris, Sér. 1 Math. 300 (1985), 331–333.

[HPS] R. Howe and I. Piatetski-Shapiro, *A counterexample to the "generalized Ramanujan conjecture" for (quasi)-split groups*, Automorphic forms, representations and L-functions (Corvallis, OR, 1977), Part 1, Proc. Sympos. Pure Math. 33 1979, 315–322.

[JL] H. Jacquet and R. P. Langlands, *Automorphic forms on* $\mathrm{GL}(2)$, Springer Lecture Notes 114 (1970).

[J] H. Jacquet, *On the residual spectrum of* $\mathrm{GL}(n)$, Lie group representations, II, Springer Lecture Notes 1041, (1984), 185–208.

[K1] R. E. Kottwitz, *Tamagawa numbers*, Ann. of Math. 127 (1988), 629–646.

[K2] ———, *On the λ-adic representations associated to some simple Shimura varieties*, Inv. Math. 108 (1992), 653–665.

[KL] R. Langlands, *Les débuts d'une formule des traces stables*, Publ. Math. Univ. Paris VII 13, Paris, 1983; R. Kottwitz, *Stable trace formula: cuspidal tempered terms*, Duke Math. J. 51 (1984), 611–650; ———, *Stable trace formula: elliptic singular terms*, Math. Ann. 275 (1986), 365–399.

[KS] R. Kottwitz and D. Shelstad, *Foundations of twisted endoscopy*, Astérisque 255 1999.

[L] R. P. Langlands, *Rank-one residues of Eisenstein series*, Israel Math. Conf. Proc. 3 (1990), 111–125.

[LL] J.-P. Labesse and R. P. Langlands, *L-indistinguishability for* $\mathrm{SL}(2)$, Canad. J. Math. 31 (1979), 726–785.

[LS] R. P. Langlands and D. Shelstad, *On the definition of transfer factors*, Math. Ann. 278 (1987), 219–271.

[MW] C. Moeglin and J.-L. Waldspurger, *Le spectre résiduel de* $\mathrm{GL}(n)$, Ann. Sci. École Norm. Sup. 22 (1989), 605–674.

[Mö] C. Moeglin, *Représentations unipotentes et formes automorphes de carré intégrable*, Forum Math. 6 (1994), 651–744.

[Mü] W. Müller, *The trace class conjecture in the theory of automorphic forms*, Ann. of Math. 130 (1989), 473–529.

[R] J. Rogawski, *Automorphic representations of unitary groups in three variables*, Annals of Math. Studies 123, Princeton University Press 1990; *The multiplicity formula for A-packets*, in The zeta functions of Picard modular surfaces, (eds. R. P. Langlands and D. Ramakrishnan), Centre de recherches mathématiques, Univ. de Montréal, 1992, 395–419.

[S] H. Saito, *Automorphic forms and extensions of number fields*, Lectures in Mathematics 8, Kyoto University, 1975.

[W] J.-L. Waldspurger, *Le lemme fondamental implique le transfert*, Comp. Math. 105 (1997), 153–236.