Notes on the Generalized Ramanujan Conjectures

by

Peter Sarnak*

(*Fields Institute Lectures - June, 2003*)

Contents:

§1. GL_n

Ramanujan's original conjecture is concerned with the estimation of Fourier coefficients of the weight 12 holomorphic cusp form Δ for $SL(2,\mathbb{Z})$ on the upper half plane $\mathbb{H}$. The conjecture may be reformulated in terms of the size of the eigenvalues of the corresponding Hecke operators or equivalently in terms of the local representations which are components of the automorphic representation associated with Δ. This spectral reformulation of the problem of estimation of Fourier coefficients (or more generally periods of automorphic forms) is not a general feature. For example, the Fourier coefficients of Siegel modular forms in several variables carry more information than just the eigenvalues of the Hecke operators. Another example is that of half integral weight cusp forms on $\mathbb{H}$ where the issue of the size of the Fourier coefficients is equivalent to special instances of the Lindelof Hypothesis for automorphic L-functions (see [Wal], [I-S]). As such, the general problem of estimation of Fourier coefficients appears to lie deeper (or rather farther out of reach at the present time). In these notes we discuss the spectral or representation theoretic generalizations of the Ramanujan Conjectures (GRC for short).

We begin with some general comments. In view of Langlands Functoriality Conjectures (see [A1]) all automorphic forms should be encoded in the GL_n automorphic spectrum. Moreover, Arthur's recent conjectural description of the discrete spectrum for the decomposition of a general

*Department of Mathematics, Princeton University and the Courant Institute of Mathematical Sciences.

group [A2],[A3] has the effect of reducing the study of the spectrum of a classical group for example, to that of GL_n. Form this as well as the point of view of L-functions, GL_n plays a special role. Let F be a number field, $\mathbb{A}_F$ its ring of adeles, v a place of F (archimedian or finite) and F_v the corresponding local field. Let GL_n be the group of $n \times n$ invertible matrices and $GL_n(\mathbb{A}_F)$, $GL_n(F)$, $GL_n(F_v) \cdots$ be the corresponding group with entries in the indicated ring. The abelian case GL_1, is well understood and is a guide (though it is way too simplistic) to the general case. The constituents of the decomposition of functions on $GL_1(F)\backslash GL_1(\mathbb{A}_F)$ or what is the same, the characters of $F^*\backslash\mathbb{A}_F^*$, can be described in terms of class field theory. More precisely, if W_F is Weil's extension of the Galois group $GAL(\bar{F}/F)$ then the 1-dimensional representations of $F^*\backslash\mathbb{A}_F^*$ correspond naturally to the 1-dimensional representations of W_F (see [Ta]). As Langlands has pointed out [Lang1] it would be very nice for many reasons, to have an extended group L_F whose n-dimensional representations would correspond naturally to the automorphic forms on GL_n. The basic such forms are constituents of the decomposition of the regular representation of $GL_n(\mathbb{A}_F)$ on $L^2(Z(\mathbb{A}_F)GL_n(F)\backslash GL_n(\mathbb{A}_F), w)$. Here Z is the center of GL_n and w is a unitary character of $Z(\mathbb{A}_F)/Z(F)$. In more detail, the L^2 space consists of functions $f : GL_n(\mathbb{A}_F) \longrightarrow \mathbb{C}$ satisfying $f(\gamma z g) = w(z)f(g)$ for $\gamma \in GL_n(F)$, $z \in Z(\mathbb{A}_F)$ and

$$\int_{Z(\mathbb{A}_F)GL_n(F)\backslash GL_n(\mathbb{A}_F)} |f(g)|^2\, dg \, < \, \infty\,. \tag{1}$$

Notwithstanding, the success by Harris-Taylor and Henniart [H-T] giving a description in the local case of the representations of $GL_n(F_v)$ in terms of n-dimensional representations of the Deligne-Weil group W_F', or the work of Lafforgue in the case of $GL_n(F)$ where F is a function field over a finite field, it is difficult to imagine a direct definition of L_F in the number field case. My reason for saying this is that L_F would have to give through its finite dimensional representations, an independent description of the general Maass cusp form for say $GL_2(\mathbb{A}_\mathbb{Q})$ (see [Sa] for a recent discussion of these). These are eigenfunctions of elliptic operators on infinite dimensional spaces with presumably highly transcendental eigenvalues. Arthur in his definition [A3] of L_F gets around this difficulty by using among other things the GL_n cusp forms as building blocks for the construction of the group. With this done, he goes on to give a much more precise form of the general functoriality conjectures.

We turn to GL_n and a description of the generalized Ramanujan Conjectures. According to the general theory of Eisenstein series $L^2(Z(\mathbb{A}_F)GL_n(F)\backslash GL_n(\mathbb{A}_F), w)$ decomposes into a discrete part and a continuous part. The discrete part coming from residues of Eisenstein series, as well as the continuous part coming from Eisenstein series, are described explicitly in [M-W]. They are given in terms of the discrete spectrum on GL_m, $m < n$. This leaves the cuspidal spectrum as the fundamental part. It is defined as follows:

$$\begin{aligned} &L^2_{\text{cusp}}(Z(\mathbb{A}_F)GL_n(F)\backslash GL_n(\mathbb{A}_F), w) \\ &\qquad = \{f \text{ satisfying (1) and } \int_{N(F)\backslash N(\mathbb{A}_F)} f(ng)dn = 0 \\ &\text{for all unipotent radicals } N \text{ of proper parabolic subgroups } P \text{ of } G(F)\} \end{aligned} \tag{2}$$

The decomposition into irreducibles of $GL_n(\mathbb{A}_F)$ on L^2_{cusp} is discrete and any irreducible constituent π thereof is called an automorphic cusp form (or representation). Now such a π is a tensor product $\pi = \bigotimes_v \pi_v$, where π_v is an irreducible unitary representation of the local group $GL_n(F_v)$. The problem is to describe or understand which π_v's come up this way. For almost all v, π_v is unramified, that is π_v has a nonzero K_v invariant vector, where K_v is a maximal compact subgroup of $GL_n(F_v)$. If v is finite then $K_v = GL_n(O(F_v))$, $O(F_v)$ being the ring of integers at v. Such "spherical" π_v can be described using the theory of spherical functions (Harish-Chandra, Satake) or better still in terms of the Langlands dual group ${}^LG$. For $G = GL_n$, ${}^LG = GL(n, \mathbb{C})$ (or rather the connected component of ${}^LG$ is $GL(n, \mathbb{C})$ but for our purposes here this will suffice) and an unramified representation π_v is parameterized by a semi-simple conjugacy class

$$\alpha(\pi_v) = \begin{bmatrix} \alpha_1(\pi_v) & & & & \\ & \ddots & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \alpha_n(\pi_v) \end{bmatrix} \in {}^LG \tag{3}$$

as follows: Let B be the subgroup of upper triangular matrices in GL_n. For $b \in B(F_v)$ and $\mu_1(v), \dots, \mu_n(v)$ in $\mathbb{C}$ let χ_μ be the character of $B(F_v)$

$$\chi_\mu(b) = |b_{11}|_v^{\mu_1} |b_{22}|_v^{\mu_2} \cdots |b_{nn}|_v^{\mu_n}. \tag{4}$$

Then $\psi_\mu = \mathrm{Ind}_{B(F_v)}^{G(F_v)} \chi_\mu$ yields a spherical representation of $G(F_v)$ (the induction is normalized unitarily and at $\mu's$ for which it is reducible we take the spherical constituent). ψ_μ is equivalent to $\psi_{\mu'}$ with μ and μ' considered $\mod \mathbb{Z}2\pi i/\log N(v)$, iff $\mu = \sigma\mu'$ where σ is a permutation. In this notation $\alpha(\pi_v)$ corresponds to $\psi_{\mu(v)}$ by setting $\alpha_j(\pi_v) = N(v)^{\mu_j(v)}$ for $j = 1, 2, \dots n$. The trivial representation of $G(F_v)$, or constant spherical function corresponds to

$$\mu = \left(\frac{n-1}{2}, \frac{n-3}{2} \cdots \frac{1-n}{2}\right) \tag{5}$$

In terms of these parameters the local L-function $L(s, \pi_v)$ corresponding to such an unramified

π_v has a simple definition:

$$\begin{aligned} L(s,\pi_v) &= \det\left(I - \alpha(\pi_v)\, N(v)^{-s}\right)^{-1} \\ &= \Pi_{j=1}^{n}\left(1 - \alpha_j(\pi_v)\, N(v)^{-s}\right)^{-1} \end{aligned} \tag{6}$$

if v is finite, and

$$L(s,\pi_v) = \Pi_{j=1}^{n}\, \Gamma_v(s - \mu_j(\pi_v))\,,$$

$$\text{with } \Gamma_v(s) = \pi^{-s/2}\,\Gamma\left(\frac{s}{2}\right)\,, \text{ if } F_v \cong \mathbb{R} \text{ and } \Gamma_v(s) = (2\pi)^{-s}\,\Gamma(s) \text{ if }, \; F_v \cong \mathbb{C}\,. \tag{7}$$

More generally if $\rho : {}^LG \longrightarrow GL(\nu)$ is a representation of ${}^LG$ then the local L-function is defined by

$$L(s,\pi_v,\rho) = \det\left(I - \rho(\alpha(\pi_v))N(v)^{-s}\right)^{-1}. \tag{8}$$

We digress and discuss some local harmonic analysis for more general groups. Let $G(F_v)$ be a reductive group defined over a local field F_v. We denote by $\widehat{G(F_v)}$ the unitary dual of $G(F_v)$, that is the set of irreducible unitary representations of $G(F_v)$ up to equivalence. $G(F_v)$ has a natural topology, the Fell topology, coming from convergence of matrix coefficients on compact subsets of $G(F_v)$. Of particular interest is the tempered subset of $\widehat{G(F_v)}$ which we denote by $\widehat{G(F_v)}_{\text{temp}}$. These are the representations which occur weakly (see [Di]) in the decomposition of the regular representation of $G(F_v)$ on $L^2(G(F_v))$ or what is the same thing $\mathrm{Ind}_H^{G(F_v)}1$ where $H = \{e\}$. If $G(F_v)$ is semi-simple then the tempered spectrum can be described in terms of decay of matrix coefficients of the representation. For ψ a unitary representation of $G(F_v)$ on a Hilbert space H, these are the functions $F_w(g)$ on $G(F_v)$ given by $F_w(g) = \langle\psi(g)w, w\rangle_H$ for $w \in H$. Clearly, such a function is bounded on $G(F_v)$ and if ψ is the trivial representation (or possibly finite dimensional) then $F_w(g)$ does not go to zero as $g \to \infty$ (we assume $w \neq 0$). However, for other ψ's these matrix coefficients do decay (Howe-Moore [H-M])† and the rate of decay is closely related to the "temperdness" of ψ and is important in applications. In particular, ψ is tempered iff its K_v-finite matrix coefficients are in $L^{2+\epsilon}(G(F_v))$ for all $\epsilon > 0$.

For spherical representations (and in fact for the general ones too) one can use the asymptotics at infinity of spherical functions (that is K_v bi-invariant eigenfunctions of the Hecke algebra) to determine which are tempered. For $GL_n(F_v)$ this analysis shows that the π_v defined in (3) and (4) is tempered iff

$$|\alpha_j(\pi_v)| = 1 \text{ for } j = 1,2,\ldots n \text{ if } v \text{ finite} \tag{9}$$

and

$$\Re(\mu_j(\pi_v)) = 0 \text{ if } v \text{ archimedian}\,. \tag{10}$$

†if say $G(F_v)$ is simple

For ramified representations of $G(F_v)$ one can give a similar description of the tempered representations in terms of Langlands parameters (see Knapp-Zuckerman [K-Z] for v archimedian).

To complete our digression into more general local groups $G(F_v)$ we recall property T. $G(F_v)$ has property T if the trivial representation is isolated in $\widehat{G(F_v)}$. Kazhdan in introducing this property showed that if $G(F_v)$ is simple and has rank at least two then it satisfies property T. One can quantify this property in these cases (as well as in the rank one groups which satisfy property T) by giving uniform estimates for the exponential decay rates of any non-trivial unitary representation of $G(F_v)$. In [Oh], Oh gives such bounds which are in fact sharp in many cases (such as for $SL_n(F_v)$ $n \geq 3$ and $Sp_{2n}(F_v)$). In the case that $F_v \cong \mathbb{R}$, Li [Li1] determines the largest $p = p(G(F_v))$ for which every non-trivial representation of $G(F_v)$ is in $L^{p+\epsilon}(G(F_v))$ for all $\epsilon > 0$. Besides the isolation of the trivial representation in $\widehat{G(F_v)}$ it is also very useful to know which other representations are isolated. For v archimedian and π_v cohomological (in the sense of Borel and Wallach [B-W]) Vogan [Vo1] gives a complete description of the isolated points.

We return now to the global setting with $G = GL_n$ and formulate the main Conjecture.

GENERALIZED RAMANUJAN CONJECTURE FOR GL_n.

Let $\pi = \bigotimes_v \pi_v$ be an automorphic cuspidal representation of $GL_n(\mathbb{A}_F)$ with a unitary central character, then for each place v, π_v is tempered.

Remarks

(1) At the (almost all) places at which π_v is unramified the Conjecture is equivalent to the explicit description of the local parameters satisfying (9) and (10).

(2) For analytic applications the more tempered (ie the faster the decay of the matrix coefficients) the better. It can be shown (compare with (28) of §2) that the π_v's which occur cuspidally and automorphically are dense in the tempered spectrum, hence GRC if true is sharp.

(3) Satake [Sat] appears to be the first to observe that the classical Ramanujan Conjecture concerning the Fourier coefficients of $\Delta(z)$ can be formulated in the above manner. GRC above generalizes both these classical Ramanujan-Petersson Conjectures for holomorphic forms of even integral weight as well as Selberg's 1/4 eigenvalue conjecture for the Laplace spectrum of congruence quotients of the upper half plane [Se]. In this representation theoretic language the latter is concerned with π_∞ which are unramified and for which $\pi = \bigotimes_v \pi_v$ is an automorphic cuspidal representation of $GL_2(\mathbb{A}_\mathbb{Q})$.

There are some special but important cases of π's for which the full GRC is known. These are contained in cases where π_v for v archimedian is of special type. For $GL_2(\mathbb{A}_{\mathbb{Q}})$ and π_∞ being holomorphic discrete series (that is the case of classical holomorphic cusp forms of even integral weight) RC was established by Deligne. For a recent treatment see the book by Conrad [Con]. The proof depends on $\Gamma_0(n)\backslash\mathbb{H}$ being a moduli space for elliptic curves (with level structure) and this leads eventually to an identification of $\alpha_j(\pi_p)$, $j = 1, 2$ in terms of arithmetic algebraic geometric data. Specifically as eigenvalues of Frobenius acting on ℓ-adic cohomology groups associated with a related moduli problem. The RC, ie $|\alpha_1(\pi_p)| = |\alpha_2(\pi_p)| = 1$, then follows from the purity theorem (the Weil Conjectures) for eigenvalues of Frobenius, which was established by Deligne.

Recently, Harris and Taylor [H-T] following earlier work of Clozel have established GRC for an automorphic cusp form π on $GL_n(\mathbb{A}_F)$ for which the following are satisfied:

(i) F is a CM field.

(ii) $\widetilde{\Pi} \cong \Pi^c$ (ie a contragredient - galois conjugate condition.)

(iii) Π_∞ (∞ here being the product over all archimedian places of F) has the same infinitesimal character over $\mathbb{C}$ as the restriction of scalars from F to $\mathbb{Q}$ of an algebraic representation of GL_n over $\mathbb{C}$. In particular, π_∞ is a special type of cohomological representation.

(iv) For some finite place v of F, π_v is square-integrable (that is its matrix coefficients are square integrable).

The proof of the above is quite a tour-de-force. It combines the trace formula (see Arthur's Lectures) and Shimura varieties and eventually appeals to the purity theorem. To appreciate some of the issues involved consider for example F an imaginary quadratic extension of $\mathbb{Q}$. F has one infinite place v_∞ for which $F_{v_\infty} \simeq \mathbb{C}$. Hence, automorphic forms for $GL_n(F)$ live on quotients of the symmetric space $SL_n(\mathbb{C})\big/SU(n)$ which is not Hermitian. So there is no apparent algebro-geometric moduli interpretation for these quotient spaces. The basic idea is to transfer the given π on $GL_n(\mathbb{A}_F)$ to a π' on a Shimura variety (see Milne's lectures for definitions of the latter). The Shimura varieties used above are arithmetic quotients of unitary groups (see example 3 of Section 2). The transfer of π to π' is achieved by the trace formula. While the complete functorial transfers are not known for the general automorphic form, enough is known and developed by Harris, Taylor, Kottwitz and Clozel to deal with the π in question. Conditions (i), (ii) and (iv) are used to ensure that π corresponds to a π' on an appropriate unitary group, while condition (iii) ensures that at the archimedian place, π' is cohomological. The last is essential in identifying the eigenvalues of π'_v (v-finite) in terms of Frobenius eigenvalues.

In most analytic applications of GRC all π's enter and so knowing that the Conjecture is valid for special π's is not particularly useful. It is similar to the situation with zeros of the Riemann

Zeta Function and L-functions where it is not information about zeros on $\Re(s) = \frac{1}{2}$ that is so useful, but rather limiting the locations of zeros that are to the right of $\Re(s) = \frac{1}{2}$. We describe what is known towards GRC beginning with the local bounds. If $\pi = \bigotimes_v \pi_v$ is automorphic and cuspidal on $GL_n(\mathbb{A}_F)$ then π and hence π_v is firstly unitary and secondly generic. The latter asserts that π_v has a Whittaker model (see Cogdell's lectures [Co]). That π_v is general for π cuspidal follows from the Fourier Expansions on $GL_n(\mathbb{A}_F)$ of Jacquet, Piatetsky-Shapiro and Shalika [J-PS-S]. Now, Jacquet and Shalika [J-S] show that for π_v generic the local Rankin-Selberg L-function of π_v with its contragredient $\tilde{\pi}_v$,

$$L(s, \pi_v \times \tilde{\pi}_v) \;=\; \det\left(I - \alpha(\pi_v) \,\otimes\, \alpha\,(\tilde{\pi}_v) N(v)^{-s}\right)^{-1} \tag{11}$$

is analytic in $\Re(s) > 1$.

This leads directly to bounds towards GRC. Specifically, in the most important case when π_v is unramified, (11) implies that

$$|\log_{N(v)} |\,\alpha_j(\pi_v)|\,| \;<\; \frac{1}{2} \quad \text{for } j = 1, \ldots n \text{ and } v \text{ finite} \tag{12}$$

and

$$|\Re(\mu_j(\pi_v))\,| \;<\; \frac{1}{2} \quad \text{for } j = 1, 2, \ldots n \text{ and } v \text{ archimedian}\,. \tag{13}$$

Within the context of generic unitary representations of $GL_n(F_v)$, (12) and (13) are sharp. Recall that the trivial representation corresponds to μ as given in (5), so that for $n = 2$ (12) and (13) recover the trivial bound. However, for $n \geq 3$ these bounds are non-trivial. For $n = 3$, (12) and (13) correspond to the sharp decay rates for matrix coefficients of non-trivial representations of $SL_3(F_v)$ mentioned earlier. For $n > 3$, the bound (12) and (13) are much stronger (the trivial bound being $\frac{n-1}{2}$).

For many applications these local bounds fall just short of what is needed (this is clear in the case $n = 2$). One must therefore bring in further global information. The global Rankin-Selberg L-function is the key tool. In fact, it was already used by Rankin and Selberg in the case $n = 2$, $F = \mathbb{Q}$ and v finite, for such a purpose. The extension of their analysis to general n and F was observed by Serre [Ser]. However, this method which uses twisting by quasi-characters α^s (a technique which we now call deformation in a family (see [I-S]) of L-functions in this case the parameter being s) and a theorem of Landau [La], has the drawback of only working for v-finite and also it deteriorates in quality as the extension degree of F over $\mathbb{Q}$ increases. The last being a result of the increasing number of Gamma factors in the complete L-function (see [I-S]). In [L-R-S], the use of the Rankin-Selberg L-functions in a different way and via deformation in another family was developed. It has the advantage of applying to the archimedian places as well as being uniform in its applicability. It leads to the following bounds towards GRC. Let $\pi = \bigotimes_v \pi_v$ be an automorphic cuspidal representation of $GL_n(\mathbb{A}_F)$.

For v finite and π_v unramified and $j = 1, \dots n$

$$| \log_{N(v)} \alpha_j(\pi_v)| \le \frac{1}{2} - \frac{1}{n^2+1} \,. \tag{14}$$

For v archimedian and π_v unramified and $j = 1, 2, \dots n$

$$|\Re(\mu_j(\pi_v))\,| \le \frac{1}{2} - \frac{1}{n^2+1} \,. \tag{15}$$

In [M-S] these bounds are extended to include analogous bounds for places v at which π_v is ramified.

We describe briefly this use of the global Rankin-Selberg L-function. Let π be as above and v_0 a place at which π_{v_0} is unramified. For χ a ray class character of $F^* \backslash \mathbb{A}_F^*$ which is trivial at v_0, we consider the global Rankin-Selberg L-functions

$$\begin{aligned} \Lambda\,(s, \pi \times (\tilde{\pi} \times \chi)) &:= \prod_v L(s, \pi_v \times (\tilde{\pi}_v \times \chi_v)) \\ &= L(s, \pi_{v_0} \times \tilde{\pi}_{v_0})\, L_{S_0}(s, \pi \times (\tilde{\pi} \times \chi))\,, \end{aligned} \tag{16}$$

where $L_S(s)$ denotes the partial L-function obtained as the product over all places except those in S and $S_0 = \{v_0\}$. Now, according to the theory of the Rankin-Selberg L-function ([J-PS-S], [Sh], [M-W]) the left-hand side of (16) is analytic for $0 < \Re(s) < 1$. In particular, if $0 < \sigma_0 < 1$ is a pole of $L(s, \pi_{v_0} \times \tilde{\pi}_{v_0})$ (which will be present according to (9), (10) and (11) if GRC fails for π_{v_0}) then

$$L_{S_0}(\sigma_0, \pi \times \tilde{\pi} \times \chi) = 0 \text{ for all } \chi \text{ with } \chi \text{ trivial at } v_0 \,. \tag{17}$$

Thus, we are led to showing that $L_{S_0}(\sigma_0, \pi \times \tilde{\pi} \times \chi) \neq 0$ for some χ in this family. To see this one averages these L-functions over the set of all such χ's of a large conductor q. The construction of χ's satisfying the condition at v_0 is quite delicate (see [Roh]). In any event, using techniques from analytic number theory for averaging over families of L- functions, together with the positivity of the coefficients of $L(s, \pi \times \tilde{\pi})$, one shows that these averages are not zero if $N(q)$ is large enough and σ_0 is not too small. Combined with (17) this leads to (14) and (15).

The bounds (14) and (15) are the best available for $n \ge 3$. For $n = 2$ much better bounds are known and these come from the theory of higher tensor power L-functions. Recall that for $G = GL_n$, ${}^LG_0 = GL(n, \mathbb{C})$. In the case of $n = 2$ and $k \ge 1$ let $\mathrm{sym}^k : {}^LG_0 \longrightarrow GL(k+1, \mathbb{C})$ be the representation of $GL(2, \mathbb{C})$ on symmetric k-tensors (ie the action on homogeneous polynomials of degree k in x_1, x_2 by linear substitutions). The corresponding local L-function associated to an automorphic cusp form π on $GL_2(\mathbb{A}_F)$ and the representation sym^k of ${}^LG_0$ is given in (8). The global L-function with appropriate definitions at ramified places is given as usual by

$$\Lambda(s, \pi, \mathrm{sym}^k) = \prod_v L(s, \pi_v, \mathrm{sym}^k) \,. \tag{18}$$

Langlands [Lang2] made an important observation that if $\Lambda(s, \pi, \mathrm{sym}^k)$ is analytic in $\Re(s) > 1$ for all $k \geq 1$ (as he conjectured it to be) then a simple positivity argument yields GRC for π.[‡] Moreover, his general functoriality conjectures assert that $\Lambda(s, \pi, \mathrm{sym}^k)$ should be the global L-function of an automorphic form Π_k on $GL_{k+1}(\mathbb{A}_F)$. Hence, the functoriality conjectures imply GRC. There have been some striking advances recently in this direction. The functorial lift $\pi \longrightarrow \Pi_k$ of GL_2 to GL_{k+1} is now known for $k = 2, 3$ and 4. The method of establishing these lifts is based on the converse theorem (see Cogdell's lectures [Cog]). This asserts that Π is automorphic on $GL_n(\mathbb{A}_F)$ as long as the L-functions $\Lambda(s, \pi \times \pi_1)$ are entire and satisfy appropriate functional equations for all automorphic forms π_1 on $GL_m(\mathbb{A}_F)$ for $m \leq n-1$ (one can even allow $m < n-1$ if $n \geq 3$). In this way automorphy is reduced to establishing these analytic properties. This might appear to beg the question, however for $k = 2$ (and $\pi = \Pi_2$ on GL_3 as above) the theory of theta functions and half integral weight modular forms, combined with the Rankin-Selberg method yields the desired analytic properties of $\Lambda(s, \pi, \mathrm{sym}^2)$ (Shimura [Shi]). For $k = 3, 4$ the analytic properties were established by Kim and Shahidi [K-S] , [K]. They achieve this using the Langlands-Shahidi method which appeals to the analytic properties of Eisenstein series on exceptional groups (up to and including E_8, so that this method is limited) to realize the functions $\Lambda(s, \Pi_k \times \pi')$ above in terms of the coefficients of Eisenstein series along parabolic subgroups. The general theory of Eisenstein series and their meromorphic continuation (Langlands) yields in this way the meromorphic continuation and functional equations for these $\Lambda(s, \Pi_k \times \pi')$. The proof that they are entire requires further ingenious arguments. Their work is precise enough to determine exactly when Π_k is cuspidal (which is the case unless π is very special, that being it corresponds to a two-dimensional representation of the Weil group W_F in which case GRC for π is immediate). Now, using that Π_k for $1 \leq k \leq 4$ is cuspidal on $GL_{k+1}(\mathbb{A}_F)$ and forming the Rankin-Selberg L-functions of pairs of these leads to $\Lambda(s, \pi, \mathrm{sym}^k)$ being analytic for $\Re(s) > 1$ and $k \leq 9$. From this one deduces that for π as above, cuspidal on $GL_2(\mathbb{A}_F)$ and π_v unramified (if π_v is ramified on $GL_2(F_v)$ then it is tempered) that

$$|\log_{N(v)} |\,\alpha_j(\pi_v)|\,| \leq \frac{1}{9} \qquad \text{for } j = 1, 2 \text{ and } v \text{ finite} \tag{19}$$

and

$$|\Re(\mu_j(\pi_v))| \leq \frac{1}{9} \qquad \text{for } j = 1, 2 \text{ and } v \text{ archimedian}\,. \tag{20}$$

There is a further small improvement of (19) and (20) that has been established in the case $F = \mathbb{Q}$ [Ki-Sa]. One can use the symmetric square L-function in place of the Rankin-Selberg

[‡]This approach to the local statements involved in GRC via the analytic properties of the global L-functions associated with large irreducible representations of ${}^LG$ has been influential. In Deligne's proof of the Weil Conjectures mentioned earlier, this procedure was followed. In that case, ${}^LG$ is replaced by the monodromy representation of the fundamental group of the parameter space for a family of zeta functions for whose members the Weil Conjectures are to be established. The analytic properties of the corresponding global L-functions follows from Grothendieck's cohomology theory.

L-function in (16). This has the effect of reducing the "analytic conductor" (see [I-S] for the definition and properties of the latter). Applying the technique of Duke and Iwaniec [D-I] at the finite places and [L-R-S 2] at the archimedean place, one obtains the following refined estimates. For $n \leq 4$ and π an automorphic cusp form on $GL_n(\mathbb{A}_{\mathbb{Q}})$, or if $n = 5$ and $\pi = \mathrm{sym}^4\psi$ with ψ a cusp form an $GL_2(\mathbb{A}_{\mathbb{Q}})$, we have

$$|\log_p|\alpha_j(\pi_p)|\,| \leq \frac{1}{2} - \frac{1}{1 + \frac{n(n+1)}{2}}\,, p \text{ finite} \tag{21}$$

and

$$|\Re(\mu_j(\pi_\infty))| \leq \frac{1}{2} - \frac{1}{1 + \frac{n(n+1)}{2}}\,, p = \infty\,. \tag{22}$$

In particular, if we apply this to a cusp from ψ on $GL_2(\mathbb{A}_{\mathbb{Q}})$ we get

$$|\log_p \alpha_j(\psi_p)| \leq \frac{7}{64}\,, \text{ for } j = 1, 2 \text{ and } p < \infty \tag{23}$$

and

$$|\Re\mu_j(\psi_\infty)| \leq \frac{7}{64}\,, \text{ for } j = 1, 2\,. \tag{24}$$

(24) is equivalent to the following useful bound towards Selberg's 1/4 conjecture concerning the first eigenvalue of the Laplacian $\lambda_1(\Gamma(N)\backslash\mathbb{H})$ for a congruence quotient of the upper half plan $\mathbb{H}$.

$$\lambda_1(\Gamma(N)\backslash\mathbb{H}) \geq \frac{975}{4096} = 0.238\ldots\,. \tag{25}$$

§2. GENERAL G

Let G be a reductive linear algebraic group defined over F. The principle of functoriality gives relations between the spectra of $G(F)\backslash G(\mathbb{A}_F)$ for different G's and F's. In particular, in cases where versions of this principle are known or better yet where versions of the more precise conjectures of Arthur are known, one can transfer information towards the Ramanujan Conjectures from one group to another. For example, if D is a quaternion algebra over F, then the Jacquet-Langlands correspondence [Ge] from $D^*(F)\backslash D^*(\mathbb{A}_F)$ into $GL_2(F)\backslash GL_2(\mathbb{A}_F)$ allows one to formulate a precise GRC for D as well as to establish bounds towards it using (19) and (20). In fact, if $D\big/\mathbb{Q}$ is such that $D \otimes \mathbb{R} \cong H(\mathbb{R})$, the Hamilton quaternions, then the transfer to $GL_2(\mathbb{A}_{\mathbb{Q}})$ yields only π's for which π_∞ is a holomorphic representation of $GL_2(\mathbb{R})$. Hence for such D's the full GRC is known by Deligne's results mentioned in Section 1. Our main interest however is in G's for which $G(F_v)$ is non-compact for at least one archimedean place v of F. The remarks above about quaternion algebras apply to division algebras of degree n over F using the correspondence to $GL_n(\mathbb{A}_F)$ established by Arthur and Clozel [A-C]. Another example is that of unitary groups G over F in 3-variables and the transfer established by Rogawski [Ro] of the non-lifted forms (from $U(2) \times U(1)$) on $G(\mathbb{A}_F)$ to $GL_3(\mathbb{A}_E)$ where E is a quadratic extension of F (we discuss this example further in example 3 below). In all of the above examples the forms are lifted to GL_n and after examining for cuspidality (14) and (15) yield the best approximations to GRC for the corresponding G. We note that in the cases above, functoriality is established using the trace formula.

For a general semi-simple G (for the rest of this Section we will assume that G is semi-simple) defined over F, the Ramanujan Conjecture can be very complicated. It has been known for some time, at least since Kurokawa [Ku], that there are non-tempered automorphic cuspidal representations for groups such as $GSp(4)$. So the naive generalization of the GL_n GRC is not valid. Today the general belief is that such non-tempered representation are accounted for by functorial lifts from smaller groups.

One approach to GRC for more general G and which is along the lines of the original Ramanujan Conjecture, is to formulate the problem in a cruder form which is well-suited for analytic applications of the spectrum. For the latter, one wants to know the extent to which the local representations appearing as components of a global automorphic representation are limited. Put another way, which local representations in $\widehat{G(F_v)}$ can be excited arithmetically? Let $\pi = \underset{v}{\otimes} \pi_v$ be an automorphic representation appearing in $L^2(G(F)\backslash G(\mathbb{A}_F))$. That is, π occurs cuspidally or as a residual Eisenstein series or as part of a unitary integral of Eisenstein series. We will not distinguish the part of the spectrum in which these occur. This is one sense in which we seek cruder information. Now fix a place w of F and define the subset $\widehat{G(F_w)}_{\mathrm{AUT}}$ of $\widehat{G(F_w)}$, to be the closure in the Fell topology of the set of π_w's for which $\pi = \underset{v}{\otimes} \pi_v$ occurs in $L^2(G(F)\backslash G(\mathbb{A}_F))$. This closure process is the second sense in which we seek cruder information. We call $\widehat{G(F_w)}_{\mathrm{AUT}}$ the automorphic dual of G at w. More generally, if S is a finite set of places of F we define $\widehat{G(S)}_{\mathrm{AUT}}$ to be the closure in $\widehat{G(S)}$ of $\underset{w\in S}{\otimes} \pi_w$ as π varies over all π in $L^2(G(F)\backslash G(\mathbb{A}_F))$ and $G(S) = \Pi_{w\in S} G(F_w)$. Similarly, one can define $\widehat{G_{\mathrm{AUT}}}$ to be the corresponding closure in $\Pi_v \widehat{G(F_v)}$. By approximation theorems for adele groups we can describe these sets in terms of congruence subgroups as follows.[§] Let S_∞ be the set of archimedean places of F. Then $\widehat{G(S_\infty)}_{\mathrm{AUT}}$ is the closure of all $\underset{w\in S_\infty}{\otimes} \Pi_w$ in $\widehat{G(S_\infty)}$ which occur in $L^2(\Gamma\backslash G(S_\infty))$ where Γ varies over all congruence subgroups of $G(O_F)$, O_F being the ring of integers of F. Similarly, if S is a finite set of places containing S_∞ then $\widehat{G(S)}_{\mathrm{AUT}}$ is the closure in $\underset{v\in S}{\Pi} \widehat{G(F_v)}$ of all $\underset{w\in S}{\otimes} \pi_w$ which occur in $L^2(\Gamma\backslash G(S))$, as Γ varies over all congruence subgroups of the S-arithmetic group $G(O_S)$, with O_S being the ring of S integers of F. We can

[§]At least if G is simply connected and F simple otherwise the description is more complicated.

now state the basic problem for G.

GENERALIZED RAMANUJAN PROBLEM (GRP):

To determine for a given G defined over F, the sets $\widehat{G(F_v)}_{\mathrm{AUT}}$ and more generally $\widehat{G}_{\mathrm{AUT}}$.

We emphasize that the local data $\widehat{G(F_v)}_{\mathrm{AUT}}$ is determined by the global group G. Also, while the set of π_w's in $\widehat{G(F_w)}$ that arise as the w component of an automorphic π in $L^2(G(F)\backslash G(\mathbb{A}_F))$ is typically very difficult to describe, the closure process in the definition of the automorphic dual makes this task much simpler. Moreover, the above formulation allows one to measure progress towards GRP by giving set theoretic upper and lower bounds for $\widehat{G}_{\mathrm{AUT}}$. Non-trivial upper bounds are what are most useful in applications while various methods for constructing automorphic forms (some of which are discussed in the examples below) produce lower bounds. We denote by $\widehat{G(F_v)}^{\mathrm{sph}}_{\mathrm{AUT}}$, the spherical part of $\widehat{G(F_v)}_{\mathrm{AUT}}$.

Let G be defined over F and let H be a semi-simple subgroup of G also defined over F. Then $\widehat{H(F_v)}_{\mathrm{AUT}}$ and $\widehat{G(F_v)}_{\mathrm{AUT}}$ and more generally $\widehat{H(S)}_{\mathrm{AUT}}$ and $\widehat{G(S)}_{\mathrm{AUT}}$, satisfy some simple functorial properties.

If $\sigma \in \widehat{H(F_v)}_{\mathrm{AUT}}$ then

$$\mathrm{Ind}_{H(F_v)}^{G(F_v)}\,\sigma \,\subset\, \widehat{G(F_v)}_{\mathrm{AUT}}\,. \tag{26}$$

If $\beta \in \widehat{G(F_v)}_{\mathrm{AUT}}$ then

$$\mathrm{Res}_{H(F_v)}^{G(F_v)}\,\beta \,\subset\, \widehat{H(F_v))}_{\mathrm{AUT}}\,. \tag{27}$$

The induction and restriction computations involved in (26) and (27) are purely local. Their precise meaning is that any irrducible ψ which is contained (weakly) on the left is contained on the right-hand side of the inclusions. These inclusions were proven in [B-S] and [B-L-S1] for $F_v = \mathbb{R}$ and in general (that is for finitely many places at a time) in [C-U]. The proof of (26) depends on realizing the congruence subgroups of $H(F)$ as geometric limits (specifically as infinite intersections) of congruence subgroups of $G(F)$ and applying the spectral theory of such infinite volume quotients. In [Ven] a characterization of such intersections of congruence subgroups of $G(F)$ is given. (27) is established by approximating diagonal matrix coefficients of $\mathrm{Res}_{H(F_v)}^{G(F_v)}\,\beta$ by matrix coefficients of elements in $\widehat{H(F_v)}_{\mathrm{AUT}}$. This is done by constructing suitable sequences of H cycles, in a given congruence quotient of G, which become equidistributed in the limit. The last can be done either using Hecke operators or using ergodic theoretic techniques associated with unipotent flows.

(26) and (27) may be used to give upper and lower bounds for $\hat{G}_{\mathrm{AUT}}$. For example, if $H = \{e\}$ and $\sigma = 1$ then (26) applies and since $\overline{\mathrm{Ind}_{\{e\}}^{G(F_v)} 1} = \widehat{G(F_v)}_{\mathrm{temp}}$, we obtain the general lower bound

$$\widehat{G(F_v)}_{\mathrm{AUT}} \supset \widehat{G(F_v)}_{\mathrm{temp}} \,. \tag{28}$$

Next, we illustrate by way of examples, some bounds towards GRP that have been established using current techniques.

<u>*Example 1. SL*</u>

Let $G = SL_2$ over $\mathbb{Q}$. The local components of the unitary eisenstein integrals involved in the spectral decomposition of $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q}})$ satisfy GRC at all places. Moreover, the only residue of the Eisenstein series is the trivial representation. Hence, the Ramanujan and Selberg Conjectures for the cuspidal spectrum of $SL_2(\mathbb{A}_{\mathbb{Q}})$ are equivalent to

$$\widehat{SL_2(\mathbb{Q}_v)}_{\mathrm{AUT}} = \{1\} \cup \widehat{SL_2(\mathbb{Q}_v)}_{\mathrm{temp}}\,, \text{ for all places } v \text{ of } \mathbb{Q}\,. \tag{29}$$

For this case (23) and (24) give the best known upper bounds towards (29).

Let $G = SL_3$ over $\mathbb{Q}$. Again, there are no poles of the Eisenstein series yielding residual spectrum other than the trivial representation. However, there is an integral of non-tempered unitary Eisenstein series contributing to $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q}}))$. These correspond to the eisenstein series on the maximal $(2,1)$ parabolic subgroup of G taken with the trivial representation on its Levi. In particular for any place v of $\mathbb{Q}$, $\widehat{G(\mathbb{Q}_v)}_{\mathrm{AUT}}$ contains the following non-tempered spherical principal series (we use the parameters in (4) above):

$$\mathrm{Cont}(v) = \left\{\mu_t = \left(\frac{1}{2} + it,\, -2it,\, -\frac{1}{2} + it\right) \Big|\, t \in \mathbb{R}\right\} \subset \widehat{G(\mathbb{Q}_v)}^{\mathrm{sph}}\,. \tag{29$'$}$$

If (29) is true then the rest of the Eisenstein series contribution to SL_3, consists of tempered spectrum. Hence using (12) and (13) we see that the cuspidal GRC for $SL_3(\mathbb{A}_{\mathbb{Q}})$ is equivalent to:

For any place v of $\mathbb{Q}$

$$\widehat{G(\mathbb{Q}_v)}_{\mathrm{AUT}} = \{1\} \cup \mathrm{Cont}(v) \cup \widehat{G(\mathbb{Q}_v)}_{\mathrm{temp}}\,. \tag{29$''$}$$

The best upper bound on $\widehat{G}_{\mathrm{AUT}}$ in this case is given by (21) and (22) which assert that for any place v of $\mathbb{Q}$

$$\widehat{G(\mathbb{Q}_v)}^{\mathrm{sph}}_{\mathrm{AUT}} \subset \Big\{ \mu \in \widehat{G(\mathbb{Q}_v)}^{\mathrm{sph}} \Big| \mu = (1,0,-1);\ \mu = \left(\frac{1}{2}+it,\ -2it,\ -\frac{1}{2}+it\right)$$
$$t \in \mathbb{R};\ \mu \text{ such that } |\Re(\mu_j)| \le \frac{5}{14}\,. \Big\} \tag{29'''}$$

Using [M-W] one can make a similar analysis for SL_n, $n \ge 4$.

Example 2. Orthogonal Group

Let f be the quadratic form over $\mathbb{Q}$ in $n+1$ variables given by

$$f(x_1, x_2, \dots x_{n+1}) = 2x_1x_{n+1} + x_2^2 + \cdots + x_n^2\,. \tag{30}$$

Let $G = SO_f$ be the special orthogonal group of $(n+1)\times(n+1)$ matrices preserving f. G is defined over $\mathbb{Q}$ and is given explicitly by

$$G = \left\{ g \in SL_{n+1} \,\middle|\, g^t \begin{pmatrix} & & 1 \\ & I_{n-1} & \\ 1 & & \end{pmatrix} g = \begin{pmatrix} & & 1 \\ & I_{n-1} & \\ 1 & & \end{pmatrix} \right\}. \tag{31}$$

Thus $G(\mathbb{Q}_\infty) = G(\mathbb{R}) \cong SO_{\mathbb{R}}(n,1)$ which has real rank 1. The corresponding symmetric space $G(\mathbb{R})/K$ with $K \cong SO_{\mathbb{R}}(n)$ is hyperbolic n-space. Let $M(\mathbb{R}), N(\mathbb{R})$ and $A(\mathbb{R})$ be the subgroups of $G(\mathbb{R})$

$$A(\mathbb{R}) = \left\{ \begin{pmatrix} a & & \\ & I_{n-1} & \\ & & a^{-1} \end{pmatrix} \,\middle|\, a \in \mathbb{R}^* \right\} \tag{32}$$

$$N(\mathbb{R}) = \left\{ \begin{pmatrix} 1 & -u^t & -\frac{1}{2}\langle u,u\rangle \\ & I_{n-1} & u \\ & & 1 \end{pmatrix} \,\middle|\, u \in \mathbb{R}^{n-1} \right\} \tag{33}$$

$$M(\mathbb{R}) = \left\{ \begin{pmatrix} 1 & & \\ & h & \\ & & 1 \end{pmatrix} \,\middle|\, h^t h = I_{n+1} \right\} \tag{34}$$

Then $P(\mathbb{R}) = M(\mathbb{R})A(\mathbb{R})N(\mathbb{R})$ is a parabolic subgroup of G with Levi factor MA and unipotent radical N. The spherical unitary dual of $G(\mathbb{R})$ may be described in terms of the principal series. For $s \in \mathbb{C}$ let

$$\pi_s = \mathrm{Ind}_{M(\mathbb{R})A(\mathbb{R})N(\mathbb{R})}^{G(\mathbb{R})}\, 1_M \otimes |a|^s\,. \tag{35}$$

(For $s \in \mathbb{C}$ for which π_s is reducible we take the spherical constitutent for π_s.) In this normalization $s = \frac{n-1}{2} := \rho_n$ corresponds to the trivial representation and the tempered spherical representations consist of π_s with $s \in i\mathbb{R}$. For $-\rho_n \le s \le \rho_n$, π_s is unitrizable and these constitute the complementary series. Moreover, π_s is equivalent to π_{-s}. These yield the entire spherical unitary dual, that is

$$\widehat{G(\mathbb{R})}^{\mathrm{sph}} = \left\{\pi_s \mod \pm 1 \,\middle|\, s \in i\mathbb{R} \cup [-\rho_n, \rho_n] \right\} . \tag{36}$$

Here $i\mathbb{R}$ identified with $\widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{temp}}$ and $(0, \rho_n]$ identified with the non-tempered part of $\widehat{G(\mathbb{R})}^{\mathrm{sph}}$. Towards the GRP for G_f we have the following inclusions ($n \geq 3$), see [B-S]:

$$i\mathbb{R} \cup \{\rho_n, \rho_n - 1, \ldots, \rho_n - [\rho_n]\} \subset \widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{AUT}} \subset i\mathbb{R} \cup \{\rho_n\} \cup \left[0, \rho_n - \frac{7}{9}\right] . \tag{37}$$

In particular, for $n \geq 4$, $\widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{AUT}}$ contains non-tempered points besides the trivial representation. (37) is deduced from (26) and (27) as follows. Let H be the subgroup of G stabilizing x_2. H together with $\sigma = 1$ satisfies the assumptions in (26). Hence

$$\widehat{G(\mathbb{R})}_{\mathrm{AUT}} \supset \mathrm{Ind}_{H(\mathbb{R})}^{G(\mathbb{R})} 1 . \tag{38}$$

The space $G(\mathbb{R})/H(\mathbb{R})$ is an affine symmetric space and for general such spaces the induction on the right-hand side of (38) has been computed explicitly by Oshima (see [O-M] and [Vo2]). For the case at hand, one has

$$\mathrm{Ind}_{H(\mathbb{R})}^{G(\mathbb{R})} \supset \{\rho_n, \rho_n - 1 \ldots, \rho_n - [\rho_n]\} \cup i\mathbb{R} . \tag{39}$$

This gives the lower bound in (37).

To see the upper bound, first note that for $n = 3$ we have

$$\widehat{G(\mathbb{R})}^{\mathrm{sph}} \subset i\mathbb{R} \cup \{1\} \cup \left[0, \frac{2}{9}\right] . \tag{40}$$

This follows by passing from this SO_f to its spin double cover which at $\mathbb{Q}_\infty$ is $SL_2(\mathbb{C})$ and then invoking the bound (20) for $GL_2(\mathbb{A}_E)$ where E is an imaginary quadratic extension of $\mathbb{Q}$. If $n > 3$ we let H be the subgroup of G which stabilizes $x_2, \ldots, x_{n-2}$. Then $H = G_{f'}$ with f' a form in 4 variables of signature $(3, 1)$. Thus according to (40)

$$\widehat{H(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{AUT}} \subset i\mathbb{R} \cup \{1\} \cup \left[0, \frac{2}{9}\right] . \tag{41}$$

Now apply the restriction principle (27) with the pair G and H as above and with β a potential non-tempered element in $\widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{AUT}}$. Computing the local restriction $\mathrm{Res}_{H(\mathbb{R})}^{G(\mathbb{R})} \beta$ and applying (41) leads to the upper bound in (37).

One is led to a precise GRC for G at $\mathbb{Q}_\infty$:

CONJECTURE: Let G be as in (31) then

$$\widehat{G(\mathbb{R})}^{\text{sph}}_{\text{AUT}} = i\mathbb{R} \cup \{\rho_n, \rho_n - 1, \dots, \rho_n - [\rho_n]\}\,. \tag{42}$$

Example 3. Unitary Group

Let $SU(2,1)$ be the special unitary group of 3×3 complex matrices of determinant equal to one. That is, such matrices preserving the Hermitian form $|z_1|^2 + |z_2|^2 - |z_3|^2$. If $g \in SU(2,1)$ and $g = \begin{bmatrix} A & b \\ c^* & d \end{bmatrix}$ with A 2×2, b and c 2×1 and d 1×1 complex matrices, then g acts projectively on

$$\begin{aligned} B^2 &= \left\{(z_1,z_2) \,|\, |z_1|^2 + |z_2|^2 < 1\right\} \text{ by} \\ gz &= (\langle z, c\rangle + d)^{-1}(Az+b)\,. \end{aligned} \tag{43}$$

In this way $B^2 \simeq SU(2,1)/K$, with $K = S(U(2)\times U(1))$, is the corresponding Hermitian symmetric space. The biholomorphic action (43) extends to the closed ball $\overline{B^2}$. If $e_1 = (1,0) \in \overline{B^2}|B^2$ then its stabilizer $P = \{g \in SU(2,1)|ge_1 = e_1\}$ is a parabolic subgroup of $SU(2,1)$. Let Γ be co-compact lattice in $SU(2,1)$. It acts discontinuously on B^2 and we form the compact quotient $X_\Gamma = \Gamma\backslash B^2$ which is a compact, complex Kahler surface. We examine the Betti numbers $b_j(\chi_\Gamma)$ for $j = 0,1,2,3$ and 4. According to the Gauss-Bonnet-Chern formula

$$\chi(X_\Gamma) = b_0 - b_1 + b_2 - b_3 + b_4 = \operatorname{Vol}(\Gamma\backslash SU(2,1)) \tag{44}$$

with dg being a suitable fixed normalized Haar measure on $SU(2,1)$. By duality this yields

$$\operatorname{Vol}(\Gamma\backslash SU(2,1)) = b_2 - 2b_1 + 2\,. \tag{45}$$

It follows that if $\operatorname{Vol}(\Gamma\backslash SU(2,1))$ goes to infinity then so does $b_2(X_\Gamma)$. Thus for large volume X_Γ will have cohomology in the middle dimension. The behavior of $b_1(X_\Gamma)$ is subtle and in algebraic surface theory this number is known as the irregularity of X_Γ. It can be calculated from the decomposition of the regular representation of $SU(2,1)$ on $L^2(\Gamma\backslash SU(2,1))$, for a discussion see Wallach [Wa]. We indicate briefly how this is done. The representation $\operatorname{Ind}_P^{SU(2,1)} 1$ (nonunitary induction) of $SU(2,1)$ is reducible. Besides containing the trivial representation as a subrepresentation it also contains two irreducible subquotients π_0^+ and π_0^- (see [J-W]). $\pi_0^\pm$ are non-tempered unitary representations of $SU(2,1)$, in fact their K-finite matrix coefficients lie in $L^p(SU(2,1))$ for $p > 4$, but not in L^4. Let $m_\Gamma(\pi_0^+)$ and $m_\Gamma(\pi_0^-)$ be the multiplicities with which π_0^+(respectively π_0^-) occur in the decomposition of $L^2(\Gamma\backslash SU(2,1))$. For the example at hand, these multiplicities are equal (which is reflection of X_Γ being Kahler). The following is a particular case of Matsushima's formula (see Borel-Wallach [B-W]) which gives the dimensions of various cohomology groups of a general locally symmetric space $\Gamma\backslash G/K$ in terms of the multiplicities with which certain π's in $\hat{G}$ occur in $L^2(\Gamma\backslash G)$.

$$b_1(X_\Gamma) \;=\; m_\Gamma(\pi_0^+) \,+\, m_\Gamma(\pi_0^-) \;=\; 2m_\Gamma(\pi_0^+)\,. \tag{46}$$

We examine the above in the case that Γ is a special arithmetic lattice. Let E be an imaginary quadratic extension of $\mathbb{Q}$ and let D be a degree 3 division algebra central over E and which carries an involution α of the second kind. That is the restriction of α to E is Galois conjugation $E/\mathbb{Q}$. Let G be the $\mathbb{Q}$-algebraic group whose $\mathbb{Q}$ points $G(\mathbb{Q})$ equal to $\{g \in D^* | \alpha(g)g = 1 \text{ and } Nrd(g) = 1\}$. Here Nrd is the reduced norm on D. G is the special unitary group $SU(D,f)$ where f is the 1-dimensional (over D) Hermitian form $f(x,y) = \alpha(x)y$. Localizing G at $\mathbb{Q}_\infty = \mathbb{R}$ we obtain $G(\mathbb{R})$ which is a special unitary group in 3-variables and which we assume has signature $(2,1)$. That is $G(\mathbb{R}) \simeq SU(2,1)$. In this case $G(\mathbb{Q})\backslash G(\mathbb{A}_\mathbb{Q})$ is compact and we consider its automorphic dual and specifically $\widehat{G(\mathbb{R})}_{\mathrm{AUT}}$. The key to obtaining information about $\widehat{G}_{\mathrm{AUT}}$ is the explicit description by Rogawski [Ro] of the spectrum of $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}_\mathbb{Q}))$ in terms of certain automorphic forms on $GL_3(\mathbb{A}_E)$ (see his Chapter 14 which discusses inner forms). Not surprisingly the Π's on $GL_3(\mathbb{A}_E)$ arising this way satisfy conditions similar to (i), (ii) and (iii) on page 6. If $\pi = \bigotimes_v \pi_v$ is an automorphic representation of $G(\mathbb{A}_\mathbb{Q})$ and is not 1-dimensional then the lifted form $\Pi = \bigotimes_w \Pi_w$ is cuspidal on $GL_3(\mathbb{A}_E)$. The relation between Π_w for $w|v$ and π_v is given explicitly. Thus the GRC for G takes the simplest form: If π is not 1-dimensional then π_v is tempered for all places v of $\mathbb{Q}$. Moreover (14) and (15) yield corresponding non-trivial bounds on $\widehat{G}_{\mathrm{AUT}}$.

We fixate on the representations $\pi_0^\pm$ in $\widehat{G(\mathbb{R})}$. (15) implies that

$$\pi_0^\pm \;\notin\; \widehat{G(\mathbb{R})}_{\mathrm{AUT}} \tag{47}$$

(see [B-C]).

This upper-bound on GRC for this G implies afortiori that $m_\Gamma(\pi_0^-)$ are zero for any congruence subgroup Γ of $G(\mathbb{Z})$. This combined with (46) has the following quite striking vanishing theorem as a consequence (and was proved in this way by Rogawski)

$$b_1(X_\Gamma) \;=\; 0, \text{ for } \Gamma \text{ any congruence subgroup of } G(\mathbb{Z})\,. \tag{48}$$

In particular, these arithmetic surfaces X_Γ have no irregularities and all their non-trivial cohomology is in the middle degree and its dimension is given by the index (45).

The vanishing theorem (48) is of an arithmetic nature. It is a direct consequence of restrictions imposed by the Ramanujan Conjectures. It should be compared with vanishing theorems which are direct consequences of Matsushima's formula, by which we mean the vanishing of certain cohomology groups of general locally symmetric spaces $X_\Gamma = \Gamma\backslash G/K$, independent of Γ. The vanishing, resulting from the fact that none of the potential π's which contribute to Matsushima's formula are unitary. A complete table of the cohomological unitary representations and the corresponding vanishing degrees for general real groups G, is given in [V-Z].

Example 4. Exceptional groups

The theory of theta functions and its extension to general dual pairs provides a powerful method for constructing "lifted" automorphic forms and in particular non-tempered elements in $\widehat{G}_{\text{AUT}}$. Briefly, a reductive dual pair is a triple of reductive algebraic groups H, H' and G with H and H' being subgroups of G which centralize each other. If π is a representation of G then the analysis of the restriction $\pi\big|_{G\times G'}$ (here $\pi((g,g')) = \pi(gg'))$ can lead to a transfer of representations on H to H' (or visa-versa). The classical case of theta functions is concerned with G being the symplectic group and $\pi = w$, the oscillator representation. That w is automorphic was shown in Weil [We] while the general theory in this setting is due to Howe [Ho]. Recent works ([Ka-Sav], [R-S2], [G-G-J]) for example show that this rich theory can be extended to other groups G such as exceptional groups with π being the minimal representation. For such suitably split G the minimal representation is shown to be automorphic by realizing it as a residue of eisenstein series [G-R-S]. For an account of the general theory of dual pairs and the minimal representation see [Li2].

For example, the dual pair $O(n,1) \times SL_2$ in a suitable symplectic group may be used to give another proof of the lower bound in (37). Restricting the oscillator representation to this dual pair one finds that holomorphic discrete series of weight k on SL_2 correspond to point $\rho_n - k$ in (37), see Rallis-Schiffmann [R-S1] and [B-L-S2].

We illustrate these methods with a couple of examples of exceptional groups. Let G be the automorphism group of the split Cayley algebra over $\mathbb{Q}$ (see [R-S] for explicit descriptions of the group as well as various data associated with it). G is a linear algebraic group defined over $\mathbb{Q}$ and is split of type G_2. It is semi-simple, it has rank 2 and as a root system for a maximal split torus we can take $\Delta = \{\pm(e_1 - e_2), \pm(e_1 - e_3), \pm(e_2 - e_3), \pm(2e_1 - e_2 - e_3), \pm(2e_2 - e_1 - e_3), \pm(2e_3 - e_1 - e_2)\}$ in $V = \{(a,b,c) | a + b + c = 0\}$ and with the standard pairing $\langle\, , \rangle$. Here e_1, e_2, e_3 are the standard basis vectors. The corresponding Weyl group W is of order 12. It is generated by reflections along the roots and preserves $\langle\, , \rangle$. The long root $\beta_1 = 2e_1 - e_2 - e_3$ together with the short root $\beta_6 = -e_1 + e_2$ are a basis and determine corresponding positive roots $\beta_1, \beta_2, \ldots, \beta_6$, see Figure 1. Up to conjugacy G has 3 parabolic subgroups; P_0 the minimal parabolic subgroup, P_1 the maximal parabolic corresponding to β_1 and P_2 the maximal parabolic corresponding to β_6. P_j factorizes as $L_j N_j$ with L_j the Levi factor and N_j the unipotent factor. Here L_1 and L_2 are isomorphic to GL_2. We examine the automorphic dual $\widehat{G}_{\text{AUT}}$ associated with the spectrum of $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q}}))$ and specifically $\widehat{G(\mathbb{R})}_{\text{AUT}}$.

We recall the classification by Vogan [Vo3] of the unitary spherical dual $\widehat{G(\mathbb{R})}^{\text{sph}}$. The maximal compact subgroup K of $G(\mathbb{R})$ is $SU(2) \times SU(2)$. The corresponding Riemannian symmetric space $G(\mathbb{R})/K$ is 8-dimensional. For $j = 0, 1, 2$ let $P_j(\mathbb{R}) = M_j(\mathbb{R})A_j(\mathbb{R})N_j(\mathbb{R})$ be the Langlands decomposition of the parabolic subgroup $P_j(\mathbb{R})$. $M_0(\mathbb{R})A_0(\mathbb{R})$ is a split Cartan subgroup of $G(\mathbb{R})$ and we identify the dual Lie algebra of A_0, denoted $\mathfrak{a}^*_{\mathbb{R}}$, with $V = \{(a_1, a_2, a_3) \in \mathbb{R}^3 | a_1+a_2+a_3 = 0\}$. The corresponding root system $\Delta(\mathfrak{g}, \mathfrak{a})$ is Δ. Here $M_0(\mathbb{R})$ is the dihedral group D_4 while $M_1(\mathbb{R})$ and $M_2(\mathbb{R})$ are isomorphic to $SL_2(\mathbb{R})$. For χ a unitary character of $\mathbb{A}_0(\mathbb{R})$ let $I_{P_0(\chi)}$ be the spherical constituent of $\mathrm{Ind}^{G(\mathbb{R})}_{P_0(\mathbb{R})}(1_{M_0(\mathbb{R})} \otimes \chi)$. For $j = 1$ or 2 and χ_j a unitary character of $A_j(\mathbb{R})$ and $0 < \sigma \leq \frac{1}{2}$ a complementary series representation of $M_j(\mathbb{R})$, let $I_{P_j}(\sigma, \chi)$ be the spherical constituent of $\mathrm{Ind}^{G(\mathbb{R})}_{P_j(\mathbb{R})}(\sigma \otimes \chi_j)$. The representations $I_{P_0}(\chi)$ are tempered and as we vary over all unitary χ these exhaust $\widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{temp}}$. The representations $I_{P_j}(\sigma, \chi)$ are nontempered and unitary, they together with the tempered representations exhaust all the nonreal part of $\widehat{G(\mathbb{R})}^{\mathrm{sph}}$ (ie the spherical representations with nonreal infintesimal characters). The rest of $\widehat{G(\mathbb{R})}^{\mathrm{sph}}$ may be described as a subset of $\mathfrak{a}^*_{\mathbb{R}}$ with $\alpha \in \mathfrak{a}^*_{\mathbb{R}}$ corresponding to $\mathrm{Ind}^{G(\mathbb{R})}_{P_0(\mathbb{R})}(1_{M_0(\mathbb{R})} \otimes \exp(\alpha(\cdot)))$. According to Vogan [Vo3] the set of such α's which are unitary is the shaded region in Figure 1.

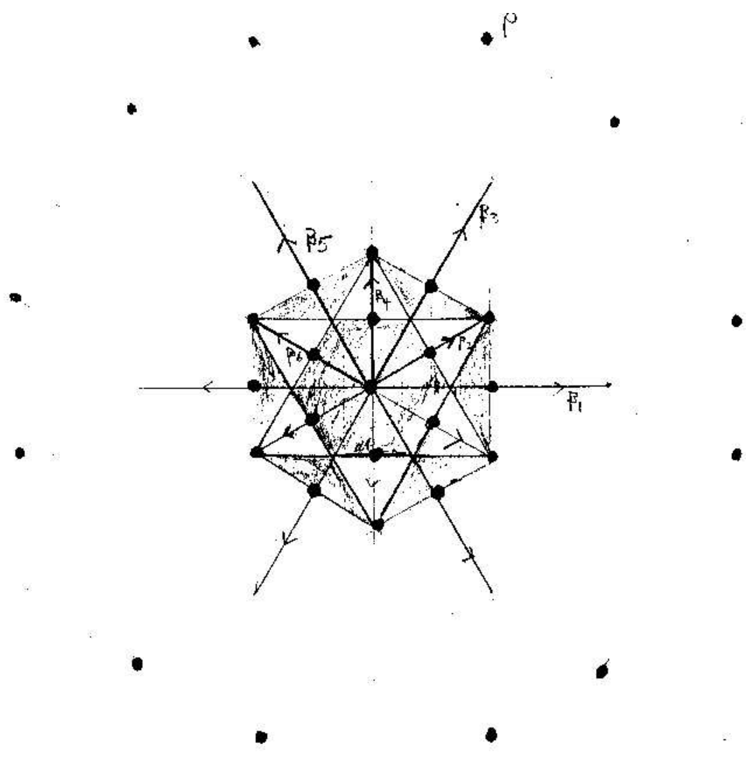

Figure 1.

The shaded area together with the outside dots yield the spherical unitary dual of G_2. The dots (that is $\rho, \beta_6, \beta_1/2, \beta_6/2$ and 0 and their images under W) are in $\widehat{G_2(\mathbb{R})}^{\text{sph}}_{\text{AUT}}$.

Note that points in $\mathfrak{a}^*_{\mathbb{R}}$ equivalent under W correspond to the same point in $\widehat{G(\mathbb{R})}^{\text{sph}}$. The point ρ is half the sum of the positive roots and corresponds to the trivial representation of $G(\mathbb{R})$. Clearly it is isolated in $\widehat{G(\mathbb{R})}$ (as it should be since $G(\mathbb{R})$ has property T). Also note that for $0 \le \sigma \le 1/2$ and $j = 1, 2$, $I_{P_j}(\sigma, \mathbf{1})$ (which is real) corresponds to the point $\sigma\beta_j$ in $\mathfrak{a}^*_{\mathbb{R}}$.

We turn to $\widehat{G(\mathbb{R})}^{\text{sph}}_{\text{AUT}}$. Let

$$C_0 = \{I_{P_0}(\chi) | \chi \text{ is unitary}\} \tag{49}$$

$$C_1 = \left\{ I_{P_1}\left(\frac{1}{2}, \chi\right) | \chi \text{ is unitary} \right\} \tag{50}$$

and

$$C_2 = \left\{ I_{P_2}\left(\frac{1}{2}, \chi\right) | \chi \text{ is unitary} \right\}. \tag{51}$$

We have the following lower bound

$$\widehat{G(\mathbb{R})}^{\text{sph}}_{\text{AUT}} \supset C_0 \cup C_1 \cup C_2 \cup \{\beta_4\} \cup \{\rho\}. \tag{52}$$

Note that the set of points on the right-hand side of (52) meets $\mathfrak{a}^*_{\mathbb{R}}$ in the set of dotted points in Figure 1.

We explain the containment (52). Firstly, the point $\{\rho\}$ is self evident. Since $C_0 = \widehat{G(\mathbb{R})}^{\text{sph}}_{\text{temp}}$ its inclusion in (52) follows from (28). One can show the containment of C_1 of C_2 by a variation of (26) where we allow H to be a parabolic subgroup, specifically P_1 and P_2 in this case. However, the theory of Eisenstein series demonstrates this more explicitly. Form the Eisenstein series $E_{P_1}(g, s)$ on $G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q}})$ corresponding to P_1 and with the trivial representation on $M_1^{(1)}$ (where $L_1 = M_1^{(1)} A_1$). E_{P_1} has a meromorphic continuation in s and is analytic on $\Re(s) = 0$ where it furnishes continuous spectrum in $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q}}))$. The corresponding spherical parameters fill out C_1 and place them in $\widehat{G}^{\text{sph}}_{\text{AUT}}$. Similarly the continuous spectrum corresponding to the Eisenstein series E_{P_2} (with the trivial representation on $M_2^{(1)}$) yields C_2. The remaining point $\{\beta_4\}$ in (52) is more subtle. Again one can see that it is in $\widehat{G(\mathbb{R})}_{\text{AUT}}$ using (26). The Lie sub-algebra of $\mathfrak{g}$ generated by $\mathfrak{a}$ and the root vectors corresponding to the six long roots is of type A_2. The corresponding subgroup H of G is SL_3 and is defined over $\mathbb{Q}$. $H(\mathbb{R})$ and $G(\mathbb{R})$ are both of rank 2 and they share the split torus $A_0(\mathbb{R})$. Choosing β_1 and β_5 as simple positive rootes of $\Delta(\mathfrak{h}, \mathfrak{a})$, we find that $\rho_H = \beta_3$. Now,

$$\beta_4 \in \mathrm{Ind}_{H(\mathbb{R})}^{G(\mathbb{R})}\left(\mathrm{Ind}_{H(\mathbb{Z})}^{H(\mathbb{R})} 1\right). \tag{48'}$$

This can be shown by considering the density of $H(\mathbb{Z})$ points in expanding regions in $G(\mathbb{R})$, first by examining $H(\mathbb{Z})$ as a lattice in $H(\mathbb{R})$ and second by using (48′) (see [Sa2]). The key points being that β_2 (or β_4) is an extreme point of the outer hexagon and that $\rho_G = 2\rho_H - \beta_2$. From (48′) and (26) it follows that $\beta_4 \in \widehat{G(\mathbb{R})}_{\mathrm{AUT}}$. As before, the Eisenstein series provide a more explicit automorphic realization of β_4. In fact it occurs as a residue (and hence in the discrete spectrum) of the minimal parabolic Eisenstein series $E_{P_0}(g,s)$ (here s denotes two complex variables). See for example [K2].

The above account for the lower bound (52). It is interesting that there are other residual and even cuspidal spectrum which contribute to various points on the right-hand side of (52). The Eisenstein series $E_{P_1,\pi}(g,s)$, where π is an automorphic cuspidal representation on $M^{(1)} \cong PGL_2$, has a pole at $s = 1/2$ if the special value $L\left(\frac{1}{2}, \pi, \mathrm{sym}^3\right)$ of the symmetric cube L-function, is not zero, see [K2]. If $\pi_\infty(\pi = \underset{v}{\otimes}\, \pi_v)$ is spherical and tempered then the corresponding residue on $G(\mathbb{A}_\mathbb{Q})$ produces a point in C_2 (for example if π_∞ is spherical corresponding to a Maass cusp form with eigenvalue 1/4 then the corresponding point in $\widehat{G(\mathbb{R})}^{\mathrm{sph}}_{\mathrm{AUT}}$ is $I_{P_2}\left(\frac{1}{2}, 1\right) = \beta_2/2$, that is the point in the middle of the side of the inner hexagon). Similarly, the Eisenstein series $E_{P_2,\pi}(g,s)$, where π is an automorphic cuspidal representation on $M_2^{(1)} \simeq PGL_2$ has a pole at $s = \frac{1}{2}$ if $L(\left(\frac{1}{2}, \pi\right) \neq 0$ (see [K2]). If π_∞ is spherical and tempered the residue produces a point in C_1 (this time the eigenvalue 1/4 produces the point $\beta_1/2$ - ie the midpoint of the outer hexagon).

It is a deeper fact that $\{\beta_4\}$ and a dense subset of points in C_1 can be produced cuspidally. In [G-G-J], Gan-Gurevich and Jiang show that $S_3 \times G$ can be realized as a dual pair in $H = \mathrm{Spin}(8) \ltimes S_3$. Restricting the automorphic minimal representation of $H(\mathbb{A})$ ([G-R-S]) to $S_3 \times G$ yields a correspondence between automorphic forms on S_3 and G. The spherical representation β_4 of $G(\mathbb{R})$ is a constituent of this restriction. Moreover, by comparing what they construct with the multiplicities of the residual spectrum, they show that β_4 occurs as an archimedian component of a cusp form in $L^2_{\mathrm{cusp}}(G(\mathbb{Q})\backslash G(\mathbb{A}_\mathbb{Q}))$. A dense set of points in C_1 corresponding to cuspidal representations was constructed by Rallis and Schiffmann [R-S2] using the oscillator representation of w. They realize $G \times \widetilde{SL_2}$ as a subgroup of Sp_{14}. While this does not form a dual pair they show that nevertheless restricting w to $G \times \widetilde{SL_2}$ yields a correspondence between forms on $\widetilde{SL_2}$ and G. In particular, suitable cuspidal representations σ of $\widetilde{SL_2}$ are transferred to automorphic cusp forms $\pi(\sigma)$ on $G(\mathbb{A}_\mathbb{Q})$ and the corresponding $\pi(\sigma)_\infty$ lies in C_1 (assuming that σ_∞ is tempered). For example, choosing σ appropriately, one can produce the point $\beta_1/2$ in $\mathfrak{a}^*_\mathbb{R}$ cuspidally. In [G-G-J] and [G-G] the authors compute the Arthur parameters (see (53) below) explicitly corresponding to these cuspidal automorphic forms on $G(\mathbb{A}_\mathbb{Q})$. They find an excellent agreement with the Arthur Conjectures for G.

Our discussion above shows that the lower bound (52) is achieved by various parts of the spectrum. Unfortunately, I don't know of any nontrivial upper bounds for $\widehat{G(\mathbb{R})}^{\text{sph}}_{\text{AUT}}$ (either for this G or any other exceptional group). An interesting start would be to establish that β_4 is isolated in $\widehat{G(\mathbb{R})}_{\text{AUT}}$. The natural conjecture here about this part of the automorphic dual of G is that the inclusion (52) is an equality.

The above are typical examples of the use of dual pairs in constructing automorphic representations and in particular non-tempered ones. As a final example we consider the case of a group of type F. We fixate on the problem of cohomology in the minimal degree. Let $F_{4,4}(\mathbb{R})$ be the real split group of type F_4 and of rank 4 (see the description and notation in Helgason [He]). The corresponding symmetric space $F_{4,4}(\mathbb{R})\big/Sp(3) \times Sp(1)$ has dimension 28. For Γ a co-compact lattice in $F_{4,4}(\mathbb{R})$ the cohomology groups $H^j(\Gamma, \mathbb{C})$ vanish for $0 < j < 8$, $j \neq 4$ (see [V-Z]). For $j = 4$ the cohomology comes entirely from parallel forms (ie from the trivial representation in Matsushima's formula) and so $\dim H^4(\Gamma, \mathbb{C})$ is constant (ie independent of Γ). So the first interesting degree is 8. According to Vogan [Vo1] there is a non-tempered cohomological representation ψ, which is isolated in $\widehat{F_{4,4}(\mathbb{R})}$ and which contributes to $H^8(\Gamma, \mathbb{C})$. Now let G be an algebraic group defined over $\mathbb{Q}$ (after restriction of scalars) such that $G(\mathbb{Q}_\infty) \simeq F_{4,4}(\mathbb{R}) \times$ compact and with $G(\mathbb{Q})\backslash G(\mathbb{A}_\mathbb{Q})$ compact. Using the classification in [Ti] one can show (see [B-L-S2]) that G contains a symmetric $\mathbb{Q}$ subgroup H such that $H(\mathbb{Q}_\infty) \simeq \text{Spin}_\mathbb{R}(5,4) \times$ compact. According to Oshima's computation of the spectra of the affine symmetric space $F_{4,4}(\mathbb{R})\big/\text{Spin}_\mathbb{R}(5,4)$ one finds that ψ occurs discretely in $\text{Ind}_{\text{Spin}_\mathbb{R}(5,4)}^{F_{4,4}(\mathbb{R})} 1$. Hence according to (26), $\psi \in \widehat{G(\mathbb{Q}_\infty)}_{\text{AUT}}$. Since ψ is isolated in $F_{4,4}(\mathbb{R})$ it follows that ψ occurs in $L^2(\Gamma\backslash F_{4,4}(\mathbb{R}))$ for a suitable congruence subgroup Γ of $G(\mathbb{Z})$. Using the classification of lattices Γ (see [Ma]) in $F_{4,4}(\mathbb{R})$ one can show in this way that for any such lattice Γ and any $N > 0$ there is a subgroup Γ' of finite index in Γ such that $\dim H^8(\Gamma, \mathbb{C}) > N$ (see [B-L-S2]). For a survey of results concerning nonvanishing of cohomology in the minimal degree see [Li-Sc].

This concludes our list of examples. We return to the general G. In Example 2, the upper bound (37) implies the useful fact that the trivial representation **1** is isolated in $\widehat{G_f(\mathbb{R})}_{\text{AUT}}$. It was conjectured by Lubotzky and Zimmer that this feature is true in general. That is if G is a semi-simple group defined over F then the trivial representation is isolated in $\widehat{G(F_v)}_{\text{AUT}}$ for any place v of F (they called this property τ). Of course, if v is a place at which $G(F_v)$ has property T then there is nothing to prove. Clozel [Cl1] has recently settled this property τ conjecture, this being the first general result of this kind. One proceeds by exhibiting (in all cases where $G(F_v)$ has rank 1 for some place v) an F subgroup H of G, for which the isolation property is known for $\widehat{H(F_v)}_{\text{AUT}}$ and hence by the restriction principle (27) this allows one to deduce the isolation property for $\widehat{G(F_v)}_{AUT}$. For example if G is isotropic then such an H isomorphic to SL_2 (or PGL_2) can be found. Hence by (19) and (20) the result follows. If G is anisotropic then he shows that G contains an F subgroup H isomorphic to $SL(1, D)$ with D a division algebra of prime degree over F or $SU(D, \alpha)$, a unitary group corresponding to a division algebra D of prime degree over a quadratic extension E of F, and with α an involution of the second kind (cf

Example 3 above). Thus one needs to show that the isolation property holds for these groups. For $SL(1,D)$ this follows the generalized Jacquet-Langlands correspondence [A-C] and the bounds (14) and (15) (for $GL_p(\mathbb{A}_F)$ with p-prime there are no non-trivial residual of Eisenstein series so the discrete spectrum is cuspidal). For the above unitary groups $G = SU(D,\alpha)$ of prime degree Clozel establishes the base change lift from G over F to G over E (this being based on earlier works by Kottwitz, Clozel and Labesse). Now G over E is essentially $SL(1,D)$ over E so one can proceed as above. As Clozel points out, it is fortuitous that these basic cases that one lands up with, are among the few for which one can stabilize the trace formula transfer at present.

It is of interest (see comment 3 of Section 3) to know more generally which π_v's are isolated in $\widehat{G(F_v)}_{\mathrm{AUT}}$? In this connection a natural conjecture is that if $G(F_v)$ is of rank 1 then every non-tempered point of $\widehat{G(F_v)}_{\mathrm{AUT}}$ is isolated.

At the conjectual level, Arthur's Conjectures [A2] give very strong restrictions (upper bounds) on $\hat{G}_{\mathrm{AUT}}$. While these conjectures involve the problematic group L_F, they are functorial and localizing them involves the concrete group¶ L_{F_v} and its representations. In this way these conjectures impose explicit restrictions on the automorphic spectrum. For example, if G is a split group over F then the local components π_v of an automorphic representation π occuring discretely in $L^2(G(F)\backslash G(\mathbb{A}_F))$ must correspond to certain Arthur parameters. In the unramified case these are morphisims of the local Weil group times $SL_2(\mathbb{C})$ into ${}^LG$ satisfying further properties. That is

$$\psi : W_{F_v} \times SL(2,\mathbb{C}) \longrightarrow {}^LG \tag{53}$$

such that

(i) $\psi\big|_{W_{F_v}}$ is unramified and $\psi(\mathrm{Frob}_v)$ lies in a maximal compact subgroup of ${}^LG$ (c.f. (9)).

(ii) If j is the unramified map of $W_{F_v} \longrightarrow SL(2,\mathbb{C})$ which sends Frob_v to $\begin{bmatrix} N(v)^{1/2} & 0 \\ 0 & N(v)^{1/2} \end{bmatrix}$, then the corresponding Arthur parameter is the conjugacy class ψ (Frob_v, j (Frob_v)) in ${}^LG$.

(54)

Thus the $SL(2,\mathbb{C})$ factor in (49) allows for non-tempered parameters but they are highly restricted.

In many of these split examples these local restrictions are probably even sharp and hence yield precise conjectures for $\widehat{G(F_v)}_{\mathrm{AUT}}$. We note however that it is by no means clear that the upper bounds imposed on $\widehat{G}_{\mathrm{AUT}}$ by Arthur's parameters are consistent for example with the lower bound (26) which must hold for all subgroups H. Establishing this would be of interest. As Clozel [Cl3]

¶ L_{F_v} is simply W_{F_v} if v is archimedian and is $W_{F_v} \times SU(2,\mathbb{R})$ if v is finite.

has shown, the Arthur Conjectures together with (26) and (27) (in the form extended to $\widehat{G(S)}_{\mathrm{AUT}}$) lead to some apparently non-obvious statements and structures for unipotent representations of local groups. Assuming a general twisted form of the "Fundamental Lemma" (see Hales' lectures) Arthur [A4] using the trace formula, gives a precise transfer of automorphic forms from classical orthogonal and symplectic groups to the corresponding general linear group. Hence, if and when this fundamental lemma is established, one will be able to combine this transfer with the bounds of Section 1 to get new sharp upper bounds for $\widehat{G}_{\mathrm{AUT}}$ with G classical. Armed with this one might be able to apply (27) for G exceptional and H a suitable large classical subgroup to obtain upper bounds for $\widehat{G}_{\mathrm{AUT}}$ when G is of exceptional type. In the meantime, the functorial transfer of generic representations from classical groups to GL_n is known (Cogdell, Kim, Piatetsky-Shapiro and Shahidi [C-K-PS-S]) and so this could probably be used, at least to give upper bounds for the generic part of $\widehat{G}_{\mathrm{AUT}}$, when G is classical. Such upper bounds should of course be compared with the generic unitary dual which has recently been determined for classical groups by Lapid, Muić and Tadić [L-M-T].

§3. APPLICATONS

The Ramanujan Conjectures and their generalizations in the form that we have described them and especially the upper bounds, have varied applications. We give a brief list of some recent ones.

1. For GL_2/F there are applications to the problem of estimation of automorphic L-functions on their critical lines and expecially to the fundmanetal "sub-convexity" problem. See [I-S] and [Sa] for recent accounts as well as for a description of some of the applications of sub-convexity.

2. The problem of counting asymptotically integral and rational points on homogeneous varieties for actions by semi-simple and reductive groups as well as the equi-distribution of "Hecke Orbits" on homogeneous spaces, depends directly on the upper bounds towards GRC. For recent papers on these topics see [Oh], [C-O-U], [G-O], [S-T-T] and [G-M] and also [Sa2].

3. There have been many works concerning geometric constructions of cohomology classes in arithmetic quotients of real and complex hyperbolic spaces. Bergeron and Clozel have shown that the injectivity of the inclusion and restriction of cohomology classes associated with $H < G$ (here H and G are $SO(n,1)$ or $SU(m,1)$) can be understood in terms of the isolation properties of these cohomological representations in $\widehat{G(F_\infty)}_{\mathrm{AUT}}$. This allows for an elegant and unified treatment of the constructions of cohomology classes as well as for reaching extensions thereof. They have also established the isolation property for some unitary groups. See [Be] and the references therein.

4. Mueller and Speh [M-S] have recently established the absolute convergence of the spectral side of the Arthur trace formula for GL_n. Their proof requires also the extension of (14) and (15) to ramified representations of $GL_n(F_v)$, which they provide. Their work has applications to the construction of cusp forms on GL_n and in particular to establish that Weyl's Law holds for the cuspidal spectrum.

5. An older application to topics outside of number theory is to the construction of highly connected but sparse graphs ("Ramanujan Graphs"). These applications as well as ones related to problems of invariant measures are described in the monograph of Lubotsky [Lu]. The proterty τ conjecture mentioned in Sectin 2 is related to such applications.

For a discussion of the automorphic spectral theory of $GL_2(\mathbb{A}_{\mathbb{Q}})$ in classical language see [Sa]. The recent article [Cl2] is close in flavor to these notes and should be consulted as it goes into more detail at various places.

ACKNOWLEDGEMENTS: I thank N. Gurevich, E. Lapid, J.S. Li and H. Oh for comments and discussions concerning these notes.

§4. REFERENCES

[A] J. Arthur, *BAMS* Vol. 40, No 1, (2002), 39-53.

[A2] J. Arthur, *Asterisque*, 171-172, (1989), 13-71.

[A3] J. Arthur, *CAN. Math. Bull.*, **45**, (2002), 466-482.

[A4] J. Arthur, Lecture Series I.A.S., (2001).

[A-C] J. Arthur and L. Clozel, *Ann. Math. Studies*, (1989).

[Be] N. Bergeron, *IMRN*, **20**, (2003), 1089-1122.

[B-S] M. Burger and P. Sarnak, *Invent. Math.*, **106**, (1991), 1-11.

[B-L-S1] M. Burger, J.S. Li and P.Sarnak, *BAMS*, (1992), 253-259.

[B-L-S2] M. Burger, J.S. Li and P. Sarnak, "Ramanujan duals and automorphic spectrum III," (in preparation).

[B-W] A. Borel and N. Wallach, *Ann. Math. Studies*, **94**, (1980).

[B-C] N. Bergeron and L. Clozel, *C.R. Acad. Sci.* Paris Ser. I, **334** (2002), 995-998.

[Cl1] L. Clozel, *Invent. Math*, **151**, (2003), 297-328.

[Cl2] L. Clozel, "Spectral theory of automorphic forms," (2003), (preprint).

[C-U] L. Clozel and E. Ullmo, "Equidistribution de points de Hecke,", (2001), (preprint).

[C-O-U] L. Clozel, H. Oh and E. Ullmo, *Invent. Math.*, **144**, (2001), 327-351.

[Co] J. Cogdell, Fields Institute Lectures on GL_n and converse theorems, (2003).

[Con] B. Conrad, "Modular forms, cohomology and the Ramanujan Conjectures," *CUP*, (to appear).

[C-K-PS-S] J. Cogdell, H. Kim, I. Piatetsky-Shapiro and F. Shahidi, Publ. IHES, **93**, (2001), 5-30.

[Cl3] "Combinatorial consequences of Arthur's conjectures and the Burger-Sarnak method," (2003), (preprint).

[de-W] D. de George and N. Wallach, *Annals of Math.*, **107**, (1978), 133-150.

[D-I] W. Duke and H. Iwaniec, "Automorphic forms and analytic number theory," (1989), *CRM*, Montreal, (1990), 43-47.

[Di] J. Dixmier, "Le C^*-algebras et leurs representations," Paris, (1969).

[G-M] D. Goldstein and A. Mayer, *Forum Math.*, **15**, (2003), 165-189.

[G-O] W. Gan and H. Oh, *Compositio Math.*, Vol. 323, (2003), 323-352.

[Ge] S. Gelbart, *Annal. of Math. Studies*, Vol. 83, (1975).

[G-G] W. Gan and N. Gurevich, "Non-tempered A packets of G_2: Liftings from $\overline{SL_2}$," (2003), (preprint).

[G-G-J] W. Gan, N. Gurevich and D. Jiang, *Invent. Math.*, **149**, (2002), 225-265.

[G-R-S] D. Ginzburg, D. Soudry and S. Rallis, *Israel Jnl. Math.*, **100**, (1997), 61-116.

[H-T] M. Harris and R. Taylor, *Annal. of Math. Studies*, **151**, (2002).

[H-M] R. Howe and C. Moore, *Jnl. Funct. Anal.*, **32**, (1979), 72-96.

[Ho] R. Howe, Notes on the oscillator representation, (unpublished).

[He] S. Helgason, "Differential Geometry, Lie Groups and Symmetric Spaces," A.P., (1978).

[I-S] H. Iwaniec and P. Sarnak, *GAFA*, (2000), 705-741.

[J-PS-S] H. Jacquet, I. Piatetsky-Shapiro and J. Shalika, *Am. Jnl. of Math.*, **103**, (1981), 499-558.

[J-W] Johnson and N. Wallach, *T.A.M.S.*, **229**, (1977), 137-173.

[K-S] H. Kim and F. Shahidi, *Ann. of Math.*, **155**, (2002), 837-893.

[K] H. Kim, *JAMS*, Vol. 16, **1**, (2002), 139-183.

[K2] H. Kim, *Can. J. Math.*, Vol. 48, (1996), 1245-1272.

[Ku] N. Kurokawa, *Invent. Math.*, **49**, (1978), 149-165.

[Ka-Sav] D. Kazhdan and G. Savin, *Israel Math. Conf. Proc.*, **2**, (1990), 209-223.

[Ki-Sa] H. Kim and P. Sarnak, Appendix to [K].

[K-Z] A. Knapp and G. Zuckerman, *Ann. of Math.*, **116**, (1982), 389-500.

[Li1] J.S. Li, *Math. Appl.*, **327**, Kluwer, (1995), 146-169.

[Li2] J.S. Li, "Minimal Representations and Reductive Dual Pairs," *IAS*, Park City Math. Series, Vol. 8, (2000).

[Lang1] R. Langlands in A.M.S., *Proc. Sym. Pure Math.*, Vol. XXXIII, part 2, (1979), 205-246.

[Lang2] R. Langlands, Springer Lecture Notes, Vol. 170, (1970), 18-86.

[Lu] A. Lubotsky, "Discrete groups, expanding graphs and invariant measures," Burkhauser, (1994).

[L-R-S1] W. Luo, Z. Rudnick and P. Sarnak, *Proc. Sym. Pure Math.*, **66**, Vol. II, (1999), 301-311.

[L-R-S2] W. Luo, Z. Rudnick and P. Sarnak, *GAFA*, **5**, (1995), 387-401.

[Lan] E. Landau, *Nachr. Ge Wiss Göttingen*, (1915), 209-243.

[Li-Sc] J.S. Li and J. Schwermer, "Automorphic representations and cohomology of arithmetic grows," In Challanges for the 21$^{\text{st}}$ Century, *World Scientific*, (2001), 102-138.

[L-M-T] E. Lapid, G. Muić and M. Tadić, "On the generic unitary dual of quasi-split classical groups," (2003), preprint.

[Ma] G. Margulis, *Ergeb. Math.*, Grenzeb(3), Vol. 17, Springer, (1989).

[M-S] W. Mueller and B. Speh, "Absolute convergence of the spectral side of the Arthur trace formula for GL_n," (2003), (preprint).

[M-W] C. Moeglin and L. Waldpurger, *Ann. Ecole Norm. Sup*, **22**, (1989), 605-674.

[Oh] H. Oh, "Harmonic Analysis and Hecke Operators," in rigidity in dynamics and geometry, (2002).

[O-M] T. Oshima and T. Matsuki, *Adv. Studies in Pure Math*, **4**, (1984), 331-390.

[Ro] J. Rogawski, *Ann. Math. Studies*, **123**, (1990).

[R-S1] S. Rallis and Schiffmann, *BAMS*, **83**, (1977), 267-276.

[R-S2] S. Rallis and G. Schiffmann, *Amer. Jnl. of Math.*, **111**, (1989), 801-849.

[Roh] D. Rohrlich, *Invent. Math.*, **97**, (1989), 383-401.

[Sa] P. Sarnak, www.math.princeton.edu/˜sarnak file baltimore.

[Sa2] P. Sarnak, In Proceedings of the *I.C.M.*, Kyoto, (1990), Vol. 1, 459-471.

[Sel] A. Selberg, *Proc. Symp. Pure Math.*, VIII, (1965), 1-15.

[Ser] J. P. Serre, Letter to Jacquet, (1981).

[Sh] F. Shahidi, "Automorphic L-functions: a survey," in book edited by Clozel Mike AP, (1988), 49-109.

[Sat] I. Satake, *Proc. Symp. AMS*, **9**, (1971), 258-264.

[Shi] G. Shimura, *Proc. London Math. Soc.*, **31**, (1975), 79-98.

[S-T-T] J. Shalika, R. Takloo-Bighash and Y. Tschinkel, "Rational points and automorphic forms," (2002), (preprint).

[Ta] J. Tate, *Proc. Symp. Pure Math.*, Vol. XXXIII, part 2, (1979), 3-26.

[Ti] J. Tits, *Proc. Symp. Pure Math.*, **IX**, (1967), 33-62.

[Ven] T.N. Venkataramana, *IMRN*, No. 15, (1999), 835-838.

[Vo1] D. Vogan, "Isolated unitary representations," (1994), (preprint).

[Vo2] D. Vogan, *Advanced Studies in Pure Math*, **14**, (1988), 191-221.

[Vo3] D. Vogan, *Invent. Math.*, **116**, (1994), 677-791.

[V-Z] D. Vogan and G. Zuckerman, *Compositio Math.*, **53**, (1984), 51-90.

[Wa] N. Wallach, *BAMS*, Vol. 82, No. 2, (1976), 171-195.

[Wal] L. Waldspurger, *J. Math. Pures et Appl.*, **60**, (1981), 365-384.

[We] A. Weil, *Acta Math.*, **111**, (1964), 143-211.

October 15, 2003