Analytic Number Theory and Families of Automorphic L-functions

Philippe Michel (version of May 24, 2006)

Contents

IAS/Park City Mathematics Series
Volume 00, 0000

Analytic Number Theory and Families of Automorphic L-functions

Philippe Michel

Foreword

These notes describe several applications of methods from analytic number theory to the theory of automorphic L-functions with a special focus on the methods involving families. Families of L-functions occur naturally in Analytic Number Theory. Indeed, many arithmetical quantities can be evaluated via harmonic analysis on some appropriate space of automorphic forms: the original object, φ say, is decomposed spectrally into a sum of eigencomponents along a family $\mathcal{F} = \{\pi\}$ of automorphic representations. Then a "simple" method to evaluate the original object is to combine a trivial averaging over $\mathcal{F}$ with non-trivial bounds for each eigencomponent φ_π; in many instance the bounds can be deduced from the (non-trivial) analytic properties of an L-function attached to each individual π, $L(\pi, s)$ say. A typical example following this principle is the problem of counting the number of primes in a given arithmetic progression: the standard method, which goes back to Dirichlet, is to express the characteristic function of the arithmetic progression as a sum over the family $\mathcal{F} = \{\chi(\mod q)\}$ of characters with the given modulus. Then the possibility of accurately counting the primes in this progression depends directly on the existence of a non-trivial zero free region for the L-function of each character; moreover, the quality of the counting is linked directly to the width of the zero-free region. Of course, establishing the individual analytic properties of L-functions is one of the main goals of Analytic Number Theory; however, in these lectures we will not focus (directly) on the individual aspects but rather on the global properties of the underlying families of L-functions $\{L(\pi, s)\}_{\pi\in\mathcal{F}}$. A first motivation is that for most advanced problems, the above, rather direct, method may not be sufficient, either because the individual analytic properties of the L-functions are not available unconditionally (like the Generalized Riemann Hypothesis) or, even if they were available, they wouldn't be strong enough (even the Generalized Riemann Hypothesis has some limitations !). A good part of the analytic number theory of the last

[1]Mathématiques, Université Montpellier II CC 051, 34095 Montpellier Cedex 05, FRANCE..
E-mail address: `michel@math.univ-montp2.fr`.
1991 *Mathematics Subject Classification.* 11F70.
The author was partially supported by the Institut Universitaire de France and the ACI Jeunes Chercheurs "Arithmétique des fonctions L".

century was devoted to the development of techniques capable of providing good substitutes for unproved hypotheses (zero density theorems, for instance are used as a substitute to GRH). Another (probably as important) motivation is the recent realization that global analytic estimates for a family of L-functions can be used to give non-trivial estimates for each of its members (or sometimes for other related L-functions). The method of moments or its refinement, the *amplification* technique, is able to provide sharp individual estimates for critical values of L-functions and is a good example of this principle. A striking application of the method of moments is the recent improvement by Conrey/Iwaniec of Burgess's 40 years standing bound for the values of $L(\chi, s)$, for χ a quadratic character and s on the critical line. Interestingly, the bound follows from a bound for the third moment of the central value L-function for a family of GL_2-modular forms (including Eisenstein series), rather than for a family of Dirichlet characters. The fact that individual estimates can be obtained from global ones is another (somewhat coarse) manifestation of the powerful principle that much insight can be gained for an individual object, if one is able to deform it into a family and to obtain enough global information on the deformation space: this principle was beautifully illustrated, in the past, in Deligne's proof of the Weil conjectures (GRH for L-functions over finite fields) and in the more recent solution by Wiles of Fermat's Last Theorem.

The lectures are organized as follows:

- In the first lecture, we review various aspects of the analytic theory of *individual* automorphic L-functions; more precisely, we describe their functional equation, the standard conjectures and what is known unconditionally: this includes the bounds for their local parameters (the Ramanujan/Petersson Conjecture, RPC), the location of their zeros (the Generalized Riemann Hypothesis, GRH), and the size of their values on the critical line (the Generalized Lindelöf Hypothesis, GLH). We prove the so-called *convexity bound*, which is the best result towards GLH known in general, introduce the Subconvexity Problem (ScP), which is the problem of improving the convexity bound, and describe how families can be used to solve it. At the end of the lecture, we illustrate the usefulness of families of automorphic forms for individual estimates with the Theorem of Luo/Rudnick/Sarnak, which gives the best general (and non-trivial) approximation to the RPC.
- The second lecture gives a short description of the analytic theory of GL_2-automorphic forms from the classical viewpoint: so far, it is essentially in the classical setting, that the most advanced aspects of the analytic theory have been developed (a notable exception is the recent subconvexity bound for L-functions for Hilbert modular forms of Cogdell/Piateski-Shapiro/Sarnak). We compare automorphic L-functions with their classical counterparts. We also give various "trace formulas," which are the main tool for performing averaging over families; these formulas transform an averaging over a family into a "dual" side that putatively is more tracktable. Particularly important to the GL_2 theory is the Petersson/Kuznetsov formula, which expresses the average of the Fourier coefficients of a modular form in terms of sums of Kloosterman sums. Note that for analytic purposes, this formula is more powerful than the Selberg trace formula, for instance. However, trace formulas alone are not sufficient (mainly because they are involutory) and so

they must be supplemented by various techniques of transformation or of estimation, some of which will be described in the following lectures.

- In the third lecture, we discuss an important method involving families: the large sieve method and the large sieve type inequalities. These inequalities are relatively coarse but also very robust and they have many interesting applications, one of which is the problem of bounding non-trivially the dimension of the space of holomorphic forms of weight one and given conductor. By an elegant application of the large sieve, Duke was the first to make progress on this question. We also show how such inequalities can be used to produce zero density estimates that are good substitutes for GRH.
- The fourth lecture is the most significant: it gives an overview of the various ingredients used in the resolution of the **Subconvexity Problem** of GL_1, GL_2 and $GL_2 \times GL_2$, L-functions which is by now to a large extend solved.
  1. The methods of Weyl and Burgess work well for L-functions of degree one but are hardly extendible to L-functions of higher degree.
  2. The method of moments and the amplification method, which are built on families, are to date the most general methods available for solving the Subconvexity Problem. The amplification method may have other applications: for instance, we will use it to give easy improved bounds for the dimension of the space of holomorphic forms of weight one.
  3. Next, we show how these methods reduce the Subconvexity Problem for modular L-functions to another one: the **Shifted Convolution Problem, SCP**. It consists of bounding non-trivially partial sums of Rankin/Selberg type, but with an additional non-trivial additive twist. We describe two somewhat independant methods for solving the **ScP** (in fact, these methods are related via the Petersson/Kuznetzov trace formula): an elementary approach that builds on the δ-symbol and relies ultimately on non-trivial bounds for Kloosterman sums, and a spectral approach, inspired by the Rankin-Selberg unfolding method.
  4. The latter approach uses the full force of the theory of Maass forms (even if one is only interrested in L-functions of holomorphic forms) and requires a non-trivial bound for their local parameter at infinity. It also depends on good bounds for integrals of triple products of modular forms: such bounds can be obtained by various rather advanced techniques, which we have no time to describe here.

Finally, as an illustration, we collect all these methods together in the proof of the Subconvexity Problem for Rankin-Selberg L-functions (which are of degree 4). Interestingly, this case "closes the circle," since the proof (of this individual estimate), starts with families and ends up after several transformations with another set of non-trivial individual estimates for another set of modular forms (of course quite different from the one we started from): namely the non-trivial approximations to the **RPC** for Maass forms and the non-trivial (subconvex) bounds for L-functions of degree 1 and 2 (which have been proved before). Thus one may suspect that the complete use of the interplay between individual type bounds and averaged bounds is far from finished.

- In the fifth and last lecture we discuss many applications of the Subconvexity Problem. Our aim is to convince the reader, by means of examples, that a subconvex bound is not simply another improvement over some existing exponent, but has an intrisic geometrical or arithmetical meaning. For instance, we show how in several cases the convexity bound matches exactly another bound obtained by other generic methods of geometric or arithmetic nature (the Riemann/Roch or the Minkowski Theorem). Subconvexity bounds are also used to establish a variety of equidistribution results, ranging from the distribution of lattice points on the sphere, to the Heegner points, and the context of Quantum Chaos.

There are several important topics connected with families of L-functions that can be handled by similar techniques and which, for lack of time and energy, we have not treated in these lectures. One is the problem of proving the existence of an L-function in a given family that does not vanish at some special (meaningful) point. An example however, is given at the end of the first lecture. This kind of problem can be handled by many methods (such as the mollification method) and has many applications in various fields. The other topic is the "Katz/Sarnak philosophy," which is a net of far reaching conjectures (going far beyond GRH) describing the local distribution of zeros of families of L-functions in terms of the eigenvalues of random matrices of large rank. Although this field is highly conjectural, it is nevertheless important, as it reveals beautiful inner structures in families and provides a unified framework for various phenomena occurring in analytic number theory; moreover, it strongly suggests that further progress on GRH will come from the use of families.

These notes are an extended version of a series of five lectures given during the Park City Mathematical Institute in july 2002[2]. I hope that these notes have retained the informal style of the lectures. In particular, few proofs have been given in full detail: we hope that this will serve to capture better the main ideas. To fill the many remaining gaps, the reader will need to look further at the existing literature. The other lectures of the present volume should be fully sufficient to cover the part relevant to the general theory of automorphic forms. For a more complete introduction to the general methods of analytic number theory, several good books are available: for the basics, the reader may consult Davenport's "Multiplicative Number Theory" and Tenenbaum's "Introduction a la Théorie Analytique et Probabiliste des Nombres" (also available in english) and for an introduction to advanced topics Bombieri's little (big) book "Le Grand Crible dans la Theorie Analityque des Nombres". For us, the ultimate reference is Kowalski/Iwaniec's book "Analytic Number Theory" (the series of H. Iwaniec's lectures at Rutgers University): these notes contain the most complete available account of the methods used in modern analytic number theory, presented in the short and elegant style characterizing its authors.

[2]ADDED IN PROOF (2006): since then, there has been a lot of progress made regarding several topics discussed in these lectures. One of the most striking recent progress, in our opinion, are the Berstein/Reznikov and Venkatesh new approaches to the subconvexity problem [**Ve, BR3**]. Their methods differ from the ones discussed in these lectures in that the subconvexity problem is approached in terms of bounds for *periods* of automorphic forms rather than of bounds for central values of L-functions and this has many advantages. For instance, in [**Ve**], many cases of the subconvexity problem in the level aspec are proven for L-functions related to GL_1 and GL_2 automorphic representation over an arbitrary fixed number field !

I would like to thank the PCMI audience for its attention, its questions, and its constructive criticism. I also thank the organizers, P. Sarnak and F. Shahidi and the PCMI staff for the quality of the organization. Some parts of these notes were written in june 2002 in Caltech under the invitation of D. Ramakrishnan and I thank this institution for its hospitality and excellent working conditions. Thanks are due to J. Cogdell, L. Clozel, E. Fouvry, G. Harcos, M. Harris, H. Iwaniec, E. Kowalski, D. Ramakrishnan, P. Sarnak, J.-P. Serre, F. Shahidi and H. Stark for their explanations, comments and corrections prior to, during, and after the writing of this text, and to E. Royer and L. B. Pierce who have undertaken the painful task of correcting many errors, inaccuracies and misprints of earlier versions; the ones that may remain or that I may have added in the final version are my responsibility.

LECTURE 1
Analytic Properties of Individual L-functions

1.1. Automorphic L-functions

1.1.1. Principal L-functions

Let $\pi = \otimes\pi_p$ be an automorphic cuspidal (irreducible) representation of $GL_d(\mathbf{A_Q})$ with unitary central character (we denote by $\mathcal{A}_d^0(\mathbf{Q})$ the set of all such representations). By the general theory (Hecke, Gelbart/Jacquet, Godement/Jacquet: see Cogdell's lectures in this volume [**Co2**]), π admits an L-function:

$$L(\pi, s) = \prod_{p<\infty} L_p(\pi, s) = \sum_{n\geqslant 1} \frac{\lambda_\pi(n)}{n^s}.$$

This is an Euler product absolutely convergent for $\Re e s$ sufficiently large where for each (finite) prime p, the inverse of the local factor $L_p(\pi, s)$ is a polynomial in p^{-s} of degree $\leqslant d$:

$$L_p(\pi, s)^{-1} = L(\pi_p, s)^{-1} = \prod_{i=1}^{d}(1 - \frac{\alpha_{\pi,i}(p)}{p^s}).$$

The L-function of π is completed by a local factor at the infinite place, given by a product of d Gamma factors:

$$L_\infty(\pi, s) = L(\pi_\infty, s) = \prod_{i=1}^{d} \Gamma_{\mathbf{R}}(s - \mu_{\pi,i}),\ \Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\Gamma(s/2);$$

the coefficients $\{\alpha_{\pi,i}(p)\}_{i=1\dots d}$ (resp. $\{\mu_{\pi,i}\}_{i=1\dots d}$) will be called the local parameters of π at p (resp. ∞). The completed L-function $L_\infty(\pi, s)L(\pi, s)$ has the following analytic properties:

- $L_\infty(\pi, s)L(\pi, s)$ has a meromorphic continuation to the complex plane with at most two simple poles; the latter occur only if $d = 1$ and $\pi = |.|^{it}$ for some $t \in \mathbf{R}$, in which case $L(\pi, s) = \zeta(s + it)$ and the poles are at $s = -it, 1 - it$.
- $L_\infty(\pi, s)L(\pi, s)$ satisfies a functional equation of the form

$$q_\pi^{s/2} L_\infty(\pi, s)L(\pi, s) = w(\pi) q_\pi^{(1-s)/2} L_\infty(\tilde{\pi}, 1-s)L(\tilde{\pi}, 1-s), \tag{1.1}$$

where $q_\pi \geqslant 1$ is an integer (the *arithmetic conductor* of π) supported at the finite ramified places for π, $w(\pi)$ is a complex number of modulus 1 (the

root number) and $\tilde{\pi}$ denotes the contragredient of π. In particular, one has $q_{\tilde{\pi}} = q_\pi$, and for any place v $L_v(\tilde{\pi}, s) = \overline{L_v(\pi, \overline{s})}$.)
- $L_\infty(\pi, s)L(\pi, s)$ is bounded in vertical strips (and in fact has exponential decay) and is of finite order away from its poles (if any).

It is convenient to encapsulate the main parameters attached to π in a single quantity that occurs in many problems as a normalizing factor; for that purpose, Iwaniec and Sarnak [**IS2**] introduced the *analytic conductor* of π: it is a function over the reals given by

$$t \in \mathbf{R} \to Q_\pi(t) = q_\pi \prod_{i=1}^{d}(1 + |it - \mu_{\pi,i}|) =: q_\pi Q_{\pi_\infty}(t).$$

For the rest of these lectures, we denote $Q_\pi(0)$ by Q_π.

Remark 1.1. We mostly concentrate on cuspidal L-functions because they form the building blocks for L-functions of general automorphic representations. Indeed, given $d_1, \ldots, d_r$ with $d_1 + \cdots + d_r = d$, the Langlands theory of Eisenstein series associates to an r-tuple $(\pi_1, \ldots, \pi_r)$ of (not necessarily unitary) cuspidal representations a distinguished automorphic representation of $GL_d(\mathbf{A_Q})$, denoted $\pi = \pi_1 \boxplus \cdots \boxplus \pi_r$, the *isobaric sum* of the π_i, $i = 1 \ldots, r$. By construction, cuspidal representations are isobaric and it is a result of Shalika that the π_i appearing in the construction of π (the *constituents of* π) are unique up to permutation (see [**Co2**]). Then the L-function of π is given by the product

$$L(\pi_1 \boxplus \cdots \boxplus \pi_r, s) = \prod_{i=1}^{r} L(\pi_i, s).$$

Langlands also proved that any automorphic representation π is nearly equivalent to an isobaric sum π' (i.e. for almost every place v, $\pi_v \backsimeq \pi'_v$), and as a consequence $L(\pi, s)$ and $L(\pi', s)$ coincide up to finitely many local factors.

1.1.2. L-functions of pairs

Another class of L-functions fundamental to the whole theory are the Rankin/-Selberg type L-functions $L(\pi \otimes \pi', s)$ associated to pairs of automorphic representations $(\pi, \pi') \in \mathcal{A}_d^0(\mathbf{Q}) \times \mathcal{A}_{d'}^0(\mathbf{Q})$; their theory was initiated by Rankin and Selberg in the case of classical modular forms [**Ran, Se**]. For general automorphic forms, the analytic theory of L-functions of pairs was initiated and developped in several papers by Jacquet, Piatetsky-Shapiro and Shalika [**J, JS2, JPSS2**] and completed in works of Shahidi, Moeglin/Waldspurger and Gelbart/Shahidi [**Sha1, Sha2, Sha3, MW1, GeSh**]; we refer again to [**Co2**] for a detailled exposition of their construction and the derivation of their basic properties. Given $(\pi, \pi') \in \mathcal{A}_d^0(\mathbf{Q}) \times \mathcal{A}^0(d')$, the Rankin/Selberg type L-function $L(\pi \otimes \pi', s)$ is a Dirichlet series

$$L(\pi \otimes \pi', s) = \prod_p L_p(\pi \otimes \pi', s) = \sum_{n \geqslant 1} \frac{\lambda_{\pi \otimes \pi'}(n)}{n^s},$$

absolutely convergent for $\Re e s$ large enough. This is an Euler product of degree dd' with local factors of the form

$$L_p(\pi \otimes \pi', s) = L(\pi_p \otimes \pi'_p, s) = \prod_{i=1}^{dd'}(1 - \frac{\alpha_{\pi \otimes \pi', i}(p)}{p^s})^{-1}$$

at the finite places, and at the infinite one,

$$L_\infty(\pi\otimes\pi',s)=L(\pi_\infty\otimes\pi'_\infty,s)=\prod_{i=1}^{dd'}\Gamma_{\mathbf{R}}(s-\mu_{\pi\otimes\pi',i}).$$

Moreover, at places v for which π_v is unramified, $L_v(\pi\otimes\pi',s)$ has the explicit expression

$$L_p(\pi\otimes\pi',s)=\prod_{i=1}^{d}\prod_{i'=1}^{d'}(1-\frac{\alpha_{\pi,i}(p)\alpha_{\pi',i'}(p)}{p^s})^{-1} \tag{1.2}$$

at $v=p$ a finite place, and at the infinite place,

$$L_\infty(\pi\otimes\pi',s)=\prod_{i=1}^{d}\prod_{i'=1}^{d'}\Gamma_{\mathbf{R}}(s-\mu_{\pi,i}-\mu_{\pi',i'}). \tag{1.3}$$

These completed L-functions have the following analytic properties, which are proved in the above cited papers (again see [**Co2**]):

- $L_\infty(\pi\otimes\pi',s)L(\pi\otimes\pi',s)$ has a meromorphic continuation to $\mathbf{C}$ with at most two simple poles; the latter occur if and only if $\pi'\simeq\tilde{\pi}\otimes|\det|^{it}$ for some $t\in\mathbf{R}$ and are located at $s=-it,1-it$.
- $L_\infty(\pi\otimes\pi',s)L(\pi\otimes\pi',s)$ satisfies a functional equation of the form

$$q_{\pi\otimes\pi'}^{s/2}L_\infty(\pi\otimes\pi',s)L(\pi\otimes\pi',s)= w(\pi\otimes\pi')q_{\pi\otimes\pi'}^{(1-s)/2}L_\infty(\tilde{\pi}\otimes\tilde{\pi}',1-s)L(\tilde{\pi}\otimes\tilde{\pi}',1-s). \tag{1.4}$$

  where $q_{\pi\otimes\pi'}\geqslant 1$ is an integer and $w(\pi\otimes\pi')$ has modulus one.
- The "completed" L-function $L_\infty(\pi\otimes\pi',s)L(\pi\otimes\pi',s)$ is bounded in vertical strips as $|\Im m s|\to+\infty$ (with, in fact, exponential decay) and is of finite order away from its poles (if any).

The integer $q_{\pi\otimes\pi'}$ is by definition the conductor of (the pair) $\pi\otimes\pi'$; it is supported on the primes dividing $q_\pi q_{\pi'}$, and in fact one has the following upper bound [**BH**]

$$q_{\pi\otimes\pi'}\leqslant q_\pi^{d'}q_{\pi'}^{d}/(q_\pi,q_{\pi'}),$$

which now is an easy consequence of the local Langlands correspondance for GL_d. We denote, for $t\in\mathbf{R}$, by

$$Q_{\pi\otimes\pi'}(t)=q_{\pi\otimes\pi'}\prod_{i=1}^{dd'}(1+|it-\mu_{\pi\otimes\pi',i}|),$$

the analytic conductor of $\pi\otimes\pi'$. An archimedean analog of the bound given above for the conductors shows (and follow easily from (1.3) if π_∞ is unramified) that one has

$$Q_{\pi\otimes\pi'}(t)\ll_{d,d'}Q_\pi^{d'}Q_{\pi'}(t)^d. \tag{1.5}$$

L-functions of pairs are defined more generally for automorphic representations. In particular, for isobaric representations $\pi=\pi_1\boxplus\cdots\boxplus\pi_r$ and $\pi'=\pi'_1\boxplus\cdots\boxplus\pi'_{r'}$, one has

$$L(\pi\otimes\pi',s)=\prod_{i=1\dots r}\prod_{j=1\dots r'}L(\pi_i\otimes\pi'_j,s).$$

Remark 1.2. In particular for π unitary cuspidal and Π an isobaric sum of unitary representations, the multiplicity of π as a component of Π is given by $-\mathrm{ord}_{s=1}L(s,\pi\otimes\tilde{\pi})$.

In fact, L-functions of pairs are expected to be automorphic. More precisely a special case of the Langlands Functoriality Conjecture predicts the following

Functoriality Conjecture for Pairs. *Given $\pi = \pi_1 \boxplus \cdots \boxplus \pi_r$ and $\pi' = \pi'_1 \boxplus \cdots \boxplus \pi'_{r'}$ two isobaric sums of unitary cuspidal representations of $GL_d(\mathbf{A_Q})$ and $GL_{d'}(\mathbf{A_Q})$ respectively, there exists an isobaric sum of unitary cuspidal representations $\pi \boxtimes \pi'$ of $GL_{dd'}(\mathbf{A_Q})$ such that*

$$L(\pi \otimes \pi', s) = L(\pi \boxtimes \pi', s).$$

Moreover, one has the distributive formula

$$\pi \boxtimes \pi' = \boxplus_{i=1\dots r} \boxplus_{j=1\dots r'} \pi_i \boxtimes \pi'_j.$$

Consequently, $L(\pi \otimes \pi', s)$ should factor as a product of principal cuspidal L-functions,

$$L(\pi \otimes \pi', s) = \prod_{i=1\dots r''} L(\pi''_i, s)$$

where $\pi''_i \in \mathcal{A}^0_{d''_i}(\mathbf{Q})$, $i = 1\dots r''$ with $d''_1 + \cdots + d''_{r''} = dd'$, and the analytic properties that we will discuss below regarding principal L-functions should extend to L-functions of pairs. See [**Ra1**] for further discussions about this conjecture.

By now this conjecture is known for a few cases, namely in the trivial but important case of $d = 1$ (twists by characters), and for $d = 2$, $d' = 2, 3$ by the works of (among others) Weil, Shimura, Gelbart/Jacquet, Cogdell/Piatetski-Shapiro, Ramakrishnan, Kim and Shahidi (see [**Co2**] Lect. 6).

1.1.3. The Ramanujan/Petersson Conjecture

It is of fundamental importance for the study of automorphic forms (and for many applications of an arithmetic nature) to have good bounds for the local parameters $\alpha_{\pi,i}(p),\ \mu_{\pi,i},\ i = 1\dots d$. By the general theory it is known that $L(\pi, s)$ is absolutely convergent for $\Re es > \theta$ for some $\theta \geqslant 1$. Since $L(\pi, s)$ is an Euler product, it does not vanish in this domain and since $L(\pi_\infty, s)L(\pi, s)$ is holomorphic, this implies that

$$\log_p |\alpha_{\pi,i}(p)|,\ \Re e\mu_{\pi,i} \leqslant \theta, i = 1\dots d$$

(clearly we have assumed that π is not equal to $|.|^{it}$ for $t \in \mathbf{R}$). As was discovered by Rankin and Selberg, the analytic properties of $L(\pi\otimes\tilde{\pi}, s)$ are very usefull in providing stronger bounds. A fundamental property of $L(\pi\otimes\tilde{\pi}, s)$ is the non-negativity of the Dirichlet coefficients $\lambda_{\pi\otimes\tilde{\pi}}(n)$: set

$$\log L(\pi\otimes\tilde{\pi}, s) := \sum_{n\geqslant 1} \frac{\ell_{\pi\otimes\tilde{\pi}}(n)}{n^s};$$

Lemma 1.1.1. *For $n \geqslant 1$, $\ell_{\pi\otimes\tilde{\pi}}(n) \geqslant 0$; in particular, for any $n \geqslant 1$,*

$$\lambda_{\pi\otimes\tilde{\pi}}(n) \geqslant 0.$$

This lemma is easily proved for n coprime with q_π by (1.2); that it holds for every n follows from the structure of the admissible representations of GL_d and the expression of their local factors of pairs (see [**JPSS2, HR, RS2**]).

Remark 1.3. This positivity property extends to isobaric sums of unitary cuspidal representations

As explained in [**Co2**] 4.4, this property together with the non-vanishing of the local factors $L_v(\pi\otimes\pi', s)$ and the fact that $L_\infty(\pi\otimes\tilde{\pi}, s)L(\pi\otimes\tilde{\pi}, s)$ has no pole for $\Re es > 1$, implies by Landau's Lemma that $L^{(q_\pi)}(\pi\otimes\tilde{\pi}, s)$ is absolutely convergent for $\Re es > 1$, and non-vanishing in this domain; moreover, by Cauchy/Schwarz, the same is true of the L-series $L^{(q_\pi)}(\pi, s)$ and $L^{(q_\pi q_{\pi'})}(\pi\otimes\pi', s)$ for all pairs (π, π').

One can deduce the following bounds for the local parameters:

Proposition 1.1. *For $\pi\in\mathcal{A}_d^0(\mathbf{Q})$ and all $i\in\{1,\dots,d\}$ one has*

$$\Re e\mu_{\pi,i}\leqslant\frac{1}{2}\text{ and }\log_p|\alpha_{\pi,i}(p)|\leqslant\frac{1}{2},\text{ for all primes }p. \tag{1.6}$$

Moreover, for an unramified place, one has

$$|\Re e\mu_{\pi,i}|\leqslant\frac{1}{2}\text{ and }|\log_p|\alpha_{\pi,i}(p)||\leqslant\frac{1}{2}. \tag{1.7}$$

Proof. For each place v, the local factor $L_v(\pi\otimes\tilde{\pi}, s)$ has no pole for $\Re es > 1$, since this would contradict the non-vanishing of $L(\pi\otimes\tilde{\pi}, s)$ and the holomorphy of $L_\infty(\pi\otimes\tilde{\pi}, s)L(\pi\otimes\tilde{\pi}, s)$ in this domain. In particular one can deduce that

$$\log_p|\alpha_{\pi\otimes\tilde{\pi},i}(p)|,\ \Re e\mu_{\pi\otimes\tilde{\pi},i}\leqslant 1,\ i=1\dots dd'.$$

Now the expression of the local factor at an unramified place (1.2) implies (1.7). The bounds (1.6) for the remaining (ramified) places follow from the structure of the admissible representations of $GL_d(\mathbf{Q}_v)$ and the expression of local factors of pairs; for more detail we refer to the Appendix of [**RS2**]. □

In fact, the bounds (1.6) and (1.7) can be obtained with the stronger inequality $<1/2$ by purely local arguments, using the fact that the local components of π are *generic* (see [**JS2, HR**]). However, as we shall see, improving beyond $1/2$ requires global arguments. We refer to the bounds given in Proposition 1.1 as the "trivial" bounds for the local parameters of π. More generally, given $0\leqslant\theta\leqslant 1/2$, one can consider the following:

Hypothesis $\mathrm{H}_d(\theta)$. *For all $d'\leqslant d$, all $\pi\in\mathcal{A}_d^0(\mathbf{Q})$ and all $i\in\{1,\dots,d'\}$, one has*

$$\Re e\mu_{\pi,i}\leqslant\theta\text{ and }\log_p|\alpha_{\pi,i}(p)|\leqslant\theta\text{ for all primes }p. \tag{1.8}$$

Moreover, for an unramified place, one has for all $i\in\{1,\dots,d\}$

$$|\Re e\mu_{\pi,i}|\leqslant\theta\text{ and }|\log_p|\alpha_{\pi,i}(p)||\leqslant\theta. \tag{1.9}$$

Remark 1.4. Usually these bounds are presented for the unramified places only in the form of (1.9); however, for several technical purposes the corresponding bounds (1.8) for the ramified places are useful.

The bound $\mathrm{H}_d(1/2)$ is probably far from the truth: assuming that the Functoriality Conjecture for Pairs holds, applying (1.6) to $\pi^{\boxtimes k}$ gives the bounds $\mathrm{H}_d(1/2^k)$ for π at the unramified places (it holds for the ramified one too), and letting $k\to+\infty$, one obtains

Ramanujan/Petersson Conjecture (RPC). $\mathrm{H}_d(0)$ *is true.*

Unconditionally, Proposition 1.1 can be strengthened to the following (non-trivial) bound:

Theorem 1.1. *For any $d \geqslant 1$, $\mathrm{H}_d(\theta)$ is true for*

$$\theta = \theta_d = \frac{1}{2} - \frac{1}{d^2+1}.$$

We present below a proof due to Serre [**Ser3**][1] that works for the non-archimedean places and uses the most basic analytic properties of $L(\pi \otimes \tilde{\pi}, s)$; there are also two alternative methods developped by Duke/Iwaniec and Luo/Rudnick/Sarnak [**DuI, DuI2, LRS, LRS2**], which interrestingly both build on *families* of twists $L(\chi.\pi \otimes \tilde{\pi}, s)$ by appropriate characters. At the end of this lecture, we will present the method of Luo/Rudnick/Sarnak in the case of the archimedean place (but the method works for every place and can be extended to number fields as well).

Proof. The theorem is a consequence of the following refinment (due to Landau) of Landau's Lemma [**La**] (see also [**BR**]):

Theorem 1.2. *Let $L(s) = \sum_n \lambda(n) n^{-s}$ be a Dirichlet L-series with non-negative coefficients $\lambda(n)$, and convergent for $\Re e s$ sufficiently large. Assume that $L(s)$ has meromorphic continuation to $\mathbf{C}$ with at most poles of finite order at $s = 0, 1$; assume also that $L(s)$ is of bounded order in the half-plane $\Re e s \geqslant -1$, and that it satisfies a functional equation of the form*

$$q^s L_\infty(s) L(s) = w q^{1-s} L_\infty(1-s) L(1-s)$$

for some constants w, $q > 0$, where

$$L_\infty(s) = \prod_{i=1}^{d} \Gamma(\alpha_i s + \beta_i),$$

for some $d \geqslant 1$ and $\alpha_i \geqslant 0$, $\beta_i \in \mathbf{C}$ for $i \in \{1, \dots, d\}$. Setting $\eta = \sum_{i=1}^d \alpha_i$, one has as $x \to +\infty$,

$$\sum_{n \leqslant x} \lambda(n) = P(\log x) x + O_{\varepsilon, L}(x^{\frac{2\eta-1}{2\eta+1}+\varepsilon})$$

for all $\varepsilon > 0$, where P is some polynomial of degree $\mathrm{ord}_{s=1} L(s)$ and depends only on L.

Applying this to $L(\pi \otimes \tilde{\pi}, s)$, one obtains

$$\sum_{n \leqslant N} \lambda_{\pi \otimes \tilde{\pi}}(n) = c_\pi N + O_{\varepsilon, \pi}(N^{\frac{d^2-1}{d^2+1}+\varepsilon}), \tag{1.10}$$

for some $c_\pi > 0$; this yields

$$\lambda_{\pi \otimes \tilde{\pi}}(N) = O_{\pi, \varepsilon}(N^{1-\frac{2}{d^2+1}+\varepsilon})$$

by subtracting the $N-1$ sum from the N sum. Taking $N = p^k$ (with $k \to +\infty$), one deduces, for all p and all $i \in \{1, \dots, d^2\}$, the bound

$$|\alpha_{\pi \otimes \tilde{\pi}, i}(p)| \leqslant p^{1-\frac{2}{d^2+1}};$$

in particular, for p unramified, one has $|\alpha_{\pi, i}(p)|^2 \leqslant p^{1-\frac{2}{d^2+1}}$. A more carefull analysis of the ramified factors implies that the latter bound is valid at the remaining non-archimedean places. □

[1] which might even go back to Rankin

1.2. Zero-free regions for L-functions

The most important problem of the analytic theory of L-functions is the Generalized Riemann Hypothesis, giving the optimal zero-free region of an automorphic L-function:

Generalized Riemann Hypothesis (GRH). *Given $\pi \in \mathcal{A}_d^0(\mathbf{Q})$, the product $L_\infty(\pi,s)L(\pi,s)$ does not vanish for $\Re es \neq 1/2$.*

We have already seen that zero-free regions for $L(\pi\otimes\tilde{\pi},s)$ are usefull for providing bounds for the local parameters of π. Note that GRH is also expected for $L(\pi\otimes\tilde{\pi},s)$, since the latter is expected to be automorphic; for instance, one can show that **GRH** for $L(\pi\otimes\tilde{\pi},s)$ implies the bounds of (1.8) with $\theta = 1/8$ for π. However, the main application of zero-free regions is to provide good (in fact optimal) control of the sums $\sum_{p\leqslant x}\lambda_\pi(p)$, when p ranges over the prime numbers: for instance, under **GRH** (for $L(\pi,s)$), one has

$$\sum_{p\leqslant x}\lambda_\pi(p) = \delta_{\pi=triv}x + O_d(x^{1/2}\log^2(Q_\pi x)),\ as\ x\to+\infty.$$

In this section we review the most basic approximations to GRH (zero-free regions) and their most classical applications. We refer to [**Ser2**] for some other arithmetic applications.

1.2.1. The Hadamard/de la Vallée-Poussin method

Even in the case of Riemann's zeta function, very little is known about this conjecture (but we have theoretical and extensive numerical evidence to support it). At the end of the 19th century, J. Hadamard [**H**] and Ch. de la Vallée-Poussin [**VP**] proved (independently) that $\zeta(s)\neq 0$ for $\Re es = 1$. The non-vanishing of ζ at the edge of the critical strip turns out to be equivalent to the **Prime Number Theorem (PNT)**:

Prime Number Theorem. *As $x\to+\infty$, one has*

$$\sum_{p\leqslant x}1 = \frac{x}{\log x} + o(\frac{x}{\log x}).$$

A little later, de la Vallée-Poussin [**VP2**] extended the non-vanishing of ζ to an explicit region inside the critical strip:

Theorem 1.3. *There exists a $c>0$ such that $\zeta(s)$ does not vanish for*

$$\Re es \geqslant 1 - \frac{c}{\log(|t|+2)}. \tag{1.11}$$

Remark 1.5. This zero-free region gives the **PNT** as above with an error term of the form $O(x\exp(-c\sqrt{\log x}))$ for some absolute positive c; the Riemann Hypothesis gives the (optimal) error term $O(x^{1/2}\log x)$.

Suppose one needs to show that for a given π, $L(\pi,s)$ does not vanish at $s = \sigma_0 + it_0$ for some $t_0\in\mathbf{R}$ and some $\sigma_0\leqslant 1$ but close to 1; the Hadamard/De la Vallée-Poussin method is based on the possibility of constructing an auxiliary Dirichlet series D with non-negative coefficients, convergent for $s>1$, which is divisible by $L(\pi,s+it_0)$ to an order larger than the order of the pole of D at $s=1$. In the case of Riemann's ζ, this is achieved with the product

$$D(s) = \zeta(s)^3\zeta(s+it_0)^4\zeta(s+2it_0).$$

Indeed, from the trigonometric inequality $3+4\cos(t_0)+\cos(2t_0)\geqslant 0$, one sees easily that $-D'(s)/D(s)$ has non-negative Dirichlet coefficients, hence the coefficients of $D(s)$ are non-negative. Alternatively, one can also use

$$D(s)=\zeta(s)^3\zeta(s+it_0)^2\zeta(s-it_0)^2\zeta(s+2it_0)\zeta(s-2it_0). \tag{1.12}$$

Suppose that $\zeta(1+it_0)=0$ (and $t_0\neq 0$ since we know that ζ does not vanish on $[0,1]$); then $D(s)$ has a pole of order 3 at $s=1$ (coming from the ζ factors), compensated by a zero of order at least 4, and thus $D(s)$ has no poles on the real axis, which contradicts Landau's Lemma. The extension of the non-vanishing of $\zeta(s)$ to the zero-free region (1.11) follows from a more quantitative analysis involving the logarithmic derivative of $D(s)$, and is possible because 4 is strictly larger than 3 (see below).

It is possible to extend the method to more general L-functions; the case of Dirichlet characters is straightforward and was carried out by Landau, Gronwall and Titchmarsh (see [**Dav**] for instance), except for a new difficulty that we discuss below. The case of general automorphic L-functions was treated by Moreno (for $d=2$), Perelli et al. (in a somewhat axiomatic setting), and by Hoffstein/Ramakrishnan in general ([**Mor2, Mor3, CMP, HR**]); the proofs in this generality are deeper, as they involve the analytic properties of Rankin/Selberg L-functions. Indeed, the non-vanishing is obtained by considering the auxiliary products:

$$\begin{aligned} D(s)=\zeta(s)L(\pi\otimes\tilde\pi,s)^2L(\pi,s+it_0)^2L(\tilde\pi,s-it_0)^2&\\ \times L(\pi\otimes\pi,s+2it_0)L(\tilde\pi\otimes\tilde\pi,s-2it_0),& \end{aligned} \tag{1.13}$$

if $\tilde\pi\neq\pi\otimes|.|^{2it_0}$ or

$$D(s)=\zeta(s)L(\pi,s+it_0)^2L(\pi\otimes\pi,s+2it_0), \tag{1.14}$$

if $\tilde\pi=\pi\otimes|.|^{2it_0}$. These Dirichlet series have non-negative coefficients: these are each the L-function of pairs $L(\Pi\otimes\tilde\Pi,s)$, where Π is the (self-dual) isobaric sum representation (cf. Remark 1.3):

$$\Pi=1\boxplus\pi\otimes|.|^{it_0}\boxplus\tilde\pi\otimes|.|^{-it_0}\ (\text{resp. }\Pi=1\boxplus\pi\otimes|.|^{it_0}\ ).$$

Again, if $L(\pi,1+it_0)=0$ then $D(s)$ has no pole at $s=1$, contradicting Landau's Lemma. To extend the non-vanishing to regions inside the critical strip, it is more efficient to consider the logarithmic derivative $D'(s)/D(s)$ combined with Hadamard's factorization theorem and positivity arguments: a general version of the de la Vallée-Poussin method is given in the following lemma of Goldfeld/Hoffstein-/Lieman [**GHL**]:

Lemma 1.2.1. *Let $D(s)$ be a Dirichlet series with non-negative coefficients, absolutely convergent for $\Re es>1$, which is also an Euler product, say, of degree d. Suppose that $D(s)$ satisfies a functionnal equation of the form (1.1),*

$$q_D^{s/2}D_\infty(s)D(s)=w(D)q_D^{(1-s)/2}D_\infty(1-s)D(1-s);$$

with a pole a $s=1$ of order $m>0$; assume as well that $D_\infty(s)D(s)$ is of order 1 as $|s|\to+\infty$. Then there is a constant $c_{d,m}>0$ depending only on d,m such that $D(s)$ has at most m zeros in the interval $[1-c_{d,m}/\log(Q_D)]$. Here Q_D denotes the analytic conductor of D.

Suppose we are in the first case ($\tilde{\pi} \neq \pi \otimes |.|^{2it_0}$) and that for a constant $c > 0$, $L(\pi, \sigma + it_0)$ vanishes for $\sigma \in [1 - c/\log(Q_\pi(|t_0| + 2)), 1]$. Then $D(s)$, as given by (1.13), would have a pole of order 4 in this interval, while it has only a pole of order 3 at $s = 1$, thus contradicting Lemma 1.2.1. When $\tilde{\pi} = \pi \otimes |.|^{2it_0}$ we consider $D(s)$ given as (1.14) which has a pole of order 4 at $s = 1$. In that case it is possible that $L(\pi, s + it_0)$ has a zero $\sigma \in [1 - c/\log(Q_\pi(|t_0| + 2)), 1]$; however, Lemma 1.2.1 prevents it from having two (counted with multiplicity). Thus we have:

Theorem 1.4. *There exists a constant $c_d > 0$ (depending on d only) satisfying: for any $\pi \in \mathcal{A}_d^0(\mathbf{Q})$ such that $\tilde{\pi}$ is not equivalent to $\pi \otimes |.|^{it_0}$ for any $t_0 \in \mathbf{R}$, $L(\pi, s)$ has no zeros in the region*

$$\Re e s \geqslant 1 - \frac{c_d}{\log(Q_\pi(|\Im m s| + 2))}; \tag{1.15}$$

if π is self-dual ($\pi \simeq \tilde{\pi}$), then $L(\pi, s)$ has no zeros in the region (1.15), except for an hypothetical simple real zero $\beta_\pi \in [1 - c_d/\log(Q_\pi), 1)$. If such a zero occurs, it is called the exceptional (or Landau/Siegel) zero.

It is remarkable how little the matter of enlarging Hadamard/de la Vallée-Poussin's region has progressed during the past century. Even for Riemann's zeta, the most significant progress goes back to the 50's: the Vinogradov/Korobov zero-free region

$$\zeta(s) \neq 0 \text{ for } \Re e s \geqslant 1 - \frac{c}{\log(|t| + 2)^{2/3} \log\log(|t| + 3)^{1/3}}.$$

This zero-free region was established by means of far reaching exponential sums techniques invented by Vinogradov and perfected by his students; the exponent $2/3$ has not moved since [**Ko, Vi**]. On the other hand, many efforts have been made to find unconditional substitutes for **GRH** (*i.e.* density theorems: these state roughly that, given a family of L-functions, very few of its elements have zeros close to the edge of the critical strip $\Re e s = 1$). We will barely touch on these questions in these lectures. Note also that, althought relatively elementary, the Hadamard/de la Vallée-Poussin method was used by Deligne as a key initial step in his proof of the Weil conjectures: *i.e.* **GRH** for L-functions over finite fields (see [**De3**] and also [**Mor1**]).

To conclude this section, we wish to mention that the method of Eisenstein series (the Langlands/Shahidi method) provides an alternative for showing the non-vanishing of L-functions near the edge of the critical strip. In the mid-seventies Jacquet/Shalika used this method to show that $L(\pi, s) \neq 0$ on $\Re e s = 1$ (before the general theory of Rankin/Selberg L-functions was completed) [**JS1**]. In [**Sha1**], Shahihi proved the non-vanishing of $L(\pi \otimes \pi', s)$ on $\Re e s = 1$ for general (π, π') by this method (this is also accessible to the Hadamard/de la Vallée-Poussin method). More importantly, the method of Eisenstein series applies to L-functions that cannot be handled by the Hadamard/de la Vallée-Poussin method: a notable example is the symmetric 9-th power, $\mathrm{sym}^9\pi$, of a $\pi \in \mathcal{A}_2^0(\mathbf{Q})$ with trivial central character. In [**KiSh2**], it is shown that $L(\mathrm{sym}^9\pi, s) \neq 0$ in $\{\Re e s \geqslant 1\} \backslash [1, 71/67]$. Moreover, it was shown recently by Sarnak (in the simplest case of Riemann's zeta), how this method can be made effective to provide zero-free regions inside the critical strip [**Sa5**].

1.2.2. The Landau/Siegel zero

Before enlarging the standart region for L-functions, the most urgent task is to rule out the exceptional zero in Theorem 1.4 for self-dual representations. This is one of the deepest and most important problems of analytic number theory: the possible existence of this single zero, close to 1, limits severely the strength of several techniques. As we will see below, the hardest case is for $d = 1$ and for $\pi = \chi$ a quadratic character. For that reason, we discuss some applications related to Dirichlet character L-functions and how the exceptional zero limits them.

Zero-free regions are used to show cancellations in the sums of the $\lambda_\pi(p)$ over the prime numbers. For Dirichlet characters, for instance, one can use these cancellations to analyse the distribution of primes in arithmetic progression. The standart zero-free region (1.15) yields:

Prime Number Theorem in Arithmetic Progressions. *Given a, q two coprime integers and $x \geqslant 1$, set*

$$\psi(x; q, a) = \sum_{\substack{p^k \leqslant x \\ p \equiv a(q)}} \log p.$$

As $x \to +\infty$, one has

$$\psi(x; q, a) = \frac{x}{\varphi(q)} + \mathrm{Err}(x; q, a) \tag{1.16}$$

with

$$\mathrm{Err}(x; q, a) = O(\frac{x}{\varphi(q)} x^{\beta_q - 1}) + O(x \exp(-c\sqrt{\log x}));$$

here β_q denotes the largest exceptional zero of the real characters of modulus q and $c > 0$ is an absolute constant.

Since $\beta_q < 1$, it follows that for fixed q and $x \to +\infty$,

$$\psi(x; q, a) \sim x/\varphi(q). \tag{1.17}$$

However, for more advanced applications, one needs an asymptotic uniform for q varying in some interval depending on x. From the expression of $\mathrm{Err}(x; q, a)$ above, it is clear that a β_q close to 1 limits the size of the available range for q. A consequence of Dirichlet Class Number formula[2] is that

$$1 - \beta_q \gg 1/\sqrt{q}(\log q)^2, \tag{1.18}$$

where the implied constant is absolute and explicit (see below). Hence (1.17) holds uniformly for $q = o(\sqrt{\log x})$, which is a rather short range in applications; on the other hand, if no exceptional zero exists, the asymptotic hold in the much larger (althought still small) range $q \ll_A (\log x)^A$ for any $A \geqslant 1$. Finally, under **GRH** one has $\mathrm{Err}(x; q, a) \ll x^{1/2} \log^2(qx)$ and (1.16) is an asymptotic formula uniformly for $q \ll o(x^{1/2}/\log x)$.

Given a primitive quadratic Dirichlet character $\chi(\mod q)$, the existence of the exceptional zero β_χ was studied (among other people) by Heilbronn, Landau and Siegel.

A first remark (basically due to Hecke and Landau) is that β_χ being far from $s = 1$, is essentially equivalent to $L(\chi, 1)$ being large: to see this, consider

$$D(s) = \zeta(s) L(\chi, s),$$

[2](1.18) can also be obtained by purely analytic means

which has non-negative coefficients. If $L(\chi, s)$ has an exceptional zero (i.e. satisfying $0 < 1 - \beta_\chi \ll 1/\log q$), one considers for some $x \geqslant 1$ the complex integral along the line $\Re e s = 2$:

$$\frac{1}{2\pi i}\int_{(2)} D(s+\beta_\chi)\Gamma(s)x^s ds = \sum_{n\geqslant 1}\frac{(1*\chi)(n)}{n^{\beta_\chi}}e^{-n/x}$$

(where $*$ denotes the Dirichlet convolution of arithmetic functions). Since $0 \leqslant (1*\chi)(n) \leqslant (1*1)(n) = \tau(n)$, the righthand side is bounded below by $\geqslant e^{-1/x} \gg 1$ and bounded above by $\ll \exp(O(\log x/\log q))\log^2 x$. On the other hand, moving the line of integration to $\Re e s = -\beta_\chi + 1/2 < 0$, we pass a simple pole at $s = 1-\beta_\chi$ with residue $\Gamma(1-\beta_\chi)x^{1-\beta_\chi}L(\chi, 1)$ and no pole at $s = 0$ since $D(\beta_\chi) = 0$; moreover, the resulting integral is bounded by $O(q^A x^{-1/2})$ for some absolute $A(=1)$ (see Section 1.3 for instance). Taking $x = q^{2A+1}$, we infer that

$$1 \ll \frac{L(\chi,1)}{1-\beta_\chi} \ll \log^2 q, \tag{1.19}$$

where the implied constants are absolute and explicit.

Remark 1.6. If $L(\chi, s)$ has no zero in $[1 - c/\log q, 1]$, the same argument shows that $L(\chi, 1) \gg (\log q)^{-1}$ by taking $\beta_\chi := 1 - c/\log q$ and noting that $D(\beta_\chi) < 0$.

In fact several results tend to show that if such a zero ever exists, it should be *unique*. For example, one has:

Theorem (Landau/Page). *There exists an (effective) constant $c > 0$ such that for any $Q \geqslant 1$, the product*

$$\prod_{q\leqslant Q}\prod_{\chi(q)} L(\chi, s)$$

has at most one real zero within the interval $[1 - c/\log Q, 1]$; here the inner product runs over the primitive real characters of modulus q.

Proof. In view of Theorem 1.4 we may restrict our attention to the quadratic χ's. Suppose that some quadratic χ_1 of modulus $\leqslant Q$ has an exceptional zero in the interval $[1-c/\log Q, 1]$; now consider any quadratic $\chi \neq \chi_1$ of modulus $\leqslant Q$. Then if c is sufficiently small (but fixed), it follows, from the Lemma 1.2.1 applied to

$$D(s) = \zeta(s)L(\chi, s)L(\chi_1, s)L(\chi_1\chi, s),$$

that $L(\chi, s) \neq 0$ in the interval $[1 - c/\log Q, 1]$. □

Another property of the Landau/Siegel zero, when it gets too close to 1, is that it has the property to "repel" the other zeros (real and complex) from the line $\Re e s = 1$. This is illustrated in the following quantitative version (due to Linnik) of the exceptional zero repulsion phenomenon discovered by Deuring and Heilbronn:

Exceptional Zero-Repulsion. *There exists (effective) constants $c_1, c_1' > 0$ such that for any $T \geqslant 2$ and any $q \geqslant 1$, if for some quadratic $\chi_{ex}(\bmod q)$, $L(\chi_{ex}, s)$ has an exceptional zero*

$$\beta_{ex} \in [1 - c_1/\log qT, 1],$$

then the product $\prod_{\chi(q)} L(\chi, s)$ has no other zero in the domain

$$\Re e s \geqslant 1 - \frac{c_1' |\log\big((1-\beta_{\chi_{ex}})\log qT\big)|}{\log qT},\ |\Im m s| \leqslant T;$$

here the product runs over all the characters of modulus q (including χ_{ex}).

Finally, Siegel's famous theorem (in fact, a sharpening of a former result of Landau) shows (ineffectively) that the exceptional zero cannot be too close to 1

Siegel. *Let χ be a primitive quadratic Dirichlet character of modulus q; for any $\varepsilon > 0$ there is a constant $c(\varepsilon) > 0$ such that $L(\chi, s)$ has no zeros in the interval*

$$[1 - c(\varepsilon)/q^{\varepsilon}, 1].$$

Proof. Given $0 < \varepsilon < 1/16$, one can assume that there exists some primitive quadratic character, χ_{ex} (say) of modulus q_{ex}, having a real zero β_{ex} in the interval $]1-\varepsilon, 1[$ (otherwise we are done). This hypothetical zero is used to show that, given any other primitive quadratic character $\chi \neq \chi_{ex}$ of modulus q, $L(\chi, s)$ has no zero within a distance $c(\varepsilon)q^{-4\varepsilon}$ of 1. For this, one considers the auxilliary product

$$D(s) = \zeta(s)L(\chi_{ex}, s)L(\chi, s)L(\chi\chi_{ex}, s),$$

which has non-negative coefficients (since $(1+\chi_{ex}(n))(1+\chi(n)) \geqslant 0$). We proceed as above and consider the integral

$$I = \frac{1}{2\pi i}\int_{(2)} D(s+\beta_{\chi})\Gamma(s)x^s ds = \sum_{n\geqslant 1} \frac{(1*\chi*\chi_{ex}*\chi\chi_{ex})(n)}{n^{\beta_{ex}}} e^{-n/x}.$$

By positivity, the latter sum is bounded below by $\geqslant e^{-1/x} \gg 1$. We shift the contour to $\Re es = 1/2 - \beta_{ex} < 0$; in the process, we pass a simple pole at $s = 1 - \beta_{ex}$ with residue $\Gamma(1-\beta_{ex})x^{1-\beta_{ex}}L(\chi_{ex}, 1)L(\chi\chi_{ex}, 1)L(\chi, 1)$ and no pole at $s = 0$ (since $D(\beta_{ex}) = 1$). The resulting integral is bounded by $O(qq_{ex}x^{-1/2+1/16})$; Hence, we derive

$$1 \ll I = \Gamma(1-\beta_{ex})x^{1-\beta_{ex}}L(\chi_{ex}, 1)L(\chi\chi_{ex}, 1)L(\chi, 1) + O(\frac{qq_{ex}}{x^{1/2-1/16}}).$$

We take $x = q^3$ and use the bound $L(\chi\chi_{ex}, 1) \ll \log(qq_{ex})$ (which is easily derived by a contour shift), to infer that

$$(\log q)^{-1}q^{-3\varepsilon} \ll_{\chi_{ex},\varepsilon} L(\chi, 1),$$

where the implied constant depends on χ_{ex} and ε. The conclusion follows from (1.19). We have followed the elegant presentation of Goldfeld [**Go1**] $\square$

Remark 1.7. Although Siegel's Theorem does not strictly eliminate the exceptional zero, it is rather sharp. (For instance, it implies that the asymptotic (1.17) holds uniformly for q in the range $q \ll_A (\log x)^A$ for any $A \geqslant 1$, the implied constant depending only on A.) However, the above constant $c(\varepsilon)$ depends also on the hypothetical χ_{ex}, and in particular, cannot be computed effectively (for any $\varepsilon < 1/2$). This is the major drawback of the Theorem, since, as we shall see below, the question of the effectivity can be a very important issue.

Remark 1.8. Later, Tatuzawa gave a slightly different formulation of this result[**Ta**]: *for any $\varepsilon > 0$, there exists an effectively computable constant $c(\varepsilon) > 0$ such that for all quadratic characters χ, with but at most one exception, $L(\chi, s)$ has no zero in the interval*

$$[1 - c(\varepsilon)/q_{\chi}^{\varepsilon}, 1].$$

1.2.2.1. *Dirichlet Class Number Formula and the Class Number Problem.* Dirichlet gave several proofs of the non-vanishing of $L(\chi, 1)$ for quadratic characters; one is a direct consequence of his Class Number Formula. Given $\chi\ (\mod q)$, a primitive odd (resp. even) quadratic character, one has

$$L(\chi, 1) = \frac{2\pi h_K}{w_K q^{1/2}},\ (\text{resp.}\ = \frac{\log \varepsilon_K h_K}{q^{1/2}}) \tag{1.20}$$

where $K = \mathbf{Q}(\sqrt{\chi(-1)q})$ denotes the associated quadratic field, w_K the order of the group of units in K (if K is imaginary), $\varepsilon_K > 1$ the fundamental unit (if K is real) and $h_K := |\mathrm{Pic}(O_K)|$ denotes the Class Number. In particular, since $h_K \geqslant 1$ (and $\varepsilon_K \geqslant \sqrt{q}/2$ if K is real), one has $L(\chi, 1) \gg q^{-1/2}$; this and (1.19) imply (1.18).

A famous conjecture of Gauss (formulated in the language of binary integral quadratic forms) is that there are finitely many imaginary quadratic fields with given class number; assuming this conjecture and given some $h \geqslant 1$, the *Class Number Problem* asks for the list of all imaginary quadratic fields K with class number $|\mathrm{Pic}(O_K)| = h$. Thus the Class Number Formula provides a way to approach Gauss's conjecture by analytic methods.

Eventually, Gauss's conjecture was solved by Heilbronn by proving a weak form of Siegel's Theorem; for instance, if follows from (1.19) and (1.20) that, for any $\varepsilon > 0$,

$$|\mathrm{Pic}(O_K)| \gg_\varepsilon q^{1/2-\varepsilon};$$

in particular $|\mathrm{Pic}(O_K)| \to +\infty$ as $q = |\mathrm{Disc}(O_K)| \to +\infty$.

However, because of the lack of effectivity in Siegel's theorem, this does not solve the *Class Number Problem*; at best, it follows from Tatuzawa's version discussed above that in principle (*i.e.* with a sufficient amount of computer assistance), a list of all imaginary quadratic fields with Class Number h can be given, up to possibly one missing field.

For $h = 1$ and 2, the Class Number Problem was solved independently and at the same time by Baker and Stark (in 1966 for $h = 1$ and in 1970 for $h = 2$), each one using quite different methods. In fact, it was recognized later (by Birch and Stark) that an earlier solution (going back to 1952) by Heegner of the Class Number One Problem was correct. For general h, the Class Number Problem has been solved, in principle (*i.e.* with a sufficient amount of computer assistance), by the conjunction of the works of Gross/Zagier and Goldfeld [**Go2, GZ**]. Their solution comes from a weak (but effective) lower bound for $L(\chi_K, 1)$ (see [**O**] for the derivation this very explicit version).

Theorem (Goldfeld/Gross/Zagier). *Let K be an imaginary quadratic field with discriminant $-q$. Then*

$$|\mathrm{Pic}(O_{K)})| \geqslant \frac{1}{55}\prod_{p|q}(1+\frac{1}{p})^{-6}(1+\frac{1}{\sqrt{p}})^{-2}(\log q). \tag{1.21}$$

Thus with a sufficient amount of computer assistance, the **CNP** can be solved for each h. This has been worked out effectively for all $h \leqslant 16$ and for all odd $h \leqslant 23$ [**Ar**] and recently for all $h \leqslant 100$ [**Wal**]. The proof of (1.21) splits into two parts: firstly, Goldfeld [**Go2**] showed that if there exists a GL_2-L-function vanishing with order at least 3 at 1/2 then (1.21) holds; then, later, Gross/Zagier [**GZ**] established the existence of such an L-function (in fact the L-function of a modular elliptic curve), as a consequence of their formula connecting the height of Heegner points to derivatives of L-series. We refer to the survey by D. Goldfeld for an account of the proof of this magnificient result and an update on the new cases of the Class Number Problem that have been treated so far [**1**].

Recently Iwaniec/Sarnak proposed a very interesting approach to rule out the exceptional zero of Dirichlet characters. This approach, which unfortunately we cannot describe here, builds in an essential way on families of automorphic L-functions and amounts to showing that for sufficiently many primitive holomorphic cusp forms f of some auxiliary level q', the central L-value $L(f,1/2)$ is large. A very interesting point here is that the auxiliary level q' depends weakly on K: q' may be as large as an arbitrary large power of $|\mathrm{Disc}(K)|$, and, for instance, if q' is prime, it must be inert in K. Unfortunately, this approach has not been successful so far, although by the existing techniques one seems to be tantalizingly close to the solution. For more information on this approach we refer to [**IS3**], and for some of the most advanced ingredients that may be useful to this approach see [**DFI5, DFI6**].

1.2.3. The Landau/Siegel zero for automorphic L-functions of higher degree.

For general d, the analog of the Landau/Page theorem holds [**HR**]:

Theorem 1.5. *Given $d \geqslant 1$, there is a computable constant c_d such that for $Q \geqslant 2$, there is at most one $\pi \in \mathcal{A}_d^0(\mathbf{Q})$ with $Q_\pi \leqslant Q$ and $L(\pi,s)$ having a zero in the interval $[1-c_d/\log Q, 1[$. Moreover, such π is self-dual and the zero is simple.*

It follows, for example, from an application of Lemma 1.2.1 to

$$D(s) = L(\Pi \otimes \Pi, s) = \zeta(s)L(\pi_1,s)^2L(\pi,s)^2L(\pi_1\otimes\pi,s)^2L(\pi_1\otimes\pi_1,s)L(\pi\otimes\pi,s),$$

and where π,π_1 are self-dual with $Q_\pi, Q_{\pi_1} \leqslant Q$ and $\Pi = 1 \boxplus \pi \boxplus \pi_1$.

Remark 1.9. On the other hand, the analog of Siegel's theorem doesn't seem to be known for general automorphic L-functions, but see below

It is quite remarkable, however, that in a significant number of cases the very existence of a Landau/Siegel zero can be ruled out.

- The first cases are due to Stark [**St2**], who has shown that the Artin L-function $\zeta_F(s)$ of a Galois extension $F/\mathbf{Q}$ has no Landau/Siegel zero unless it contains a quadratic field.
- Hoffstein/Lockhart proved the analog of Siegel's theorem for the symmetric square L-functions $L(\mathrm{sym}^2\pi,s)$, for $\pi \in \mathcal{A}_2^0(\mathbf{Q})$ [**HL**], and shortly afterwards Goldfeld, Hoffstein and Liemann [**GHL**] eliminated the exceptional zero effectively of the adjoint square lift L-function, $L(Ad\pi,s)$, for non-dihedral π. (When π is dihedral, $L(Ad\pi,s)$ is divisible by the L-function of a quadratic character, and then Siegel's theorem is available.)
- Hoffstein/Ramakrishnan proved that there is no exceptional zero for $L(\pi,s)$, $\pi \in \mathcal{A}_2^0(\mathbf{Q})$ and for many π of degree 3 [**HR**]; then Banks [**Ban**] completed their argument and so there is no exceptional zero for $L(\pi,s)$, $\pi \in \mathcal{A}_3^0(\mathbf{Q})$.
- Ramakrishnan/Wang [**RaW**] proved that $L(\pi\otimes\pi',s)$ and $L(\mathrm{sym}^2\pi\otimes\mathrm{sym}^2\pi,s)$ for $\pi,\pi' \in \mathcal{A}_2^0(\mathbf{Q})$ have no exceptional zero, excepted when the corresponding L-functions are divisible by the L-functions of quadratic characters. In these latter cases (which are completely characterized), an exceptional zero,

if it ever occurs, comes only from the quadratic character L-functions (so that one may apply Siegel's theorem).

Again, the method to eliminate the exceptional zero of some $L(s)$ is to construct a Dirichlet series $D(s)$ with non-negative coefficients that is divisible by $L(s)$ to an order *strictly larger* than the order of the pole of $D(s)$ at $s = 1$, and to then apply Lemma 1.2.1. The construction of $D(s)$ depends on the recent progress made in the direction of the Langlands functoriality conjecture. For instance, in the case of $L(\mathrm{sym}^2\pi, s)$ for a self-dual $\pi \in \mathcal{A}_2^0(\mathbf{Q})$, Goldfeld, Hoffstein and Liemann used

$$D(s) = \zeta(s)L(\mathrm{sym}^2\pi, s)^2 L(\mathrm{sym}^2\pi \otimes \mathrm{sym}^2\pi, s) = \zeta(s)L(\mathrm{sym}^2\pi, s)^3 L(\mathrm{sym}^2(\mathrm{sym}^2\pi), s)$$

and is was proved by Bump/Ginzburg (and it follows from the recent results of Kim/Shahidi) that the last factor has a simple pole at $s = 1$.

More generally, Hoffstein/Ramakrishnan [**HR**] proved that:

Theorem 1.6. *If the Functoriality Conjecture for Pairs is true then principal L-functions of degree $d > 1$ have no exceptional zero at all !*

Proof. Given $\pi \in \mathcal{A}_d^0(\mathbf{Q})$ with $d > 1$ and self-dual. The functoriality conjecture for pairs implies that $\pi \boxtimes \pi$ has a (cuspidal) constituent $\tau \neq 1, \pi$. Setting $\Pi = 1 \boxplus \tau \boxplus \pi$ and $D(s) = L(\Pi \otimes \tilde{\Pi}, s)$, one sees that $D(s)$ has a pole of order 3 at $s = 1$ and is divisible by $L(\pi, s)^2 L(\pi \boxplus \tau, s) L(\pi \boxplus \tilde{\tau}, s)$; the conjecture again implies that π is a constituent of $\pi \boxplus \tau$ and $\pi \boxplus \tilde{\tau}$ (by Remark 1.2). So $D(s)$ is divisible by $L(\pi, s)^4$ and one concludes by Lemma 1.2.1. $\square$

For a more complete account on Landau/Siegel zeros and for some methods to attack the problem, we refer to the surveys [**IS3, Ra2**].

1.3. Bounds for L-functions on the critical line

An important consequence of **GRH** is the Generalized Lindelöf Hypothesis, which predicts a sharp upper bound for the size of $L(\pi, s)$ when s is on the critical line:

Generalized Lindelöf Hypothesis (GLH). *For any $\varepsilon > 0$, and $\Re e s = 1/2$, one has $|L(\pi, s)| \ll Q_\pi(t)^\varepsilon$, the implied constant depending on ε.*

In full generality, the best one can prove (unconditionally) is the following:

Convexity Bound. *For any $\varepsilon > 0$, and $\Re e s = 1/2$, we have*

$$|L(\pi, s)| \ll Q_\pi(t)^{1/4+\varepsilon}, \tag{1.22}$$

the implied constant depending on ε and d.

As we shall see below, we refer to this bound as the *convexity bound*.

Remark 1.10. While we will often refer to this bound as the *trivial bound* for $L(\pi, s)$, it is, in this generality, not quite a trivial result: its proof uses implicitly the heavy machinery of L-functions of pairs and a trick of Iwaniec to bypass the fact that the Ramanujan/Petersson Conjecture is not known in general. Note also that there are several L-functions, like Rankin/Selberg L-functions or triple product L-functions, that are expected (but not proved so far) to be automorphic. In these cases, the convexity bound is important for certain arithmetic applications and may indeed be available, but at the cost of good bounds toward the **RPC**, such as the deep results of Kim and Shahidi [**KiSh, Ki**] in direction of the Langlands functoriality conjecture; an exampple is the Rankin-Selberg L-function $L(\mathrm{sym}^2\pi \otimes \pi', s)$ of the symmetric square of a $GL_{2,\mathbf{Q}}$ representation π times a $GL_{2,\mathbf{Q}}$ representation π'.

Remark 1.11. From (1.26) below, one deduces easily (Rankin's trick) that

$$\sum_{n\leqslant x}|\lambda_\pi(n)| \leqslant x^{1+\varepsilon}\sum_{n\geqslant 1}\frac{|\lambda_\pi(n)|}{n^{1+\varepsilon}} \ll_\varepsilon (xQ_\pi)^\varepsilon x, \tag{1.23}$$

for any $x \geqslant 1$ and any $\varepsilon > 0$; this is interpreted as the Ramanujan/Petersson Conjecture on average, and in many situations this bound is as good as **RPC**.

Proof. Our first step is the upper bound

$$L(\pi, s) \ll_{\varepsilon,d} Q_\pi(t)^\varepsilon \tag{1.24}$$

for $\Re es = 1$ and any $\varepsilon > 0$ with the implied constant depending only on ε and d. By the functional equation (1.1) and Stirling's formula

$$|\Gamma(s)| \asymp |s|^{\Re es-1/2}\exp(-\frac{\pi}{2}|s|),$$

one has, when $\Re es = 0$,

$$L(\pi, s) \ll_{\varepsilon,d} Q_\pi(t)^{1/2+\varepsilon}. \tag{1.25}$$

Then for $\Re es = 1/2$, (1.22) follows from (1.25), (1.24), and the Phragmen-Lindelöf principle.

To prove (1.24), we use the following argument of Iwaniec [**I3**]. It was designed to handle the case of the symmetric square L-function of a Maass form (for which **RPC** is not available) and was extended by Molteni to more general automorphic L-functions [**Mol**]. For simplicity, we present only a rough form of the method, yet it is sufficient to handle the basic automorphic L-functions; we refer to [**Mol**] for other related results. By convexity, it is sufficient to prove that for any $\varepsilon, \delta > 0$,

$$\sum_{\substack{n\geqslant 1\\(n,S)=1}}\frac{\mu^2(n)|\lambda_\pi(n)|}{n^{1+\delta}} \ll_{\varepsilon,\delta,d} Q_\pi^\varepsilon, \tag{1.26}$$

where S is a (sufficiently large) product of primes containing the first few primes and the ramified primes for π. Indeed, one has for $\Re es \gg 1$,

$$L(\pi, s) = (\sum_{(n,S)=1}\mu^2(n)\frac{\lambda_\pi(n)}{n^s})L_S(\pi, s)\prod_{p\notin S}\frac{L_p(\pi, s)}{(1+\frac{\lambda_\pi(p)}{p^s})},$$

say; by Proposition 1.1 the two rightmost factors converge absolutely and are uniformly bounded for $\Re es \geqslant 1+\delta$ the implied constant depending only on δ and d (granted that $|S|$ is sufficiently large with respect to d). By the trivial bound, and the positivity of the Dirichlet coefficients $\lambda_{\pi\otimes\tilde{\pi}}(n)$, one has

$$\mu^2(nS)|\lambda_\pi(n)| \leqslant 1 + \mu^2(nS)|\lambda_\pi(n)|^2 = 1 + \mu^2(nS)\lambda_{\pi\otimes\tilde{\pi}}(n) \leqslant 1 + \lambda_{\pi\otimes\tilde{\pi}}(n).$$

Since $L(\pi\otimes\tilde{\pi}, s)$ is uniformly bounded for $\Re es \geqslant 3$, the functional equation (1.4) and the convexity principle imply that

$$L(\pi\otimes\tilde{\pi}, 1+\delta) \ll_\delta Q^A_{\pi\otimes\tilde{\pi}}$$

for some absolute $A > 0$. Moreover, one can prove (see [**BH, RS2**]) that $Q^A_{\pi\otimes\tilde\pi} \leqslant Q^B_\pi$ for some B depending on A and d (in fact $B = 2dA$ is sufficient); it follows that

$$\sum_{(n,S)=1} \mu^2(n)\frac{|\lambda_\pi(n)|}{n^{1+\delta}} \leqslant \zeta(1+\delta) + L(\pi\otimes\tilde\pi, 1+\delta) \ll_{\delta,d} Q^B_\pi. \tag{1.27}$$

This is the initial bound and we are going to improve it by bootstrapping. By multiplicativity of the arithmetic function $\lambda_\pi(n)$, one has, for n_1 and n_2 two squarefree integers, $\lambda_\pi(n_1)\lambda_\pi(n_2) = \lambda_\pi(m)^2\lambda_\pi(n_1n_2/m^2)$, where $m = (n_1, n_2)$. Hence,

$$\Big(\sum_{(n,S)=1} \mu^2(n)\frac{|\lambda_\pi(n)|}{n^{1+\delta}}\Big)^2 \leqslant \sum_m \frac{\mu^2(m)|\lambda_\pi(m)|^2}{m^{2+2\delta}})\Big(\sum_{(n,S)=1}\mu^2(n)\tau(n)\frac{|\lambda_\pi(n)|}{n^{1+\delta}}\Big)$$
$$\ll_{\delta,d} \Big(\sum_{(n,S)=1}\mu^2(n)\frac{|\lambda_\pi(n)|}{n^{1+\delta/2}}\Big) \ll_{\delta,d} Q^B,$$

the latter inequality being deduced from Prop. 1.1 and (1.27). Hence, (1.27) holds (up to changing the implied constant) with B replaced by $B/2$. Iterating the process, we get (1.26), after $O(|\log(B/\varepsilon)|)$ steps. □

1.3.1. The Subconvexity Problem

In view of (1.22) one has the:

Subconvexity Problem (ScP). *Find $\delta > 0$ (depending on d only) such that for any automorphic cuspidal representation π of degree d, and $\Re es = 1/2$, one has*

$$L(\pi, s) \ll_d Q_\pi(\Im ms)^{1/4-\delta}, \tag{1.28}$$

the implied constant depending on d.

Alternatively, one can also measure the size of $L(\pi, s)$ with respect to three quantities separately: the "height" of s, $\sqrt{1/4+|t|^2}$, the "arithmetic" conductor q_π, or the "conductor at the infinite place": $q_\infty := \prod_{i=1\dots d}(1+|\mu_{\pi,i}|)$. Thus one can consider three weakened variants of the **ScP** and seek a subconvex exponent for only one of these parameters, the others remaining fixed – eventually one can also ask for polynomial control on the remaining parameters, if one is greedy. These variants are called the *Subconvexity Problem in the t-aspect, the level-aspect or the ∞-aspect*, respectively.

As we shall see in Lecture 5, the solution of the **ScP** in any of these aspects has striking consequences, but let us start first with a very basic and very practical corollary of (sub)convexity: consider a fixed, smooth, compactly supported function V on $\mathbf{R}_{>0}$. For $X \geqslant 1$ we consider the problem of bounding the sum of length $\simeq X$,

$$\Sigma_V(\pi, X) = \sum_{n\geqslant 1} \lambda_\pi(n)V(\frac{n}{X});$$

our goal is to improve on the "trivial" bound

$$\Sigma_V(\pi, X) \ll_{\varepsilon,V} (Q_\pi X)^\varepsilon X,$$

which follows from (1.23). By the inverse Mellin transform, we have

$$\Sigma_V(\pi, X) = \frac{1}{2\pi i}\int_{(3)} L(\pi, s)\hat V(s)X^s ds, \text{ where } \hat V(s) = \int_0^\infty V(x)x^s\frac{dx}{x}.$$

Now one may shift the contour to $\Re es = 1/2$, passing no pole in the process (unless $L(\pi, s) = \zeta(s)$), and we obtain eventually (integrating by parts several times in $\hat{V}(s)$ to gain convergence) that for any $A > 0$,

$$(1.29) \quad \Sigma_V(\pi, X) \ll_{V,A} X^{1/2} \sup_{\Re es=1/2} \frac{|L(\pi, s)|}{|s|^A}$$
$$\ll_{V,A,\varepsilon} X^{1/2} \sup_{\Re es=1/2} \frac{Q_\pi^{1/4-\delta+\varepsilon}(t)}{|s|^A} \ll_{V,A,\varepsilon} Q_\pi^\varepsilon X^{1/2} Q_\pi^{1/4-\delta},$$

for some $\delta \geqslant 0$. When $\delta = 0$ (*i.e.* for the convexity bound), we already see an improvement over the trivial bound $(Q_\pi X)^\varepsilon X$ for X in the range $X \gg Q_\pi^{1/2+\varepsilon}$, and a subconvex exponent would provide an improvement for X in the larger range $X \gg Q_\pi^{1/2-2\delta}$; ultimately, **GLH** would provide an improvement for all $X \gg_\varepsilon Q_\pi^\varepsilon$ for any $\varepsilon > 0$.

Remark 1.12. This kind of improvement over (1.23) is the manifestation of the oscillation of the coefficients $\lambda_\pi(n)$ in the given ranges of X: indeed, it follows from **GRH** for $L(\pi \otimes \tilde{\pi}, s)$ that for any $X \geqslant 1$ and any $\varepsilon > 0$,

$$\sum_{n \sim X} |\lambda_\pi(n)||V(\frac{n}{X})| \gg_{\varepsilon,V} (Q_\pi X)^{-\varepsilon} X.$$

1.3.2. Interlude: How to compute $L(\pi, s)$ inside the critical strip ?

As we are interested in the behavior of $L(\pi, s)$ within the critical strip $0 < \Re es < 1$, we should first have a manageable expression for $L(\pi, s)$ to work with; this is not immediate since $s = \sigma + it$ is not within the region of absolute convergence. However, one can obtain such an expression by standard means using the functional equation and contour shifts: the result is called (sometimes inappropriately) an "approximate functional equation". There are various possibilities for achieving this, the choice depending on the problem considered; we follow here a derivation borrowed from [**DFI8**] (but see also the elegant version of [**Ha1**]). One starts with the following integral: let $G(u)$ be an even function, normalized by $G(0) = 1$, holomorphic and bounded in the vertical strip $|\Re eu| \leqslant 4$. For $X > 0$, we set

$$I_\pi(s, X) = \frac{1}{2\pi i} \int_{(2)} q_\pi^{(s+u)/2} L(\pi_\infty, s+u) L(\pi, s+u) X^u G(u) \frac{du}{u}$$
$$= q_\pi^{s/2} \sum_{n \geqslant 1} \frac{\lambda_\pi(n)}{n^s} \frac{1}{2\pi i} \int_{(2)} L(\pi_\infty, s+u) (\frac{n}{\sqrt{q_\pi} X})^{-u} G(u) \frac{du}{u}.$$

Because of the exponential decay of $q_\pi^{(s+u)/2} L(\pi_\infty, s+u) L(\pi, s+u)$ in vertical strips we may shift the line of integration to $\Re eu = -2$, passing a pole at $u = 0$ (here we are assuming that $L(\pi, s) \neq \zeta(s + it)$ for any $t \in \mathbf{R}$, otherwise there is an extra harmless contribution), with residue $q_\pi^{s/2} L(\pi_\infty, s) L(\pi, s)$. On the resulting integral, we apply the functional equation (1.1) and make the change of variable $u \leftrightarrow -u$ to obtain

$$q_\pi^{s/2} L(\pi_\infty, s) L(\pi, s) = I_\pi(s, X) + w(\pi) I_{\tilde{\pi}}(1 - s, X^{-1}).$$

Now, taking the Dirichlet series expression of $L(\pi,s)$ and dividing by $q_\pi^{s/2}L(\pi_\infty,s)$, one finds
(1.30)

$$L(\pi,s)=\sum_{n\geqslant 1}\frac{\lambda_\pi(n)}{n^s}V_{s,\pi_\infty}(\frac{n}{X\sqrt{Q_\pi(t)}})+w(\pi,s)\sum_{n\geqslant 1}\frac{\lambda_{\tilde\pi}(n)}{n^{1-s}}V_{1-s,\tilde\pi_\infty}(\frac{nX}{\sqrt{Q_\pi(t)}})$$

where

$$V_{s,\pi_\infty}(y)=\frac{1}{2\pi i}\int_{(2)}\frac{L(\pi_\infty,s+u)}{L(\pi_\infty,s)}Q_{\pi_\infty}(t)^{-u/2}y^{-u}G(u)\frac{du}{u}. \tag{1.31}$$

and

$$w(\pi,s)=w(\pi)q_\pi^{\frac{1-2s}{2}}\frac{L(\tilde\pi_\infty,1-s)}{L(\pi_\infty,s)}.$$

Note that $w(\pi,s)$ is well defined when $1-s$ is not a pole of $L(\tilde\pi_\infty,s)$ (in particular when $\sigma<1/2+1/(d^2+1)$). Note also that $w(\pi,s)$ has modulus one on $\Re es=1/2$ (which justifies our notation). To be more specific, we assume that $\Re es=1/2$ and we take

$$G(u)=(\cos(\frac{\pi u}{A}))^{-Ad},$$

for some parameter $A>4$. In (1.31) we shift the line of integration to $\Re eu=B$ with either $B=A$, if $y\geqslant 1$, or with B satisfying $0>B>\sup_{i=1\dots d}\Re e\mu_{\pi,i}-1/2$ if $y<1$; in the latter case, we hit a pole at $u=0$ with residue 1. Differentiating j times, we see that

$$V^{(j)}_{s,\pi_\infty}(y)\ll_j \delta_{\substack{B<0\\ j=0}}+y^{-B-j}\int_{(B)}|\frac{L(\pi_\infty,s+u)}{L(\pi_\infty,s)}Q_{\pi_\infty(t)}^{-B/2}||G(u)||u|^j\frac{|du|}{|u|}.$$

By Stirling's formula one sees that

$$|\frac{L(\pi_\infty,s+u)}{L(\pi_\infty,s)}Q_{\pi_\infty(t)}^{-B/2}|\ll_{d,B}\exp(\frac{\pi}{4}d|u|),$$

hence we infer that for any $j\geqslant 0$ and any $0<\eta<1/2-\sup_{i=1\dots d}\Re e\mu_{\pi,i}$,

$$V^{(j)}_{s,\pi_\infty}(y)=\delta_{\substack{y<1\\ j=0}}+O(\frac{y^{-j+\eta}}{(1+y)^A}), \tag{1.32}$$

the implied constant depending on η,j,A. When A is large, $V_{s,\pi_\infty}(y)$ becomes very small as $y\rightarrow+\infty$, so that the two sums of (1.30) are essentially of length $X\sqrt{Q_\pi(t)}$ and $\sqrt{Q_\pi(t)}/X$, respectively. In particular, with the most symmetric choice $X=1$, we obtain two sums of length $\approx\sqrt{Q_\pi(t)}$, and for $\Re es=1/2$ we retrieve, by (1.23), the convexity bound (1.22); this shows that a subconvexity bound is the result of a cancellation in the sum of Hecke eigenvalues $\lambda_\pi(n)$ when n is close to $\sqrt{Q_\pi(t)}$ in the logarithmic scale.

Remark 1.13. Although we will not use it in these lectures, it is good to know that there are cases where an asymmetric representation (*i.e.* $X\neq 1$) is better adapted: in particular, when one has not enough control on the "analytic" root number $w(\pi,s)$, one may take $X>1$ to reduce the length of the second sum and thus its influence; of course, the price one pays is then a longer first sum (see [**Vdk**] for an example).

1.3.3. The subconvexity problem and families of L-functions

Families of L-functions enter naturally into the Subconvexity Problem via the method of moments. Suppose one wants to solve some aspect of the Subconvexity Problem for a given $L(\pi_0, 1/2)$; the approximate functional equation reduces essentially this question to a non-trivial bound for linear forms (in the Hecke eigenvalues $\lambda_{\pi_0}(n)$) of the type

$$\Sigma_V(\pi_0, X) = \sum_n \frac{\lambda_{\pi_0(n)}}{n^{1/2}} V(\frac{n}{X}),$$

where V is rapidly decreasing and $X \simeq \sqrt{Q_{\pi_0}}$. Suppose one can put the given π_0 into a "natural" family $\mathcal{F} = \{\pi\}$ endowed with some probability measure $\mu_{\mathcal{F}}$; "natural" implying (among other things) that the analytic conductor Q_{π_0} is close to the average analytic conductor, *i.e.* $Q_{\mathcal{F}} := E(Q_\pi, \mu_{\mathcal{F}}) \approx Q_{\pi_0}$. The method of moments consists of bounding the L^∞ norm of the random variable $\pi \mapsto \Sigma_V(\pi, X)$ by its L^{2k} norm and then estimating the latter. That is, one wants a bound for the $2k$-th moment:

$$|\Sigma_V(\pi_0, X)|^{2k} \mu_{\mathcal{F}}(\pi_0) \leqslant E(|\Sigma_V(\pi, X)|^k, \mu_{\mathcal{F}}) = \int_{\mathcal{F}} |\Sigma_V(\pi, X)|^k d\mu_{\mathcal{F}}(\pi). \tag{1.33}$$

Under **GLH**, one has for any $\varepsilon > 0$

$$E(|\Sigma_V(\pi, X)|^k, \mu_{\mathcal{F}}) \ll_{\varepsilon,k} E(Q_\pi^\varepsilon, \mu_{\mathcal{F}}) \ll_{\varepsilon,k} (Q_{\pi_0})^\varepsilon. \tag{1.34}$$

However, for sufficiently "nice" families and for sufficiently small k's, such a bound can be obtained unconditionally. In these nice situations, we obtain by positivity, (1.33), and (1.34), that

$$|\Sigma_V(\pi_0, X)| \ll_{\varepsilon,k} \frac{Q_{\mathcal{F}}^{(1+\varepsilon)/k}}{\mu_{\mathcal{F}}(\{\pi_0\})} \ll_{\varepsilon,k} \frac{Q_{\pi_0}^{(1+\varepsilon)/k}}{\mu_{\mathcal{F}}(\{\pi_0\})};$$

this bound breaks the trivial estimate, provided that one can take k strictly bigger than $-4\log(\mu_{\mathcal{F}}(\pi_0))/\log(Q_{\pi_0})$.

1.3.3.1. *What do we mean by "nice family" ?* In practice one may proceed slightly differently: instead of averaging the $2k$-th moment of the linear form $\Sigma_V(\pi, X)$ over $\mathcal{F}$, one starts by computing the k-th power $\Sigma_V(\pi_0, X)^k$. Since the arithmetic function $n \to \lambda_\pi(n)$ is multiplicative, one finds that

$$\Sigma_V(\pi_0, X)^k = \sum_{n \sim X^k} \frac{\tau_{k,\pi_0}(n)}{\sqrt{n}} \lambda_{\pi_0}(n) = \Sigma'_{V,\pi_0}(\pi_0, X^k)$$

(say); here $\tau_{k,\pi_0}(n)$ is some arithmetical function that depends (mildly) on π_0. In fact, this function looks essentially like the standard k-th divisor function (hence is bounded by $\ll_\varepsilon n^\varepsilon$ for any $\varepsilon > 0$)[3]. Next one average over $\mathcal{F}$ the *square* of the following linear form of length X^k:

$$\Sigma'_{V,\pi_0}(\pi, X^k) = \sum_{n \ll X^k} \frac{\tau_{k,\pi_0}(n)}{\sqrt{n}} \lambda_\pi(n).$$

[3] This is pretty straightforward to check for $d = 1, 2$, but for higher d, a good control on $\tau_{k,\pi_0}(n)$ seems to require at least (1.23).

Hence, by inverting the summations, one obtains

$$E(|\Sigma'_{V,\pi_0}(\pi,X^k)|^2,\mu_{\mathcal{F}})=\sum_{m,n\ll X^k}\frac{\tau_{k,\pi_0}(m)\tau_{k,\pi_0}(n)}{\sqrt{mn}}\int_{\mathcal{F}}\lambda_\pi(m)\overline{\lambda_\pi(n)}d\mu_{\mathcal{F}}(\pi).$$

Thus "nice family" means that for m,n less than X^k, the expectations of the random variables $\pi\to\lambda_\pi(m)\overline{\lambda_\pi(n)}$ are well controlled. In practice, these expectations are to be "close" to the Dirac symbol $\delta_{m=n}$, a property of the family that we name *approximate orthogonality*. One of the purposes of the next lectures is to describe several families of arithmetic objects satisfying the "approximate orthogonality" property for m and n in appropriate ranges. Assuming that approximate orthogonality holds for $\mathcal{F}$ and $m,n\ll X^k$ (which is the hard step, especially when dealing with the **ScP**), one can then derive a bound of the form

$$\mu_{\mathcal{F}}(\{\pi_0\})|\Sigma'_V(\pi_0,X)|^{2k}\ll\sum_{m,n\ll X^k}\frac{\tau_{k,\pi_0}(m)\tau_{k,\pi_0}(n)}{\sqrt{mn}}\delta_{m=n}\ll_{\varepsilon,k}(Q_{\pi_0}X)^{\varepsilon}.$$

1.4. Appendum: bounds for local parameters via families of L-functions

In this section we complete the proof of Theorem 1.1, following Luo/Rudnick/Sarnak [**LRS**]. Hopefully this will provide the first concrete example in these lectures of the usefulness of families in analytic number theory.

We present the bound for the local parameters at infinity assuming that π_∞ is unramified (*i.e.* spherical). By unitarity of π, it is sufficient to show that

$$\mathrm{Max}\,\Re e\mu_{\pi,i}\leqslant\theta_d=\frac{1}{2}-\frac{1}{d^2+1}.$$

From (1.3), we see that $L_\infty(\pi\otimes\tilde{\pi},s)$ has a pole at $s=\beta_0:=2\,\mathrm{Max}\,\Re e\mu_{\pi,i}$, hence it is sufficient to show that $L(\pi\otimes\tilde{\pi},s)$ does not vanish for $s>2\theta_d=1-\frac{2}{d^2+1}$, since $(s-1)L_\infty(\pi\otimes\tilde{\pi},s)L(\pi\otimes\tilde{\pi},s)$ is entire there. Actually, instead of considering this L-function alone, one considers a whole family of L-functions: namely the $L(\chi.\pi\otimes\tilde{\pi},s)$, where $\chi.\pi:=\pi\otimes\chi$ ranges over the twists of π by caracters of $\mathbf{A}_{\mathbf{Q}}^\times/\mathbf{Q}^\times$ trival at ∞. These correspond to *even* primitive Dirichlet characters, which by an abuse of notation, we still denote by χ. We take the χ to be even, of prime moduli q not dividing q_π. Since $q\not|\,q_\pi$ and $\chi.\pi\not\simeq\pi$, the $L(\chi.\pi\otimes\tilde{\pi},s)$ are entire. Moreover, we have the following facts (see [**LRS**]):

$$L_\infty(\chi.\pi\otimes\tilde{\pi},s)=L_\infty(\pi\otimes\tilde{\pi},s);\tag{1.35}$$

$$L(\chi.\pi\otimes\tilde{\pi},s)=\sum_{n\geqslant1}\frac{\chi(n)\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^s};\tag{1.36}$$

$$q_{\chi.\pi\otimes\tilde{\pi}}=q_{\pi\otimes\tilde{\pi}}q^{d^2/2};\tag{1.37}$$

$$w(\chi.\pi\otimes\tilde{\pi})=w(\pi\otimes\tilde{\pi})\chi(q_{\pi\otimes\tilde{\pi}})(\frac{G_\chi}{\sqrt{q}})^{d^2},\tag{1.38}$$

where G_χ denotes the Gauss sum. Hence by (1.35), the bound $\beta_0\leqslant2\theta_d$ follows from:

Proposition 1.2. *For any $\beta > 2\theta_d$, one has*

$$\sum_{q\sim Q}\sum_{\substack{\chi(q)\\ \chi\neq\chi_0, even}} L(\chi.\pi\otimes\tilde{\pi},\beta) \gg_{\pi,\beta} \frac{Q^2}{\log Q},$$

where q ranges over primes. In particular, for any $\beta > 2\theta_d$ there exists χ as above such that $L(\chi.\pi\otimes\tilde{\pi},\beta)\neq 0$.

Proof. Applying the functional equation and a method similar to those of section 1.3.2, one infers from (1.36), (1.37) and (1.38) that, for $0<\beta<1$, one has

$$L(\chi.\pi\otimes\tilde{\pi},\beta) = \sum_{n\geqslant 1}\frac{\chi(n)\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^\beta}V_1(\frac{n}{Y}) + \frac{w(\pi\otimes\tilde{\pi})\chi(q_{\pi\otimes\tilde{\pi}})}{(q_{\pi\otimes\tilde{\pi}}q^{d^2})^\beta}\sum_{n\geqslant 1}\frac{\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^{1-\beta}}\overline{\chi}(n)(\frac{G_\chi}{\sqrt{q}})^{d^2}V_2(\frac{nY}{q_{\pi\otimes\tilde{\pi}}q^{d^2}})$$

where $Y\geqslant 1$ is some parameter and $V_1(x),V_2(x)$ are smooth test functions satisfying

$$V_1(x),V_2(x) = O_A(x^{-A}) \quad \text{as } x\to+\infty,$$
$$V_1(x) = 1+O_A(x^A),\ V_2(x)\ll_\varepsilon 1+x^{1-\beta_0-\beta-\varepsilon} \quad \text{as } x\to 0, \tag{1.39}$$

for all $A\geqslant 0$ and all $\varepsilon>0$. Next, one averages the approximate functional equation over even primitive characters of prime moduli $q\sim Q$. Then

$$\sum_{q\sim Q}\sum_{\substack{\chi\neq\chi_0\\ \chi\text{ even}}} L(\chi.\pi\otimes\tilde{\pi},\beta)$$

is decomposed as the sum of two terms T_1+T_2, say. Using the orthogonality relations

$$\sum_{q\sim Q}\sum_{\substack{\chi\neq\chi_0\\ \chi\text{ even}}}\chi(n) = \begin{cases} 0 & if\ n\equiv 0(q)\\ \frac{q-1}{2}-1 & if\ n\equiv\pm1(q)\\ -1 & \text{otherwise,}\end{cases}$$

one finds that

$$T_1 = \sum_{q\sim Q}\sum_{\substack{\chi\neq\chi_0\\ \chi\text{ even}}}\sum_{n\geqslant 1}\frac{\chi(n)\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^\beta}V_1(\frac{n}{Y})$$
$$= \sum_{q\sim Q}\frac{q-1}{2}\sum_{n\equiv\pm1(q)}\frac{\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^\beta}V_1(\frac{n}{Y}) - \sum_{q\sim Q}\sum_{n\not\equiv 0,\pm1(q)}\frac{\lambda_{\pi\otimes\tilde{\pi}}(n)}{n^\beta}V_1(\frac{n}{Y})$$

The contribution from $n=1$ in the first sum above is

$$\sum_{q\sim Q}\frac{q-1}{2}V_1(\frac{1}{Y}) = \sum_{q\sim Q}\frac{q-1}{2} + O_A(Q^2Y^{-A})$$

for any $A\geqslant 0$. The sum over $n\equiv 1(q)$, $n\neq 1$ contributes as an error term

$$\sum_{q\sim Q}\frac{q-1}{2}\sum_{d\geqslant 1}\frac{\lambda_{\pi\otimes\tilde{\pi}}(1+dq)}{(1+dq)^\beta}V_1(\frac{1+dq}{Y}) \ll Q\sum_n\frac{\lambda_{\pi\otimes\tilde{\pi}}(n)n^\varepsilon}{n^\beta}|V_1(\frac{n}{Y})| \ll_{\varepsilon,\pi} QY^{1-\beta+\varepsilon},$$

for any $\varepsilon > 0$; here we have used that $\lambda_{\pi\otimes\tilde\pi}(n) \geqslant 0$ and that the number of representations of n of the form $n = 1 + dq$ with $d, q \geqslant 1$ is $O_\varepsilon(n^\varepsilon)$, and also (1.39). The remaining terms in T_1 are bounded similarly and we find that

$$T_1 = \sum_{q\sim Q} \frac{q-1}{2} + O(QY^{1-\beta+\varepsilon}).$$

The term T_2 contains the following moments of Gauss sums:

$$M_{d^2}(n) = \sum_{\substack{\chi\neq\chi_0 \\ \chi \text{ even}}} \chi(q_{\pi\otimes\tilde\pi})\overline{\chi}(n)\left(\frac{G_\chi}{\sqrt{q}}\right)^{d^2}. \tag{1.40}$$

This sum is zero if $q|n$ and otherwise equals

$$\frac{q-1}{2}\{Kl_{d^2}(r;q) + Kl_{d^2}(-r;q)\} - (-1)^{d^2}$$

where $r \equiv n\overline{q_{\pi\otimes\tilde\pi}} (\mod q)$ and $Kl_{d^2}(r;q)$ denotes the hyper-Kloosterman sum

$$Kl_{d^2}(r;q) = \sum_{x_1x_2\ldots x_{d^2}\equiv r(q)} e\left(\frac{x_1 + x_2 + \cdots + x_r}{q}\right).$$

The latter sum was bounded by $\ll_d q^{\frac{d^2-1}{2}}$ by Deligne as a consequence of his resolution of the Weil Conjectures [**De4**]. Hence

$$M_{d^2}(n) \ll_d q^{\frac{d^2+1}{2}}.$$

This bound is the key saving and shows considerable oscillation of the root numbers of $L(\chi.\pi\otimes\tilde\pi, s)$. Using this bound, the inequalities $\beta_0 < 1$, $\beta > 0$, and the bounds for V_2 in (1.39), one obtains

$$T_2 = \sum_{q\sim Q} \frac{w(\pi\otimes\tilde\pi)}{(q_{\pi\otimes\tilde\pi} q^{d^2})^\beta} \sum_{n\geqslant 1} M_{d^2}(n) \frac{\lambda_{\pi\otimes\tilde\pi}(n)}{n^{1-\beta}} \overline{\chi}(n)\left(\frac{G_\chi}{\sqrt{q}}\right)^{d^2} V_2\left(\frac{nY}{q_{\pi\otimes\tilde\pi} q^{d^2}}\right) \ll Q^{1+\frac{d^2+1}{2}} Y^{-\beta},$$

and hence

$$T_1 + T_2 = \sum_{q\sim Q} \frac{q-1}{2} + O(QY^{1-\beta+\varepsilon}) + O(Q^{1+\frac{d^2+1}{2}} Y^{-\beta})$$

which gives Proposition 1.2, on choosing $Y = Q^{\frac{d^2+1}{2}}$.

□

Remark 1.14. This method has several advantages over the first proof. Firstly, it can be adapted to deal with the non-archimedean places as well. Secondly, it can be extended to automorphic forms over general number fields and provides a bound of the same quality $\theta_d = \frac{1}{2} - \frac{1}{d^2+1}$, *independently* of the base field, while the first method instead gives a degenerating (although still non-trivial) bound $\frac{1}{2} - \frac{1}{md^2+1}$, where m is the degree of the base field [**LRS2**].

LECTURE 2
A Review of Classical Automorphic Forms

In this section we review the theory of $GL_{2,\mathbf{Q}}$-automorphic forms from the classical point of view, following Maass and Selberg, and collect various estimates and technical formulas that will be used later. Some of this material is borrowed from the exposition given in Sections 4/5/6 of [**DFI8**].

2.1. Spaces of Holomorphic and Maass forms

The group $SL_2(\mathbf{R})$ acts on the upper half-plane by fractional linear transformations

$$\gamma z = \frac{az+b}{cz+d}, \text{ if } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

For $\gamma \in SL_2(\mathbf{R})$, we define the multipliers

$$j(\gamma, z) = (cz+d),$$

$$j_\gamma(z) = \frac{cz+d}{|cz+d|} = \exp(i \arg(cz+d)),$$

and for any integer $k \geqslant 0$, two actions (of weight k) on the space of functions $f : \mathbf{H} \to \mathbf{C}$, given by

$$f_{|^k\gamma}(z) = (cz+d)^{-k} f(\gamma z).$$

$$f_{|_k\gamma}(z) = j_\gamma(z)^{-k} f(\gamma z).$$

For $q \geqslant 1$, we denote by Γ the congruence subgroup $\Gamma_0(q) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}),\ c \equiv 0(q)\}$; then any character $\chi(mod\ q)$ defines a character of Γ by the formula

$$\chi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \chi(d) = \overline{\chi}(a).$$

2.1.1. Holomorphic forms

For $k \geqslant 1$, we denote $\mathcal{S}_k(q,\chi)$ the space of holomorphic cusp forms of weight k, level q, and nebentypus χ (*i.e.* the space of holomorphic functions $F : \mathbf{H} \to \mathbf{C}$ that satisfy

$$F_{|^k\gamma}(z) = \chi(\gamma)F(z) \tag{2.1}$$

for every $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, and that vanish at every cusp). This space is finite dimensional and is equipped with the Petersson inner product:

$$\langle F, G\rangle_k = \int_{\Gamma\backslash\mathbf{H}} F(z)\overline{G}(z)y^k \frac{dxdy}{y^2}.$$

Such a form has a Fourier expansion at ∞,

$$F(z) = \sum_{n\geqslant 1} \rho_F(n)n^{\frac{k}{2}}e(nz). \tag{2.2}$$

2.1.2. Maass forms

A function $f : \mathbf{H} \to \mathbf{C}$ is said to be Γ-automorphic of weight k and nebentypus χ iff it satisfies

$$f_{|_k\gamma}(z) = \chi(\gamma)f(z) \tag{2.3}$$

for all $\gamma \in \Gamma$. Following Maass, one introduces the two differential operators of order one

$$R_k := \frac{k}{2} + y(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}), L_k := \frac{k}{2} + y(i\frac{\partial}{\partial x} - \frac{\partial}{\partial y}),$$

which transform smooth automorphic functions of weight k into smooth automorphic functions of weight $k+2$ and $k-2$ respectively. Thus the Laplace operator of weight k, given by

$$\Delta_k = -R_{k-2}L_k - \frac{k}{2}(1-\frac{k}{2})\mathrm{Id} = y^2(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}) - iky\frac{\partial}{\partial x},$$

acts on smooth automorphic functions of weight k. The Laplacian is a self-adjoint operator with respect to Petersson's inner product

$$\langle f, g\rangle = \int_{\Gamma\backslash\mathbf{H}} f(z)\overline{g}(z)\frac{dxdy}{y^2},$$

which is bounded from above by $-\frac{k}{2}(1-\frac{k}{2})$, *i.e.*

$$\langle \Delta_k f, f\rangle \leqslant -\frac{k}{2}(1-\frac{k}{2})\langle f, f\rangle. \tag{2.4}$$

In particular, Δ_k admits a self-adjoint extension to the L^2-space of square-integrable automorphic functions, $\mathcal{L}_k(q,\chi)$ (say); moreover, this space has a complete spectral resolution, which we describe below.

2.1.2.1. *Eisenstein series.* An important class of automorphic functions is the set of Eisenstein series (although these are not square-integrable): Eisenstein series are indexed by the *singular cusps* $\{\mathfrak{a}\}$ and are given by the (absolutely convergent for $\Re es > 1$) series

$$E_{\mathfrak{a}}(z,s) = \sum_{\gamma\in\Gamma_{\mathfrak{a}}\backslash\Gamma} \overline{\chi}(\gamma)j_{\sigma_{\mathfrak{a}}^{-1}\gamma}(z)^{-k}(\Im m(\sigma_{\mathfrak{a}}^{-1}\gamma z))^s$$

where $\sigma_{\mathfrak{a}}$ is the scaling matrix of the cusp $\mathfrak{a}$. Recall that the scaling matrix $\sigma_{\mathfrak{a}}$ is the unique matrix (up to right translations) such that

$$\sigma_{\mathfrak{a}}\infty = \mathfrak{a},\ \ \sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \Gamma_\infty = \Big\{\pm\begin{pmatrix} 1 & b \\ & 1 \end{pmatrix},\ b\in\mathbf{Z}\Big\},$$

where $\Gamma_{\mathfrak{a}}$ denotes the stabilizer of the cusp $\mathfrak{a}$, and that a cusp $\mathfrak{a}$ is *singular* whenever

$$\chi(\sigma_{\mathfrak{a}}\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}\sigma_{\mathfrak{a}}^{-1}) = 1, or\ (-1)^k.$$

Eisenstein series $E_{\mathfrak{a}}(z,s)$ are eigenfunction of Δ_k with eigenvalues $\lambda(s) = s(1-s)$,

$$\Delta_k E_{\mathfrak{a}}(z,s) + s(1-s)E_{\mathfrak{a}}(z,s) = 0.$$

Selberg proved that they have a meromorphic continuation (in s) to the complex plane with no pole in the domain $\Re es \geqslant 1/2$, except for a simple pole at $s = 1$ when $k = 0$ and $\chi(q)$ is trivial, and that they satisfy a functional equation relating their value at s and $1-s$. It turns out that the analytically continued Eisenstein series $\{E_{\mathfrak{a}}(z, 1/2+it),\ t \in \mathbf{R}\}_{\mathfrak{a}}$ realize the (continuous) spectral decomposition of the subspace $\mathcal{E}_k(q,\chi)$ of $\mathcal{L}_k(q,\chi)$ generated by *the incomplete Eisenstein series*,

$$E_{\mathfrak{a}}(z|\psi) = \sum_{\gamma\in\Gamma_{\mathfrak{a}}\backslash\Gamma} \overline{\chi}(\gamma) j_{\sigma_{\mathfrak{a}}^{-1}\gamma}(z)^{-k}\psi(\Im m(\sigma_{\mathfrak{a}}^{-1}\gamma z)),$$

where ψ ranges over the smooth compactly supported functions on $\mathbf{R}^+$. The orthogonal complement of $\mathcal{E}_k(q,\chi)$ is called the cuspidal subspace and we denote it by $\mathcal{C}_k(q,\chi)$: it turns out that the spectrum of $\Delta_{k|}\mathcal{C}_k(q,\chi)$ (the cuspidal spectrum) is discrete and has a basis composed of real analytic, square integrable eigenfunctions of Δ_k (such functions are called Maass cusp forms). It follows that any $f \in \mathcal{L}_k(q,\chi)$ admits the following spectral decomposition (see [**I4**] for the proof when $k = 0$):

$$f(z) = \sum_{j\geqslant 1}\langle f, u_j\rangle u_j(z) + \sum_{\mathfrak{a}} \frac{1}{4\pi}\int_{\mathbf{R}} \langle f, E_{\mathfrak{a}}(*, 1/2+it)\rangle E_{\mathfrak{a}}(*, 1/2+it)dt, \tag{2.5}$$

where $\{u_j(z)\}_{j\geqslant 1}$ denotes an orthonormal basis of $\mathcal{C}_k(q,\chi)$ composed of Maass cusp forms.

We describe in more detail the structure of $\mathcal{C}_k(q,\chi)$. For $t \in \mathbf{C}$, we denote by $S_k(q,\chi,it)$ the Δ_k-eigenspace associated to the eigenvalue

$$\lambda = \lambda(it) := (1/2+it)(1/2-it).$$

From (2.4), we already know that $\lambda \geqslant \frac{k}{2}(1-\frac{k}{2})$ and this lower bound is in fact attained, because of the following map

$$F(z) \mapsto y^{k/2}F(z), \tag{2.6}$$

which maps $S_k(q,\chi)$ isometrically onto $S_k(q,\chi,\frac{k}{2}(1-\frac{k}{2}))$. More generally, for any $k' \equiv k(2)$, $k' \leqslant k$, the map

$$F(z) \in S_{k'}(q,\chi) \mapsto \prod_{\substack{k'\leqslant l<k \\ l\equiv k(2)}} R_l[y^{k'/2}F(z)] \in S_k(q,\chi,\frac{k'-1}{2})$$

defines, up to multiplication by some explicit scalar, an isometric isomorphism between both spaces.

On the other hand, if it is not of the form $\frac{k'-1}{2}$ for any $k' \equiv k(2)$, $k' \leqslant k$, then $\lambda(it) \geqslant 0$ (*i.e.* $t \in [-i/2, i/2] \cup \mathbf{R}$). This follows from the fact that if $k \equiv \kappa(2)$ for $\kappa = 0$ or 1, the map

$$f(z) \mapsto \prod_{\substack{\kappa\leqslant l<k \\ l\equiv k(2)}} R_l f(z)$$

defines an isomorphism (which is in fact an isometry up to some explicit scalar) between the spaces $S_\kappa(q,\chi,it)$ and $S_k(q,\chi,it)$. This shows that the study of general Maass forms of weight k can be reduced either to that of Maass forms of weight $\kappa=0,1$ with positive Δ_k-eigenvalue, or to that of holomorphic forms of weight $\geqslant 1$; however, for several technical purposes it may be useful to consider holomorphic forms of weight k in terms of Maass forms of weight k.

Selberg's conjecture (which is the Ramanujan/Petersson conjecture for $GL_{2,\mathbf{Q}}$ at the infinite place) is the statement that whenever $\lambda>0$, then $\lambda\geqslant 1/4$ (*i.e.* $t\in\mathbf{R}$). Note that for $k=1$ this holds trivially by (2.4), while for $k=0$ the best result toward this conjecture so far is $|\Im m t|\leqslant 7/64$ [**KiSa**].

2.1.2.2. *Fourier expansion.* By periodicity $f(z+1)=f(z)$, and by separation of variables, one shows that a Maass cusp form f with Δ_k-eigenvalue $\lambda=(1/2+it)(1/2-it)$ has a Fourier expansion at infinity of the form

$$f(z)=\sum_{\substack{n=-\infty\\ n\neq 0}}^{+\infty}\rho_f(n)W_{\frac{n}{|n|}\frac{k}{2},it}(4\pi|n|y)e(nx), \tag{2.7}$$

where $W_{\alpha,\beta}(y)$ denotes the Whittaker function; the $\{\rho_f(n)\}_{n\in\mathbf{Z}-\{0\}}$ are called the Fourier coefficients of f. The Eisenstein series have a similar Fourier expansion,
(2.8)

$$E_{\mathfrak{a}}(z,1/2+it)=\delta_{\mathfrak{a}}y^{1/2+it}+\varphi_{\mathfrak{a}}(1/2+it)y^{1/2-it}+\sum_{\substack{n=-\infty\\ n\neq 0}}^{+\infty}\rho_{\mathfrak{a}}(n,t)W_{\frac{n}{|n|}\frac{k}{2},it}(4\pi|n|y)e(nx),$$

where $\delta_{\mathfrak{a}}=0$ unless $\mathfrak{a}=\infty$, $\delta_\infty=1$, and $\varphi_{\mathfrak{a}}(1/2+it)$ is the entry $(\infty,\mathfrak{a})$ of the scattering matrix. Recall that the Whittaker function $W_{\alpha,it}(2y)$ is the unique solution of the differential equation

$$W''(y)+(\lambda y^{-2}+2\alpha y^{-1}-1)W(y)=0$$

that decays exponentially as $y\rightarrow+\infty$; more precisely, we have $W_{\alpha,it}(y)\sim y^\alpha e^{-y/2}$. In the particular case where $it=\alpha-1/2$, we have an exact formula

$$W_{\alpha,\alpha-1/2}(y)=y^\alpha e^{-y/2}. \tag{2.9}$$

In particular, for $F(z)\in S_k(q,\chi)$, if we denote $f(z)=y^{k/2}F(z)\in S_k(q,\chi,\frac{k-1}{2})$, one has the following relation between the Fourier coefficients:

$$\rho_F(n)=(4\pi)^{k/2}\rho_f(n).$$

2.1.2.3. *The reflection operator.* Next we introduce the reflection operator X, which acts on functions by

$$(Xf)(z)=f(-\overline{z}).$$

The operator X sends forms of weight $-k$ isometrically to forms of weight k and satisfy $X^2=1$; moreover, X commutes with the Laplacian, *i.e.* $\Delta_{-k}X=X\Delta_k$, so that $XS_k(q,\chi,it)=S_{-k}(q,\chi,it)$. In fact, if we denote by $Q_{it,k}$ the operator

$$Q_{it,k}:=\delta(it,k)X\prod_{\substack{-k<l\leqslant k\\ l\equiv k(2)}}L_l,$$

where

$$\delta(it,k)=\frac{\Gamma(1/2+it-k/2)}{\Gamma(1/2+it+k/2)},$$

then $Q_{it,k}$ is null if it is of the form $it = \frac{k'-1}{2}$ for some $k' \equiv k(2)$, $k' \leqslant k$, and defines an isometric involution on $S_k(q,\chi,it)$ if it is not. In the latter case, if f is an eigenfunction of $Q_{it,k}$ with eigenvalue $\delta_f = \pm 1$, one has the following symmetry:

$$\rho_f(-n) = \delta_f \frac{\Gamma(1/2+it+k/2)}{\Gamma(1/2+it-k/2)}\rho_f(n).$$

2.2. Hecke operators

For $n \geqslant 1$ the Hecke operator T_n is defined on automorphic functions of weight k by

$$(T_n f)(z) = \frac{1}{\sqrt{n}} \sum_{ad=n} \chi(a) \sum_{b(d)} f(\frac{az+b}{d}).$$

These operators form a commutative algebra: more precisely, one has the relation

$$T_m T_n = \sum_{d|(m,n)} \chi(d) T_{mn/d^2}.$$

Since T_n commutes with Δ_k, T_n acts on each eigenspace $S_k(q,\chi,it)$ and on the Eisenstein subspace. Moreover the T_n with $(n,q)=1$ are normal on $\mathcal{L}_k(q,\chi)$; more precisely one has for any f, g,

$$\langle T_n f, g\rangle = \overline{\chi}(n)\langle f, T_n g\rangle,$$

and in particular, one can choose an orthonormal basis composed of common eigenvalues of Δ_k and of the T_n for $(n,q)=1$. Such a distinguished basis will be called a Hecke eigenbasis. For f a Hecke cusp form, we denote by $\lambda_f(n)$ its n-th Hecke eigenvalue; one has

$$\lambda_f(n) = \chi(n)\overline{\lambda_f}(n),\ \lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \chi(d)\lambda_f(mn/d^2),$$

and by Möbius inversion

$$\lambda_f(mn) = \sum_{d|(m,n)} \mu(d)\chi(d)\lambda_f(m/d)\lambda_f(n/d)$$

for $(mn,q) = 1$. Note that the two relations above hold for all m, n if f is an eigenvalue of all the Hecke operators. Applying the Hecke operators to the Fourier expansion (2.7), one finds also the simple proportionality relations: for $n \geqslant 1$,

$$\rho_f(n) = \rho_f(1)\lambda_f(n)n^{-1/2}, \rho_f(-n) = \rho_f(-1)\lambda_f(n)n^{-1/2}. \tag{2.10}$$

By Atkin/Lehner/Li theory, the whole Hecke algebra can be diagonalized further inside the subspaces of new-forms $S_k^n(q,\chi,it)$ say. Recall that by definition $S_k^n(q,\chi,it)$ is the orthogonal complement inside $S_k(q,\chi,it)$ of the subspace generated by the forms

$$f(dz),\ f \in S_k(q'q^*,\chi',it), dq'|q/q^*,\ q'q^* < q$$

where q^* is the conductor of the primitive character χ^* underlying χ and χ' is its induced character. We recall as well that the strong multiplicity one property holds in $S_k^n(q,\chi,it)$. This means that a new Hecke eigenform (*i.e.* belonging to $S_k^n(q,\chi,it)$) is determined up to scalars by all but finitely many Hecke eigenvalues; as a consequence a new Hecke eigenform is automatically an eigenform of all the T_n. We denote by $S_k^p(q,\chi,it)$ the set of *primitive* forms of $S_k^n(q,\chi,it)$: *i.e.* the set of eigenforms for the full Hecke algebra normalized by $\rho_f(1) = 1$. Primitive forms form orthogonal basis of $S_k^n(q,\chi,it)$ and by the strong multiplicity one property primitive forms in $S_k^n(q,\chi,it)$ are also eigenforms of the operator $Q_{it,k}$, hence one has $(\rho_f(-1)=\delta_f\rho_f(1))$. Primitive forms are also "quasi-eigenvalues" of the Atkin-Lehner-Li operators W_{q_1}, which are defined, for each $q_1|q$ such that $(q_1,q/q_1)=1$, by the $|_k$ action of some matrices W_{q_1} of determinant q_1, which normalize $\Gamma_0(q)$. We have $W_{q_1}:S_k(q,\chi,it)\mapsto S_k(q,\overline{\chi}_{q_1}\chi_{q/q_1},it)$ where $\chi=\chi_{q_1}\chi_{q/q_1}$ is the factorization of χ into characters of moduli q_1 and q/q_1. Then for f a primitive form, one has $W_{q_1}f=w_f(q_1)g$, where g is primitive and $w_f(q_1)$ has modulus one (see [**AL, ALi**]).

2.3. Classical L-functions vs. Automorphic L-functions

Given $f\in S_k^p(q,\chi,it)$ a primitive form, Hecke associated to it a classical L-function

$$L(f,s)=\sum_{n\geqslant 1}\frac{\lambda_f(n)}{n^s}=\prod_p(1-\frac{\lambda_f(p)}{p^s}+\frac{\chi_f(p)}{p^{2s}})^{-1}$$

and proved its basic analytic properties (see [**I5**]). There are also other L-functions related to classical modular forms, namely the L-functions $L(f\times g,s)$ associated to pairs (f,g) which were initially studied by Rankin and Selberg, and the symmetric square L-function $L(\mathrm{sym}^2f,s)$, which was first defined and studied by Shimura ([**Sh**]). These are Euler product of degree 4 and 3 respectively and one of their main advantages is that the coefficients of these Dirichlet series are easily expressed in terms of the Hecke eigenvalues of f (and g), as one can see from the expressions below. However, the major drawback of considering these L-functions from the classical viewpoint is their functional equation, which may be quite complicated and hard to prove in full generality, especially when levels are divisible by high powers; to be convinced, the reader may want to have a look at the paper of W. Li [**Li2**], which derives the functional equation for Rankin/Selberg L-functions by classical means.

On the other hand, a well known recipe associates to f an automorphic representation, π_f (say), with the same conductor [**Ge**]. Its local parameters are related to the Hecke or Laplace eigenvalues as follows:

$$\lambda_f(p)=\alpha_{\pi_f,1}(p)+\alpha_{\pi_f,2}(p),\ \alpha_{\pi_f,1}(p)\alpha_{\pi_f,2}(p)=\chi(p)$$

at the non-archimedean places (in particular $\lambda_{\pi_f}(n)=\lambda_f(n)$); for the archimedean places one has the identities

$$\begin{aligned}
-\mu_{f,1}&=\frac{1-\delta_f}{2}+it\quad,-\mu_{f,2}=\tfrac{1-\delta_f}{2}-it,\quad |\Im m t|<1/2\\
-\mu_{f,1}&=\frac{1-\delta_f}{2}+it\qquad,-\mu_{f,2}=-it,\qquad t\in\mathbf{R}\\
-\mu_{f,1}&=\frac{k-1}{2}\qquad\quad,-\mu_{f,2}=\tfrac{k+1}{2}
\end{aligned}$$

depending whether f is a Maass form of weight $k=0$ or $k=1$ (here δ_f is the eigenvalue for the operator $Q_{it,k}$), or is holomorphic of weight $k\geqslant 1$, respectively. In particular, under $\mathrm{H}_2(\theta)$ one has

$$\lambda_f=\frac{1}{4}+t^2\geqslant\frac{1}{4}-\theta^2,\ \text{and}\ |\lambda_f(n)|\leqslant\tau(n)n^\theta. \tag{2.11}$$

Recall that for holomorphic forms, **RPC** was proved by Deligne and by Deligne/Serre [**De, De2, DS**]: hence $f \in S_k^p(q,\chi)$ one has

$$|\lambda_f(n)| \leqslant \tau(n). \tag{2.12}$$

In general, by the work of Kim/Shahidi, Kim and Kim/Sarnak [**KiSh, KiSh2, Ki, KiSa**], one knows that $\mathrm{H}_2(7/64)$ holds.

From the above identification, it follows that the classical Hecke L-function matches the automorphic one: $L(\pi_f, s) = L(f, s)$. Following the general theory of automorphic forms, one can also form the automorphic Rankin/Selberg L-function $L(\pi_f \otimes \pi_g, s)$, which is an Euler product of degree 4 and whose local factors match the local factors of $L(f \times g, s)$ at every prime not dividing the conductors of f and g. Also, by a result of Gelbart/Jacquet [**GeJ1**], there exists an automorphic representation of GL_3, the symmetric square of π_f, noted $\mathrm{sym}^2\pi_f$, whose L-function, $L(\mathrm{sym}^2\pi_f, s)$ has the same local factors as $L(\mathrm{sym}^2 f, s)$ at every prime not dividing q_f. These automorphic L-functions have *a priori* a less explicit definition (in particular they do not match their classical counterpart in general), but they are in many aspects the most natural objects to study. In particular, their functional equation takes a more natural form and is proved in full generality. As long as the analytic techniques have not been completely translated to the automorphic setting, it will be useful to take advantages of both aspects; the switch between the automorphic and the classical aspects is formalized via the following factorizations:

$$L(\pi_f.\chi, s) = F(\pi_f\chi, s)L(f \times \chi, s), \text{ with } L(f \times \chi, s) = \sum_n \frac{\lambda_f(n)\chi(n)}{n^s}, \tag{2.13}$$

$$L(\pi_f \otimes \pi_g, s) = F(\pi_f \otimes \pi_g, s)L(f \times g, s), \quad \text{with } L(f \times g, s) = L(\chi_f\chi_g, 2s)\sum_n \frac{\lambda_f(n)\lambda_g(n)}{n^s}, \tag{2.14}$$

$$L(\mathrm{sym}^2\pi_f, s) = F(\mathrm{sym}^2\pi_f, s)L(\mathrm{sym}^2 f, s), \quad \text{with } L(\mathrm{sym}^2 f, s) = L(\chi_f^2, 2s)\sum_n \frac{\lambda_f(n^2)}{n^s}. \tag{2.15}$$

Here the fudge factors F are finite Euler products supported at the ramified primes and are well controlled by the known bounds for the local parameters of π_f and π_g. In particular, it follows from $H_2(\theta)$ (for any fixed $\theta < 1/4$) that

$$(q_f q_g)^{-\varepsilon} \ll_\varepsilon F(\pi_f\chi, s), F(\pi_f \otimes \pi_g, s), F(\mathrm{sym}^2\pi_f, s) \ll_\varepsilon (q_f q_g)^\varepsilon$$

for any $\varepsilon > 0$, uniformly for $\Re es \geqslant 1/2 - \delta$, where δ is some positive absolute constant.

We can then utilize the analytic results on automorphic L-functions of the first lecture to derive upper and lower bounds for several classically defined quantities, such as the inner product $\langle f, f\rangle$. From the method of the proof of (1.19), the Siegel type theorem for $L(\pi_f \otimes \tilde{\pi}_f, s)$ (which is known, see Section 1.2.3), and from (1.24) (applied to $\pi = \mathrm{sym}^2\pi_f \otimes \overline{\chi_f}$), one has for any $t \in \mathbf{R}$,

$$(Q_{\pi_f}(1 + |t|))^{-\varepsilon} \ll_\varepsilon \mathrm{res}_{s=1} L(\pi_f \otimes \tilde{\pi}_f, s) \ll_\varepsilon (Q_{\pi_f}(1 + |t|))^{\varepsilon}. \tag{2.16}$$

The same bounds holds for its classical counterpart $L(f\times\overline{f},1+it)$. In particular, one has the following form of **RPC** on average:

$$\sum_{n\leqslant N}|\lambda_f(n)|^2\ll_\varepsilon (Q_fN)^\varepsilon N. \tag{2.17}$$

Moreover, the residue at $s=1$ of $L(f\times\overline{f},s)$ is related to the inner product $\langle f,f\rangle=|\rho_f(1)|^{-2}$ (see [**DFI8**] Sect. 19):

$$\mathrm{res}_{s=1}L(f\times\overline{f},s)=\begin{cases}\frac{32\pi^3\langle f,f\rangle}{\mathrm{vol}(X_0(q))|\Gamma(\frac12+it_f+\frac{k}{2})|^2} & \text{, if } f \text{ is a Maass form of weight k=0,1}\\ \frac{4\pi^2(4\pi)^k\langle f,f\rangle}{\mathrm{vol}(X_0(q))\Gamma(k)}, & \text{if } f \text{ is holomorphic.}\end{cases}$$

It follows that

$$(Q_f)^{-\varepsilon}q|\Gamma(\frac{1}{2}+it_f+\frac{k}{2})|^2\ll_\varepsilon\langle f,f\rangle\ll_\varepsilon (Q_f)^\varepsilon q|\Gamma(\frac{1}{2}+it_f+\frac{k}{2})|^2, \tag{2.18}$$

if f is a Maass form of weight $k=0,1$, or

$$(Q_f)^{-\varepsilon}q\frac{\Gamma(k)}{(4\pi)^k}\ll_\varepsilon\langle f,f\rangle\ll_\varepsilon (Q_f)^\varepsilon q\frac{\Gamma(k)}{(4\pi)^k} \tag{2.19}$$

if f is holomorphic.

2.3.1. Voronoi's summation formula

From the automorphy relations (2.1), one can deduce the following Voronoi-type summation formula; we display it for holomorphic forms in a simple case (see [**KMV2**] for more general formulae).

Lemma 2.3.1. *Let $W:\mathbf{R}^{*+}\to\mathbf{C}$ be a smooth function with compact support. Let $c\equiv 0(q)$ and a be an integer coprime to c. For $g\in\mathcal{S}_k(q,\chi)$ we have:*

$$c\sum_{n\geqslant 1}\sqrt{n}\rho_g(n)e(n\frac{a}{c})W(n)=2\pi i^k\overline{\chi}(a)\sum_{n\geqslant 1}\sqrt{n}\rho_g(n)e(-n\frac{\overline{a}}{c})\int_0^\infty W(x)J_{k-1}(\frac{4\pi\sqrt{nx}}{c})dx.$$

Proof. The proof follows from the automorphic relation

$$g(\frac{az+b}{cz+d})=(cz+d)^k\overline{\chi}(a)g(z)$$

applied to $z_t=-d/c+i/(ct)$ for $t\in\mathbf{R}_{>0}$. Then taking the Mellin transform of this equality, one deduces that the L-function

$$L(g,\frac{a}{c},s)=\sum_{n\geqslant 1}\rho_g(n)n^{-1/2-s}e(\frac{an}{c})$$

satisfies a functional equation with the same factors at infinity, conductor and root number, but related to $L(g,\frac{-\overline{a}}{c},1-s)$. The formula follows then from the inverse Mellin transform, the J Bessel function appearing as the inverse Mellin transform of the ratio of the Gamma factors. □

Remark 2.1. S. D. Miller and W. Schmidt [**MiS**] have derived Voronoi's summation formula by a somewhat different method based on the boundary distribution attached to a modular form. This new approach is smooth and provides a uniform treatment for all types of GL_2-forms; moreover, it extends naturally to automorphic forms of higher degree.

One can derive a similar identity for Eisenstein series: let $E(z,s)$ be the standard Eisenstein series for the full modular group:

$$E(z,s)=\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(1)}(\Im m\gamma z)^s,$$

and let g be its derivative at $1/2$; g has the following Fourier expansion

$$g(z)=\frac{\partial}{\partial s}E(z,s)_{|s=1/2}=y^{1/2}\log y+\sum_{n\geqslant 1}\tau(n)n^{-1/2}4\sqrt{ny}\cos(2\pi nx)K_0(2\pi ny).$$

Under the assumptions of Lemma 2.3.1, one has

$$c\sum_{n\geqslant 1}\tau(n)e\Big(\frac{an}{c}\Big)W(n)=2\int_0^{+\infty}(\log\frac{\sqrt{x}}{c}+\gamma)F(x)dx$$

$$+\sum_{n\geqslant 1}\tau(n)\int_0^{+\infty}\bigg(e\Big(-\frac{\overline{a}n}{c}\Big)(-2\pi Y_0)\Big(\frac{4\pi\sqrt{nx}}{c}\Big)+e\Big(\frac{\overline{a}n}{c}\Big)(4K_0)\Big(\frac{4\pi\sqrt{nx}}{c}\Big)\bigg)W(x)dx.$$

2.4. Trace formulae

In the sequel we will average heavily over families of automorphic forms; in order to do this, we use several trace formulae that are consequences of the spectral decomposition of the underlying spaces. The simplest of these formulae are the Dirichlet orthogonality relations for the characters of a finite group G: in particular, for $G=\mathbf{Z}/q\mathbf{Z}$ one has

$$\sum_{k(q)}\frac{1}{q}e(\frac{k(m-n)}{q})=\delta_{m\equiv n(q)} \tag{2.20}$$

and for $G=(\mathbf{Z}/q\mathbf{Z})^*$ one has

$$\sum_{\chi(q)}\frac{1}{\varphi(q)}\chi(m)\overline{\chi}(n)=\delta_{\substack{m\equiv n(q)\\(mn,q)=1}}. \tag{2.21}$$

For modular forms, the most natural candidate seems to be Selberg's trace formula; however, for analytic purposes, the following Petersson-Kuznetzov type formulae are much more efficient. We give two versions, one in the holomorphic case and one in the Maass case:

Theorem 2.1. *For $k\geqslant 2$, let $B_k(q,\chi)$ denotes an orthogonal basis of $S_k(q,\chi)$. Then for any $m,n\geqslant 1$ one has*

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\sum_{f\in B_k(q,\chi)}(mn)^{1/2}\overline{\rho_f(m)}\rho_f(n)=\delta_{m=n}+\Delta(m,n)$$

with

$$\Delta(m,n):=2\pi i^{-k}\sum_{\substack{c\equiv 0(q)\\c>0}}\frac{S_\chi(m,n;c)}{c}J_{k-1}(\frac{4\pi\sqrt{mn}}{c})$$

and $S_\chi(m,n;c)$ being the (twisted) Kloosterman sum

$$S_\chi(m,n;c)=\sum_{x(c),(x,c)=1}\overline{\chi}(x)e(\frac{m\overline{x}+nx}{c}).$$

Remark 2.2. There is no such formula for holomorphic forms of weight 1.

We note the following property of Bessel functions (see [**Wats**], p. 206)

$$J_\kappa(x) = e^{ix}W_\kappa(x) + e^{-ix}\overline{W}_\kappa(x), \tag{2.22}$$

where

$$W(x) = \frac{e^{i(\frac{\pi}{2}\kappa - \frac{\pi}{4})}}{\Gamma(\kappa+\frac{1}{2})}\sqrt{\frac{2}{\pi x}}\int_0^\infty e^{-y}\big(y(1+\frac{iy}{2x})\big)^{\kappa-\frac{1}{2}}dy.$$

When κ is a positive integer, we derive (using the Taylor expansion for $J_\kappa(x)$ if $0 < x \leqslant 1$, or the above integral expression for $W(x)$ if $x \geqslant 1$) the following bounds for the derivatives of W:

$$x^j W^{(j)}(x) \ll \frac{x}{(1+x)^{3/2}} \tag{2.23}$$

for any $j \geqslant 0$, the implied constant depending on j and κ.

In the case of Maass forms, we borrow the following version of the Kuznetsov formula from [**DFI8**]. Given $B_k(q,\chi) = \{u_j\}_{j\geqslant 1}$ an orthonormal basis of $\mathcal{C}_k(q,\chi)$ composed of Maass cusp forms with eigenvalues $\lambda_j = 1/4 + t_j^2$ and Fourier coefficients $\rho_j(n)$; or any real number r and any integer k, we set

$$h(t) = h(t,r) = \frac{4\pi^3}{|\Gamma(1-\frac{k}{2}-ir)|^2}.\frac{1}{\cosh\pi(r-t)\cosh\pi(r+t)} \tag{2.24}$$

Theorem 2.2. *For any positive integers m,n and any real r, one has*

$$\sqrt{mn}\sum_{j\geqslant 1} h(t_j)\overline{\rho}_j(m)\rho_j(n) + \sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}} h(t)\overline{\rho}_{\mathfrak{a}}(m,t)\rho_{\mathfrak{a}}(n,t)dt \tag{2.25}$$
$$= \delta_{m=n} + \sum_{c\equiv 0(q)}\frac{S_\chi(m,n;c)}{c}I(\frac{4\pi\sqrt{mn}}{c}),$$

where $I(x)$ is the Kloosterman integral

$$I(x) = I(x,r) = -2x\int_{-i}^{i}(-i\zeta)^{k-1}K_{2ir}(\zeta x)d\zeta.$$

In fact, this formula is not quite sufficient for all purposes. In order to have a function on the right hand side with rapid decay as c grows, one can perform an extra averaging over r (see [**DFI8**] Sect. 14). Given A a fixed large real number, we set

$$q(r) = \frac{r\sinh 2\pi r}{(r^2+A^2)^8}\big(\cosh\frac{\pi r}{2A}\big)^{-4A}. \tag{2.26}$$

Integrating $q(r)h(t,r)$ over r we form

$$\mathcal{H}(t) = \int_{\mathbf{R}} h(t,r)q(r)dr \text{ and } \mathcal{I}(x) = \int_{\mathbf{R}} I(x,r)q(r)dr. \tag{2.27}$$

Correspondingly, one has

$$\sqrt{mn}\sum_{j\geqslant 1}\mathcal{H}(t_j)\overline{\rho}_j(m)\rho_j(n) + \sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}}\mathcal{H}(t)\overline{\rho}_{\mathfrak{a}}(m,t)\rho_{\mathfrak{a}}(n,t)dt, \tag{2.28}$$
$$= c_A\delta_{m=n} + \sum_{c\equiv 0(q)}\frac{S_\chi(m,n;c)}{c}\mathcal{I}(\frac{4\pi\sqrt{mn}}{c})$$

where c_A is the integral of $q(r)$ over $\mathbf{R}$. We collect below the following estimates for $\mathcal{I}$ and $\mathcal{H}$ (see [**DFI8**] sect.14 and 17).
For t real or purely imaginary, one has

$$\mathcal{H}(t) > 0,\ \mathcal{H}(t) \gg (1+|t|)^{k-16} e^{-\pi t}. \tag{2.29}$$

For all $j \geqslant 0$, we have

$$x^j \mathcal{I}^{(j)}(x) \ll_j \left(\frac{x}{1+x}\right)^{A+1} (1+x)^{1+j}. \tag{2.30}$$

Such formulae are obtained by considering two Poincaré series $P_m(z)$ and $P_n(z)$, and by computing their scalar products $\langle P_m, P_n\rangle$ in two different ways. The first one is direct and based on the definition of the Poincaré series: the resulting expression involves Kloosterman sums. The other way is by applying the spectral decomposition (2.5), Parseval's formula, and the fact that $\langle P_m, f\rangle$ is proportional to $\overline{\rho}_f(m)$. In the case of Maass forms, there is some flexibility in the choice of the Poincaré series and one could obtain similar formulae with more general test functions in place of $h(t)$ or $\mathcal{H}(t)$, but this primitive version is sufficient for the applications given in these lectures. As was shown by Kuznetsov, it is also possible to replace $\mathcal{I}$ by a fairly general test function; this in turn enables one to connect the distribution of Kloosterman sums to automorphic forms. We refer to [**DI, Du, Pr**] for more general versions of Kuznetsov's formula, in particular, for forms of half-integral weight, and also to the book of Cogdell and Piatetski-Shapiro, where a representation theoretic derivation of Petersson/Kuznetsov's formula is given [**CoPS**].

Using Weil's bound for Kloosterman sums

$$|S_\chi(m,n;c)| \leqslant 2^{\omega(c)} (m,n,c)^{1/2} c^{1/2}, \tag{2.31}$$

and the following bounds for the Bessel functions

$$J_{k-1}(x) \ll \min\left(1, \frac{x^{k-1}}{(k-1)!}\right) \ll \left(\frac{x}{k}\right)^\sigma,\ for\ \sigma \in [0,1],$$

or the bounds for $\mathcal{I}$ given in (2.30), one can derive the following approximated versions of the above identities. For holomorphic forms of weight $k \geqslant 2$, one has (here the constant implied is absolute):

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}} \sum_{f \in B_k(q,\chi)} \sqrt{mn}\, \overline{\rho_f(m)} \rho_f(n) = \delta_{m=n} + O\left(\tau(q(m,n))(m,n,q)^{1/2} \log^2(1+mn) \frac{(mn)^{1/4}}{q\sqrt{k}}\right); \tag{2.32}$$

for Maass forms of weight $k = 0$ or 1, one has:

$$\sqrt{mn} \sum_{j \geqslant 1} \mathcal{H}(t_j) \overline{\rho}_j(m) \rho_j(n) + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{\mathbf{R}} \mathcal{H}(t) \overline{\rho}_{\mathfrak{a}}(m,t) \rho_{\mathfrak{a}}(n,t)\, dt = c_A \delta_{m=n} + O_A\left(\tau(q(m,n))(m,n,q)^{1/2} \frac{(mn)^{1/4}}{q}\right). \tag{2.33}$$

Hence with an appropriate averaging, the Fourier coefficients $\{\sqrt{n}\rho_f(n)\}_{f \in B(q,\chi)}$ are approximately orthogonal in the sense of section 1.3.3.

Remark 2.3. Applying (2.32) or (2.33) to $m = n$ and using positivity, we conclude that for f a cusp form, one has

$$\sqrt{n}\rho_f(n) \ll_{\varepsilon,f} n^{1/4+\varepsilon} \tag{2.34}$$

for any $\varepsilon > 0$. We will refer to this bound as a *Kloosterman type bound*; by applying it to a primitive form f and using the multiplicative properties of Hecke eigenvalues, we conclude that

$$|\lambda_f(n)| \leqslant \tau(n)n^{1/4},$$

which is $H_2(1/4)$ for the finite places. There is a similar argument for the infinite place (with a somewhat different test function $\mathcal{H}(t)$), which proves $H_2(1/4)$ for the infinite place (this is Selberg's bound). Note also that any improvement over the trivial bound for Kloosterman sums yields $H_2(\theta)$ for some $\theta < 1/2$.

LECTURE 3
Large sieve inequalities

3.1. The large sieve

In this lecture we describe in greater detail the concept of quasi-orthogonality for Hecke eigenvalues in families of automorphic forms: with the notations and assumptions of Section 1.3.3, quasi-orthogonality means that

$$E(\overline{\lambda_\pi}(m)\lambda_\pi(n), \mu_{\mathcal{F}}) = \delta_{m=n} + \mathrm{Err}_{\mathcal{F}}(m,n)$$

where the error term $\mathrm{Err}_{\mathcal{F}}(m,n)$ becomes small, as $Q_{\mathcal{F}}$ get large and m,n remain in some range depending on $Q_{\mathcal{F}}$. The orthogonality relations (2.20), (2.21) for characters are an instance of quasi-orthogonality since, for $1 \leqslant m,n \leqslant q$,

$$m \equiv n(q) \Leftrightarrow m = n.$$

The Kuznetsov-Petersson trace formula is another example for the range $1 \leqslant mn < kq^4$, in view of (2.32). However, sharp bounds for individual $\mathrm{Err}_{\mathcal{F}}(m,n)$ cannot be obtained with sufficient uniformity in m,n to be truly useful for several important applications (like the Subconvexity Problem), so one is inclined to mollify the problem and to consider quasi-orthogonality on average. Thus we switch to the problem of bounding, for $N \geqslant 1$ as large as possible, the hermitian quadratic form over $\mathbf{C}^N$

$$\mathcal{Q}_{\mathcal{F}}((a_n)_{n\leqslant N}) := E(|\sum_{n\leqslant N} a_n\lambda_\pi(n)|^2, \mu_{\mathcal{F}}) = \sum_{m,n\leqslant N} \overline{a_m} a_n E(\overline{\lambda_\pi}(m)\lambda_\pi(n), \mu_{\mathcal{F}})$$

in terms of $||(a_n)|| = \sum_{n\leqslant N} |a_n|^2$. Such bounds that are valid for all $(a_n)_{n\leqslant N}$ are called *large sieve inequalities*. For example, a version of quasi-orthogonality on average would be a bound of the form

$$\int_{\mathcal{F}} |\sum_{n\leqslant N} a_n\lambda_\pi(n)|^2 d\mu_{\mathcal{F}}(\pi) \ll_\varepsilon Q_{\mathcal{F}}^{\varepsilon} \sum_{n\leqslant N} |a_n|^2, \tag{3.1}$$

valid for all $\vec{a} \in \mathbf{C}^N$ and for N as large as possible; such a bound means that the averaged contribution of the $\mathrm{Err}_{\mathcal{F}}(m,n)$ is not much larger than the contribution from the diagonal terms $\delta_{m=n}$.

Remark 3.1. There is a natural limitation on the possible size of N in (3.1). To fix ideas, suppose that $\mathcal{F}$ is finite and is equipped with the uniform measure: if we take $a_n = \overline{\lambda_{\pi_0}(n)}$ for some $\pi_0 \in \mathcal{F}$ in (3.1), one obtains

$$\sum_{n \leqslant N} |\lambda_{\pi_0}(n)|^2 \ll_\varepsilon Q_{\mathcal{F}}^\varepsilon |\mathcal{F}|,$$

which puts a natural limit on N, since one expects (by Rankin/Selberg theory) that

$$\sum_{n \leqslant N} |\lambda_{\pi_0}(n)|^2 \gg_\varepsilon (NQ_{\pi_0})^{-\varepsilon} N.$$

3.1.1. Why large sieve ?

The denomination *large sieve inequality* is a bit misleading since the inequality itself has apparently little to do with sieving. The term goes back to the work of Linnik [**Lin1**], who showed how such inequalities could be very helpful in several problems related to the sieve. The purpose of this lecture is to provide several examples of this sieving technique.

Given a (finite) family $\mathcal{F} = \{\pi\}$ and a subset $\mathcal{E} \subset \mathcal{F}$, a primary objective of the sieve is to improve on the trivial bound

$$|\mathcal{E}| \leqslant |\mathcal{F}|$$

(for the purpose of showing that the complement $\mathcal{F} \backslash \mathcal{E}$ is not empty, for instance). In various cases it is possible to use a general inequality like (3.1): suppose one can find $N \geqslant 1$ (large) and a sequence $(a_n^{\mathcal{E}})_{n \leqslant N}$, depending on the set $\mathcal{E}$ we are interested in, normalized by

$$\sum_{n \leqslant N} |a_n^{\mathcal{E}}|^2 = 1$$

and satisfying, for any $\pi \in \mathcal{E}$, the lower bound

$$|\sum_{n \leqslant N} a_n^{\mathcal{E}} \lambda_\pi(n)| \geqslant N^\alpha$$

for some $\alpha > 0$. This inequality means that the linear form $\sum_{n \leqslant N} a_n^{\mathcal{E}} \lambda_\pi(n)$ takes large values at the π's belonging to $\mathcal{E}$[1] and so acts as a detector for $\mathcal{E}$. Applying the large sieve inequality (3.1) to the sequence $(a_n^{\mathcal{E}})_{n \leqslant N}$ (with $\mu_{\mathcal{F}}$ the uniform measure, for instance), one obtains

$$\frac{|\mathcal{E}|}{|\mathcal{F}|} N^{2\alpha} \leqslant \sum_{\pi \in \mathcal{E}} \frac{1}{|\mathcal{F}|} |\sum_{n \leqslant N} a_n^{\mathcal{E}} \lambda_\pi(n)|^2 \leqslant \sum_{\pi \in \mathcal{F}} \frac{1}{|\mathcal{F}|} |\sum_{n \leqslant N} a_n^{\mathcal{E}} \lambda_\pi(n)|^2 \ll_\varepsilon (Q_{\mathcal{F}})^\varepsilon,$$

hence

$$|\mathcal{E}| \ll_\varepsilon Q_{\mathcal{F}}^\varepsilon |\mathcal{F}| N^{-2\alpha},$$

which is non-trivial at least if N is large enough.

[1] and presumably not for $\pi \in \mathcal{F} \backslash \mathcal{E}$

3.1.2. Large sieve inequalities for characters

Simple examples of large sieve type inequalities are the ones for additive or multiplicative characters of modulus q; these follow easily from the Dirichlet orthogonality relations:

$$\sum_{a(q)} \frac{1}{q} \Big| \sum_{n \leqslant N} a_n e(\frac{an}{q}) \Big|^2 \ll (1 + \frac{N}{q}) \sum_{n \leqslant N} |a_n|^2, \tag{3.2}$$

$$\sum_{\chi(q)} \frac{1}{\varphi(q)} \Big| \sum_{n \leqslant N} a_n \chi(n) \Big|^2 \ll (1 + \frac{N}{q}) \sum_{\substack{n \leqslant N \\ (n,q)=1}} |a_n|^2. \tag{3.3}$$

These inequalities exhibit quasiorthogonality on average over $n \leqslant N \ll q$. Deeper and stronger are their extensions (due to Bombieri) when an extra averaging over the modulus q is performed ([**Bo**]):

Theorem 3.1. *For $N, Q \geqslant 1$ and any $\vec{a} \in \mathbf{C}^N$, we have*

$$\sum_{q \leqslant Q} \sum_{\substack{a(q) \\ (a,q)=1}} \Big| \sum_{n \leqslant N} a_n e(\frac{an}{q}) \Big|^2 \ll (Q^2 + N) \sum_{n \leqslant N} |a_n|^2; \tag{3.4}$$

here the summation is restricted to primitive characters and where the implied constant is absolute.

Theorem 3.2. *For $N, Q \geqslant 1$ and any $\vec{a} \in \mathbf{C}^N$, we have*

$$\sum_{q \leqslant Q} \sideset{}{^\times}\sum_{\chi(q)} \frac{q}{\varphi(q)} \Big| \sum_{n \leqslant N} a_n \chi(n) \Big|^2 \ll (Q^2 + N) \sum_{n \leqslant N} |a_n|^2; \tag{3.5}$$

here the summation is restricted to primitive characters and the implied constant is absolute.

This inequality means that on average over the moduli $q \leqslant Q$, quasiorthogonality holds when $N \ll Q^2$. Note that the proof of the additive version uses spectral analysis on $(\mathbf{R}, +)$ (*i.e.* Fourier analysis) and the duality principle presented below; interestingly, the multiplicative version is obtained from the additive one: the switch from additive to multiplicative characters is done by means of Gauss sums.

The most important consequence of this inequality (together with Siegel's Theorem) is the:

Theorem (Bombieri/Vinogradov). *For any $A \geqslant 1$, there exists $B(= 3A + 23) > 0$ such that for any $Q \leqslant x^{1/2}/\log^B x$, one has*

$$\sum_{q \leqslant Q} \operatorname*{Max}_{a(q),(a,q)=1} \Big| \psi(x; q, a) - \frac{1}{\varphi(q)} \psi(x; 1, 1) \Big| \ll_A x/(\log x)^A$$

where the implied constant depend on A only (but is not effective).

This inequality means that the primes less than x are very well distributed in arithmetic progressions of modulus q, on average over moduli q up to $Q = x^{1/2}/\log^{B+1} x$ (see [**BFI1**] for an elegant proof). For many applications, this is nearly as strong as **GRH** (the only difference being that, under **GRH**, B can be taken to equal 2). The Bombieri/Vinogradov theorem is a key ingredient (among others) in the proof of Chen's celebrated theorem which states that every sufficiently large even integer is the sum of a prime and an integer with at most 2 prime factors.

Remark 3.2. By using the spectral analysis on modular forms (based on the Petersson/Kuznetsov formula and its further developments by Deshouillers and Iwaniec [**DI**]), and the dispersion method of Linnik, Fouvry/Iwaniec obtained for the first time results of Bombieri/Vinogradov type on the distribution of arithmetic sequences (including the primes) in arithmetic progressions over special (*i.e. well factorable*), but very large, moduli (*i.e. with $q \leqslant Q$ with $Q = x^{\beta}$ and $\beta > 1/2$*) [**FoI1, FoI2**]; this is *beyond* the possibilities of **GRH** ! The ideas and methods of these papers were further polished and magnified by Bombieri/Friedlander/Iwaniec and Fouvry in [**BFI1, BFI2, BFI3, Fo1, Fo2**]: among the various Bombieri/Vinogradov type theorems beating **GRH**, one can quote the following two results from [**BFI2, BFI1**]. For any fixed a, for some $A > 0$ and for any $Q \leqslant x^{1/2}$,

$$\sum_{\substack{q\leqslant Q\\(a,q)=1}} \bigl|\psi(x;q,a) - \frac{1}{\varphi(q)}\psi(x;1,1)\bigr| \ll_a x/(\log x)^A.$$

At the expense of replacing the absolute value above by a more flexible function, it is possible to pass far beyond the critical exponent $1/2$ (and **GRH**): for any $\varepsilon > 0$ and any $A \geqslant 1$, one has

$$\sum_{\substack{q\leqslant Q\\(a,q)=1}} \lambda(q)\bigl(\psi(x;q,a) - \frac{1}{\varphi(q)}\psi(x;1,1)\bigr) \ll_{a,A,\varepsilon} x/(\log x)^A,$$

for any $Q \leqslant x^{4/7-\varepsilon}$; here $\lambda(q)$ denotes any bounded arithmetical function satisfying an extra technical assumption (*i.e. is well factorability*), which is not limiting in applications.

One cannot end this section without mentioning the large sieve inequality for real Dirichlet characters of Heath-Brown [**HB**].

Theorem 3.3. *For any complex numbers a_n we have*

$$\sideset{}{^\flat}\sum_{q\leqslant Q}\bigl|\sideset{}{^\flat}\sum_{n\leqslant N} a_n(\frac{n}{q})\bigr|^2 \ll (QN)^{\varepsilon}(Q+N)\sideset{}{^\flat}\sum_{n\leqslant N}|a_n|^2 \tag{3.6}$$

with any $\varepsilon > 0$, the implied constant depending only on ε. Here $\sum^\flat$ indicates restriction to positive odd squarefree integers.

Although this inequality looks very similar to the previous ones (observe that the number of primitive real characters of modulus $\leqslant Q$ is $\gg Q$), its proof is much more involved; this is a powerful estimate, and we refer to [**HB, So, IM**] for some applications.

3.1.3. Large sieve inequalities for modular forms

In the context of modular forms, the natural analog of the character values $\chi(n)$ are the Fourier coefficients $\rho_f(n)$; not surprisingly, large sieve inequalities exist and are consequences of the Petersson/Kuznetzov formulae of the previous section. In the case of holomorphic forms, one has the following [**DFI3**] :

Proposition 3.1. *For $(a_n)_{n\leqslant N}$ a sequence of complex numbers and $k \geqslant 2$, one has*

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\sum_{f\in B_k(q,\chi)}\Bigl|\sum_{n\leqslant N} a_n n^{1/2}\rho_f(n)\Bigr|^2 = \bigl(1+O(\frac{N\log N}{qk})\bigr)\sum_{n\leqslant N}|a_n|^2. \tag{3.7}$$

Thus (3.7) shows quasiorthogonality of the Fourier coefficients on average over $n \leqslant N \ll qk$ (note that $qk/12$ is approximately the size of $B_k(q,\chi)$). For nonlacunary sequences (a_n), the above inequality is much stronger than the one obtained from the individual estimates (2.32), (2.33) of the previous lecture. The proof of (3.7) is obtained through Petersson's formula, which transform the above quadratic form into another one with Kloosterman sums (weighted by Bessel functions). The proof follows by opening the Kloosterman sums and by reduction to the large sieve inequality (3.2) for additive characters.

Proof. By Theorem. 2.1, the lefthand side of (3.7) equals

$$\sum_{n\leqslant N}|a_n|^2+2\pi i^{-k}\sum_{c\equiv 0(q)}\frac{1}{c}\sum_{m,n\leqslant N}\overline{a_m}a_nS_\chi(m,n;c)J_{k-1}(\frac{4\pi\sqrt{mn}}{c}).$$

To apply the large sieve inequality (3.2) we need to separate the variables m and n in the Bessel function: we do this by means of the Taylor expansion (which we use for $0\leqslant x/2\leqslant 1$),

$$J_{k-1}(x)=\sum_{l\geqslant 0}\frac{(-1)^l}{l!(l+k-1)!}(\frac{x}{2})^{2l+k-1}.$$

We assume first that $2\pi N/q\leqslant 1$; then by opening the Kloosterman sum, and applying Cauchy/Schwarz and (3.2), one has for $c\equiv 0(q)$ and $l\geqslant 0$,

$$\frac{1}{c}\sum_{m,n\leqslant N}\overline{a_m}a_nS_\chi(m,n;c)(\frac{2\pi\sqrt{mn}}{c})^{2l}\leqslant\frac{1}{c}\sum_{\substack{x(c)\\(x,c)=1}}\Big|\sum_{n\leqslant N}a_n(\frac{2\pi n}{c})^le(\frac{nx}{c})\Big|^2\ll\sum_{n\leqslant N}|a_n|^2.$$

On the other hand, by Weil's bound (2.31) one has

$$\frac{1}{c}\sum_{m,n\leqslant N}\overline{a_m}a_nS_\chi(m,n;c)(\frac{2\pi\sqrt{mn}}{c})^{2l}\leqslant\frac{N\tau^2(c)}{c^{1/2}}\sum_{n\leqslant N}|a_n|^2.$$

It follows that (since $k\geqslant 2$)

$$\sum_{c\equiv 0(q)}\frac{1}{c}\sum_{m,n\leqslant N}\overline{a_m}a_nS_\chi(m,n;c)J_{k-1}(\frac{4\pi\sqrt{mn}}{c})$$

$$\ll\frac{(2\pi)^{k-1}}{(k-1)!}\sum_{n\leqslant N}|a_n|^2\sum_{c\equiv 0(q)}\frac{N^{k-1}}{c^{k-1}}\min(1,\frac{N\tau^2(c)}{c^{1/2}})\ll\frac{(2\pi)^{k-1}}{(k-1)!}\frac{N\log^2(N)}{q}.$$

Hence (3.7) is proved for $N\leqslant q/2\pi$. The remaining case is proved by the following trick which goes back to **[I1]**: one choose p a prime such that $pq\ll 2\pi N\leqslant pq$; the basis $\frac{1}{[\Gamma_0(q):\Gamma_0(pq)]^{1/2}}\mathcal{B}_k(q,\chi)$ can be embedded into some orthonormal basis $\mathcal{B}_k(pq,\chi)$ of $S_k(pq,\chi)$. The remaining case follows from the inequality

$$\sum_{f\in B_k(q,\chi)}\Big|\sum_{n\leqslant N}a_nn^{1/2}\rho_f(n)\Big|^2\leqslant[\Gamma_0(q):\Gamma_0(pq)]\sum_{f\in B_k(pq,\chi)}\Big|\sum_{n\leqslant N}a_nn^{1/2}\rho_f(n)\Big|^2.$$

□

In the case of Maass forms, one has a similar inequality (by comparison with Weyl's law (4.20)) due to Deshouillers/Iwaniec [**DI**]: for $k=0,1$ one has, for any $\varepsilon>0$

$$\sum_{|t_j|\leqslant T}\frac{1}{\mathrm{ch}(\pi t_j)}\sum_{f\in B_k(q,\chi,it_j)}\Big|\sum_{n\leqslant N}a_n n^{1/2}\rho_f(n)\Big|^2\ll_\varepsilon (T^2+\frac{N^{1+\varepsilon}}{q})\sum_{n\leqslant N}|a_n|^2. \tag{3.8}$$

The proof of (3.8) is similar to the proof of (3.7) above with Petersson's formula replaced by the Petersson/Kuznetzov formula and with a more involved analysis of the Bessel transforms, we refer to [**DI**] for a proof.

Remark 3.3. It is interesting to take $k=1$ and $T=0$ in (3.8) above: then by positivity and the identification (2.6), one obtains a large sieve inequality for holomorphic forms of weight 1:

$$\sum_{f\in B_1(q,\chi)}\Big|\sum_{n\leqslant N}a_n n^{1/2}\rho_f(n)\Big|^2\ll_\varepsilon (1+\frac{N^{1+\varepsilon}}{q})\sum_{n\leqslant N}|a_n|^2. \tag{3.9}$$

We give below another more direct and more elementary proof (due to William Duke) of this inequality which does not use the spectral decomposition of Δ_1.

3.2. The Duality Principle

A crucial ingredient for the proof of many large sieve type inequalities is the following duality principle:

Proposition 3.2. *Let $N\geqslant 1$ and $\mathcal{F}$ be a finite set, $(\lambda_\pi(n))_{\substack{\pi\in\mathcal{F}\\ 1\leqslant n\leqslant N}}$ any complex numbers and $\Delta>0$ a real number. Then the following are equivalent:*

1. *For all $(b_\pi)_{\pi\in\mathcal{F}}\in\mathbf{C}^{\mathcal{F}}$,*
$$\sum_{n\leqslant N}\Big|\sum_{\pi}\lambda_\pi(n)b_\pi\Big|^2\leqslant\Delta\sum_{\pi\in\mathcal{F}}|b_\pi|^2.$$
2. *For all $(a_n)_{1\leqslant n\leqslant N}\in\mathbf{C}^N$,*
$$\sum_{\pi\in\mathcal{F}}\Big|\sum_{n\leqslant N}\lambda_\pi(n)a_n\Big|^2\leqslant\Delta\sum_{n\leqslant N}|a_n|^2.$$

Proof. (1) states that the L_2 norm of the linear operator between Hilbert spaces

$$\mathbf{C}^{\mathcal{F}}\mapsto\mathbf{C}^N:(b_\pi)_{\pi\in\mathcal{F}}\mapsto(a_n)_{n\leqslant N}:=(\sum_{\pi\in\mathcal{F}}b_\pi\lambda_\pi(n))_{n\leqslant N}$$

is bounded by Δ. (2) states that the adjoint of the above operator is bounded by Δ. The equivalence of (1) and (2) is then obvious. □

The duality principle is quite powerful and flexible and can be applied in a number of contexts where harmonic analysis is a priori missing. We illustrate this with the derivation of two inequalities. The first is a direct derivation (due to W. Duke) of the large sieve inequality (3.9) for holomorphic modular forms of weight one. The second (due to Duke and Kowalski [**DK**]) deals with quite general families of automorphic forms of arbitrary rank.

3.2.1. The large sieve inequality for forms of weight one

Theorem 3.4. *For any $\varepsilon > 0$, $N \geqslant 1$ and any $(a_n)_{n\leqslant N} \in \mathbf{C}^N$,*

$$\sum_{f\in S_1^p(q,\chi)} \Big| \sum_{n\leqslant N} \lambda_f(n) a_n \Big|^2 \ll_\varepsilon q^\varepsilon (q+N) \sum_{n\leqslant N} |a_n|^2. \tag{3.10}$$

Proof. We provide here an elementary proof of (a slightly stronger form of) (3.9):

$$\sum_{f\in B_1(q,\chi)} \Big| \sum_{n\leqslant N} a_n n^{1/2} \rho_f(n) \Big|^2 \ll (1+\frac{N}{q}) \sum_{n\leqslant N} |a_n|^2.$$

By duality, it is sufficient to prove that

$$\sum_{n\leqslant N} \Big| \sum_{f\in B_1(q,\chi)} \rho_f(n)\sqrt{n} b_f \Big|^2 \ll (1+\frac{N}{q}) \sum_f |b_f|^2$$

Set $g = \sum_{f\in B_1(q,\chi)} b_f f$; then $\sum_{f\in B_1(q,\chi)} \rho_f(n)\sqrt{n} b_f = n^{1/2}\rho_g(n)$ and the inequality will follows from the inequality

$$\sum_{n\leqslant N} |n^{1/2}\rho_g(n)|^2 \leqslant (1+\frac{N}{q})\langle g,g\rangle = (1+\frac{N}{q}) \sum_f |b_f|^2.$$

We use the following argument of Iwaniec:

$$\sum_n |n^{1/2}\rho_g(n)|^2 \exp(-4\pi n y) = \int_0^1 |g(x+iy)|^2 dx$$

and hence

$$\sum_{n\leqslant N} |n^{1/2}\rho_g(n)|^2 \ll \sum_n |n^{1/2}\rho_g(n)|^2 \int_{1/N}^{+\infty} \exp(-4\pi n y) y \frac{dy}{y^2} = \int_{P(1/N)} |g(z)|^2 y \frac{dxdy}{y^2}$$

where $P(1/N) = \{z \in \mathbf{H},\ 0 \leqslant \Re e z < 1, \Im m z > 1/N\}$. The number of fundamental domains for $\Gamma_0(q)$ that are needed to cover $P(1/N)$ is bounded by (which we leave as an exercise)

$$\operatorname*{Max}_{z\in P(1/N)} |\{\gamma \in \Gamma_0(q), \gamma z \in P(1/N)\}| \leqslant (1+10\frac{N}{q}).$$

Hence

$$\sum_{n\leqslant N} |n^{1/2}\rho_g(n)|^2 \ll (1+\frac{N}{q})\langle g,g\rangle.$$

□

Remark 3.4. By the same method and Weyl's law, one could prove (3.8), without the use of the Petersson-trace formula.

3.2.2. Large sieve inequalities for automorphic forms of higher rank

For quite general families of automorphic L-functions, large sieve inequalities similar to (3.5) are almost a formal consequence of the above duality principle and the existence of Rankin-Selberg theory.

Theorem 3.5. *Let $\mathcal{F}$ be a finite subset of $\mathcal{A}_d^0(\mathbf{Q})$ and $Q_{\mathcal{F}} = \sup_{\pi\in\mathcal{F}} Q_\pi$. We assume that any $\pi \in \mathcal{F}$ satisfies (1.8) for some fixed $\theta < 1/4$, and that two distinct $\pi, \pi' \in \mathcal{F}$ are never twist of each other by a character of the form $|\det|^{it}$ for some $t \in \mathbf{R}$; then for all $N \geqslant 1$ and all $(a_n) \in \mathbf{C}^N$, one has*

$$\sum_{\pi\in\mathcal{F}} \Big|\sum_{n\leqslant N} a_n\lambda_\pi(n)\Big|^2 \ll_{d,\varepsilon} (Q_{\mathcal{F}}N)^\varepsilon (N + |\mathcal{F}|^2 Q_{\mathcal{F}}^d) \sum_n |a_n|^2.$$

Remark 3.5. While such a bound is not as strong as quasiorthogonality in the sense of (3.1), it is still non-trivial (when N is sufficiently large) by comparison with the trivial bound $(Q_{\mathcal{F}}N)^\varepsilon|\mathcal{F}|N\sum_n |a_n|^2$.

Proof. By the duality principle, it is sufficient to prove that for $(b_\pi)_{\pi\in\mathcal{F}}$ one has

$$\sum_{n\leqslant N} \Big|\sum_{\pi\in\mathcal{F}} b_\pi\lambda_\pi(n)\Big|^2 \ll_{d,\varepsilon} (Q_{\mathcal{F}}N)^\varepsilon (N + |\mathcal{F}|^2 Q_{\mathcal{F}}^d) \sum_\pi |b_\pi|^2.$$

To do this we multiply the left hand side by $g(n)$ where g is a smooth non-negative function, compactly supported in $]0, N+1]$, which takes the value 1 on $[1, N]$. Opening the square, one has to bound the sum

$$\sum_{\pi,\pi'} b_\pi\bar{b}_{\pi'} \sum_n g(n)\lambda_\pi(n)\overline{\lambda_{\pi'}(n)} = \frac{1}{2\pi i}\int_{(3)} \hat{g}(s) \sum_n \lambda_\pi(n)\overline{\lambda_{\pi'}(n)} n^{-s} ds. \tag{3.11}$$

The fact that all $\pi, \pi' \in \mathcal{F}$ satisfy (1.8) for some $\theta < 1/4$ enables us to show:

1. the following factorization holds: for $Res > 3$,

$$\sum_n \lambda_\pi(n)\overline{\lambda_{\pi'}(n)} n^{-s} = H(\pi, \tilde{\pi}', s) L(\pi \otimes \tilde{\pi}', s)$$

   for $H(\pi, \pi', s)$ some Euler product, which, in the the domain $\Re es \geqslant 1/2$, is absolutely convergent and uniformly bounded by $C(\varepsilon, d) Q^\varepsilon_{\pi\otimes\tilde{\pi}'}$, for any $\varepsilon > 0$.
2. the convexity bound

$$L(\pi \otimes \pi', s) \ll_{\varepsilon,d} Q^{1/4+\varepsilon}_{\pi\otimes\tilde{\pi}'}$$

   for $\Re es = 1/2$, and at $s = 1$, the bound

$$\mathrm{res}_{s=1} L(\pi \otimes \pi', s) \ll_{\varepsilon,d} Q^\varepsilon_{\pi\otimes\pi'},$$

the latter residue being non-zero iff $\pi = \pi'$. A shift of the s-contour to the line $\Re es = 1/2$ in (3.11) yields

$$\frac{1}{2\pi i}\int_{(3)} \hat{g}(s) \sum_n \lambda_\pi(n)\overline{\lambda_{\pi'}(n)} n^{-s} ds = \hat{g}(1)\mathrm{res}_{s=1} L(\pi \otimes \pi', s) + O_{\varepsilon,d,\theta}(N^{1/2} Q^{1/4+\varepsilon}_{\pi\otimes\tilde{\pi}'}).$$

Hence, from the bound $Q_{\pi\otimes\tilde{\pi}'} \ll_d (Q_\pi Q_{\pi'})^d \leqslant Q_{\mathcal{F}}^{2d}$ we obtain that

$$\sum_{\pi,\pi'} b_\pi\bar{b}_{\pi'} \sum_n g(n)\lambda_\pi(n)\overline{\lambda_{\pi'}(n)} \ll_{\varepsilon,d,\theta} Q_{\mathcal{F}}^\varepsilon (N + N^{1/2}|\mathcal{F}| Q_{\mathcal{F}}^{d/2}) \sum_\pi |b_\pi|^2.$$

$\square$

3.3. Some applications of the large sieve

As we have seen in section 3.1.1, one of the main applications of the large sieve inequalities is to provide nontrivial upper bounds for the cardinality of sets having some arithmetic structure.

3.3.1. Linnik's theorem and its generalizations

Linnik's original motivation [**Lin1**] was the classical problem:

Question. *Given q a prime number, what is the size of the least prime $p_{\mathrm{q}}(q)$ (resp. $p_{\mathrm{nq}}(q)$) that is a quadratic (resp. a nonquadratic)-residue mod q?*

It follows from **GRH** that

$$p_{\mathrm{q}}(q),\ p_{\mathrm{nq}}(q) \ll (\log q)^A$$

for some absolute constant $A(=2)$; unconditionally one has the much weaker upper bound from Burgess's Theorem 4.3 of the next lecture,

$$p_{\mathrm{q}}(q),\ p_{\mathrm{nq}}(q) \ll_\varepsilon q^{1/4+\varepsilon}.$$

By an additional sieving trick, Vinogradov improved this bound for the least non-quadratic residue to

$$p_{\mathrm{nq}}(q) \ll_\varepsilon q^{1/(4\sqrt{e})+\varepsilon}.$$

As one of the first application of the large sieve, Linnik proved that for A sufficiently large, almost all primes q have their least prime quadratic (resp. nonquadratic)-residue smaller than $(\log q)^A$. More precisely, one has (in a slightly different form than Linnik's original statement)

Theorem 3.6. *For $A, Q \geqslant 1$, the number of primes $q \leqslant Q$ such that*

$$p_{\mathrm{q}}(q),\ \text{or } p_{\mathrm{nq}}(q) \geqslant (\log q)^A$$

is bounded by $\ll_A Q^{B/A}$ for some absolute constant B.

Proof. Let $\mathcal{M}_A(Q)$ be the set of such primes q. By definition, for any $q \in \mathcal{M}_A(Q)$ and any prime $p < (\log q)^A$ one has $(\frac{p}{q}) = -1$ (resp. 1). For $N \geqslant Q$ we let $\mathcal{N}$ be the set of integers less than N, having all their prime factors $< (\log q)^A$; hence for $n \in \mathcal{N}$ and $q > (\log Q)^A$ one has $(\frac{n}{q}) = (-1)^{\Omega(n)}$ (resp. 1). We set $a_n = (-1)^{\Omega(n)}$ (resp. 1) if $n \in \mathcal{N}$ and 0 otherwise and we apply[2] (3.5), obtaining

$$|\mathcal{M}_A(Q)||\mathcal{N}|^2 \leqslant \sum_{\substack{q \leqslant Q \\ q \text{ prime}}} |\sum_{n \leqslant N} a_n (\frac{n}{q})|^2 \ll (Q^2+N)|\mathcal{N}|,$$

and hence

$$|\mathcal{M}_A(Q)| \ll (Q^2+N)/|\mathcal{N}|.$$

It is elementary to show (see [**Te**] III.5 Proof of Thm.2 for instance) that $|\mathcal{N}| \gg_A N^{1-1/A}$, and the result follows by taking $N = Q^2$. □

The question about the size of the least non-quadratic residue can be generalized to various contexts: for instance in the case of modular forms, one has the following:

[2] for instance, but a weaker inequality would be sufficient if one were not concerned with the size of B.

Question. *Given f and g two distinct primitive holomorphic forms, what is the smallest possible $n = N(f, g)$ for which $\lambda_f(n) \neq \lambda_g(n)$?*

Serre [**Ser2**] observed that under **GRH**, one has $N(f, g) \ll \log^A(Q_f Q_g)$ for some absolute constant $A(= 4)$, and we will discuss some weaker unconditional bounds for $N(f, g)$ in the last lecture. The following result of Duke and Kowalski [**DK**] is the analog of Thm. 3.6:

Theorem 3.7. *Let $S_2^p(\leqslant Q)$ be the set of primitive holomorphic forms with trivial nebentypus, weight 2 and level $q \leqslant Q$. Then there exists an absolute constant $B > 0$ such that for any $A > 1$, the number of pairs $(f, g) \in S_2^p(\leqslant Q)^2$ such that $N(f, g) \geqslant \log(Q_f Q_g)^A$ is bounded by*

$$|S_2^p(\leqslant Q)|^{3/2+B/A}.$$

In particular, if A is sufficiently large, the probability that two forms are not distinguished by their first $\log(Q_f Q_g)^A$ Hecke eigenvalues goes to 0 as $Q \to +\infty$.

The proof of this statement is similar to the previous one; one uses Theorem 3.5 for the automorphic representations π_f, $f \in S_2^p(\leqslant Q)$. However, the fact that one does not have (unconditionally) a good lower bound for

$$\sum_{\substack{n \in \mathcal{N} \\ n \leqslant N}} |\lambda_f(p)|^2$$

creates some difficulty. The latter is solved with the help of the identity

$$\lambda_f(p)^2 - \lambda_f(p^2) = 1,$$

which shows that at least one of the two quantities $\lambda_f(p)^2$ or $\lambda_f(p^2)$ is large. Using this idea requires a further modification of Linnik's original method, for which we refer to [**DK**], and in particular, another application of Thm 3.5, but this time for the family of symmetric squares $\{\mathrm{sym}^2 \pi_f,\ f \in S_2^p(\leqslant Q)\}$.

3.3.2. Duke's bound for $|S_1^p(q, \chi)|$

Another recent striking application of the large sieve inequalities is to the problem of estimating the size of the set of primitive holomorphic forms of weight one (of some given nebentypus). There is no Selberg trace nor Petersson's type formula on the space generated by these forms, hence apparently no manageable formula for its dimension. However, an application of the trace formula to the whole space of automorphic forms of weight 1 yields a generic bound (Weyl's law),

$$|S_1^p(q, \chi)| \ll \mathrm{vol}(X_0(q)). \tag{3.12}$$

With a more careful choice of the test function this bound can be improved by a factor $\log q$ (Sarnak) and this seems to be the best one can do without exploiting further the specific (arithmetic) nature of the forms of weight one.

On the other hand, Deligne and Serre [**DS**] attached to each $f \in S_1^p(q, \chi)$ an irreducible 2-dimensional Galois representation $\rho_f : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \mapsto GL_2(\mathbf{C})$, unramified outside the primes dividing q and satisfying for any prime $p \not| q$,

$$\mathrm{tr}(\rho_f(\mathrm{Frob}_p)) = \lambda_f(p),\ \det(\rho_f(\mathrm{Frob}_p)) = \chi(p).$$

In particular (since the image of ρ_f is finite), f satisfies the Ramanujan-Petersson conjecture: $|\lambda_f(p)| \leqslant 2$. The form f can then be further classified according to the image of ρ_f via the projection $GL_2(\mathbf{C}) \mapsto PGL_2(\mathbf{C})$: the latter is isomorphic either to a dihedral group $D_{2n},\ n \geqslant 1$, or to one of the *exotic* groups A_4, S_4, A_5. Accordingly, one says that f is of *dihedral, tetrahedral, octahedral* or *icosahedral* type, respectively.

The forms of dihedral type are well understood: up to some twist, they are given by theta series associated to (non-real) characters of the ideal class group of imaginary quadratic orders. In particular one has the upper bound

$$|S_1^{Dihedral}(q,\chi)| \ll_\varepsilon q^{1/2}\log q; \tag{3.13}$$

while Siegel's theorem and the known estimates for the size of the non-maximal orders of imaginary quadratic fields (see for example [**Cor**]) give, for χ is quadratic, the lower bound

$$|S_1^{Dihedral}(q,\chi)| \gg_\varepsilon q^{1/2-\varepsilon}$$

for any $\varepsilon > 0$.

The forms of exotic type are much more mysterious. In [**Ser1**], Serre studied in depth the possible Hecke eigenvalues of exotic forms (for quadratic nebentypus) and found that they belong to an explicit finite set; this lead him to conjecture that exotics forms are very rare: more precisely, that

$$|S_1^{Exotic}(q,\chi)| \ll_\varepsilon q^\varepsilon,$$

for any $\varepsilon > 0$. Duke made the first significant step in the direction of this conjecture by improving the trivial bound by a factor $q^{1/11}$ for a quadratic nebentypus χ; a little later, S. Wong extended Duke's argument to nebentypus χ [**Wo**]:

Theorem 3.8. *One has*

$$|S_1^{exotic}(q,\chi)| \ll_\varepsilon q^{11/12+\varepsilon},$$

for any $\varepsilon > 0$.

Proof. Duke's method uses two conflicting properties of the Hecke eigenvalues of exotic forms.

1. *Rigidity*: for $p \not| q$, the $\lambda_f(p)$ are frobenius eigenvalues associated to a 2-dimensional complex galois representation ρ_f; the study of the possible lifting of the exotic subgroups to $GL_2(\mathbf{C})$ provides a lot of information on the character table of ρ_f hence on the $\lambda_f(p)$ ([**Ser1, Wo**]). In particular, one can the following linear relation, valid for any $f \in S_1^{Exotic}(q,\chi)$ and any prime $p \not| q$,

$$\overline{\chi}^6(p)\lambda_f(p^{12}) - \overline{\chi}^4(p)\lambda_f(p^8) - \overline{\chi}(p)\lambda_f(p^2) = 1. \tag{3.14}$$

   It implies that the $\lambda_f(p)$ are algebraic integers that can take only finitely many values (depending on the order of χ).
2. *Orthogonality*: the quasiorthogonality relations among eigenvalues are encapsulated in (3.10) and reflect their modular nature.

One chooses a_n as follows:

$$a_n = \begin{cases} \overline{\chi}^6(p), \text{ for } n = p^{12} \leqslant N,\ (p,q)=1 \\ -\overline{\chi}^4(p), \text{ for } n = p^8 \leqslant N^{8/12},\ (p,q)=1 \\ -\overline{\chi}(p), \text{ for } n = p^2 \leqslant N^{2/12},\ (p,q)=1 \\ 0 \text{ else.} \end{cases}$$

By (3.14) and the prime number theorem, one has for any $f\in S_1^{Exotic}(q,\chi)$

$$\sum_{n\leqslant N}a_n\lambda_f(n)\gg\frac{N^{1/12}}{\log N}-O(\log q)$$

(this amounts to counting the primes coprime to q and $\leqslant N^{1/12}$), and

$$\sum_{n\leqslant N}|a_n|^2\ll\frac{N^{1/12}}{\log N}.$$

Thus

$$|S_1^{Exotic}(q,\chi)|\frac{N^{2/12}}{\log^2 N}\ll\sum_{f\in S_1^p(q,\chi)}\Big|\sum_{n\leqslant N}\lambda_f(n)a_n\Big|^2\ll_\varepsilon q^\varepsilon(q+N),\frac{N^{1/12}}{\log N}.$$

and hence the bound follows by taking $N=q$.

$\square$

Remark 3.6. J. Ellenberg made the following observation [**El**]: for q is squarefree and $f\in S_1(q,\chi)$, and p any prime dividing q, the local representation of the inertia subgroup at p is isomorphic to

$$(\rho_f)_{|I_p}\simeq 1\oplus(\rho_\chi)_{|I_p}.$$

In particular, $\rho_f(I_p)$ is isomorphic to its image in $PGL_2(\mathbf{C})$ which is a cyclic subgroup contained in A_4,S_4 or A_5; hence χ^{60} is unramified at any $p|q$ hence is trivial. It follows that for q squarefree, the number of $\chi(q)$ that are nebentypus of a form of weight 1 and level q is bounded by $O_\varepsilon(q^\varepsilon)$ for any $\varepsilon>0$. Hence if we denote by $S_1^{Exotic}(q)$ the number of exotic forms of level q (regardless of their nebentypus), one has

$$|S_1^{Exotic}(q)|\ll_\varepsilon q^{11/12+\varepsilon}.$$

3.3.3. Zero density estimates

Given $\pi\in\mathcal{A}_d^0(\mathbf{Q})$, one can show by standard methods that the number of zeros of $L(\pi,s)$ within the critical strip and with height less by $T\geqslant 1$ is

(3.15)
$$N(\pi,T)=|\{\rho,\ L(\pi,\rho)=0,\ 0\leqslant\Re e\rho,\ |\Im m\rho|\leqslant T\}|\sim T\log(Q_\pi T),\ T\to+\infty.$$

Of course, **GRH** predicts that all such zeros are on the critical line; however, for many applications, it is sufficient to know that few zeros are close to $\Re es=1$. To measure this, we set for $\alpha\geqslant 1/2$,

$$R(\alpha,T):=\{\rho\in\mathbf{C},\ \Re e\rho\geqslant\alpha,\ |\Im m\rho|\leqslant T\},$$

$$\mathcal{Z}(\pi;\alpha,T)=\{\rho,\ L(\pi,\rho)=0,\rho\in R(\alpha,T)\}\text{ and }N(\pi;\alpha,T):=|\mathcal{Z}(\pi;\alpha,T)|.$$

A *zero density estimate* is a bound for $N(\pi;\alpha,T)$ that improves on (3.15) for $\alpha>1/2$. One can mollify the problem further – again this is sufficient for many applications – and ask for a non-trivial bound for $N(\pi,\sigma,T)$ *on average* over a family $\mathcal{F}$.

There are many sorts of zero density estimates; in this section, we describe a very general version that is obtained with the large sieve; for families of Dirichlet characters, this approach started with Linnik and was developed in the works of Barban, Bombieri, Montgomery and others [**Bar, Bo, Mon**]. For example, zero density estimates have been applied by Linnik and then by others to study the following:

Question. *Given* $1 \leqslant a < q$ *two coprimes integers, give a bound for the smallest prime* $> a$*,* $p(q,a)$ *say, congruent to* a *modulo* q*.*

Clearly $p(q,a) \gg_\varepsilon q^{1-\varepsilon}$. The **GRH** implies[3] $p(q,a) \ll_\varepsilon q^{2+\varepsilon}$, while the Hadamard/de la Vallée-Poussin zero-free region and Siegel's theorem give the much weaker bound $p(q,a) \ll_\varepsilon \exp(q^\varepsilon)$ for any $\varepsilon > 0$ (the implied constant being ineffective). A formidable achievement is the theorem of Linnik [**Lin2**], from 1944, which gave an unconditional bound, qualitatively as good as the **GRH** bound:

Theorem (Linnik). *There exists an absolute constant* $A > 2$ *such that*

$$p(q,a) \ll q^A;$$

moreover, A *and the constant implied are effectively computable.*

Remark 3.7. Subsequently, several people have computed admissible values of A (among others, Chen, Jutila and Graham) and currently the sharpest exponent is due to Heath-Brown and isfairly close to 2: $A = 5.5$.

Basically, Linnik's proof combines three main ingredients: the Hadamard/de la Vallée-Poussin zero free region, the quantitative form (due to Linnik) of the Deuring/Heilbronn phenomenon, see Lecture 1) and the following zero density estimate (also due to Linnik):

Theorem 3.9. *For* q *and* $T \geqslant 1$*, and* $\alpha \geqslant 1/2$ *one has*

$$\sum_{\chi(q)} N(\chi;\alpha,T) \ll (qT)^{c(1-\alpha)}, \tag{3.16}$$

where c *and the constant implied in* $\ll$ *are absolute and effectively computable.*

Some comments are in order concerning this last bound: by (3.15), (3.16) is non-trivial only when α is close to 1 (*i.e.* $> 1 - 1/c$). Other zero density estimates are non-trivial for any fixed $\alpha > 1/2$: for instance, by using (3.3), Bombieri proved ([**Bo**] Theorem. 18)

$$\sum_{\chi(q)} N(\chi;\alpha,T) \ll T^A q^{3\frac{1-\alpha}{2-\alpha}} \log^B q,$$

for some absolute A, B. This is much sharper than the Linnik density theorem when α is small but weaker when α is very close to 1 (*i.e.* $1-\alpha = O(1/\log qT)$), because of the $\log^B q$ factor. This latter feature and the exceptional zero repulsion phenomenon are critical to balance the influence of the exceptional zero.

Recently there have been various extensions of zero density estimates to general families of automorphic forms; see for instance [**KM1, KM3, Lu3**]. For example, one has the following analog of (3.16) [**KM3**]:

Theorem 3.10. *With the notations and assumptions of Theorem 3.5, one has, for* $\alpha > 3/4$ *and* $T \geqslant 1$*,*

$$\sum_{\pi\in\mathcal{F}} N(\pi;\alpha,T) \ll T^B |Q_{\mathcal{F}}|^{c(1-\alpha)},$$

[3] Note that the Bombieri/Vinogradov theorem implies the bound $p(q,a) \ll_\varepsilon q^{2+\varepsilon}$ for almost all $q \leqslant Q$ as $Q \to +\infty$.

where $Q_{\mathcal{F}} = \operatorname{Max}_{\pi \in \mathcal{F}} Q_\pi$*; here the exponents* B *and* c *and the constant implied in* $\ll$ *depend only on* d, θ *and on the ratio* $\frac{\log Q_{\mathcal{F}}}{\log |\mathcal{F}|}$.

Proof. Since our objective is to describe a basic application of the large sieve, we merely sketch a proof of the cruder bound

$$\sum_{\pi \in \mathcal{F}} N(\pi; \alpha, T) \ll_\varepsilon T^B Q_{\mathcal{F}}^{\varepsilon + c(1-\alpha)},$$

for any $\varepsilon > 0$; the replacement of ε by 0 uses much more refined techniques (namely mollification methods).

The first thing needed is a *detector of zeros*: a basic method to construct such a detector is to form a Dirichlet polynomial

$$D_N(\pi, s) := \sum_{1 \leqslant n \leqslant N} a_n \frac{\lambda_\pi(n)}{n^s}$$

that takes large values on the zeros of $L(\pi, s)$ contained in $R(\alpha, T)$ and such that the coefficients $(a_n)_{n \leqslant N}$ (which have to be independent of π) are not too large. There are several ways to construct efficient zero detecting polynomials (via mollifiers); since we do not seek the best possible results we will choose a very crude one. Pick $\rho \in R(\alpha, T)$ a zero of $L(\pi, s)$ and consider the contour integral

$$\frac{1}{2\pi i} \int_{(3)} L(\pi, s+\rho)\Gamma(s)\Big(\frac{N}{\log^2 N}\Big)^s ds = e^{-\frac{\log^2 N}{N}} + \sum_{2 \leqslant n} \frac{\lambda_\pi(n)}{n^\rho} e^{-\frac{n \log^2 N}{N}};$$

since $L(\pi, \rho) = 0$, one may shift the above contour to $\Re e s = 1/2 - \Re e \rho < 0$ without hitting any poles. By the convexity bound, one finds

$$e^{-\frac{\log^2 N}{N}} + \sum_{n \geqslant 2} \frac{\lambda_\pi(n)}{n^\rho} e^{-\frac{n \log^2 N}{N}} \ll_{\varepsilon, d, \theta} (Q_\pi)^\varepsilon (Q_\pi)^{1/4} N^{1/2-\alpha} = o_{\varepsilon, d, \theta}(1)$$

granted that N is sufficiently large (*i.e.* $N \geqslant (Q_\pi)^{1/(2\alpha-1)}$ is sufficient, for instance). By (1.23) one shows, by trivial estimation, that the tail $(n \geqslant N)$ becomes negligible when N is large and one deduces for such N's and for any zero in $R(\alpha, T)$ that

$$|D_N(\pi, \rho)| := \Big|\sum_{2 \leqslant n \leqslant N} e^{-\frac{n \log^2 N}{N}} \frac{\lambda_\pi(n)}{n^\rho}\Big| \geqslant 1/2; \tag{3.17}$$

thus $D_N(\pi, s)$ is our zero detecting linear form.

For the next step, one uses the large sieve inequality, or, more precisely, a variant of it. Given $\mathcal{Z} \subset R(\alpha, T)$, we say that $\mathcal{Z}$ is well-spaced if $\Im m(\rho - \rho') \geqslant 1$ whenever $\rho \neq \rho' \in \mathcal{Z}$. To each $\pi \in \mathcal{F}$, we associate some (possibly empty) well-spaced finite set $\mathcal{Z}(\pi) \subset R(\alpha, T)$, and given some Dirichlet polynomial $\sum_{n \leqslant N} a_n \lambda_\pi(n) n^{-s}$, we seek bounds for the averaged mean square of the values of this polynomial at $\rho \in \mathcal{Z}(\pi)$:

$$\sum_{\pi \in \mathcal{F}} \sum_{\rho \in \mathcal{Z}(\pi)} \Big|\sum_{n \leqslant N} \frac{a_n}{n^\rho}\Big|^2.$$

By considering the pairs (π, ρ), $\pi \in \mathcal{F}$, $\rho \in \mathcal{Z}(\pi)$ as a family $\mathcal{F}'$ one can give large sieve type inequalities for such sums, either by applying the duality principle or (as in [**Mon**]) by using an existing large sieve inequality for $\mathcal{F}$; for instance, under the assumptions of Theorem 3.5, one can show that

$$\sum_{\pi\in\mathcal{F}}\sum_{\rho\in\mathcal{Z}(\pi)}|\sum_{n\leqslant N}a_n\frac{\lambda_\pi(n)}{n^\rho}|^2\ll_{d,\varepsilon}(Q_{\mathcal{F}}N)^\varepsilon(N+|\mathcal{F}|^2Q_{\mathcal{F}}^d)\sum_n\frac{|a_n|^2}{n^{2\alpha}}. \tag{3.18}$$

Exercise. *Prove (3.18) by using the duality principle and the proof of Theorem 3.5.*

To deduce zero density estimates, we apply this bound to a zero detecting polynomial and to the set of zeros of π, denoted $\mathcal{Z}(\pi;\alpha,T)$, contained in $R(\alpha,T)$. Note that the well-spacedness assumption for $\mathcal{Z}(\pi)$ is not a severe restriction: by standard methods one can show that for any $T\in\mathbf{R}$, the number of zeros of $L(\pi,s)$ inside the critical strip and satisfying $|\Im m\rho-T|\leqslant 1$ is bounded by $O_d(\log Q_\pi(|T|+3))$. In particular, $\mathcal{Z}(\pi;\alpha,T)$ can be subdivided into $O_d(\log Q_\pi(|T|+3))$ well-spaced subsets. To keep control of the range on n, we need a dyadic decomposition of the zero detecting polynomial $D_N(\pi,s)$ given in (3.17): we divide the summation over the n variable into $O(\log N)$ dyadic subintervals of the form $[M,2M[$ for $M=2^\nu$, $0\leqslant\nu\ll\log N$, obtaining

$$|D_N(\pi,s)|^2\ll(\log N)^2\operatorname*{Max}_M|\sum_{\substack{n\sim M\\ n\leqslant N}}a_n\frac{\lambda_\pi(n)}{n^s}|^2.$$

We now apply (3.18) to the polynomials $\sum_{n\sim M}a_n\frac{\lambda_\pi(n)}{n^s}$ and obtain by (3.17),

$$\sum_{\pi\in\mathcal{F}}N(\pi;\alpha,T)\ll_{\varepsilon,d,\theta}(\log N)^2(\log Q_\pi(|T|+3))(Q_{\mathcal{F}}N)^\varepsilon(N+|\mathcal{F}|^2Q_{\mathcal{F}}^d)N^{1-2\alpha}$$
$$\ll_{\varepsilon,d,\theta}N^{2\varepsilon+2(1-\alpha)},$$

granted than $N\geqslant|\mathcal{F}|^2Q_{\mathcal{F}}^d$. We conclude by taking

$$N=\operatorname{Max}(Q_{\mathcal{F}}^{1/(2\alpha-1)},|\mathcal{F}|^2Q_{\mathcal{F}}^d).$$

□

Such estimates can be applied to fairly general families of L-functions.

An interesting family is the set of dihedral modular forms attached to class group characters of an imaginary quadratic order of discriminant $-D<0$. Using the above precise density estimate and an analog of the zero-repulsion phenomenon in this context, one can deduce a polynomial bound for the smallest Euler prime of discriminant $-D$ (a direct analog of Linnik's theorem above).

Theorem 3.11. *There is an absolute constant A (effectively computable) such that for any $D\geqslant 1$, the smallest prime (p_D say) of the form*

$$p=m^2+Dn^2,\ m,n\in\mathbf{Z}$$

is bounded by

$$p_D\ll D^A,$$

where the implied constant is absolute and effective.

Another type of application concerns the distribution of the values of automorphic L-functions at the edge of the critical strip (at $s=1$ for instance).

Indeed, a zero density estimate (like the one above) gives, for almost all $\pi\in\mathcal{F}$ a rather large zero-free region. Having such a good zero-free region at hand, one has for such π's, a good approximation of $L(\pi,1)^z$ by a very short Dirichlet polynomial (for z an arbitrary fixed complex number). One can use these approximations (together with the trivial bound for the π in the exceptional subset) and quasi-orthogonality relation for $\mathcal{F}$ (in small ranges of m,n) to derive an asymptotic formula for the averaged moment:

$$M(\mathcal{F},z) := \sum_{\pi\in\mathcal{F}} L(\pi,1)^z.$$

(The contribution of the exceptional π's for which the zero-free region is not large enough is bounded trivially.) Eventually, these computations show that the random variable $\pi\in\mathcal{F}\rightarrow L(\pi,1)$ has a limiting distribution function when $|\mathcal{F}|\rightarrow+\infty$, and in some cases, one can also study its extreme values. This kind of application was initiated by Barban in the case of (real) Dirichlet L-functions (improving and simplifying earlier results of Bateman, Chowla and Erdos). Later, the questions of the distribution function and of the extreme values of $L(\chi,1)$ were studied further by Montgomery/Vaughan and more recently in the exhaustive work of Granville/Soundararajan [**MoVa, GrS**].

For instance (thanks to $H_2(7/64)$) the zero density result above applies to symmetric square L-functions for families of GL_2-modular forms. Results on the distribution of the symmetric square $L(sym^2\pi,1)$ have been obtained by Luo [**Lu3**] and then by Royer [**Ro, Ro2**] by using zero density estimates, Kuznetzov/Petersson's formula and a combinatorial analysis. An example is the following result[**Ro, Ro2, RoWu**]:

Theorem 3.12. *For q prime, consider the set of primitive holomorphic forms with trivial nebentypus, $S_2^p(q,\chi_0)$, as a probability space endowed with the uniform measure. When $q\rightarrow+\infty$, the two random variables on*

$$f\rightarrow L(\mathrm{sym}^2 f,1)\ \mathrm{and} f\rightarrow L(\mathrm{sym}^2 f,1)^{-1}\in\mathbf{R}^\times_{>0}$$

admit a limiting distribution function. Moreover, for any q prime, sufficiently large, there exists f and f' in $S_2(q)$ such that

$$L(\mathrm{sym}^2 f,1)\gg(\log\log q)^3,\ L(\mathrm{sym}^2 f',1)\ll(\log\log q)^{-1},$$

where the implied constants are absolute.

Remark 3.8. This is essentially sharp: under **GRH**, one has, for any $f\in S_2^p(q,\chi_0)$,

$$(\log\log q)^{-1}\ll L(\mathrm{sym}^2 f,1)\ll(\log\log q)^3.$$

It is also possible to handle similarly an *algebraic* family of L-functions: consider

$$E_t: y^2=x^3+A(t)x+B(t),\ A,B\in\mathbf{Z}[t],$$

a non geometrically trivial family of elliptic curves over $\mathbf{Q}$. By the work of Wiles, Taylor/Wiles and Breuil/Conrad/Diamond/Taylor[**Wi, TW, BCDT**] each non-singular fiber is modular, hence (by Gelbart/Jacquet) the L-function of the symmetric square of each fiber has an analytic continuation to $\mathbf{C}$ with a possible pole at $s=1$ occurring only for the finitely many t's such that E_t is CM. One has the following ([**KM3**]):

Proposition 3.3. *there exists constants* $b, c > 0$ *such that for* $3/4 \leqslant \alpha < 1$

$$\sum_{\substack{x \in \mathbf{Z}, \Delta(x) \neq 0 \\ |x| \leqslant X}} N(\mathrm{sym}^2 E_x; \alpha, T) \ll T^b X^{c(1-\sigma)}$$

here b, c *and the implied constant depend only on* A, B.

One can use this to study the distribution of the special values $\{L(\mathrm{sym}^2 E_x, 1),\ |x| \leqslant X\}$; in this case the quasiorthogonality formulas come from the Lefschetz trace formula and Deligne's equidistribution theorems [**De3**], applied to the reduction of E_t modulo almost all primes and to its symmetric powers. Hence one can prove:

Proposition 3.4. *Assume also that* $A(t), B(t)$ *are coprime; as* $X \to +\infty$ *on the space* $[-X, X] \cap \mathbf{Z}$*, the random variables* $x \to L(\mathrm{sym}^2 E_x, 1)$ and $x \to L(\mathrm{sym}^2 E_x, 1)^{-1}$ *admit a limiting distribution function.*

Remark 3.9. The above example illustrates the fact that neither a "dense" family[4], nor a strong form of quasi-orthogonality is necessary to establish the existence of a distribution law for the values of L-functions near 1.

There are many other zero density estimates of different sorts, depending on the kind of L-functions or on the shape of the region in which zeros are considered. Also, many refinements are possible for more specific families for which stronger large sieve inequalities are available. We have not touched here on more advanced sieving techniques that would allow such results —like the mollification technique. For other kinds of zeros density estimates, the reader may look at [**Mon, Bo, HJ**] or [**KM2, HB-M**]. To date, the most striking zero density estimate is probably the following bound of Conrey/Soundararajan [**ConS**]:

Theorem. *The following bound holds:*

$$\sum_{|D| \leqslant X} N(\chi_D; 0, 0) \leqslant (c + o(1)) \sum_{|D| \leqslant X} 1,\ X \to +\infty,$$

where D *ranges over fundamental discriminants of quadratic fields,* χ_D *denotes the associated quadratic character,* $N(\chi_D; 0, 0)$ *the number of non-trivial (*i.e. *within the critical strip)* real *zeros of* $L(\chi_D, s)$*, and* c *is a constant strictly smaller than 1. In particular (since* $c < 1$*), there is a positive proportion of fundamental discriminants* D *such that* $L(\chi_D, s)$ *has* **no** *non-trivial* real *zeros !*

[4] In terms of the size of the conductors compared with the size of the family: for $|x| \leqslant X$, the conductor of $L(\mathrm{sym}^2 E_x, s)$ is typically $\approx X^{2 \deg \Delta}$ where Δ is the discriminant, while the number of such x is $\ll X$.

LECTURE 4
The Subconvexity Problem

In this lecture we describe the state of the art regarding the Subconvexity Problem (**ScP**) given in the first lecture; almost all of what is known about the problem concerns essentially L-functions of $GL_{1,\mathbf{Q}}$ and $GL_{2,\mathbf{Q}}$ automorphic forms. Remember that there are basically three aspects to consider regarding the size of an L-function associated to a modular form f, each of which one may let grow to $+\infty$: the level q_f, or the spectral parameter $|\mu_{f,1}|$ (which is essentially $(k_f-1)/2$ if f is holomorphic, or the spectral parameter $|it_f|$ if f is a Maass form of weight 0 or 1), or the height $|\Im m s|$ of the complex variable s. Today the analytic theory of L-functions is sufficiently well developed so that the following statement can be proved:

Subconvexity Theorem for GL_1, GL_2. *There exists an explicit constant $\delta > 0$ such that for any primitive Dirichlet character χ, any f and g primitive modular forms, and any s such that $\Re e s = 1/2$, one has:*

- *s-aspect: as $|s| \to +\infty$,*

$$L(\chi,s) \ll |s|^{1/4-\delta},\ L(f,s) \ll (|s|^2)^{1/4-\delta}, L(f\otimes g,s) \ll (|s|^4)^{1/4-\delta},$$

the implied constants depending on q_χ, Q_f and $Q_f Q_g$, respectively;

- *q-aspect: as $q_\chi, q_f \to +\infty$,*

$$L(\chi,s) \ll q_\chi^{1/4-\delta},\ L(f,s) \ll (q_f)^{1/4-\delta}, L(f\otimes g,s) \ll (q_f^2)^{1/4-\delta},$$

the implied constants depending on $|s|$, ($|s|, \mu_{f,1}$) and ($|s|, \mu_{f,1}$), respectively;

- *Spectral aspect: as $\mu_{f,1} \to +\infty$*

$$L(f,s) \ll (|\mu_{f,1}|^2)^{1/4-\delta}, L(f\otimes g,s) \ll (|\mu_{f,1}|^4)^{1/4-\delta},$$

the implied constants depending on ($|s|, q_f$) and ($|s|, q_f q_g, \mu_{f,1}$), respectively.

This statement (of which a few cases remain to be proved), has a long history and many contributors. The first subconvexity result is due to Weyl [**We**] for Riemann's ζ:

$$\zeta(s) \ll_\varepsilon |s|^{1/6+\varepsilon}; \tag{4.1}$$

and the proof of the s aspect for general Dirichlet L-functions is similar. However, very hard work has been done to improve Weyl's exponent of $1/6$; the current record (to date), $32/205$, is held by M. Huxley [**Hu**].

The next important progress is due to Burgess [**Bu**] who, in the 60's, solved the Subconvexity Problem for Dirichlet L-functions in the q-aspect: one has

$$L(\chi, s) \ll_\varepsilon |s|^A q_\chi^{3/16+\varepsilon}, \tag{4.2}$$

for some absolute constant A. The exponent $3/16$ resisted any improvement for the next 40 years, until the breakthrough of Conrey/Iwaniec [**ConI**] in the case of real characters: one has for any $\varepsilon > 0$

$$L(\chi, s) \ll_\varepsilon |s|^A q_\chi^{1/6+\varepsilon}. \tag{4.3}$$

It is remarkable that this exponent (which matches Weyl's original exponent for ζ) required the full power of the spectral theory of GL_2, positivity results for central values of automorphic L-functions due to Waldspurger and others, as well as the Weil conjectures over finite fields.

Things accelerated in the 80's with the first subconvexity bounds in the s-aspect for GL_2 L-functions, arising from the works of Good [**Go**] and Meurman [**Me**]. The spectral aspect for Hecke L-functions was treated by Iwaniec [**I3**] — who introduced on this occasion the fundamental *amplification method*— which we will describe below, and by Ivic, Jutila, and Peng [**Iv2, Ju4, Pe**]. The level aspect for GL_2 L-functions was treated by Duke/Friedlander/Iwaniec in a series of papers stretched over the 90's ([**DFI1**]...[**DFI8**]) culminating with the very difficult:

Theorem 4.1. *Let $f \in S_k^p(q, \chi, it)$. Assume that χ is primitive, then there exists a positive $\delta(\geqslant 1/24000)$ such that*

$$L(f, s) \ll_{k,it,s} q^{1/4-\delta}.$$

And recently the case of Rankin-Selberg L-functions has been treated, by Sarnak and Liu/Ye for the spectral aspect and Kowalski/Michel/Vanderkam and Michel for the level aspect [**Sa4, LY, KMV2, Mi**].

In the forthcoming sections, we present the various techniques that can be used to obtain the above subconvexity theorem in all cases. We only discuss the q-aspect here but we emphasize that these techniques apply equally to the other aspects with appropriate modifications.

We end this *état des lieux* with the case of more general number fields: several instances of the Subconvexity Problem for GL_2 L-functions have been solved by Petridis/Sarnak and Cogdell/Piatetsky-Shapiro/Sarnak [**PS, CoPSS**]. We highlight the latter work, as it is the key to the final solution of Hilbert's 11th problem on the representation of integers by quadratic forms in number fields[1] (see [**Co1**]):

Theorem 4.2. *Let F be a totally real number field, $\chi_\mathfrak{q}$ a ray class character of modulus $\mathfrak{q}$, and g a holomorphic Hilbert modular form over F. One has for any $\varepsilon > 0$*

$$L(g.\chi_\mathfrak{q}, s) \ll_{\varepsilon,g} N_{K/\mathbf{Q}}(\mathfrak{q})^{1/2-7/130+\varepsilon}$$

[1]Added in proofs: in a recent preprint dated of June 2005, A. Venkatesh, by a method very different of the one presented here, generalized the bound of Theorem 4.2 to arbitrary GL_2-automorphic forms over any number field and in fact generalized to a fixed arbitrary number field much of the subconvex bounds in the level aspect presented above [**Ve**]

4.1. Around Weyl's shift

The shifting method was introduced by Weyl in his proof of (4.1). More generally, this method provides non-trivial bounds for exponential sums of the form

$$\Sigma_f(N) = \sum_{m\in[1,N]} e(f(m)),$$

where f is a sufficiently regular function (*i.e.* well-approximable by polynomials for instance). Assuming first that f is a polynomial, one obtains by squaring the above sum

$$|\Sigma_f(N)|^2 = \sum_{m,n\in[1,N]} e(f(m)-f(n)) = \sum_{|l|<N} \sum_{\substack{0<n\leqslant N\\ 0<l+n\leqslant N}} e(f(l+n)-f(n)).$$

Now $f(l+x)-f(x)$ is a polynomial in x of degree reduced by one. One can then continue the above process until one reaches a sum for a polynomial of degree 1 to which one applies either the trivial bound or the geometric series summation; one can then get a non-trivial bound for the original sum $|\Sigma_f(N)|$. More generally, this method applies also to functions f that are well-approximable by polynomials, an example being $f(n) = -(it/2\pi)\log n$, which is the case occurring for ζ.

4.1.1. Burgess's method

Burgess's method is a variant of the Weyl shifting technique but in a purely arithmetic context and with significant differences[2]. In this section, we prove (4.2) when q is a prime number by an elegant variant of Burgess's original argument due to Friedlander/Iwaniec. By the approximate functional equation for $L(\chi,s)$ it is sufficient to give a bound for the character sum

$$S_V(\chi) = \sum_{n\geqslant 1} \frac{\chi(n)}{n^{1/2}} V(\frac{n}{q^{1/2}})$$

where $V(x)$ is a smooth function decaying rapidly as $x\to+\infty$. Integrating by part it is sufficient to bound the sum

$$S(\chi,N) = \sum_{n\leqslant N} \chi(n)$$

non-trivially for $q^{1/2-\delta}\leqslant N<q$, for some fixed $\delta>0$.

Theorem 4.3. *(Burgess) For all $r\geqslant 1$, one has for $N<q$*

$$S(\chi,N) \ll_r N^{1-1/r} q^{\frac{r+1}{4r^2}} (\log q)^{1+3/2r}.$$

Remark 4.1. Since $\frac{r+1}{4r^2} \sim \frac{1}{4r}$ when r is large, the above bound is non-trivial as long as $N \geqslant q^{1/4+\delta}$ for any fixed $\delta>0$. Hence $\chi(n)$ has oscillations as n ranges over very short intervals (of size up to $q^{1/4}$).

[2]A closer analog of Weyl's shift is the work of Graham-Ringrose on characters with highly factorized moduli [**GR**].

Proof. The method of Burgess consists of artificially adding summation points to the above sum —much in the spirit of a fugue where repetition of the theme with a slight shift creates the harmony. For the purpose of elegance, one introduces the piecewise linear function g that is 1 on $[1, N]$ and zero outside $[0, N+1]$. Note that

$$\hat{g}(y) = \int_{\mathbf{R}} g(x)e(-xy)dx \ll \min(N, y^{-1}, y^{-2}), \ \int_{\mathbf{R}} |\hat{g}(t)|dt \leqslant \log 3N.$$

One has (by the action of $\mathbf{Z}$ on itself),

$$S_\chi(N) = \sum_n \chi(n)g(n) = \sum_n \chi(n+ab)g(n+ab),$$

for all $a, b \geqslant 1$. Setting $AB = N$, $A, B \geqslant 1$, then

$$\begin{aligned}[A][B]S \quad &= \sum_{a\leqslant A}\sum_{b\leqslant B}\sum_n \chi(n+ab)g(n+ab)\\ &= \sum_{a\leqslant A}\chi(a)\sum_{|n|\leqslant N}\sum_{b\leqslant B}\chi(\overline{a}n+b)g(n+ab)\\ &\leqslant \sum_{a\leqslant A}\sum_{|n|\leqslant N}\Big|\sum_{b\leqslant B}\int \frac{1}{a}\hat{g}(\frac{t}{a})e(\frac{n}{a}t)\chi(\overline{a}n+b)e(bt)dt\Big|,\end{aligned}$$

by Fourier inversion. Hence

$$[A][B]S \ll \log N \sum_{a\leqslant A}\sum_{|n|\leqslant N}\Big|\sum_{b\leqslant B} e(bt_0)\chi(\overline{a}n+b)\Big|$$

for some $t_0 \in \mathbf{R}$. For $u \in \mathbf{F}_q$, we set

$$\nu(u) = |\{1 \leqslant a \leqslant A, |n| \leqslant X, \overline{a}n \equiv u(q)\}|,$$

and then

$$\begin{aligned}[A][B]S \quad &\ll \log X \sum_{u(q)}\nu(u)\Big|\sum_{b\leqslant B} e(bt_0)\chi(u+b)\Big|\\ &\leqslant \Big(\sum_{u(q)}\nu(u)\Big)^{1-1/r}\Big(\sum_{u(q)}\nu(u)^2\Big)^{1/2r}\Big(\sum_{u(q)}\Big|\sum_{b\leqslant B} e(bt_0)\chi(u+b)\Big|^{2r}\Big)^{1/2r}.\end{aligned}$$

One has

$$\sum_{u(q)}\nu(u) = [A]N,$$

and, for $AN < q/2$, one has

$$\sum_{u(q)}\nu(u)^2 = |\{1 \leqslant a_1, a_2 \leqslant A, |n_1|, |n_2| \leqslant N, a_2n_1 = a_1n_2\}| \ll AN(\log AN)^3,$$

since "$m_1 \equiv m_2(q)$" and "$m_1 = m_2$" are equivalent when $|m_1|, |m_2| < q/2$. Hence we conclude that

$$S \ll (\log q)^{1+\frac{3}{2r}}\frac{(AN)^{1-\frac{1}{2r}}}{AB}\Big(\sum_{\substack{b_1,\dots,b_r\leqslant B\\ b_1',\dots,b_r'\leqslant B}}\Big|\sum_{u(q)}\chi\big((u+b_1)\dots(u+b_r)\big)\overline{\chi}\big((u+b_1')\dots(u+b_r')\big)\Big|\Big)^{\frac{1}{2r}}.$$

At this point we use Weil's bound for algebraic exponential sums:

If $\prod_{i=1}^{r}\frac{u+b_i}{u+b_i'}$ *is not a k-th power (where we denote by k the order of χ), one has*

$$\Big|\sum_{u(q)}\chi\big((u+b_1)\dots(u+b_r)\big)\overline{\chi}\big((u+b_1')\dots(u+b_r')\big)\Big| \ll_r q^{1/2}.$$

The number of $(b_1,\dots,b_r')$ not satisfying this criterion is bounded by $\ll_r B^r$, and in this case we bound the sum trivially by q. Hence

$$\sum_{\substack{b_1,\dots,b_r\leqslant B\\ b_1',\dots,b_r'\leqslant B}}\Big|\sum_{u(q)}\chi((u+b_1)\dots)\Big| \ll B^r q + B^{2r}q^{1/2}.$$

The estimate of Theorem 4.3 follows by taking

$$A = Nq^{-1/2r},\ B = q^{1/2r}.$$

□

The proof of Theorem 4.2 follows by integration by parts with $r=4$.

4.1.2. Fouvry/Iwaniec's extension of Burgess's argument

The method of Burgess, though qualitatively very powerful, has the disadvantage of not being easily applicable to other types of arithmetic functions, and in particular, to more general cases of the Subconvexity Problem. However, recently Fouvry and Iwaniec have used a variant of Burgess's technique to solve the Subconvexity Problem for certain types of Grössencharacter L-functions attached to imaginary quadratic fields [**FoI3**]:

Theorem 4.4. *Let $D\equiv 3(4)$ be a prime number >3 and let ψ be a primitive character modulo the different of $\mathbf{Q}(\sqrt{-D})$ that takes values*

$$\psi((\alpha)) = \chi(\Re e\alpha)(\frac{\alpha}{|\alpha|})^r$$

on principal ideals (α), for χ a primitive Dirichlet character $\bmod D$, such that $\chi(-1)=(-1)^r$. Then for $\Re es = 1/2$,

$$L(\psi,s)\ll_{s,\varepsilon} D^{1/2-1/16+\varepsilon}.$$

4.2. The Amplification method

The amplification method is a tricky variant of the method of moments (see the first lecture); it was invented by Iwaniec to solve instances of the Subconvexity Problem and this is to date the most general technique for obtaining subconvex estimates: it may be used for the proof of all known cases of subconvexity (although with weaker exponents than the ones obtained by more *ad-hoc* methods).

4.2.1. Principles of Amplification

Given π_0, one wants a non-trivial bound for some given linear form in the Hecke eigenvalues,

$$\mathcal{L}(\pi_0,N) = \sum_{n\leqslant N} a_n\lambda_{\pi_0}(n),$$

the trivial bound being (at least when the sequence (a_n) is not lacunar)

$$\mathcal{L}(\pi_0, N) \ll_\varepsilon Q_{\pi_0}^\varepsilon (\sum_{n\leqslant N} |a_n|) \ll_\varepsilon Q_{\pi_0}^\varepsilon N^{1/2} (\sum_{n\leqslant N} |a_n|^2)^{1/2}.$$

To achieve this one may wish to apply the method of moments and average over an appropriate family $\mathcal{F}$ containing π_0. Taking a moment of order $2k$, the discussion at the end of the first lecture shows that, by multiplicativity of the $\lambda_\pi(n)$, this amounts essentially to bounding the mean square of some linear form of length N^k:

$$\frac{1}{|\mathcal{F}|} \sum_{\pi\in\mathcal{F}} |\sum_{n\leqslant N^k} b_n \lambda_\pi(n)|^2.$$

If N^k is not too large, one can expect that the harmonic analysis on $\mathcal{F}$ is sufficiently rich to imply quasiorthogonality in this range, namely that

$$\ll_\varepsilon Q_{\mathcal{F}}^\varepsilon \sum_{n\leqslant N^k} |b_n|^2 \ll_\varepsilon (Q_{\mathcal{F}} N)^\varepsilon (\sum_{n\leqslant N} |a_n|^2)^k. \tag{4.4}$$

We then get by positivity

$$\mathcal{L}(\pi_0, N) \ll_\varepsilon (Q_{\mathcal{F}} N)^\varepsilon |\mathcal{F}|^{1/2k} (\sum_{n\leqslant N} |a_n|^2)^{1/2}.$$

Hence, if we can take k large enough (namely $k > k_0 := \log|\mathcal{F}|/\log N$), we can improve over the trivial bound. Observe that all the large sieve type inequalities encountered so far give the bound (4.4) for k less than k_0 and no larger, thus giving no non-trivial bound for the original individual sum. This shouldn't be a surprise, since the large sieve inequalities are very general and make no use of the particular arithmetic properties of the sequence (a_n); we have experienced the strength of this generality in the last lecture and we discover its weakness now. In fact, there is no chance to go beyond the trivial bound without using some peculiar features of a_n: perhaps the most convincing example is to consider $a_n = \overline{\lambda}_{\pi_0}(n)$!

For the Subconvexity Problem, the fundamental property of (a_n) is *smoothness*, since a_n is essentially of the form

$$a_n = \frac{1}{\sqrt{n}} V(\frac{n}{\sqrt{Q_{\pi_0}}}),$$

for some smooth rapidly decaying function V. With such an extra assumption, one may hope to get a bound similar to (4.4) for $k > k_0$. Indeed, there are a priori two possibilities available for achieving this:

1. either one keeps $\mathcal{F}$ unchanged and one increases k beyond the critical moment k_0 and one tries to obtain (4.4);
2. one sticks to $k = k_0$ and one shortens $\mathcal{F}$ to decrease $|\mathcal{F}|$.

The first possibility is not very flexible since k has to be an integer so must be increased by at least one: then the problem of obtaining (4.4) for the resulting form of length N^{k_0+1} , even for smooth a_n, may well be beyond the capacities of the current technology. (However, note that if for all $\pi \in \mathcal{F}$ the linear form $L(\pi, N)$ is non-negative, there is still the possibility of increasing k by only half an integer and this may be short enough to yield sharp bounds [**ConI**].)

The second possibility of choosing a smaller subfamily of $\mathcal{F}$ containing π_0, may be available in some cases: for example, for Riemann's ζ, a subconvex bound for

$\zeta(1/2+it_0)$ follows from the upper bound

$$\int_{t_0-|t_0|^\alpha}^{t_0+|t_0|^\alpha} |\zeta(1/2+it)|^4 dt \ll_\varepsilon |t_0|^{\alpha+\varepsilon},$$

granted that one can choose α fixed < 1. However, this option is not available in some other interesting situations, such as the level aspect.

In the remaining cases, there is a third solution, *the amplification method*, which is a partial form of the first option: one keeps $k = k_0$, but one introduces an extra (short) linear form

$$A(\pi, L) := \sum_{l \leqslant L} c_l \lambda_\pi(n),$$

where the coefficients are *a priori* arbitrary. Because the (a_n) are smooth, one can hope, at least for (general) (c_l) with sufficiently small support, to get a quasiorthogonality type bound for the modified average

$$\frac{1}{|\mathcal{F}|}\sum_\pi |\sum_{n\leqslant N^{k_0}} b_n\lambda_\pi(n)|^2|\sum_{l\leqslant L} c_n\lambda_\pi(n)|^2 \ll_\varepsilon (Q_{\mathcal{F}}N)^\varepsilon(\sum_{n\leqslant N}|a_n|^2)^{k_0}\sum_{l\leqslant L}|c_l|^2. \tag{4.5}$$

Now suppose we can choose (c_l) satisfying

$$\sum_{l\leqslant L}|c_l|^2 \ll_\varepsilon (Q_{\pi_0})^\varepsilon L^{\alpha+\varepsilon}, \text{ and } \Big|\sum_{l\leqslant L}\lambda_{\pi_0}(l)c_l\Big| \gg_\varepsilon (Q_{\pi_0})^{-\varepsilon}L^{\alpha-\varepsilon}$$

for some $\alpha > 0$, one obtain by positivity that

$$\mathcal{L}(\pi_0, N) \ll_\varepsilon (Q_{\mathcal{F}}NL)^\varepsilon N^{1/2}L^{-\alpha}(\sum_{n\leqslant N}|a_n|^2)^{1/2},$$

which improves the trivial bound whenever $\log L \gg \log N$.

Such a sequence (c_l) is called an *amplifier* for π_0 since $A(\pi, L)$ evaluated at $\pi = \pi_0$ is large and thus amplifies the contribution of our preferred π_0 inside (4.5); the amplifier acts very similarly to the detector $(a_n^{\mathcal{E}})$ encountered in Section 3.1.1. The basic problem is then to construct an amplifier: when $\pi_0 = \chi_0$ is a Dirichlet character, one can choose

$$c_l = \overline{\chi}_0(l),\ l \leqslant L$$

and we clearly have an amplifier with $\alpha = 1$. When $\pi_0 = f_0 \in S^p(q, \chi, it)$ is a modular form, the most natural choice would be to take $c_l = \overline{\lambda}_{f_0}(l)$; but a difficulty crops up: since L will be a small power of Q_{f_0}, proving that

$$\sum_{l\leqslant L}|\lambda_{f_0}(l)|^2 \gg_\varepsilon (Q_{\pi_0})^{-\varepsilon}L^{1-\varepsilon}$$

would require a strong form of subconvexity for $L(f\otimes\overline{f}, s)$ (unless f_0 is dihedral). A way to get around this difficulty was found by Iwaniec and is based on the elementary identity

$$\lambda_f(p)^2 - \lambda_f(p^2) = \chi_f(p),$$

which shows that for unramified p, $\lambda_f(p)$ and $\lambda_f(p^2)$ cannot be simultaneously small. Consequently, one chooses

$$c_l = \begin{cases} -\overline{\chi}_{f_0}(p), \text{ for } n = p^2 \leqslant N,\ (p,q) = 1 \\ \lambda_{f_0}(p), \text{ for } n = p \leqslant N^{1/2},\ (p,q) = 1 \\ 0 \text{ else,} \end{cases} \tag{4.6}$$

so that for $p \leqslant L^{1/2}$, $(p,q)=1$ one has $c_{p^2}\lambda_{f_0}(p^2)+c_p\lambda_{f_0}(p)=1$. This builts a lacunar amplifier with $\alpha=1/2$.

Remark 4.2. The amplification method is completely "moral": despite appearances, there is no contradiction between the fact that $|A(\pi,L)|$ takes large values at $\pi=\pi_0$ (*i.e.* $A(\pi_0,L)\gg_\varepsilon L^{\alpha-\varepsilon}$) and the fact that on average the $|A(\pi,L)|^2$ are small (*i.e.* $\ll_\varepsilon L^\varepsilon$); indeed, an appropriate **GRH** for various Rankin-Selberg L-functions (among others, for $L(\pi\otimes\pi_0,s)$) shows that for any $\pi\neq\pi_0$, $|A(\pi,L)|\ll_{\varepsilon,d}(Q_{\mathcal{F}}L)^\varepsilon$ for any $\varepsilon>0$. But of course, by positivity, one never need to use any unproved hypotheses.

Remark 4.3. The reader should keep in mind that, the three basic strategies to attack the Subconvexity Problem are essentially formal. In particular, whichever method is chosen –higher moments, shortening or amplifying– the subconvexity bound will never come for free! The large sieve inequalities, or other general techniques may well bring us (rather easily) to the frontier between convexity and subconvexity but not further; the tough part begins now, as one tries to cross the border.

4.2.2. Improved bounds for $|S_1^p(q,\chi)|$

Our first application of the amplification method is an improvement over Duke's bound presented in the third lecture [**MiV**]:

Theorem 4.5. *For any $\varepsilon>0$, and any character of modulus q, one has*

$$|S_1^{Exotic}(q,\chi)|\ll_\varepsilon q^{6/7+\varepsilon}.$$

Remark 4.4. Remark 3.6 and this theorem imply for q squarefree

$$|S_1^{Exotic}(q)|\ll_\varepsilon q^{6/7+\varepsilon}.$$

Proof. We use the identification (2.6) between weight one holomorphic forms and weight one Maass forms with parameter at infinity $it=0$; we consider $\{u_j\}_{j\geqslant 1}$ an orthonormal basis of the space of weight 1 Maass forms containing $\{\frac{y^{1/2}f}{\langle f,f\rangle^{1/2}},\ f\in S_1^{Exotic}(q,\chi)\}$. The estimate (2.33) and positivity show that for any sequence $(c_l)_{l\leqslant L}$, one has

$$\sum_{f\in S_1^{Exotic}(q,\chi)}\frac{\mathcal{H}(0)}{\langle f,f\rangle}\Big|\sum_{\substack{l\leqslant L\\(l,q)=1}}c_l\lambda_f(l)\Big|^2=\sum_{f\in S_1^{Exotic}(q,\chi)}\frac{\mathcal{H}(0)}{\langle f,f\rangle}\Big|\sum_{\substack{l\leqslant L\\(l,q)=1}}c_l\sqrt{l}\rho_f(l)\Big|^2$$

$$\leqslant\sum_{j\geqslant 1}\mathcal{H}(t_j)\Big|\sum_{\substack{l\leqslant L\\(l,q)=1}}c_l\sqrt{l}\rho_j(l)\Big|^2+\sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}}\mathcal{H}(t)\Big|\sum_{\substack{l\leqslant L\\(l,q)=1}}c_l\sqrt{l}\rho_{\mathfrak{a}}(l,t)\Big|^2dt$$

$$\ll_\varepsilon\sum_{\substack{l\leqslant L\\(l,q)=1}}|c_l|^2+(qL)^\varepsilon\frac{L^{1/2}}{q}\Big(\sum_{\substack{l\leqslant L\\(l,q)=1}}|c_l|\Big)^2$$

If we take now (c_l) to be the coefficients given in the third lecture

$$c_l = \begin{cases} \overline{\chi}^6(p), \text{ for } l = p^{12} \leqslant L, \ (p,q) = 1 \\ -\overline{\chi}^4(p), \text{ for } l = p^8 \leqslant L^{8/12}, \ (p,q) = 1 \\ -\overline{\chi}(p), \text{ for } l = p^2 \leqslant L^{2/12}, \ (p,q) = 1 \\ 0 \text{ else},. \end{cases}$$

we obtain from (2.19), choosing $N = q^{2/(1+2\alpha)}$ (for $\alpha = 1/12$) that

$$|S^p(q,\chi)| \ll_\varepsilon q^{6/7+\varepsilon}.$$

□

Remark 4.5. This proof is an application of the amplification method where we have amplified an entire set (here $S_1^{Exotic}(q,\chi,0)$) inside a family (an orthogonal basis of $\mathcal{L}_1(q,\chi)$). One might be a bit surprised by the fact that this improvement comes from the embedding of our original family inside a much larger one; in fact, the embedding into a spectrally complete family makes Kuznetsov's formula available and the improvement arises because of the lacunarity of the amplifier: in our case the individual bounds provided by Kuznetsov's formula are stronger than the averaged bound provided by the large sieve.

4.2.3. Further Miscellaneous Applications

In fact, the amplification methods can be used in quite a variety of contexts to give non-trivial upper bounds for many arithmetic sums.

Another example concerns the L_∞-norm of L_2-normalized weight zero Maass forms with large Laplace eigenvalues on a modular curve. From its Fourier expansion, one can show that for $g \in S_0(q, it_g)$ that is L_2-normalized, one has

$$||g(z)||_\infty \ll_{\varepsilon,q} (1/4 + t_g^2)^{1/4+\varepsilon}.$$

In fact, this is a special case of a generic bound of Seeger and Sogge on the L_∞ norm of eigenfunctions of the Laplacian on a Riemann surface [**SS**]. When g is a Hecke-eigenform, Iwaniec/Sarnak [**IS1**] gave the first improvement over the generic bound using the amplification method:

Theorem 4.6. *Let $g \in S_0^p(q, it_g)$ be a primitive weight zero Maass form on $X_0(q)$ (with L_2-norm 1). Then for any $\varepsilon > 0$, one has*

$$||g(z)||_\infty \ll_{\varepsilon,q} (1/4 + |t_g|^2)^{1/4-1/24+\varepsilon}.$$

Finally, we would like to mention an unusual application of amplification to the rather technical question of bounding the following arithmetic quadratic form

$$\mathcal{B}(M,N) := \sum\sum_{\substack{m\leqslant M, n\leqslant N\\(m,n)=1}} a_m b_n e(a\frac{\overline{m}}{n})$$

(for some fixed $a \neq 0$); this sum is bounded trivially by $(MN)^{1/2}(\sum_m |a_m|^2 \sum_n |b_n|^2)^{1/2}$. By elementary techniques (Cauchy/Schwarz, the Polya/Vinogradov completion technique and the Weil bound for Kloosterman sums), it is possible to improve the trivial bound as long as $\log M/\log N$ is away from 1. In [**DFI5**], Duke/Friedlander/Iwaniec succeeded in improving on the trivial bound in the remaining range (and uniformly for $0 < |a| \leqslant MN$):

Theorem 4.7. *For any $\varepsilon > 0$, one has*

$$\mathcal{B}(M,N) \ll_\varepsilon (\sum_m |a_m|^2 \sum_n |b_n|^2)^{1/2}(|a| + MN)^{14/29}(M+N)^{1/58+\varepsilon}.$$

In fact this technical estimate is a key point in the proof of Theorem 4.1. But the most remarkable feature of this bound lies in its proof, which uses the amplification method in a completely unexpected way. Indeed, the bound follows by amplification of the contribution of the *trivial* (*a priori* invisible) character, in the averaged second moment

$$\mathcal{D}(M,N) = \sum_{m \leqslant M} \frac{1}{\varphi(m)} \sum_{\chi(m)} |\sum_{l \leqslant L} \chi(l) c_l|^2 | \sum_{\substack{n \leqslant N \\ (m,n)=1}} \chi(n) b_n e(a\frac{\overline{m}}{n})|^2$$

where χ ranges over the characters of moduli m !

4.3. Application to the Subconvexity Problem

In this section we explain how the amplification method can be used to solve Subconvexity Problems for GL_2 and $GL_2 \times GL_2$ L-functions.

We start by establishing a subconvex bound for twisted L-functions in the q-aspect, basically due to Duke/Friedlander/Iwaniec [**DFI1**] (but see also [**Ha2, Mi**]):

Theorem 4.8. *Let $g \in S_k^p(q', \chi', it')$, and let χ be a primitive Dirichlet character of modulus q. Then*

$$L(\chi.g, s) \ll_{\varepsilon,g,s} q^{1/2-1/22+\varepsilon},$$

where the implied constant depends polynomially on s and the parameters of g.

Unless otherwise specified, we assume that g is holomorphic.

4.3.1. The case of $L(g.\chi, s)$: reduction to the Shifted Convolution Problem

For simplicity, we give the argument at $s = 1/2$. By the approximate functional equation (1.30), we have

$$L(g.\chi, 1/2) = \sum_n \frac{\lambda_g(n)\chi(n)}{n^{1/2}} V(\frac{n}{q}) + w(g.\chi) \sum_n \frac{\overline{\lambda_g}(n)\overline{\chi}(n)}{n^{1/2}} \overline{V}(\frac{n}{q}),$$

were V is smooth with rapid decay. The crucial range being $n \sim q$, we will simply assume the V is supported on the interval $[1/2, 1]$; we wish to bound non-trivially the sum

$$\Sigma_V(g,\chi) := \sum_n \frac{\lambda_g(n)\chi(n)}{n^{1/2}} V(\frac{n}{q}) \ll_{k,\varepsilon} q^{1/2-\delta},$$

for some positive δ. We follow the method of [**DFI1**] (with a simplification of Sarnak): rewriting $\chi = \chi_0$ the character we started with, we take for $\mathcal{F}$ the family of Dirichlet characters of modulus q and form the averaged mean square

$$\sum_{\chi(q)} |\Sigma_V(g,\chi)|^2 | \sum_{l \leqslant L} \chi(l) c_l|^2 := \sum_{\chi(q)} | \sum_{m \leqslant Lq} \chi(m) b_m|^2 := Q((b_m)_{m \leqslant qL}),$$

say. We want to bound this sum essentially by the contribution coming from the diagonal terms *i.e.* to obtain the bound $\ll_{\varepsilon,g} (qL)^\varepsilon q \sum_{l \leqslant L} |c_l|^2$. Opening the square, one has

$$Q((b_m)) = \varphi(q) \sum_{\substack{m \equiv n(q) \\ (mn,q)=1}} \bar{b}_m b_n = \varphi(q) \sum_{\substack{a(q) \\ (a,q)=1}} |\sum_{m \equiv a(q)} b_m|^2 \leqslant \varphi(q) \sum_{a(q)} |\sum_{m \equiv a(q)} b_m|^2;$$

opening the square again and averaging over $a(\mathrm{mod} q)$, one obtains

$$Q((b_m)) \leqslant \varphi(q) \sum_{h \equiv 0(q)} \sum_{m-n=h} \overline{b_m} b_n = \varphi(q) \sum_{h \equiv 0(q)} \sum_{\ell_1, \ell_2 \leqslant L} \overline{c_{\ell_1}} c_{\ell_2} \Sigma_V(g; h, \ell_1, \ell_2)$$

where

$$\Sigma_V(g; h, \ell_1, \ell_2) = \sum_{\ell_1 m - \ell_2 n = h} \overline{\lambda_g(m)} \lambda_g(n) \frac{\overline{V}(\frac{m}{q}) V(\frac{n}{q})}{(mn)^{1/2}}.$$

This sum is much like a partial sum of Rankin-Selberg type but with an extra additive shift given by h. When $h = 0$, this is a true Rankin-Selberg type partial sum and one easily sees by (2.17) that

$$\varphi(q) \sum_{\ell_1, \ell_2 \leqslant L} c_{\ell_1} \bar{c}_{\ell_2} \Sigma_V(g; 0, \ell_1, \ell_2) \ll_{\varepsilon, k} q^{1+\varepsilon} \sum_{l \leqslant L} |c_l|^2.$$

Remark 4.6. In fact, one can get an asymptotic formula for this sum and see that this bound is nearly optimal (because of the lower bound in (2.16) for $L(\mathrm{sym}^2 g, 1)$).

When $h \neq 0$, the additive shift is non-trivial, and one expects some cancellation in the averaging of $\lambda_g(m)\lambda_g(n)$ (the trivial bound being $\ll_\varepsilon q^\varepsilon$). We will see in the next section that this is indeed the case; the outcome is

$$\Sigma_V(g; h, \ell_1, \ell_2) \ll_{\varepsilon, g} (qL)^\varepsilon L^{3/4} q^{-1/4}.$$

It follows that

$$\begin{aligned} |\Sigma_V(g, \chi_0)|^2 |\sum_{l \leqslant L} \chi_0(l) c_l|^2 &\leqslant \sum_{\chi(q)} |\Sigma_V(g, q)|^2 |\sum_{l \leqslant L} \chi(l) c_l|^2 \\ &\ll_{\varepsilon, g} q^{1+\varepsilon} \sum_{l \leqslant L} |c_l|^2 + q^{3/4+\varepsilon} L^{5/2+1/4} \sum_{l \leqslant L} |c_l|^2; \end{aligned}$$

choosing $c_l = \overline{\chi}_0(l)$ (in that case, $\alpha = 1$) and optimizing L, one obtains

$$\Sigma_V(g, \chi_0) \ll_{\varepsilon, k} q^{1/2 - 1/22 + \varepsilon}.$$

Remark 4.7. The original exponent $1/2 - 1/22 = 0.4545...$ of [**DFI1**] is not the best possible; in fact we will see in the next section that its size is directly dependent on $H_2(\theta)$. In particular, the truth of $H_2(7/64)$ due to Kim/Sarnak yields the slightly better subconvexity exponent $1/2 - 25/448 < 0.442$.

4.4. The Shifted Convolution Problem

Consider g a primitive form, $h \neq 0$, ℓ_1, ℓ_2 two coprime integers and $W(x, y)$ a bounded, smooth, compactly supported function in $[X, 2X] \times [Y, 2Y]$ for some $X, Y \geqslant 1/2$; to fix ideas we suppose also that W has well controlled derivatives, *i.e.*

$$\forall i, j \geqslant 0,\ x^i y^j W^{(i,j)}(x, y) \ll_{i,j} 1. \tag{4.7}$$

One considers the following shifted convolution sums:

$$\Sigma_W(g, \ell_1, \ell_2, h) := \sum_{\ell_1 m - \ell_2 n = h} \overline{\lambda_g}(m) \lambda_g(n) W(\ell_1 m, \ell_2 n), \tag{4.8}$$

which, using (2.17), are easily bounded by

$$\Sigma_W(g,\ell_1,\ell_2,h) \ll_\varepsilon (Q_g XY)^\varepsilon \operatorname{Max}(X/\ell_1, Y/\ell_2).$$

The **Shifted Convolution Problem (SCP)** is to improve the trivial bound for relevant choices of M,N,ℓ; what makes it possible is the non-vanishing of h, so that $\overline{\lambda_g}(m)$ do not conspire with $\lambda_g(n) = \lambda_g(\frac{\ell_1 m-h}{\ell_2})$. This problem has a long history that which goes back to Ingham with the case of the Eisenstein series $g(z) = E'(z,1/2)$ and $\lambda_g(n) = \tau(n)$ the divisor function (this instance of the **SCP** is the classical *Shifted Divisor Problem*). The strategy to solve the **SCP** is to "smooth out" the condition $\ell_1 m - \ell_2 n = h$ in Σ_W in order to "separate" the variable m from n and to estimate the resulting sums. In this section we describe two techniques for achieving this goal. These methods will provide a non-trivial bound uniformly for $\operatorname{Max}(X/\ell_1, Y/\ell_2)$ larger than some fixed (relatively large) powers of Q_g and $\ell_1\ell_2$, respectively. This will be sufficient for the present subconvexity cases. However, the question of the uniformity of these bounds with respect to the parameters of g is very interesting and will be discussed later on.

Although the Shifted Convolution Problem looks rather technical, it is ubiquitous in the resolution of the Subconvexity Problem for GL_1 and GL_2 L-functions. Indeed, all the cases of subconvexity that are presented in the Subconvexity theorem at the beginning of this lecture can be reduced to the resolution of the **SCP** for some appropriate modular form g: for instance the identity

$$L(\chi,s)^2 = \sum_{n\geqslant 1} \frac{\chi(n)\tau(n)}{n^s} = L(\chi\times E'(z,1/2),s)$$

shows that the Subconvexity Problem for Dirichlet L-function could follow from a variant of the method presented in the previous section and the solution of an instance of the **SCP**.

4.4.1. The δ-symbol Method

The δ-symbol method was developed in [**DFI1, DFI3**] as variant of the circle method. Its main purpose is to express $\delta(n)$, the Dirac symbol at 0 (restricted to the integers n in some given range: $|n| \leqslant N$), in terms of additive characters. Of course one could use the orthogonality relation

$$\delta_{n\equiv 0(q)} = \frac{1}{q}\sum_{k(q)} e(\frac{nk}{q})$$

which provides an adequate expression as long as $2N \leqslant q$; however, the conductor q of the additive characters is large compared with N and this constitute a severe limitation for many analytic applications. By exploiting the symmetry of the set of divisors[3] of n, the δ-symbol method is capable of providing an expression for $\delta(n)$ in terms of additive characters of much smaller moduli (*i.e.* basically of size $\leqslant \sqrt{n}$). The basic identity is the following: one starts with $\omega(x)$ a smooth, compactly supported, even function satisfying

$$\omega(0) = 0,\ \sum_{r\geqslant 1}\omega(r) = 1;$$

[3] if $d|n$ then $(n/d)|n$

for n an integer, the symmetry of the set of its divisors implies that

$$\delta(n)=\sum_{k|n}[\omega(k)-\omega(\frac{n}{k})].$$

We may now detect the condition $k|n$ by means of additive characters:

$$\delta(n)=\sum_{k\geqslant 1}\frac{1}{k}(\omega(k)-\omega(\frac{n}{k}))\sum_{a(k)}e(\frac{an}{k})=\sum_{r\geqslant 1}\Delta_r(n)\sum_{\substack{a(r)\\(a,r)=1}}e(\frac{an}{r}) \tag{4.9}$$

where

$$\Delta_r(n)=\sum_{k\geqslant 1}\frac{1}{kr}(\omega(kr)-\omega(\frac{n}{kr})).$$

In practice the formula is applied to $|n|\leqslant U/2$ (say), for $\omega(x)$ supported on $R/2<|x|<R$ (say) with derivatives satisfying $\omega^{(j)}(x)\ll_j R^{-j-1}$; if R is chosen equal to $U^{1/2}$, one sees that $\Delta_r(n)$ vanishes unless $1\leqslant r\leqslant \mathrm{Max}(R,U/R)=U^{1/2}$. This yields a representation of $\delta(n)$ for the integers $|n|\leqslant U/2$ in terms of additive characters of moduli $r\leqslant U^{1/2}$; moreover, one has (by the Euler-MacLaurin formula) good control on the derivatives of Δ_r:

Lemma 4.4.1. *For $j\geqslant 1$ and $|u|\leqslant U/2$ we have*

$$\Delta_r^{(j)}(x)\ll_j\frac{1}{(rR+|x|)(rR)^{j+1}}.$$

We are now in a position to describe how Duke/Friedlander/Iwaniec used the identity (4.9) to solve the Shifted Convolution Problem [**DFI1, DFI3**]. Applying (4.9) to the Dirac symbol $\delta(\ell_1 m-\ell_2 n-h)$ in (4.8), one gets rid of the constraint $\ell_1 m-\ell_2 n-h=0$, which was our primary objective. However, for technical reasons it is useful (see (4.12)) to keep partial track of this condition, namely that $\ell_1 m-\ell_2 n-h$ cannot be large. This is done by introducing a localization factor $\phi(\ell_1 m-\ell_2 n-h)$ in $\Sigma_W(g,\ell_1,\ell_2,h)$, where ϕ is a smooth function compactly supported in $[-U/2,U/2]$, satisfying $\phi^{(i)}(x)\ll_i U^{-i}$ and $\phi(0)=1$; the last condition implies that the sum $\Sigma_W(g,\ell_1,\ell_2,h)$ remains unchanged. Applying (4.9) one obtains

$$\Sigma_W(g,\ell_1,\ell_2,h)=\sum_{1\leqslant r\leqslant R}\sum_{a(r),(a,r)=1}e(\frac{-ah}{r})\sum_{m,n}\overline{\lambda_g}(m)\lambda_g(n)e(\frac{\ell_1 ma-\ell_2 na}{r})E_r(m,n,h),$$

where

$$E_r(x,y,h)=W(\ell_1 x,\ell_2 y)\phi(\ell_1 x-\ell_2 y-h)\Delta_r(\ell_1 x-\ell_2 y-h).$$

It turns out that the best choice for U is $U=\min(X,Y)=X$ (say) and consequently one has $R=U^{1/2}=X^{1/2}$, which implies that

$$\frac{\partial^i\partial^i}{\partial^i x\partial^j y}E_r(x,y,h)\ll_{i,j}\frac{1}{(rR+|\ell_1 x-\ell_2 y-h|)}\frac{\ell_1^i\ell_2^j}{X^{i+j}}(\frac{R}{r})^{i+j}. \tag{4.10}$$

We apply the Voronoi summation formula of the second lecture to both variables m and n. For simplicity, we assume that g has level[4] 1: by Lemma 2.3.1, the

[4]For non-trivial level, some more (purely technical) work is necessary; for this we refer to the papers [**KMV2, Mi**]; we also mention the paper of Harcos [**Ha2**] for a somewhat different treatment based on a variant of the δ-symbol due to Jutila.

sum is transformed into
(4.11)
$$\Sigma_W(g,\ell_1,\ell_2,h)=\sum_{r\leqslant R}\frac{(\ell_1\ell_2,r)}{r^2}\sum_{m,n}\overline{\lambda_g}(m)\lambda_g(n)S(-l_1'm+l_2'n,-h;r)I_r(m,n,h),$$

with

$$I_r(m,n,h)=(2\pi i^k)^2\int\int E_r(x,y,h)J_{k-1}(\frac{4\pi\sqrt{mx}}{r/(\ell_1,r)})J_{k-1}(\frac{4\pi\sqrt{nx}}{r/(\ell_2,r)}); \quad (4.12)$$

here we have set $l_i'=l_i/(r,l_i),\ i=1,2$. By integrating the Bessel function by parts several times (see (2.22) (2.23)), one shows using (4.10) that $I_r(m,n,h)$ is very small, unless

$$m<l_1'\frac{X}{R^{2-\varepsilon}}=l_1'X^{\varepsilon/2},n<l_2'\frac{Y}{R^{2-\varepsilon}}=l_1'X^{\varepsilon/2}\frac{Y}{X}.$$

In the remaining range, the bound (4.10) and a trivial estimate show that

$$I_r(m,n,h)\ll\frac{X(\ell_1\ell_2,r)}{\ell_1\ell_2}\log U.$$

(In fact, ϕ was introduced precisely to ensure this last estimate.)

The above considerations together with Weil's bound for Kloosterman sums

$$|S(a,b;r)|\leqslant\tau(r)(a,b,r)^{1/2}r^{1/2}$$

yield

$$\Sigma_W(g,\ell_1,\ell_2,h)\ll_{\varepsilon,g}(XY)^{\varepsilon}X^{1/4}Y^{1/2}.$$

This improves on the trivial bound as soon as $\ell_1+\ell_2\leqslant X^{3/4}/Y^{1/2}$. □

Remark 4.8. This method applies equally to Maass forms and to Eisenstein series. For instance, it applies to the Eisenstein series $g(z)=E'(z,1/2)$ of the second lecture, the main difference in that case being that $S_W(g,\dots)$ has a main term coming from the constant term of the Eisenstein series (see [**DFI3**]).

Remark 4.9. Note that any non-trivial bound for Kloosterman sums (*i.e.*

$$|S(a,b;r)|\ll(a,b,r)^{1/2}r^{\theta}$$

for some $\theta<1$ and any $\varepsilon>0$) is sufficient to give a non-trivial (although weaker) bound for the **SCP** (for example, Kloosterman's original bound with $\theta=3/4$). It is interesting to note that Kloosterman derived his bound by elementary methods that used families and moments (see [**I5**] Chap. 3).

On the other hand, it is possible to improve the given bound for the **SCP** beyond Weil's bound with more advanced techniques: starting with (4.11), it is possible to use Kuznetzov's trace formula *backwards* so as to transform (4.11) into sums of Fourier coefficients of weight 0 Maass forms of level $\ell_1\ell_2$ (taken at cusps not necessarily equal to ∞); then one can take advantage of the best known bounds for the Fourier coefficients (which come from $H_2(\theta)$) to improve the results. However, as we shall see, there is a more direct way to make the connection between shifted convolutions sums and Maass forms.

4.4.1.1. *Other Applications of the δ-symbol Method.* The delta-symbol method is pretty elementary, and as such can be used for a variety of problems:

- it was used by Duke/Iwaniec to provide non-trivial bounds for Fourier coefficients of Maass forms (although weaker than the current bounds), as well as to prove analytic continuation beyond the domain of absolute convergence of certain L-series [**DuI, DuI2, DuI3**].
- Another interesting example is the following result of N. Pitt [**Pi**], which solves the **SCP** for a GL_2 form with the simplest GL_3 Eisenstein series:

  Theorem 4.9. *Let $g \in S_k^p(1)$ be a primitive holomorphic cusp form. Then for $X > 1$ and $1 \leqslant \ell < X^{1/24}$, for any $\varepsilon > 0$ one has*

  $$\sum_{1 \leqslant n \leqslant X} \lambda_g(n)\tau_3(\ell n - 1) \ll_{\varepsilon,g} X^{71/72+\varepsilon}.$$

  The proof of this result is somewhat difficult and uses among other things good bounds for sums of Kloosterman sums, following deep results in the theory of exponential sums coming from the works of Deligne, Bombieri, Adolphson-Sperber and Katz.
- This method, combined with (the deeper) Theorem 4.7, enables one to evaluate asymptotically quite general shifted convolution type sums, *i.e.* sums of the form

  $$\sum\sum_{m_1n_1-m_2n_2=h} \alpha_{m_1}\beta_{m_2} V(\frac{m_1}{M_1}, \frac{n_1}{N_1}, \frac{m_2}{M_2}, \frac{n_2}{N_2});$$

  here $(\alpha_{m_1}), (\beta_{m_2})$ denotes two arbitrary bounded sequences of complex numbers; V is a smooth function compactly supported on $[-1,1]^4$, where the M_i, N_j close to each other (in the logarithmic scale), and (the crucial point) where the smooth variables n_1, n_2 can be taken a little smaller than the non-smooth ones. Such sums lie at the heart of the proof of Theorem 4.1.

4.4.2. The Shifted Convolution Problem via Spectral Methods

In [**Se2**], Selberg proposed an approach to handle the **SCP** using the spectral decomposition of a *product* of two modular forms; his suggestion was pursued successfully in very specific cases in works of Good, Jutila, Motohashi, and others [**Go, TV, Ju2, Ju3, Mot, JM**], but the first general treatment occurred in the work of Sarnak, who applied this technique to solve the **SCP** for Rankin-Selberg L-functions in the spectral aspect[**Sa3, Sa4**]:

Theorem 4.10. [**Sa4**] *Let $g \in S_{k'}(q', \chi', it')$ be a fixed cusp form and $f \in S_k^p(q, \chi)$ a primitive holomorphic form of weight k. Then*

$$L(f \otimes g, s) \ll_{\varepsilon,s,g,q_f} k^{1-7/165}.$$

We present this approach for $g \in S_k(q', \chi)$ holomorphic. It is then easy to see that a non-trivial bound for $\Sigma_W(g, \ell_1, \ell_2, h)$ follows from the following: for $\Re e s > 1$, we set

$$D(g, s; \ell_1, \ell_2, h) := \sum_{\ell_1 m - \ell_2 n = h} \overline{\lambda_g}(m)\lambda_g(n)(\frac{\sqrt{\ell_1\ell_2 mn}}{\ell_1 m + \ell_2 n})^{k-1}(\ell_1 m + \ell_2 n)^{-s}.$$

This series is absolutely convergent for $\Re e s > 1$ and analytic there; the next theorem show that it can be analytically continued beyond this natural barrier.

Theorem 4.11. *Assume $H_2(\theta)$ holds. For any $\varepsilon > 0$, $D(g, s; \ell_1, \ell_2, h)$ extends holomorphically to $\Re e s \geqslant 1/2 + \theta + \varepsilon$ and in this region, it satisfies the upper bound*

$$D(g, s; \ell_1, \ell_2, h) \ll_{\varepsilon, g, \theta_1} (|h|\ell_1\ell_2)^{\varepsilon}(\ell_1\ell_2)^{1/2}|h|^{1/2+\theta-\sigma}(1 + |t|)^3,$$

where $s = \sigma + it$ and the implied constant depends only on ε and g.

To pass from this bound to one for $\Sigma_W(g, \ell_1, \ell_2, h)$, we note that

$$\Sigma_W(g, \ell_1, \ell_2, h) = \frac{1}{2\pi i}\int_{(2)} D(g, s; \ell_1, \ell_2, h)\widehat{W}(h, s)ds \tag{4.13}$$

with

$$\widehat{W}(h, s) = \int_0^{+\infty} W(\frac{u + h}{2}, \frac{u - h}{2})(\frac{4u^2}{u^2 - h^2})^{\frac{k-1}{2}} u^s \frac{du}{u}.$$

Shifting the contour to $\Re e s = 1/2 + \theta + \varepsilon$, we obtain, after several integrations by parts and using (2.22) and (2.23), the bound

$$\Sigma_W(g, \ell_1, \ell_2, h) \ll_{\varepsilon, g} (Y\ell_1\ell_2)^{2\varepsilon}(\ell_1\ell_2)^{1/2}(\frac{Y}{X})^{\frac{k-1}{2}+5}Y^{1/2+\theta}. \tag{4.14}$$

Remark 4.10. Note that $H_2(\theta)$ for any $\theta < 1/2$ would be sufficient to solve the **SCP**; compare with Remark 4.9. However, improvements on $H_2(\theta)$ directly yield improved bounds for the **SCP**, hence for cases of the **ScP**. In particular $H_2(7/64)$ yields the subconvex exponent $\frac{1}{2} - \frac{7}{130}$ in place of $\frac{1}{2} - \frac{1}{22}$ in Theorem 4.8.

Proof. (**of Thm 4.11**) The proof follows from an appropriate integral representation of $D(g, s; \ell_1, \ell_2, h)$. Setting $N = q'\ell_1\ell_2$, one expresses $D(g, s; \ell_1, \ell_2, h)$ in terms of the integral of the $\Gamma_0(N)$ invariant function

$$V(z) = \Im m(\ell_1 z)^{\frac{k}{2}}\overline{g}(\ell_1 z)\Im m(\ell_2 z)^{\frac{k}{2}}g(\ell_2 z)$$

against an appropriate Poincaré series

$$U_h(z, s) = \sum_{\gamma\in\Gamma_0(N)_\infty\backslash\Gamma_0(N)} y(\gamma z)^s e(hx(\gamma z));$$

more precisely, one has by the unfolding method

$$I = \langle U_h, \overline{V}\rangle = \int_{\Gamma_0(N)\backslash\mathbf{H}} U_h(z, s)V(z)\frac{dxdy}{y^2} = \frac{\Gamma(s + k - 1)(\ell_1\ell_2)^{\frac{1}{2}}}{(2\pi)^{s+k-1}}D(g, s; h). \tag{4.15}$$

On the other hand U_h and V can be decomposed spectrally (at least formally) over a orthonormal basis for $\Gamma_0(N)$ and by Parseval we have

$$I = \sum_{j\geqslant 0}\langle U_h(., s), u_j\rangle\langle u_j, \overline{V}\rangle + \frac{1}{4\pi}\sum_{\mathfrak{a}}\int_0^\infty \langle U_h(., s), E_{\mathfrak{a}}(., \tfrac{1}{2} + it)\rangle\langle E_{\mathfrak{a}}(., \tfrac{1}{2} + it), \overline{V}\rangle dt.$$

Finally by the unfolding method one has

$$\langle U_h(., s), u_j\rangle = \frac{2^{s-1}\overline{\rho_j}(h)}{|h|^{s-1}}\Gamma(\frac{s - \frac{1}{2} + it_j}{2})\Gamma(\frac{s - \frac{1}{2} - it_j}{2})$$

so that

$$(4.16)\quad I=\sum_{j\geqslant 1}\frac{2^{s-1}\overline{\rho_j}(h)}{|h|^{s-1}}\Gamma(\frac{s-\frac{1}{2}+it_j}{2})\Gamma(\frac{s-\frac{1}{2}-it_j}{2})\langle u_j,\overline{V}\rangle+\text{Eisenstein Contr.}$$

with

$$\text{Eisenstein Contr.}=\frac{1}{4\pi}\sum_{\mathfrak{a}}\int_0^\infty\frac{2^{s-1}\overline{\rho_{\mathfrak{a},t}}(h)}{|h|^{s-1}}\Gamma(\frac{s-\frac{1}{2}+it}{2})\Gamma(\frac{s-\frac{1}{2}-it}{2})\langle E_{\mathfrak{a}}(.,\tfrac{1}{2}+it),\overline{V}\rangle dt.$$

From this decomposition one sees that $D(g,s;\ell_1,\ell_2,h)$ has analytic continuation up to $\Re es>\Re e(s-\frac{1}{2}\pm it_j)\geqslant\frac{1}{2}+\theta$.

The bound for $D(g,s;\ell_1,\ell_2,h)$ follows now by bounding each term of the spectral decomposition (4.16).

4.4.2.1. *Bounding the Fourier coefficients.* We start with $\rho_j(h)$: if u_j equals $f_j/\langle f_j,f_j\rangle^{1/2}$ for some primitive Hecke-eigenform f_j, one has $|h|^{1/2}\rho_j(h)=\pm\lambda_{f_j}(|h|)/\langle f_j,f_j\rangle^{1/2}$; hence $H_2(\theta)$ and the estimate (2.16) give

$$(4.17)\qquad \rho_j(h)\ll_\varepsilon\frac{(|h|N(1+|t_j|)))^\varepsilon}{\sqrt{N}}\mathrm{ch}(\frac{\pi t_j}{2})|h|^{\theta-1/2}$$

A similar bound for all u_j would follow from producing an explicit orthonormal basis, constructed from old and new forms as in [**ILS**] (for squarefree N). However, for general levels, the corresponding computations can be quite messy. In fact, we need this bound only on average over an orthonormal Hecke eigenbasis $B_0(N)=\{u_j\}_{j\geqslant 0}$, (4.17): one has (see [**Mi**]), for any $\varepsilon>0$,

$$(4.18)\qquad \sum_{\substack{u_j\in B_0(N)\\|t_j|\leqslant T}}\frac{|h||\rho_j(h)|^2}{\mathrm{ch}(\pi t_j)}\ll_\varepsilon(|h|NT)^\varepsilon T^2|h|^{2\theta}.$$

4.4.2.2. *Bounding the triple products $\langle u_j,\overline{V}\rangle$.* It remains to bound the integral of the triple product $\langle u_j,\overline{V}\rangle$ and, more precisely, one needs a bound with an exponential decay like $\exp(-\pi|t_j|/2)$ (as $|t_j|\to+\infty$) in order to compensate for the factor $\mathrm{ch}(\frac{\pi t_j}{2})$ from (4.17). This is a key point and it remained, until recently, the main obstacle to a general solution of the **SCP** along the lines proposed by Selberg; the exponential decay had been obtained in several specific cases (see [**Go, TV, Ju2, Ju3**]) but the first truly general treatment was given in [**Sa3**] and made explicit further in [**Sa4**]. For any u_j, one has

$$(4.19)\qquad \langle u_j,\overline{V}\rangle\ll_k||y^{k/2}g||_\infty^2\sqrt{N}(1+|t_j|)^{k+1}e^{-\frac{\pi}{2}|t_j|},$$

and the same bound holds for $\langle E_{\mathfrak{a}}(.,\frac{1}{2}+it),\overline{V}\rangle$ with t_j replaced by t. Using (4.19) together with Weyl's law,

$$(4.20)\qquad \sum_{|t_j|\leqslant T}1=\frac{\mathrm{vol}(X_0(N))}{4\pi}T^2+O_\varepsilon((1+T)^{1+\varepsilon}),$$

one obtains the averaged bound for $T \geqslant 1$:

$$(4.21)\quad \sum_{|t_j|\leqslant T} |\langle u_j, \overline{V}\rangle|^2 e^{\pi|t_j|} + \frac{1}{4\pi}\sum_{\mathfrak{a}}\int_0^T |\langle E_{\mathfrak{a}}(., \tfrac{1}{2}+it), \overline{V}\rangle|^2 e^{\pi|t|}dt \ll_{k,\varepsilon} (NT)^{\varepsilon}||y^{k/2}g||_{\infty}^4 N^2 T^{2k+4}.$$

Applying Stirling's formula, Cauchy-Schwarz, (4.18) and (4.21) in (4.16), one concludes the proof of Theorem 4.11. □

4.4.3. The SCP for Maass forms and for Eisenstein series

A similar analysis can be carried out when g is a Maass cusp form or an Eisenstein series. When g is a primitive Maass forms of weight zero with Laplace eigenvalue $1/4+t^2$, one defines (see [**Sa4**]):

$$D(g,s;\ell_1,\ell_2,h) = \sum_{\ell_1 m-\ell_2 n=h} \frac{\sqrt{mn}\overline{\rho}_g(m)\rho_g(n)}{(|\ell_1 m|+|\ell_2 n|)^s}\Big(\frac{\sqrt{|\ell_1\ell_2 mn|}}{|\ell_1 m|+|\ell_2 n|}\Big)^{2it} F_{2,1}\Big(\frac{s+2it}{2},\frac{1+2it}{2},\frac{s+1}{2};\Big(\frac{|\ell_1 m|-|\ell_2 n|}{|\ell_1 m|+|\ell_2 n|}\Big)^2\Big).$$

Theorem 4.12. *Assume $H_2(\theta)$ holds for all Maass cusp forms of weight zero with trivial nebentypus. For any $\varepsilon>0$, $D(g,s;\ell_1,\ell_2,h)$ extends holomorphically to $\sigma := \Re e s \geqslant 1/2+\theta+\varepsilon$ and in this region it satisfies the upper bound*

$$D(g,s;\ell_1,\ell_2,h) \ll_{\varepsilon,g,\theta_1} (|h\ell_1\ell_2|)^{2\varepsilon}(|\ell_1\ell_2|)^{1/2}h^{1/2+\theta-\sigma}(1+|t|)^3,$$

where $s=\sigma+it$ and the implied constant depends only on ε and g only.

For g an Eisenstein series, Theorems 4.11 and 4.12 are still valid, with the main difference being that $D(g,s;\ell_1,\ell_2,h)$ acquires a pole at $s=1$ of order at most 3: this pole comes from the constant terms of the Eisenstein series and its exact value can be computed by a (standard) regularization process, which, for example, is carried out in the paper of Tatakhjan and Vinogradov [**TV**].

Remark 4.11. However, let us mention that for g non-holomorphic, there is still a difficulty in relating the estimate of Thm. 4.12 to the original shifted convolution problem (i.e. in replacing

$$\frac{1}{(|\ell_1 m|+|\ell_2 n|)^s}\Big(\frac{\sqrt{|\ell_1\ell_2 mn|}}{|\ell_1 m|+|\ell_2 n|}\Big)^{2it} F_{2,1}\Big(\frac{s+2it}{2},\frac{1+2it}{2},\frac{s+1}{2};\Big(\frac{|\ell_1 m|-|\ell_2 n|}{|\ell_1 m|+|\ell_2 n|}\Big)^2\Big)$$

by a more general test function $W(\ell_1 m,\ell_2 n)$). This difficulty can be solved if h is small with respect to Y by expanding the hypergeometric function in its Taylor series. The general case (*i.e.* $|h|$ as large as Y) is currently being investigated with some success by G. Harcos.

4.4.4. Integral of products of eigenfunctions

As we have seen, the key point in the above approach is the problem of getting the correct exponential decay for the integral of the triple product:

$$(4.22)\quad \langle\overline{V(z)}, u_j(z)\rangle = \int_{X_0(N)} (\ell_1 y)^{k/2}\overline{g(\ell_1 z)}(\ell_2 y)^{k/2} g(\ell_2 z)u_j(z)\frac{dxdy}{y^2} \ll_g (N(1+|t_j|))^A \exp(-\frac{\pi}{2}|t_j|),$$

as $|t_j| \to +\infty$, for some constant A. In fact, in applications to the **SCP**, what one really needs is an averaged bound of the form

$$\sum_{|t_j|\leqslant T} |\langle\overline{V(z)}, u_j(z)\rangle|^2 e^{\pi|t_j|} + \cdots \ll_g (NT)^A,$$

where $\ldots$ denotes the corresponding contribution from the Eisenstein spectrum.

We will not describe in detail the methods involved in these bounds; this would bring us too far off course. We merely list the main contributors and refer to their works.

Sarnak was the first to obtain (4.22) for general modular forms [**Sa3**]; in fact, his method, which is based on the analysis of the triple product in the spherical model, is smooth and applies in quite general situations: one can replace the product $V(z) = (\ell_1 y)^{k/2}\overline{g(\ell_1 z)}(\ell_2 y)^{k/2} g(\ell_2 z)$ by an arbitrary product of automorphic forms and u_j by any primitive Maass form of appropriate weight and nebentypus. Moreover, this method works for non-arithmetic lattices, over more general number fields, and in several higher rank situations [**PS, CoPSS**].

A little later, Bernstein and Reznikov [**BR, BR2**] gave another proof of (4.22) (for g a Maass form at least, the case of holomorphic forms is treated in [**KrS**]). Their method uses techniques of complexification and analytic continuation of real analytic vectors of representations of $SL_2(\mathbf{R})$ (computed in their various models), as well as an innovative technique based on the concept of invariant Sobolev norms. Like Sarnak's, their method extends to non-arithmetic lattices and, moreover, provides an essentially optimal bound in the $|t_j|$-aspect (A can be taken arbitrarily small). More recently, and using related ideas, Krötz and Stanton [**KrS**] gave a wide generalization of the Bernstein/Reznikov method in particular, extending these techniques to automorphic forms on SL_n.

4.4.5. Questions of uniformity I.

The problem of the exponential decay in (4.22) having been solved, the next step concern the uniformity of these bounds with respect to ℓ_1, ℓ_2 or the parameters of g. This technical question is crucial for the solution of new instances of the **SCP**, and hence new instances of the **ScP**: for example, the **ScP** problem for the symmetric square $L(\mathrm{sym}^2 f, s)$ in the level aspect, and, more generally, for Rankin-Selberg L-functions $L(f \times g, s)$, where f and g both have levels varying arbitrarily.

Concerning this issue, Sarnak's method already provides reasonable (*i.e.* polynomial) control in the remaining parameters [5]: for instance when g is primitive, holomorphic of weight k and level q', the bound (4.21) gives

$$\sum_{|t_j|\leqslant T} |\langle u_j, \overline{V}\rangle|^2 e^{\pi|t_j|} + \text{Eisenstein contr.} \ll_{k,\varepsilon} (NT)^{\varepsilon} q'^4 (\ell_1\ell_2)^2 T^{2k+4}, \tag{4.23}$$

in view of the trivial bound $||y^{k/2} g(z)||_\infty \ll_k q'^{1/2}\log q'$ of [**AU**]. There is no doubt that this method can be refined further.

We pass to the methods of Bernstein/Reznikov (and its improvements by Krötz/-Stanton): the dependency with respect to the level of g was computed by Kowalski and provides stronger result ([**Kow**], [**KrS**] Thm. 8.5); namely, one has for $T \geqslant 1$

$$\sum_{|t_j|\leqslant T} |\langle u_j, \overline{V}\rangle|^2 e^{\pi|t_j|} + \text{Eisenstein contr.} \ll_k (\ell_1\ell_2 q')^2 T^{2k}, \tag{4.24}$$

[5] We will discuss here only the uniformity with respect to the level q' of g

which is stronger than (4.23) by a factor $\approx q'^2T^4$.

Although such results are quite good and solve new instances of the convexity problem, the given uniformity is not quite sufficient for the most advanced purposes. We will now discuss a third (tentative) approach to this question. This approach is far less general (and somewhat more technical), but makes full use of the arithmetic context (it applies only to arithmetic lattices). It is suggested by triple product identities of Harris/Kudla, Gross/Kudla and Böcherer/Schulze-Pillot and Watson[**HK, GK, BS, Wat**], which relate periods of products of triples of automorphic forms to central values of Rankin triple product L-functions $L(\pi_1\otimes\pi_2\otimes\pi_3, s)$.

To fix ideas, we consider these identities in the special case $\pi_1 = \pi_g,\ \pi_2 = \tilde{\pi}_g,\ \pi_3 = \pi_f$, where f and g are primitive forms of squarefree level q' with trivial nebentypus, f a Maass form of weight zero and g either a holomorphic form of weight k or a Maass form (of weight $k=0$); elaborating on the above cited works, Watson gives a completely explicit relation between the period and the central critical value of the triple L-function $L(\pi_g\otimes\pi_g\otimes\pi_f, 1/2)$. More precisely, one has

$$\frac{|\int_{X_0(q')} y^k|g(z)|^2 f(z)dxdy/y^2|^2}{\langle g,g\rangle^2\langle f,f\rangle} = \frac{\prod_{p|q'} Q_p}{q'^2}\frac{\Lambda(\pi_g\otimes\pi_g\otimes\pi_f,1/2)}{\Lambda(\mathrm{sym}^2\pi_g,1)^2\Lambda(\mathrm{sym}^2\pi_f,1)}, \tag{4.25}$$

where $\Lambda(s) = L_\infty(s)L(s)$ is the completed L-function and Q_p is an explicit, uniformly bounded, local factor depending on $\pi_{f,p},\pi_{g,p}$. From the value of the factors at infinity (see [**Wat**]) and(2.19) and (2.16), one deduces that

$$|\langle y^k|g(z)|^2, \frac{f}{\langle f,f\rangle^{1/2}}\rangle|^2 \ll_{\varepsilon,k} (Q_fQ_g)^\varepsilon(1+|t_f|)^{2k-2}\exp(-\pi|t_f|)L(\pi_g\otimes\pi_g\otimes\pi_f,1/2)$$

for any $\varepsilon>0$.

This formula shows that **GLH** provides a very sharp (probably best possible) bound in all aspects but, of course, we are looking for unconditional results. In fact, the convexity bound

$$L(\pi_g\otimes\pi_g\otimes\pi_f,1/2) \ll_{\varepsilon,k} q'^{5/4}(1+|t_f|)^{8/4}$$

is known unconditionally: it follows from the factorization

$$L(\pi_g\otimes\pi_g\otimes\pi_f,s) = L(\mathrm{sym}^2\pi_g\otimes\pi_f,s)L(\pi_f,s)$$

and the fact that the convexity bound is known for each factor. For the first factor this is a consequence of work of Molteni [**Mol**] and the truth of Hypothesis $H_2(1/9)$, due to Kim/Shahidi [**KiSh**]. This yields an individual bound stronger than (4.19) in the special case $\ell_1\ell_2 = 1$ and $N = q'$. By averaging over the primitive forms of level q', one obtains

$$\sum_{|t_f|\leqslant T} |\langle y^k|g(z)|^2, \frac{f}{\langle f,f\rangle^{1/2}}\rangle|^2 \ll_{\varepsilon,k} (Q_fQ_g)^\varepsilon q'^{9/4}(1+|t_f|)^{2k+2}, \tag{4.26}$$

still weaker than (4.24) by a factor of $T^2q'^{1/4}$. Eventually (4.26) can be improved further with subconvexity estimates, in particular the known subconvexity bounds for $L(\pi_f,s)$, but one can do much better. Indeed, by the approximate functional equations, the central values $L(\mathrm{sym}^2\pi_g\otimes\pi_f,1/2)$ and $L(\pi_f,1/2)$ are represented by linear forms in the $\lambda_f(n)$ of respective length $(1+|t_f|)^3q'^2$ and $(1+|t_f|)q'^{1/2}$;

hence, after Cauchy/Schwarz, one bounds this product of linear forms on average by the large sieve inequalities (3.8), resulting in

$$\sum_{|t_f|\leqslant T}|\langle y^k|g(z)|^2,\frac{f}{\langle f,f\rangle^{1/2}}\rangle|^2e^{\pi|t_f|}\ll_{\varepsilon,k}(q'T)^{\varepsilon}T^{2k+1/2}q'^{3/2}, \tag{4.27}$$

where the above sum ranges over primitive forms of level q' exactly. The above bound is stronger than (4.24) (in the q' aspect) by a factor of $q'^{1/2}$ and could be improved further if necessary. Of course, for applications to the **SCP** one needs to perform the average over old forms and in the case when $\ell_1\ell_2$ is not equal to one. One can reasonably expect that the contribution from the old forms is not larger than the contribution of the new forms, and that dependence in $\ell_1\ell_2$ is at most polynomial; in particular, the following bound is certainly within reach of the current techniques:

$$\sum_{|t_j|\leqslant T}|\langle u_j,\overline{V}\rangle|^2e^{\pi|t_j|}+\text{Eisenstein contr.}\ll_k(\ell_1\ell_2)^Aq'^{3/2}T^{2k+1/2}, \tag{4.28}$$

for some absolute constant A. This is more than sufficient to solve completely many new cases of the **ScP** (see below). Proving (4.28) along the above lines requires a generalization of the triple product identities. More precisely, one expects a formula similar to (4.25), relating

$$|\langle(\ell_1y)^{k/2}\overline{g(\ell_1z)}(\ell_2y)^{k/2}g(\ell_2z),f(\ell_3z)\rangle|^2,$$

where f is a primitive Maass form of level $q''|\ell_1\ell_2q'$, and where $\ell_3|\ell_1\ell_2q'/q''$, to the L-value

$$\frac{\Lambda(\pi_g\otimes\pi_g\otimes\pi_f,1/2)}{\Lambda(\mathrm{sym}^2\pi_g,1)^2\Lambda(\mathrm{sym}^2\pi_f,1)},$$

with possibly some extraneous local factors at the primes dividing $q'\ell_1\ell_2$.

4.5. Subconvexity for Rankin-Selberg L-functions

We conclude this lecture with the Subconvexity Problem for Rankin-Selberg convolutions $L(f\otimes g,s)$, more precisely, when g is fixed and one of the parameters of f varies. In this section, we describe the q aspect; in fact the other aspects are very similar to this one and, in particular, all are related to some form of the **SCP**. However, the level aspect presents an additional interesting feature, which we would like to point out. The investigations of the level aspect began with the work of Duke/Friedlander/Iwaniec [**DFI3**] for f a holomorphic form with trivial nebentypus and $g(z)$, the Eisenstein series $E'(z,1/2)$; in view of the identity

$$L(f\otimes E'(z,1/2),s)=L(f,s)^2,$$

this is equivalent to solving the **SCP** for $L(f,s)$ in the level aspect. Some generalizations to the case of g a general cusp form have been given in [**KMV2, Mi**], and we describe here the proof of the following special but interesting case:

Theorem 4.13. [**Mi**] *Let $g\in S^p_{k'}(q',\chi')$ be a holomorphic*[6] *cusp form and $f\in S^p_k(q,\chi,it)$ be a general primitive cusp form. Then one has*

$$L(f\otimes g,s)\ll_{\varepsilon,s,g}q^{1/2-1/1054}.$$

[6]When g is a Maass form, the corresponding bound has now been proven in [**HM**]

Proof. (Sketch) The conductor of $L(f\otimes g,s)$, $q_{f\otimes g}=:Q^2$ say, satisfies

$$q^2\ll_g(q/q')^2\leqslant Q^2\leqslant(qq')^2\ll_g q^2,$$

and the convexity bound is then $L(f\otimes g,s)\ll_{g,\varepsilon,f_\infty}Q^{1/2+\varepsilon}$. Thus our objective is to bound non-trivially the linear form of length $Q=q_{f\otimes g}$,

$$\Sigma_V(f\times g,Q)=\sum_n\frac{\lambda_f(n)\lambda_g(n)}{\sqrt{n}}V(\frac{n}{Q}),$$

for V some function with rapid decay; to simplify presentation we suppose that V is compactly supported in $[1,2]$. The next step is to choose an appropriate family $\mathcal{F}$ as follows

- When $f\in S_k(q,\chi)$ is holomorphic of weight $k\geqslant 2$, one can take for f an orthonormal basis of $S_k(q,\chi)$ containing $f/\langle f,f\rangle^{1/2}$, and use Petersson's formula (2.1).
- If f is a Maass form of weight $k=0$ or 1, it is natural to take for $\mathcal{F}$ an orthonormal basis of $\mathcal{C}_k(q,\chi)$ containing $f(z)/\langle f,f\rangle^{1/2}$ and use the Kuznetsov-/Petersson formula (2.25), or its smoothed version (2.28).
- Obviously, the first possibility is not available when f is holomorphic of weight one; instead one considers $F(z):=y^{1/2}f(z)$ as a Maass form of weight 1, and makes the above choice for $\mathcal{F}$.
- In fact, for technical purposes, it is useful to enlarge the family further by taking an orthonormal basis of forms of level $[q,q']$ containing $f/\langle f,f\rangle_{[q,q']}$ as an *old* form.

Remark 4.12. Actually, even for holomorphic forms of weight $k\geqslant 2$, there are some technical advantages to considering $F(z)=y^{k/2}f(z)$ inside an orthogonal basis of non-holomorphic forms of weight k: for small weights, Petersson's formula is only slowly convergent in the c variable making its use delicate. (The intrinsic reason is that the corresponding eigenvalue $\frac{k}{2}(1-\frac{k}{2})$, for small k, is not enough separated from the continuous spectrum of $\mathcal{L}_2(q,\chi)$.)

By the amplification method, it is sufficient to give a bound for the amplified second moment of

$$\Sigma_V(f\times g,Q)\times\sum_{\substack{l\leqslant L\\(l,qq')=1}}c_l\lambda_f(l)=L_{f\times g}(\vec{c}),$$

(say) for L a small power of Q, on average over $\mathcal{F}$. The multiplicative properties of $\lambda_f(n)$ and $\lambda_g(n)$ convert this product into a linear form in the Fourier coefficients of f of length QL (and into an expression which makes good sense even if f is not an eigenform):
(4.29)

$$L_{f\times g}(\vec{c})=\sum_{abe\leqslant L}c_{abe}\mu(a)\chi\chi'(a)\chi(b)\lambda_g(b)\sum_{n\geqslant 1}V(\frac{a^2bn}{Q})\frac{\lambda_g(n)}{\sqrt{a^2bn}}\sqrt{aen}\rho_f(aen).$$

Next, we form the mean square of $L_{f\times g}(\vec{c})$ over the chosen family; our objective now is the bound

$$\frac{\mathcal{H}(t_f)}{\langle f,f\rangle_{[q,q']}}|L_{f\times g}(\vec{c})|^2\leqslant\sum_{j\geqslant 1}\mathcal{H}(t_j)|L_{u_j\times g}(\vec{c})|^2+\cdots\ll_{\varepsilon,g,k}(Q_fQ_g)^\varepsilon\sum_{l\leqslant L}|c_l|^2,\tag{4.30}$$

for, by choosing the amplifier (4.6), we would then obtain (using (2.16))

$$\Sigma_V(f\times g,Q)\ll_{\varepsilon,g,t_f}(Q_fQ_g)^{\varepsilon}(\frac{[q,q']}{L^{1/2}})^{1/2}.$$

We apply (2.28) to the middle term of (4.30), converting it into the sum of a diagonal contribution (coming from the diagonal symbol $\delta_{m=n}$ in (2.28)), which is bounded by

$$\ll_{\varepsilon,g,k}(Q_fQ_g)^{\varepsilon}\sum_{l\leqslant L}|c_l|^2,$$

plus a sum of Kloosterman sums of the form

$$\begin{aligned}\text{(4.31)}\quad &\sum_{\ell_1,\ell_2\leqslant L}\overline{c_{\ell_1}}c_{\ell_2}\sum_{c\equiv 0([q,q'])}\frac{1}{c^2}c\times\\ &\qquad\sum_{m,n\geqslant 1}\frac{\overline{\lambda_g(m)}\lambda_g(n)}{\sqrt{mn}}V(\frac{m}{Q})V(\frac{n}{Q})S_\chi(\ell_1m,\ell_2n;c)\mathcal{I}(\frac{4\pi\sqrt{\ell_1\ell_2mn}}{c})\end{aligned}$$

called the offdiagonal term.(Here, to simplify presentation we have considered only the part of (4.29) where $a=b=1$ and $e=l$.) From the properties of rapid decay of $\mathcal{I}$, one sees that the crucial range for the c-sum is when $c\approx\sqrt{\ell_1\ell_2}Q$, so we assume from now on that c is within this range; to simplify further, we assume that $(c,\ell_2)=1$. At this point, one could bound all the terms trivially using Weil's bound for Kloosterman sums, however, we see quickly that such bound is too large by a factor of $q^{1/2}$ at least; thus we have to place ourselves in a more favorable position. The next move is to open the Kloosterman sum

$$S_\chi(m,n;c)=\sum_{x(c),(x,c)=1}\overline{\chi}(x)e(\frac{m\overline{x}+nx}{c}),$$

getting the additive character $e(\frac{nx}{c})$. Now Voronoi's formula applied to the n-sum transforms $e(\frac{\ell_2nx}{c})$ into $\chi'(\overline{\ell}_2)e(-\frac{n\overline{\ell_2x}}{c})$ (since $q'|[q,q']|c$) giving

$$\begin{aligned}&c\sum_{m,n\geqslant 1}\overline{\lambda_g(m)}\lambda_g(n)S_\chi(\ell_1m,\ell_2n;c)\dots\\ &\qquad=\sum_{m,n\geqslant 1}\overline{\lambda_g(m)}\lambda_g(n)\chi'(\overline{\ell}_2)S_{\chi\chi'}(\ell_1m-\overline{\ell}_2n,0;c)\mathcal{J}(m,n;c).\end{aligned}$$

Here

$$\chi'(\overline{\ell}_2)S_{\chi\chi'}(\ell_1m-\overline{\ell_2}n,0;c)=\chi(\ell_2)G_{\chi\chi'}(\ell_1\ell_2m-n;c)$$

is the Gauss sum of character $\chi\chi'$ and modulus c, and $\mathcal{J}(m,n;c)$ is some appropriate Bessel transform of $\mathcal{I}$. Thus we have considerably simplified the picture, by replacing the Kloosterman sums by the much simpler Gauss sums. By [**Mi**] Lemma 4.1, one sees that $\mathcal{J}(m,n;c)$ is essentially bounded and is very small unless $m\sim Q,\ n\sim\ell_1\ell_2Q$. Changing the variables by setting $h=\ell_1\ell_2m-n$, one is led to the following instance of the **SCP**:

$$\sum_h G_{\chi\chi'}(h;c)\sum_{\ell_1\ell_2m-n=h}\overline{\lambda_g(m)}\lambda_g(n)W(\ell_1\ell_2m,n),$$

with $W(x,y)=\mathcal{J}(\frac{x}{\ell_1\ell_2},y;c)$ and (with the notations of section 4.4) $X\sim Y\sim\ell_1\ell_2Q$.

Remark 4.13. Assume for simplicity that $\chi\chi'$ is trivial. An easy estimate shows that the global contribution of the terms above is bounded by $O_\varepsilon(q^\varepsilon L^2(\sum_{\ell\leqslant L}|c_\ell|)^2)$; this is not quite sufficient, but at least Voronoi's formula has placed us back in an acceptable position.

4.5.1. The contribution of the term $h = 0$

The contribution of the *degenerate* frequency $h = 0$ is void unless $\chi\chi'$ is the trivial character. When $\chi\chi'$ is trivial, the $h = 0$ term brings another *main* contribution that can be computed quite explicitly; this contribution is called the first *off-diagonal* main term: its existence is the first manifestation that the variables m and n have reached a critical range, since for smaller ranges, the main terms only arise from diagonal contributions. We will not describe the calculation any further (see [**DFI3, KMV2, DFI8**] for instance), but instead merely mention that if $k \not\equiv k'(2)$, or if g is holomorphic of weight $k' \geqslant 2$, this term is as small as the diagonal contribution; this fact is a consequence of an orthogonality property of the Bessel functions arising in the computation of this term. When $k \equiv k'(2)$ and g has weight 0 or 1 the Bessel functions are no longer orthogonal and the first off-diagonal main term becomes much larger than the diagonal. At this point, it is possible for the amplification method to break down... Eventually, Duke/Friedlander/Iwaniec resolved this problem and identified the true origin of the first off-diagonal main term in this case: this term is nicely compensated by the contribution from the Eisenstein spectrum (noted $+\dots$ in the left hand side of (4.30)), up to an admissible error term. The verification of the matching of both terms is carried out in [**DFI8**, Sect. 12 and 13], and uses Burgess's bound together with a delicate identification of rather different complex integrals.

4.5.2. The contribution of the term $h \neq 0$

The remaining contribution (from the "non degenerate" frequencies $h \neq 0$) is called the off-off-diagonal term and is estimated by solving the **SCP** with the methods of Section 4.4. The off-off-diagonal term turns out to be an error term when g is cuspidal, but brings another third main contribution when g is an Eisenstein series (see[**KMV1**]).

An application of Theorem 4.14 and the trivial bound for Gauss sums

$$|G_{\chi\chi'}(h;c)| \leqslant (q^*_{\chi\chi'})^{1/2}(h,c),$$

shows that the $h \neq 0$ terms of (4.31) contribute (essentially)

$$\ll_{\varepsilon,g} Q_f^\varepsilon \frac{L\sum_{l\leqslant L}|c_l|^2}{[q,q']^2}\sqrt{q^*_{\chi\chi'}}Q^{3/2+\theta}L^{3+2\theta} \ll_{\varepsilon,g} Q_f^\varepsilon (\frac{q^*_{\chi\chi'}}{q})^{1/2}q^\theta L^{4+2\theta}\sum_{l\leqslant L}|c_l|^2,$$

where $q^*_{\chi\chi'}$ is the conductor of the character $\chi\chi'$. Thus the above bound solves the Subconvexity Problem for $L(f\times g,s)$ (and g fixed) only as long as $q^*_{\chi\chi'} \leqslant q^\eta$ for some fixed $\eta < \frac{1}{1+2\theta}$; in particular, even **RPC** is not sufficient to solve our problem when χ is primitive.

The problem comes from the size of the Gauss sums $G_{\chi\chi'}(h;c)$, which get larger with $q^*_{\chi\chi'}$; hence it is clear that the oscillations of $G_{\chi\chi'}(h;c)$ must be exploited in the averaging over h: this is reasonable since h varies over rather long intervals (of length $\sim \ell_1\ell_2 Q$ when χ is primitive). This effect is best seen when one considers the extreme (but most crucial) case of $\chi\chi'$ being primitive ($q^*_{\chi\chi'} = [q,q']$): for simplicity, we examine only the contribution coming from $c = [q, q']$ (the other terms are treated similarly). Under these conditions, one has $G_{\chi\chi'}(h; c) = G_{\chi\chi'}(1; c)\overline{\chi\chi'}(h)$ and the corresponding term becomes (see the notations of Section 4.4)

$$G_{\chi\chi'}(1; c) \sum_{h\neq 0} \overline{\chi\chi'}(h)\Sigma_W(g, \ell_1\ell_2, 1, h).$$

By (4.13), (4.15) and (4.16), the above sum has the following spectral decomposition

$$G_{\chi\chi'}(1; c)\frac{1}{2\pi i} \int_{(1/2+\theta+\varepsilon)} \frac{(2\pi)^{s+k-1}2^{s-1}}{\Gamma(s+k-1)(\ell_1\ell_2)^{\frac{1}{2}}} \times$$
$$\Big(\sum_{j\geqslant 1} \Gamma(\frac{s-\frac{1}{2}+it_j}{2})\Gamma(\frac{s-\frac{1}{2}-it_j}{2})\langle u_j, \overline{V}\rangle \sum_{h\neq 0} \overline{\chi\chi'}(h)\overline{\rho_j}(h)\frac{\widehat{W}(h,s)}{|h|^{s-1}} + \dots\Big)ds,$$

where $\dots$ denotes the contribution from the continuous spectrum. We want to bound the sum $\sum_{h\neq 0} \overline{\chi\chi'}(h)\overline{\rho_j}(h)\widehat{W}(h, s)$: for simplicity we assume that u_j is of the form $u_j = f_j/\langle f_j, f_j\rangle^{1/2}$ where f_j is primitive. The above sum then equals

$$\sum_{\pm} \chi(\pm 1)\rho_j(\pm 1) \sum_{h>0} \overline{\chi\chi'}(h)\overline{\lambda_j}(h)\frac{\widehat{W}(\pm h, s)}{h^{s-1/2}}.$$

By averaging trivially over h, one has, using (2.17),

$$\Sigma_{h>0}\overline{\chi\chi'}(h)\overline{\lambda_j}(h)\frac{\widehat{W}(\pm h, s)}{h^{s-1/2}} \ll_{\varepsilon,k'} (N(1 + |t_j|))^{\varepsilon}(QL^2)^{3/2+\varepsilon}$$

which is as good as having **RPC** for the individual $\lambda_j(h)$. However, as we have seen previously, this is –just barely!– not sufficient.

One can do better by considering the h-sum in terms of the twisted L-function $L(\chi\chi' \times f_j, z)$ and by using the subconvexity bound (in the $q^*_{\chi\chi'}$-aspect) proven before. Indeed, Theorem 4.8, leads to an upper bound of the form

$$\sum_{h>0} \overline{\chi\chi'}(h)\overline{\lambda_j}(h)\frac{\widehat{W}(\pm h, s)}{h^{s-1/2}} \ll_{\varepsilon,k'} (N(1 + |t_j|))^{A}(q^*_{\chi\chi'})^{1/2-\delta}(QL^2)^{1+\varepsilon}$$

for some $A \geqslant 0$, some $\delta = 1/22 > 0$, and any $\varepsilon > 0$. The analysis of the continuous spectrum contribution follows essentially the same lines, with the subconvexity bound of Theorem 4.8 replaced by Burgess's bound for Dirichlet L-functions (4.2). With these bounds at hand, one can repeat the arguments of section 4.4.2 to find that (4.31) is bounded by

$$\ll_{\varepsilon,g} Q_f^{\varepsilon} q^{-\delta} L^{3+2A} \sum_{l\leqslant L} |c_l|^2,$$

which gives (4.30) as soon as $L \leqslant q^{\delta/(3+2A)}$, and is more than enough to solve the **SCP**. □

Remark 4.14. As we have seen, the case where χ has large conductor is deduced from (known) cases of subconvexity for L-functions of lower ranks (1 and 2); in fact this phenomenon already occurred in the work of Duke/Friedlander/Iwaniec [**DFI8**]. That a "reduction of the rank" principle exists for the Subconvexity Problem is very encouraging for its resolution in higher degrees. *A posteriori*, the possibility of this reduction could have been anticipated, firstly because of the inductive structure of the automorphic spectrum of GL_d ([**MW2**]) and secondly because this principle is already present in Deligne's proof of the Weil conjectures [**De2, De3**].

4.5.3. Questions of Uniformity II

So far we have the **ScP** for Rankin-Selberg L-functions $L(f \otimes g, s)$ when g is fixed and the level of f grows. One may wonder what happens when the level of g grows too. The discussion of Section 4.4.4 enables one to solve the **SCP** in the q and q' aspects simultaneously, as long as q' is smaller than a (small but explicit) power of q. For example, suppose that f and g have coprime levels with trivial nebentypus and that g is holomorphic. By keeping track of the dependency on q' in our estimates and by using respectively (4.23), (4.24), or (4.28), one can show that for $\Re es = 1/2$ one has

$$L(f \otimes g, s) \ll (qq')^{1/2-\delta}$$

for some positive δ (the implied constant depending on s and on the parameters at infinity of f and g), as long as $q' \leqslant q^\eta$ for some fixed η, where

$$\eta < (1/2-\theta)/(3/2+\theta), \ \eta < (1/2-\theta)/(1/2+\theta), \ \eta < (1/2-\theta)/(1/4+\theta),$$

respectively. In particular, since $\theta = 1/9$ is admissible and $(1/2-1/9)/(1/4+1/9) > 1$, the bound (4.28) would solve the Subconvexity Problem for $L(f \otimes g, s)$ in the qq'-aspect when q and q' are coprimes with no further restrictions on the relative sizes of q and q'.

LECTURE 5
Some applications of Subconvexity

In this last section, we describe several applications of subconvex bounds in arithmetic and related fields. Although analytic by nature, the (sub)convexity bound may have deep geometric or arithmetic meaning and applications.

5.1. Subconvexity vs. Riemann/Roch

We consider again the question stated in section 3.3.1, of distinguishing two modular forms by their first Hecke eigenvalues:

Question. *Given f and g two distinct primitive holomorphic forms, what is the smallest possible $n = N(f,g)$ for which $\lambda_f(n) \neq \lambda_g(n)$.*

To fix ideas we consider $f \in S_2^p(q)$ and $g \in S_2^p(q')$ two primitive holomorphic forms of weight two. Then $f(z)dz$ and $g(z)dz$ define two holomorphic differentials on the modular curve $X_0([q,q'])$, and picking $q = \exp(2\pi iz)$ the uniformizer at ∞, one has

$$f(z)dz - g(z)dz = \sum_{n\geqslant 1}(\lambda_f(n) - \lambda_g(n))q^{n-1}dq = \sum_{n\geqslant N(f,g)}(\lambda_f(n) - \lambda_g(n))q^{n-1}dq$$

so that $N(f,g) - 1$ is the order of vanishing of this differential at ∞. By the Riemann/Roch Theorem it follows that

$$\begin{aligned} N(f,g) &\leqslant \deg\Omega_{X_0[q,q']}(\mathbf{C}) + 1 = 2\mathrm{genus}(X_0[q,q']) - 1 \\ &= \frac{[q,q']\prod_{p|qq'}(1+\frac{1}{p})}{6}(1+o(1)) \ll_\varepsilon [q,q']^{1+\varepsilon} \end{aligned}$$

for any $\varepsilon > 0$. Note that at this point we have not used the fact that f and g are primitive.

On can also approach this question via Rankin-Selberg L-functions (see [**GH**] for instance). The basic idea is that the modular forms f and g can be distinguished by the different analytic properties of $L(f\otimes g,s)$ and $L(g\otimes g,s)$: the latter has a pole at $s=1$ and the former has no poles. The method amounts to making this difference explicit at the level of Hecke eigenvalues.

One chooses V a smooth non-negative function compactly supported on $[1/2, 1]$, and $N \geqslant 1$, and one evaluates the sum

$$\Sigma_V(f \times g, N) := \sum_{n \geqslant 1} \lambda_f(n)\lambda_g(n)V(\frac{n}{N}).$$

By the inverse Mellin transform,

$$\Sigma_V(f \times g, N) = \frac{1}{2\pi i} \int\limits_{(3)} \frac{L(f \times g, s)}{\zeta^{(qq')}(2s)} \hat{V}(s) N^s ds,$$

where $L(f \times g, s)$ is defined as in (2.14). Shifting the contour to $\Re es = 1/2$, we hit a pole at $s = 1$ only if $f = g$; hence for $f \neq g$ one has

$$\Sigma_V(f \times g, N) = \frac{1}{2\pi i} \int\limits_{(1/2)} \frac{L(f \times g, s)}{\zeta^{(qq')}(2s)} \hat{V}(s) ds,$$

while for $f = g$ one has

$$\Sigma_V(g \times g, N) = \mathrm{res}_{s=1} \frac{L(g \times g, s)}{\zeta^{(q')}(2s)} \hat{V}(s) N^s + \frac{1}{2\pi i} \int\limits_{(1/2)} \frac{L(g \times g, s)}{\zeta^{(q')}(2s)} \hat{V}(s) ds.$$

Now, it follows, from (2.14) and (2.16), that

$$Res_{s=1} \frac{L(g \times g, s)}{\zeta^{(q')}(2s)} \hat{V}(s) N^s = \hat{V}(1) N \mathrm{res}_{s=1} \frac{L(g \times g, s)}{\zeta^{(q')}(2s)} \gg_\varepsilon q'^{-\varepsilon} \hat{V}(1) N.$$

For simplicity, assume that q and q' are squarefree; in this case the conductor of $L(f \otimes g, s)$ equals $[q, q']^2$ and the conductor of $L(g \otimes g, s)$ equals q'^2; hence, it follows from the convexity bounds that

$$\text{(5.1)} \quad \Sigma_V(f \times g, N) = O_{\varepsilon,V}((qq')^\varepsilon N^{1/2}([q, q']^{1/2})$$
$$\text{and } \Sigma_V(g \times g, N) = \hat{V}(1) N \mathrm{res}_{s=1} \frac{L(g \times g, s)}{\zeta^{(q')}(2s)} + O_{\varepsilon,V}(q'^\varepsilon N^{1/2} q'^{1/2}).$$

Now if we take $N = N(f, g)$ we have

$$\Sigma_V(f \times g, N) = \Sigma_V(g \times g, N),$$

and it follows from (5.1) that $N(f, g) \ll_\varepsilon (qq')^\varepsilon [q, q']$. In particular the above method retrieves (essentially) the Riemann/Roch bound. More importantly any subconvex bound for $L(f \otimes g, s)$ and $L(g \otimes g, s)$ in the level aspects would improve the Riemann/Roch bound; and ultimately, **GLH** would give $N(f, g) \ll_\varepsilon (qq')^\varepsilon$ for any $\varepsilon > 0$.

Remark 5.1. When g is fixed (so that $q' = O(1)$), the Subconvexity Problem for $L(f \otimes g)$ has been solved in [**KMV2**] and yields

$$N(f, g) \ll_{q'} q^{1-1/41}.$$

Remark 5.2. Even without subconvexity, the approach via L-functions is interesting for distinguishing modular forms by means of their first Hecke eigenvalues when Riemann/Roch is not available, as is the case with Maass forms; moreover, it can be extended to automorphic forms on GL_d, for which there is no underlying Shimura variety.

5.2. Subconvexity vs. Minkowski

Let K be an imaginary quadratic field. We denote by $-D = \mathrm{Disc}(O_K)$ its discriminant, $\mathrm{Pic}(O_K)$ the ideal class group of O_K, ψ a character of $\mathrm{Pic}(O_K)$, and ψ_0 its trivial character.

Question. *Given ψ a non-trivial character of $\mathrm{Pic}(O_K)$, what is the smallest $N(\psi)$ such that there is an ideal $\mathfrak{a} \subset O_K$ of norm $N_{K/\mathbf{Q}}(\mathfrak{a}) = N(\psi)$ satisfying $\psi(\mathfrak{a}) \neq 1$?*

This question is very similar to the previous one. By Minkowski's theorem each ideal class of $\mathrm{Pic}(O_K)$ is represented by an integral ideal of norm $\leqslant (2/\pi)\sqrt{D}$, hence

$$N(\psi) \leqslant (2/\pi)\sqrt{D}.$$

On the other hand, we can proceed as before and evaluate the partial sum

$$\Sigma_V(\psi, N) := \sum_{\mathfrak{a} \subset O_K} \psi(\mathfrak{a}) V(\frac{N_{K/\mathbf{Q}}(\mathfrak{a})}{N}) = \frac{1}{2\pi i} \int_{(3)} L(\psi, s)\hat{V}(s) N^s ds;$$

here

$$L(\psi, s) = \sum_{\mathfrak{a} \subset O_K} \psi(\mathfrak{a}) N_{K/\mathbf{Q}}(\mathfrak{a})^{-s}.$$

It is well known (from Hecke), that $L(\psi, s)$ is the L-function of the theta series

$$\theta_\psi(z) := \sum_{\mathfrak{a} \subset O_K} e^{2i\pi N_{K/\mathbf{Q}}(\mathfrak{a}) z} \in M_1(D, \chi_K),$$

which is a holomorphic modular form of weight one and nebentypus χ_K, the quadratic Dirichlet character associated with K. Moreover, $\theta_\psi(z)$ is primitive, and it is cuspidal unless ψ is real, in which case $\theta_\psi(z)$ is an Eisenstein series and (Kronecker's formula) $L(\psi, s)$ factors as a product

$$L(\psi, s) = L(\chi_{1,\psi}, s) L(\chi_{2,\psi}, s)$$

corresponding to a (uniquely determined) factorization of the Kronecker symbol χ_K into a product of two primitive quadratic characters $\chi_K = \chi_{1,\psi}\chi_{2,\psi}$. In particular, for the trivial character one has $L(\psi_0, s) = \zeta(s) L(\chi_K, s)$, hence, shifting the contour to $\Re es = 1/2$ we hit a pole at $s = 1$ only when $\psi = \psi_0$. Hence we have

$$\begin{aligned} \Sigma_V(\psi, N) &= \delta_{\psi=\psi_0} \hat{V}(1) L(\chi_K, 1) N + \frac{1}{2\pi i} \int_{(1/2)} L(\psi, s)\hat{V}(s) N^s ds \\ &= \delta_{\psi=\psi_0} \hat{V}(1) L(\chi_K, 1) N + O(N^{1/2} D^{1/4-1/24000}) \qquad (5.2) \end{aligned}$$

by the subconvexity bounds (either Burgess's bound if ψ is real or Duke/Friedlander/-Iwaniec's bound of Theorem 4.1 otherwise). Taking $\psi \neq \psi_0$ and $N \leqslant N(\psi)$ we obtain by Siegel's theorem (1.19), and the identity $\Sigma_V(\psi, N) = \Sigma_V(\psi_0, N)$, the bound

$$N(\psi) \leqslant D^{1/2-1/12001};$$

here the implied constant is absolute but ineffective. Another kind of application that can be obtained along these lines is the following: suppose that $\mathrm{Pic}(O_K)$ contains a cyclic subgroup G of small index. (For example, the Cohen-Lenstra heuristics predict that there are infinitely many prime discriminants such that $\mathrm{Pic}(O_K)$

is cyclic.) One may then look for an ideal $\mathfrak{a}$ generating G of small norm; from Minkowski's bound there is such a generator with norm bounded by

$$N_{K/\mathbf{Q}}(\mathfrak{a}) \leqslant (2/\pi)\sqrt{D}.$$

Theorem 4.1 allows us to improve on the exponent $1/2$:

Theorem 5.1. *Let $G \subset \mathrm{Pic}(O_K)$ be a cyclic subgroup of index i_G. Then G is generated by an ideal of norm $\ll i_G^2 D^{1/2-1/24001}$, where the constant implied is absolute but ineffective.*

Proof. We denote by $\delta_G(.)$ the characteristic function of the generators of G; to show that there is a generator $\mathfrak{a} \subset O_K$ of norm $\leqslant N$, it is sufficient to show that

$$\Sigma_V(G,N) := \sum_{\mathfrak{a}\subset O_K} \delta_G(\mathfrak{a})V(\frac{N_{K/\mathbf{Q}}(\mathfrak{a})}{N}) \neq 0.$$

Let $\mathfrak{g} \subset G$, be a generator of G by easy Fourier analysis and Möbius inversion, one has

$$\delta_G(\mathfrak{a}) = \frac{1}{|\mathrm{Pic}(O_K)|}\sum_{\psi\in\widehat{\mathrm{Pic}(O_K)}} \psi(\mathfrak{a}) \sum_{\substack{m(|G|)\\(m,|G|)=1}} \overline{\psi}(\mathfrak{g}^m) = \frac{1}{i_G}\sum_{d||G|}\frac{\mu(d)}{d}\sum_{\psi^d_{|G}=1}\psi(\mathfrak{a}).$$

From (5.2) it follows that

$$\begin{aligned}\Sigma_V(G,N) &= \frac{1}{i_G}\sum_{d||G|}\frac{\mu(d)}{d}\sum_{\psi^d_{|G}=1}\Sigma_V(\psi,N)\\ &= \frac{1}{i_G}\frac{\varphi(|G|)}{|G|}\hat{V}(1)L(\chi_K,1)N + O(\tau(|G|)N^{1/2}D^{1/4-1/24000}),\end{aligned}$$

so by Siegel's theorem this sum is nonzero when $N \gg i_G^2 D^{1/4-1/24001}$. □

Remark 5.3. In both examples, the convexity bound essentially matches Minkowski's. Note also that Minkowski theorem is also based on a convexity argument (although quite different from the Phragmen/Lindelöf principle); moreover, via Arakelov geometry, Minkowski's theorem for number fields can also be seen as an analog of the Riemann/Roch Theorem for curves; for more on this see Szpiro's "Marabout Flash" [**Szp**].

5.3. Subconvexity and distribution of Heegner points

In this last section we describe applications of the subconvexity bound to various equidistribution problems.

5.3.1. Equidistribution of lattice points of the sphere.

Given $n \geqslant 1$, it goes back to Gauss that n is representable by the ternary quadratic form $X^2+Y^2+Z^2$ if and only if n is not of the form $4^k(8l-1)$. We denote by $R_3(n) := \{\vec{\mathbf{x}} = (x,y,z) \in \mathbf{Z}^3,\ x^2+y^2+z^2 = n\}$ (resp. $R_3^*(n) := \{\vec{\mathbf{x}} = (x,y,z) \in \mathbf{Z}^3,\ x^2+y^2+z^2 = n,\ g.c.d(x,y,z) = 1\}$) the set of representations of n as the sum of three squares (resp. of the primitive representations) and by $r_3(n)$ (resp. $r_3^*(n)$) the number of such representations. We have $r_3(n) = \sum_{d^2|n} r_3^*(n/d^2)$; on the other hand, Gauss gave a formula for the number of primitive representations in terms of the class number $h(-n)$ of primitive positive binary quadratic forms with discriminant $-n$ [**A, AB**], namely

$$r_3^*(n) = \begin{cases} 12h(-n) \text{ if } n \equiv 1, 2 (\bmod 4) \\ 8h(-n) \text{ if } n \equiv 3 (\bmod 8) \\ 0 \text{ otherwise.} \end{cases}$$

By the class number formula, Siegel's theorem, and the (more elementary) computation of class numbers of non maximal orders in imaginary quadratic fields (see [**Cor**] for instance), it follows that

$$r_3^*(n) \gg_\varepsilon n^{1/2-\varepsilon} \text{ if } r_3^*(n) > 0.$$

Hence, for the unrestricted representations one has

$$r_3(n) \gg_\varepsilon n^{1/2-\varepsilon} \text{ if } n \not\equiv 0, 4, 7 (\bmod 8). \tag{5.3}$$

Thus if $n \not\equiv 0, 4, 7(\bmod 8)$, there are many vectors in $R_3(n)$; one may then look at the distribution of their projections on the unit sphere S^2 as $n \to +\infty$. This question was studied by Linnik: by using ergodic methods, he prove the equidistribution of the projections under the extra assumption that $-n$ is a quadratic residue modulo some fixed odd prime [**Lin3**]. In [**I1**], Iwaniec removed this extraneous condition by using quite different techniques[1].

Theorem 5.2. *As n goes to $+\infty$ through integers $n \not\equiv 0, 4, 7(\bmod 8)$, the set $R_3(n)/\sqrt{n}$ becomes equidistributed on the unit sphere S^2 with respect to the standard Lebesgue measure;* i.e. *for any continuous function V on S^2*

$$\frac{1}{r_3(n)} \sum_{\vec{\mathbf{x}} \in R_3(n)} V\Big(\frac{\vec{\mathbf{x}}}{\sqrt{n}}\Big) \to \int_{S^2} V(u) d\mu. \tag{5.4}$$

Proof. By Weyl's equidistribution criterion, it is sufficient to show that for any harmonic polynomial P on $\mathbf{R}^3$ of degree $k \geqslant 1$, say, the Weyl sum

$$W(n, P) := \frac{1}{r_3(n)} \sum_{\vec{\mathbf{x}} \in R_3(n)} P\Big(\frac{\vec{\mathbf{x}}}{\sqrt{n}}\Big) \to \int_{S^2} P(u) d\mu = 0.$$

Observe that $W(n, P) = 0$ if k is odd. If k is even, one has

$$W(n, P) = \frac{\sqrt{n}^{-k/2}}{r_3(n)} \sum_{\vec{\mathbf{x}} \in R_3(n)} P(\vec{\mathbf{x}}) = \frac{\sqrt{n}^{-k/2}}{r_3(n)} r_P(n)$$

say. In view of (5.3), it is sufficient to show that

$$r_P(n) = O(n^{\frac{k+1}{2}-\delta})$$

for some absolute $\delta > 0$. The theta series

$$\Theta_P(z) := \sum_{n \geqslant 0} r_P(n) e(nz)$$

[1] This method gives, in fact, an estimate for the speed of convergence in (5.4), which apparently is not accessible by the ergodic approach.

is in fact a holomorphic modular form of weight $l = k + 3/2$ and level 4 and a cusp form when $k \geqslant 1$ (see [**I5**] Chap. 10). For a cusp form f of half-integral weight l of some level $q \equiv 0(4)$ with Fourier expansion

$$f(z) = \sum_{n \geqslant 1} \rho_f(n) n^{\frac{l}{2}} e(nz)$$

one has the following bound of Hecke type which comes from Petersson's formula (either for holomorphic forms or for half-integral weight Maass forms) (see [**I5**] chap. 5):

$$\rho_f(n) \ll_{f,\varepsilon} n^{-1/4+\varepsilon} \tag{5.5}$$

for any $\varepsilon > 0$. This bounds yields $r_P(n) \ll_{P,\varepsilon} n^{\frac{k+1}{2}+\varepsilon}$ which is barely not sufficient for our equidistribution problem. In fact, either for squarefree n or for weight $l \geqslant 5/2$, Hecke's bound is not sharp: one expects the analog of the Ramanujan/Petersson Conjecture

$$\rho_f(n) \ll_{f,\varepsilon} n^{-1/2+\varepsilon}$$

for all n and $l \geqslant 5/2$, or for squarefree n when $l = 3/2$. We see that equidistribution follows from any non-trivial improvement over the $-1/4$ exponent in (5.5); such an improvement was provided for the first time by Iwaniec [**I1**] and is given below. □

Theorem 5.3. *Let f be a holomorphic cusp form of weight $l \geqslant 3/2$, one has,*

- *for any n if $l \geqslant 5/2$,*
- *for n squarefree if $l = 3/2$*

$$\rho_f(n) \ll_{\varepsilon,f} n^{-1/4-1/28+\varepsilon} \tag{5.6}$$

for any $\varepsilon > 0$ the implied constant depending on ε and f.

Remark 5.4. The case $l \geqslant 5/2$ (which is the only case of interest for the equidistribution problem) was proved by Iwaniec for n squarefree [**I1**]; the case $l = 3/2$ was proved by Duke by using a version due to Proskurin of Petersson's formula for $3/2$-weight Maass forms [**Du**]; the reduction from squarefree n to all n follows from the Shimura lifting, as is explained in [**Sa1**] or below.

We will not give the original proof of Theorem 5.3, but rather we explain a slightly weaker bound connected with the **ScP**. This is achieved through the theory of the Shimura lift and the work of Waldspurger [**Sh, Wal**]. First we may assume that $f(z)$ is a primitive half-integral weight cusp form (in particular an eigenvalue of the Hecke operators T_{p^2} defined by Shimura). Then for d squarefree, one has

$$\sum_{n \geqslant 1} \frac{\rho_f(dn^2)}{n^s} = \rho_f(d) \prod_p (1 - \frac{\chi(p)p^{-1/2}}{p^s})(1 - \frac{\lambda_g(p)}{p^s} + \frac{\chi_g(p)}{p^{2s}})^{-1}, \tag{5.7}$$

where $\chi(p) = (\frac{-1}{p})^{l-1/2}(\frac{d}{p})$ and $\lambda_g(p)$ are the Hecke eigenvalues of some primitive holomorphic form of weight $2l - 1$ and level dividing the level of f; moreover, as long as f is orthogonal to the theta functions of one variable (for example if $l \geqslant 5/2$), then g is a cusp form [**Sh**]. Hence for $n \geqslant 1$, one has

$$\rho_f(dn^2) = \rho_f(d) \sum_{b|n} \frac{\mu(b)}{b^{1/2}} \chi(b) \lambda_g(n/b), \tag{5.8}$$

and in particular by Deligne's bound for holomorphic cusp forms ($|\lambda_g(n)| \leqslant \tau(n)$), it is sufficient to prove (5.6) for $n = d$ squarefree. In the latter case, one can proceed directly as in [**I1**] or use Waldspurger's formula, which relates the Fourier coefficient $\rho_f(d)$ to a central value of a twisted L-function. If D is the discriminant of some quadratic field K, with $D^{l-1/2} > 0$, then Waldspurger's formula has the form

$$|\rho_f(|D|)|^2 = C(f,g,D)L(g.\chi_K, 1/2),$$

where χ_K is the associated Kronecker symbol and $C(f,g,D)$, the proportionality constant, is bounded independently of D. In particular, for $D = \mathrm{Disc}(\mathbf{Q}(\sqrt{(-1)^{l-1/2}d}))$ we see from (5.8) (eventually after a Möbius inversion) and Deligne's bound, that

$$d^{1/2}\rho_f(d) \ll_{\varepsilon,f} d^{\varepsilon}|L(g.\chi_K, 1/2)|^{1/2}.$$

Hence, any improvement over the $1/4$ exponent in (5.5) is equivalent to the solution of the **ScP** for $L(g.\chi_K, 1/2)$ at the special point $s = 1/2$ in the D aspect, and **GLH** implies the analog of the Ramanujan/Petersson Conjecture. In particular, Theorem 4.8 yields

$$\rho_f(d) \ll_{\varepsilon,f} d^{-1/4-1/44+\varepsilon}$$

which is slightly weaker than (5.6). On the other hand, (5.6) (whose proof doesn't make use of L-functions) yields, for any primitive form g, the subconvex bound

$$L(g.\chi_K, 1/2) \ll_{\varepsilon,g} |D|^{1/2-1/14+\varepsilon}.$$

Remark 5.5. Recently Baruch and Mao [**BM**] have generalized Waldspurger's formula to modular forms over totally real fields. This is being used in [**CoPSS**] in the resolution of the last cases of Hilbert's 10-th problem.

5.3.2. Equidistribution of Heegner points

Given K an imaginary quadratic field, we denote by $\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$ the set of (isomorphism classes of) elliptic curves over $\overline{\mathbf{Q}}$ with complex multiplication by O_K. By the theory of complex multiplication, these curves are defined over H_K, the Hilbert class field of O_K, and $Gal(H_K/K) \simeq \mathrm{Pic}(O_K)$ acts simply transitively on $\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$; for $\sigma \in \mathrm{Pic}(O_K)$, we denote by E^{σ} the corresponding Galois action on E.

To any given place $\mathfrak{q}$ of $\overline{\mathbf{Q}}$ over some place q of $\mathbf{Q}$, we associate a map $r_{\mathfrak{q}}$ as follows:

- If $q = \infty$ is the infinite place, abusing notation, we denote by ∞ the corresponding embedding $\mathfrak{q} : \overline{\mathbf{Q}} \to \mathbf{C}$; for $E \in \mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$, the associated complex curve is the quotient of $\mathbf{C}$ by a lattice, $E(\mathbf{C}) = \mathbf{C}/(\mathbf{Z} + z_E\mathbf{Z})$, where $z_E \in \mathbf{H}$ is defined modulo $\Gamma_0(1)$. We set

$$r_\infty : \begin{array}{ccc} \mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}}) & \to & \mathcal{E}\ell\ell(\mathbf{C}) = X_0(1)(\mathbf{C}) - \{\infty\} \\ E & \to & z_E \pmod{\Gamma_0(1)}. \end{array}$$

- If $\mathfrak{q}$ is above some finite prime q, we define $r_{\mathfrak{q}}(E) := red_{\mathfrak{q}}(E)$ to be the reduction modulo $\mathfrak{q}$ of E. If q *splits* in K, then E reduces to an elliptic curve over $\overline{\mathbf{F}_q}$ with complex multiplication by O_K, and $r_{\mathfrak{q}}$ sends $\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$ bijectively to $\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{F}_q})$, the corresponding set of (isomorphism classes of) elliptic curves over $\overline{\mathbf{F}_q}$; in that case there is not much for us to say.

- On the other hand, if q is *inert* in K, then the reduction map restricted to $\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$ is not injective in general; indeed, $E \in \mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$ reduces to a supersingular elliptic curve (with complex multiplication by some maximal order in the definite quaternion algebra ramified at q). We denote by $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q}) = \{e_1, \dots e_n\}$ the (finite) set of (isomorphism classes of) supersingular elliptic curves in characteristic q (in fact, these curves are defined over $\mathbf{F}_{q^2}$); we have the map

$$r_{\mathfrak{q}} : \begin{array}{ccc} \mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}}) & \rightarrow & \mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q}) \\ E & \rightarrow & r_{\mathfrak{q}}(E). \end{array}$$

The spaces $X_0(1)(\mathbf{C})-\{\infty\}$ and $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$ each carry a natural probability measure. For the former this is simply the Poincaré measure $\mu_\infty(z) = (3/\pi)dxdy/y^2$, which is induced by the hyperbolic metric. The measure μ_q on $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$ is a bit less obvious (however, we refer to the lecture of Gross [**Gr**] and to the first section of [**BD1**] for an explanation of why this is indeed, the natural measure on this space); for any $e \in X_0^{ss}(1)(\mathbf{F}_{q^2})$, we set

$$\mu_q(e) = \frac{1/w_e}{\displaystyle\sum_{e' \in \mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})} 1/w_{e'}},$$

where $w_e = |\mathrm{End}(e)^\times|$ is the number of units of the (quaternionic) endomorphism ring of e_i. Note that μ_q is not exactly uniform, but almost (at least when q is large), since $|\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})| = n = \frac{q-1}{12} + O(1)$ and the product $w_1 ... w_n$ divides 12.

As we will see, it follows from certain cases of the **ScP** that for each $\mathfrak{q}$, the images $r_{\mathfrak{q}}(\mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}}))$ are equidistributed on the corresponding spaces relative to the corresponding measure μ_q, as $|\mathrm{Disc}(O_K)| \to +\infty$. In fact one has the following stronger equidistribution result:

Theorem 5.4. *For each K, pick $E_0 \in \mathcal{E}\ell\ell_{O_K}(\overline{\mathbf{Q}})$ and $G \subset \mathrm{Pic}(O_K)$, a subgroup of index $i \leqslant |\mathrm{Disc}(O_K)|^{1/24001}$. Then for any continuous V on $X_0(1)(\mathbf{C})$ one has, as $|\mathrm{Disc}(O_K)| \to +\infty$,*

$$W(G,V) := \frac{1}{|G|}\sum_{\sigma \in G} V(r_\infty(E_0^\sigma)) = \int_{X_0(1)} V(z) d\mu_\infty(z) + o_V(1). \tag{5.9}$$

For any place $\mathfrak{q}$ above some finite prime q that is inert in K, and for any function V on $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$ one has

$$\frac{1}{|G|}\sum_{\sigma \in G} V(r_{\mathfrak{q}}(E_0^\sigma)) = \sum_{e \in \mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})} V(e)\mu_q(e) + o_{V,q}(1). \tag{5.10}$$

Proof. By Weyl's equidistribution criterion, and in view of the spectral decomposition (2.5) for $X_0(1)(\mathbf{C})$, it is sufficient to prove (5.9) when $V(z)$ is either a Maass cusp form (which we may also assume to be a Hecke form), or the Eisenstein series $E_\infty(z, 1/2+it)$; in either case, one has $\int_{X_0(1)} V(z)d\mu_\infty(z) = 0$. By Fourier transform, one has

$$W(G,V) = \frac{1}{|G|}\sum_{\sigma \in G} V(r_\infty(E_0^\sigma)) = \sum_{\substack{\psi \in \widehat{\mathrm{Pic}(O_K)} \\ \psi_G = 1}} W(\psi, V),$$

where $W(\psi, V)$ denotes the "twisted" Weyl sum

$$W(\psi, V) = \frac{1}{|\mathrm{Pic}(O_K)|} \sum_{\sigma \in \mathrm{Pic}(O_K)} \overline{\psi}(\sigma) V(E_0^{\sigma}).$$

For $V(z) = E_\infty(z, 1/2 + it)$ Hecke has shown (see [**DFI4**] for a proof) that

$$W(\psi, E_\infty(z, 1/2+it)) = \frac{w_K}{2^{3/2+it}} \frac{L(\psi, 1/2+it)|D|^{1/4+it/2}}{\zeta(1+2it)|\mathrm{Pic}(O_K)|}.$$

In particular we deduce from Siegel's theorem, Burgess's bound (when ψ is real), and Theorem 4.1 (when ψ is complex) that

$$W(G, E_\infty(z, 1/2+it)) = O_t(\frac{|\mathrm{Pic}(O_K)|}{|G|} D^{-1/24000.5}).$$

When $V(z) = f(z) \in S_0^p(1, t)$ is a primitive weight zero Maass cusp form and $\psi = \psi_0$ is the trivial character, Maass has shown the formula

$$W(\psi_0, f) = \frac{\sqrt{2}\pi^{-1/4}|D|^{3/4}}{|\mathrm{Pic}(O_K)|} \rho_{\tilde{f}}(-D),$$

where $\rho_{\tilde{f}}(-D)$ denotes the Fourier coefficient of a Maass form of weight $1/2$ and eigenvalue $1/4 + (t/2)^2$ corresponding to f by a theta-correspondence (see [**Du**]). For such forms $\tilde{f}$, Duke proved directly that (5.6) holds, from which it follows that

$$W(\psi_0, f) = o_{f,\varepsilon}(D^{-1/28+\varepsilon}),$$

which proves (5.9) for the full orbit ($G = \mathrm{Pic}(O_K)$; eventually one could also have used Waldspurger's formula to relate non-trivial estimates for Fourier coefficients of half-integral weight Maass forms to the **ScP**.

The method of Maass, however, does not seem to generalize to the twisted Weyl sums $W(\psi, f)$ when ψ is not a real character. Recently, Zhang, in a wide generalization of the Gross/Zagier formulas, related twisted Weyl directly sums to central values of Rankin/Selberg L-functions [**Z1, Z2, Z3**]: for any character ψ of $\mathrm{Pic}(O_K)$, one has

$$|W(\psi, f)|^2 = c(f, \psi) \frac{D^{1/2} L(f \otimes \theta_\psi, 1/2)}{|\mathrm{Pic}(O_K)|^2}, \tag{5.11}$$

for some constant $c(f, \psi) \ll 1$ that is uniformly bounded in ψ; the implied constant depending on f only. Hence any subconvexity exponent for $L(f \otimes \theta_\psi, 1/2)$ in the D-aspect is sufficient to show (ineffectively) that $W(\psi, f) = o_f(D^{-\delta})$ for some $\delta > 0$; in particular, $\delta = 1/2200$ holds, and this concludes the proof of (5.9). Note that the limitation on the index of G comes from the Eisenstein spectrum.

The proof of (5.10) is very similar to (5.9), once the problem has been formulated in the appropriate context (for this we refer to [**Gr, BD1**] for more details). We denote by $M = \bigotimes_{e \in \mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})} \mathbf{Z}e$ the set of divisors supported on $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$ and by M^0 the kernel of M under the degree map:

$$\deg(\sum_e n_e.e) = \sum_e n_e.$$

M is equipped with a natural inner product given by

$$\langle e, e' \rangle = \delta_{e,e'} w_e$$

for $e, e' \in \mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$, so that M^0 is the orthogonal complement of the divisor $e_0 = \sum_e e/w_e$. Moreover, M and M^0 are acted upon by a Hecke algebra $\mathbf{T}^{(q)}$ generated by correspondences $T_p,\ p \nmid q$ of degree $p+1$, which are symmetric with respect to $\langle\ ,\ \rangle$. Their definition can by given either directly via Brandt matrices ([**Gr**]) or adelically by the identification of $\mathcal{E}\ell\ell^{ss}(\overline{\mathbf{F}_q})$ with either the set of left ideal classes of a fixed maximal order (R say) inside the definite quaternion algebra over $\mathbf{Q}$ (B say) ramified at q and ∞, or with the double cosset space $\widehat{R}^\times\backslash\widehat{B}^\times/B^\times$ (here $\widehat{B}$ (resp. $\widehat{R}$) denotes the adelization of B (resp. R)). A special case of the Jacquet/Langlands correspondence states that $M^0\otimes\mathbf{C}$ and $S_2(q)$ are isomorphic as $\mathbf{T}^{(q)}$-modules, hence $M^0\otimes\mathbf{C}$ (in fact $M^0\otimes\mathbf{R}$) admits an orthonormal basis $\{e_f\}_{f\in S_2^p(q)}$ indexed by primitive forms such that

$$T_p e_f = \sqrt{p}\lambda_f(p)e_f$$

for $p \not| q$. In particular, by the same analysis as above, in order to prove (5.10) it is sufficient to prove that for any $f \in S_2^p(q)$, one has

$$W(G,f) := \frac{1}{|G|}\langle\sum_{\sigma\in G} r_{\mathfrak{q}}(E_0^\sigma), e_f\rangle = o_f(1),$$

which follows from

$$W(\psi,f) := \frac{1}{|\mathrm{Pic}(O_K)|}\langle\sum_{\sigma\in\mathrm{Pic}(O_K)}\overline{\psi}(\sigma)r_{\mathfrak{q}}(E_0^\sigma), e_f\rangle = o_f(D^{-1/24000.5})$$

for every ψ that is trivial on G. Again, the twisted Weyl sum is related to central values of Rankin-Selberg L-functions through the formula (5.11), which in this case was proved by Gross [**Gr**] (together with the fact that the action of $\mathrm{Pic}(O_K)$ commutes with the reduction $r_{\mathfrak{q}}$ (see [**BD2**] p.120)). It follows from Theorem 4.13 that in fact $W(\psi,f) = O_f(D^{-1/2200})$.

□

Remark 5.6. The equidistribution theorem above can be widely generalized. It is a special case of general equidistribution properties for small orbits of Heegner points on Shimura curves associated to definite or indefinite quaternion algebras over $\mathbf{Q}$. In [**Z1, Z2, Z3**], Zhang provided very general formulas relating central values of Rankin/Selberg L-functions to twisted Weyl sums corresponding to the appropriate equidistribution problem. More generaly, these formulas and the corresponding subconvexity bounds show the equidistribution of (small orbits of) Heegner points associated with not necessarily non-maximal orders of large conductor, for example, in the (orthogonal) case of Heegner points associated to orders inside a *fixed* field of complex multiplication. However, this direction of inquiry can be treated by other quite different (ergodic) techniques (see [**Va, Cor, ClU**]).

5.4. Subconvexity vs. Quantum Chaos

For a more thorough description of the applications given in this section, we refer to the lectures [**Sa2, Sa6**].

5.4.1. The Quantum Unique Ergodicity Conjecture

Suppose we are giveniven X a compact Riemannian manifold and $\{\varphi_j\}_{j\geqslant 0}$ an orthonormal basis of $L^2(X)$ composed of Laplace eigenfunctions with eigenvalues ordered in increasing size: *i.e.* $\Delta\varphi_j + \lambda_j\varphi_j = 0$ with $0 = \lambda_0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \dots$. Considerations from theoretical physics (Δ is the quantization of the Hamiltonian generating the geodesic flow) led to extensive investigations of the distribution properties of φ_j in the limit as $\lambda_j \to +\infty$, and in particular, of the weak-* limits of the sequence of probability measures

$$d\mu_j = |\varphi_j(x)|^2 dx$$

(here, dx is the normalized Riemannian volume)[2]; such weak limits are called *quantum limits*. When the geodesic flow is *ergodic*, an important result of Shnirelman, Zelditch and Colin de Verdière [**Shn, Ze, C-V**] shows that, at least for a full-density subsequence $\{j_k\}_{k\geqslant 0}$, $d\mu_{j_k}$ weakly-* converges to dx. More precisely, one has for any $V \in C^\infty(X)$,

$$\frac{1}{|\{\lambda_j \leqslant \lambda\}|} \sum_{\lambda_j \leqslant \lambda} \Big| \int_X V d_{\mu_j} - \int_X V dx \Big|^2 = o_f(1); \tag{5.12}$$

this phenomenon is called *Quantum Ergodicity*. However, quantum ergodicity does not exhibit an explicit subsequence having dx as its quantum limit, nor does it exclude the possibility of having exceptional (zero density) subsequences $d\mu_{j_k}$ having a quantum limit different from dx. Such exceptional weak limits are called *strong scars* and indeed have been observed numerically in some related chaotic dynamical systems (such as billiards). In the special case of congruence hyperbolic surfaces (*i.e.* quotients of **H** by congruence subgroups associated to quaternion algebras), Rudnick and Sarnak [**RS**] have ruled out the existence of strong scars supported on a finite union of points and closed geodesics (see also [**BL**] for a recent strengthening). This lead them to conjecture that in many cases dx is the only quantum limit (*Quantum Unique Ergodicity*):

QUE. *Let X be a negatively curved compact manifold. Then $d\mu_j$ weakly converges to dx as $j \to +\infty$.*

So far, the best evidence towards **QUE** comes from the case of arithmetic surfaces and arithmetic hyperbolic 3-folds. Indeed, one can then take advantage of the extra symmetries provided by the (ergodic) action of the Hecke algebra. The study of distribution properties of explicit sequences of primitive Hecke eigenforms is sometimes called *Arithmetic Quantum Chaos*, and one of its most important conjectures is to prove **QUE** for such Hecke eigenforms: the *Arithmetic QUE* conjecture. Note that investigations of quantum limits are not limited to compact arithmetic quotients (such as Shimura curves associated to congruence subgroups associated to indefinite quaternion algebras) nor to Laplace eigenforms; for instance, **Arithmetic QUE** for modular curve is as follows:

Arithmetic QUE. *For any fixed $q \geqslant 1$, let f be a primitive weight zero Maass cusp form (resp. holomorphic cusp form) –with nebentypus trivial or not– for the group $\Gamma_0(q)$ with eigenvalue λ_f (resp. with weight k_f). Then as $\lambda \to +\infty$ (resp $k_f \to +\infty$), the measure*

$$d\mu_f(z) := \frac{|f(z)|^2}{\langle f, f\rangle}\frac{dxdy}{y^2} \ (resp.\ d\mu_f(z) := \frac{|f(z)|^2}{\langle f, f\rangle} y^{k_f} \frac{dxdy}{y^2})$$

[2] $d\mu_j$ is interpreted in quantum mechanics as the probability density for finding a particle in the state "φ_j" at the point x.

weak- converges on $X_0(q)(\mathbf{C})$ to the normalized Poincaré measure*

$$d\mu_P = \frac{1}{\mathrm{vol}(X_0(q))}\frac{dxdy}{y^2}.$$

Remark 5.7. In the case of Maass forms, asking f to be a Hecke eigenform shouldn't be too severe a restriction, when compared with the general **QUE**. Indeed, one expects that the dimension of the Laplace eigenspace with eigenvalue λ is bounded by $O_\varepsilon(\lambda^\varepsilon)$ for any ε. The latter bound seems to be difficult to reach, and there is so far essentially no better bound than the trivial bound coming from Weyl's law (compare with what is known for $\dim S_1(q,\chi)$). However, if this bound were true, then it would not be difficult to show, in the case of $X_0(1)$ for instance, that **QUE** is implied by **Arithmetic QUE**, together with a power saving estimate for the discrepancy. On the other hand, the dimension of the space of holomorphic forms of weight k is large ($\gg k$), and **QUE** certainly cannot hold for an arbitrary holomorphic form: $f(z) = \Delta^m(z)$, the m-th power of Ramanujan's function, is a weight $12m$ holomorphic form not satisfying **QUE** as $m \to +\infty$. Hence, it makes sense in this case to restrict to Hecke eigenforms. It is very possible that the condition of being a Hecke eigenform can be relaxed to the condition of being an eigenform for two Hecke operators T_p, $T_{p'}$, for distinct fixed primes p, p', not dividing q (see [**ClU**]).

Remark 5.8. In the case of holomorphic forms, **Arithmetic QUE** has the following nice consequence, due to Z. Rudnick [**Ru**]: if f is holomorphic of weight k, then f has $\simeq qk_f/12$ zeros on $X_0(q)$; this leads naturally to the question of the distribution of such zeros. It turns out that the convergence of $d\mu_f(z)$ to $d\mu_P$ implies that the zeros of f are equidistributed with respect to $d\mu_P$. A corollary is that the multiplicity of any zero of f is $o(k_f)$ as $k_f \to +\infty$, which (again) is stronger than a trivial application of Riemann/Roch. Note that in the case of Maass forms, it is not clear how to deduce from **Arithmetic QUE** a similar equidistribution for the nodal lines (*i.e.* the lines on the surface defined by the equation $\phi_j(z) = 0$).

In the non-compact case, quantum limits of the Eisenstein series can be studied as well (although their associated measure does not have finite mass). For the full modular curve $X_0(1)$, Luo/Sarnak [**LS**] proved the analog of **QUE** for $E_\infty(z, 1/2 + it)$ as $t \to +\infty$:

Theorem 5.5. *Set $d\mu_t(z) := |E_\infty(z, 1/2+it)|^2\frac{dxdy}{y^2}$. For V a continuous function compactly supported away from the cusp ∞, one has, as $t \to +\infty$:*

$$\int_{X_0(1)} V(z)d\mu_t(z) = \frac{48}{\pi}\int_{X_0(1)} V(z)\frac{dxdy}{y^2}\log t + o_V(\log t).$$

Proof. (Sketch) By density, it is sufficient to obtain the above identity for V either an incomplete Eisenstein series or a Maass/Hecke-eigenform g. Note that the former case is not trivial, since it requires both a subconvexity bound for $\zeta(1/2+it)$ (for instance (4.1)), and the Hadamard/de la Vallée-Poussin/Weyl bound (here the savings of the $\log\log t$ factor is necessary),

$$\frac{\zeta'}{\zeta}(1+it) \ll \frac{\log t}{\log\log t},$$

which follows from the Hadamard/de la Vallée-Poussin zero-free region. For $V = g$ a (primitive) Maass/Hecke-eigenform, one has to show that

$$\int_{X_0(1)} g(z)d\mu_t(z) = \int_{X_0(1)} |E_\infty(z, 1/2+it)|^2 g(z)\frac{dxdy}{y^2} = o_g(\log t); \tag{5.13}$$

by the unfolding method, one has

$$\int_{X_0(1)} g(z)d\mu_t(z) = 2\pi^{-2it}\frac{|\Gamma(\frac{1+2it}{4})|^2}{|\zeta(1+2it)|^2}\frac{\Gamma(\frac{1-2it_g-4it}{4})\Gamma(\frac{1+2it_g-4it}{4})}{|\Gamma(\frac{1+2it}{2})|^2}L(g,1/2)L(g,1/2-it).$$

By Stirling's formula and the lower bound $\zeta(1+2it) \gg (\log t)^{-1}$ (which again is a consequence of the zero-free region), the latter is bounded by

$$\ll_{\varepsilon,g} t^{\varepsilon-1/2}L(g, 1/2-it).$$

Hence, any subconvexity exponent for $L(g, 1/2-it)$ in the t-aspect, is sufficient to prove (5.13). This case of the **ScP** was solved by Meurman [**Me**] with the subconvex exponent $1/3$:

$$L(g, 1/2-it) \ll_{\varepsilon,g} |t|^{1/3+\varepsilon}.$$

□

The determination of quantum limits of cuspidal Hecke-eigenforms is deeper. In [**BL**], Bourgain and Lindenstrauss showed that the Hausdorff dimension of the support of these quantum limits is at least $1+2/9$ by mixing combinatoric techniques[3] together with number theoretic methods (interestingly the identity $|\lambda_f(p^2)-\lambda_f(p)^2| = 1$ for $p \not| q_f$ that was been used in Section 4.2.1 for the construction of amplifiers is also useful here). However, **Arithmetic QUE** can be related directly to subconvexity for a collection of L-functions: for simplicity, consider the case of the full modular curve $X_0(1)$. By the spectral decomposition (2.5) and Weyl's equidistribution criterion, it is sufficient to show that

$$\int_{X_0(1)} g(z)d\mu_f(z),\ \int_{X_0(1)} E_\infty(z, 1/2+it)d\mu_f(z) \to 0, \text{ as } \lambda_f\ (resp.\ k_f)\ \to +\infty,$$

for any primitive Maass form g and any $t \in \mathbf{R}$. By the unfolding method, one has

$$\int_{X_0(1)} E_\infty(z, 1/2+it)d\mu_f(z) = \frac{L_\infty(\pi_f\otimes\pi_f, 1/2+it)L(\pi_f\otimes\pi_f, 1/2+it)}{\langle f, f\rangle}.$$

By (2.18), (2.19), the expression for the local factor at ∞ is (depending on whether f is a Maass form or holomorphic):

$$L_\infty(\pi_f\otimes\pi_f, s) = \Gamma_{\mathbf{R}}(s)^2\prod_{\pm}\Gamma_{\mathbf{R}}(s\pm 2it_f),$$

or

$$L_\infty(\pi_f\otimes\pi_f, s) = \Gamma_{\mathbf{R}}(s)\Gamma_{\mathbf{R}}(s+1)\Gamma_{\mathbf{R}}(s+k-1)\Gamma_{\mathbf{R}}(s+k).$$

By Stirling's formula, one has

$$\int_{X_0(1)} E_\infty(z, 1/2+it)d\mu_f(z) \ll_{t,\varepsilon} (1+|t_f|)^{-1/2+\varepsilon}L(\pi_f\otimes\pi_f, 1/2+it).$$

[3]Recently, using ergodic theoretic methods, E. Lindenstrauss [**Lin**] has essentially proven that **Arithmetic QUE** holds for compact arithmetic surfaces and for non-compact ones, that any quantum limit is proportional to the hyperbolic measure. In his proof, the fact that the quantum limits have positive entropy is crucial.

For g a primitive cusp form, one uses Watson's formula (4.25),

$$\frac{|\int_{X_0(1))} g(z)d\mu_f(z)|^2}{\langle g,g\rangle} = \frac{\Lambda(\pi_f\otimes\pi_f\otimes\pi_g,1/2)}{\Lambda(\mathrm{sym}^2\pi_f,1)^2\Lambda(\mathrm{sym}^2\pi_g,1)}.$$

The local factors at ∞ are given by (see [**Wats**] for instance)

$$L_\infty(\mathrm{sym}^2\pi_f,s) = L_\infty(\pi_f\otimes\pi_f,s)/\Gamma_{\mathbf{R}}(s),$$
$$L_\infty(\pi_f\otimes\pi_f\otimes\pi_g,s) = \prod_{\{\pm\}^3}\Gamma_{\mathbf{R}}(s\pm it_f\pm it_f\pm it_g+\delta_g),$$

if f is a Maass form, and by

$$L_\infty(\pi_f\otimes\pi_f\otimes\pi_g,s) = \prod_{\pm}\Gamma_{\mathbf{R}}(s+k-1\pm it_g)\Gamma_{\mathbf{R}}(s+k\pm it_g)\Gamma_{\mathbf{R}}(s\pm it_g)\Gamma_{\mathbf{R}}(s+1\pm it_g)$$

if f is holomorphic. By Stirling's formula and (2.16), it then follows that

$$|\int_{X_0(1))} g(z)d\mu_f(z)|^2 \ll_{\varepsilon,g} (1+|t_f|)^{-1+\varepsilon}L(\pi_f\otimes\pi_f\otimes\pi_g,1/2),$$
$$|\int_{X_0(1))} g(z)d\mu_f(z)|^2 \ll_{\varepsilon,g} k^{-1+\varepsilon}L(\pi_f\otimes\pi_f\otimes\pi_g,1/2).$$

Hence **QUE** in this case follows from any subconvex bounds for $L(\pi_f\otimes\pi_f,1/2+it)$ and $L(\pi_f\otimes\pi_f\otimes\pi_g,1/2)$ in the t_f (resp. k_f) aspect. Similarly, **QUE** for other modular or Shimura curves would follow from an extension of the triple product identities and from some subconvex bounds for appropriate L-functions.

In the special case of f dihedral, Arithmetic QUE has been proven by Sarnak (in the holomorphic case) and by Liu/Ye (in the Maass case) [**Sa4, LY**]. Recall that a dihedral form, $f=f_\psi$, is a theta series associated –by Hecke and Maass– to a Grossencharacter ψ on a quadratic field. One has the following factorizations:

$$L(f,s)=L(\psi,s), L(\mathrm{sym}^2\pi_f,s)=L(\psi^2,s)L(\chi,s),$$
$$L(\mathrm{sym}^2\pi_f\otimes\pi_g,s)=L(\pi_{f_{\psi^2}}\otimes\pi_g,s)L(\chi.\pi_g,s),$$

where f_{ψ^2} is the theta series corresponding to ψ^2 (and thus has spectral parameter $2it_f$ —resp. weight $2k_f$—), and χ is the character corresponding to the quadratic field. In particular **QUE** follows in this case from the **ScP** for Hecke and Rankin-Selberg L-functions in the spectral aspect (Theorem 4.10).

The above instances of the **ScP** seem to be hard to prove in general within the framework of GL_2 analysis. At least one gets some simplifications via the factorizations

$$L(\pi_f\otimes\pi_f,s)=\zeta(s)L(\mathrm{sym}^2\pi_f,s), \tag{5.14}$$
$$L(\pi_f\otimes\pi_f\otimes\pi_g,s)=L(\pi_g,s)L(\mathrm{sym}^2\pi_f\otimes\pi_g,s);$$

in particular the **ScP** would follow (via the Gelbart/Jacquet lift) from the more general **ScP** for GL_3-L-functions and the $GL_3\times GL_2$-Rankin/Selberg L-fonctions. To achieve this goal, one must probably first develop a good analytic theory of families of GL_3 automorphic forms, similar to the GL_2 theory; in particular one needs a manageable trace type formula. Some analogs of the Petersson/Kuznetsov formula have been given for GL_3 [**BFG**], but so far they seem difficult to use effectively for refined analytic purposes.

5.4.2. QUE on the sphere

It is also possible to formulate an analog of **QUE** for the 2-dimensional sphere S^2, although the geodesic flow is not ergodic. Recall that the spectrum of Δ on S^2 is $0 < 1 < 3 < \cdots < k(k+1)/2 \ldots$, and that the corresponding eigenspace is the space of harmonic polynomials of degree k — this space has dimension $2k+1$, which is rather large. In this case, it is possible to prove a probabilistic version of **QUE**: namely, **QUE** holds for almost all orthonormal eigenbasis of Δ. However, as for the holomorphic case above, the **QUE** conjecture cannot be true in its most naive form. Recently, Böcherer/Schulze-Pillot/Sarnak formulated a deterministic version of **QUE** for the sphere [**BSSP**]: S^2 is viewed as the symmetric space attached to a definite quaternion algebras, D say,with class number one. As such, S^2 is endowed with an action of the Hecke algebra of $D^\times$ that commutes with Δ; thus the **QUE** conjecture for the sphere is: for φ_k a harmonic polynomial of degree $k \to +\infty$, which is also a Hecke-eigenform,

$$\mu_{\varphi_k} := \frac{|\varphi_k(P)|^2}{\int_{S^2} |\varphi_k(P)|^2 d\mu_{S^2}} \mu_{S^2} \to \frac{1}{\mathrm{vol}(S^2)} \mu_{S^2} \text{ weakly.}$$

Again, this conjecture can be reduced – via triple product identities due to Böcherer-/Schulze-Pillot [**BS**] – to the **ScP** in the k aspect for the L-functions (5.14), where f and g are the holomorphic primitive forms corresponding, respectively, to φ_k and to some harmonic (Hecke eigen) polynomial via the Jacquet/Langlands correspondence between GL_2 and $D^\times$.

5.4.3. The Random Wave Conjecture for Hecke-eigenforms

Another famous conjecture in quantum chaos is the Random Wave Conjecture of Berry and Hejhal. This conjecture predicts that φ_λ, viewed as a random variable on X, converges to the normal Gaussian distribution $N(0, 1/\mathrm{vol}(X)^{1/2})$ as $\lambda \to +\infty$. This is equivalent to the following moments conjecture:

Conjecture. *For any integer $m \geqslant 0$,*

$$\lim_{\lambda \to +\infty} \int_X (\varphi_\lambda)^m dx = \frac{c_m}{\sqrt{\mathrm{vol}(X)}^{m/2}},$$

where c_m is the m-th moment of the Gaussian distribution $N(0,1)$.

The cases $m = 1, 2$ are obvious. The first non-trivial case is the third moment for which $c_3 = 0$: it has been solved by Watson for $X = X_0(1)$ and $\varphi = f/\langle f, f\rangle^{1/2}$ for f a primitive Maass form, as a consequence of his triple product identity [**Wat**] and a subconvex estimate for $L(f, 1/2)$ in the spectral aspect due to Iwaniec.

Theorem 5.6. *For f a primitive Maass form on $X_0(1)$, one has*

$$\frac{3}{\pi} \int_{X_0(1)} \frac{(f(z))^3}{\langle f, f\rangle^{3/2}} \frac{dxdy}{y^2} = o(1)$$

as $t_f \to +\infty$.

Proof. By (4.25), Stirling's formula, and the value of the factors at ∞, one has

$$|\frac{3}{\pi} \int_{X_0(1)} \frac{f^3(z)}{\langle f, f\rangle^{3/2}} \frac{dxdy}{y^2}|^2 \ll \frac{\Lambda(\pi_f \otimes \pi_f \otimes \pi_f, 1/2)}{\Lambda(\mathrm{sym}^2\pi_f, 1/2)^3} \ll_\varepsilon |t_f|^\varepsilon |t_f|^{-3/2+\varepsilon} L(\pi_f \otimes \pi_f \otimes \pi_f, 1/2).$$

One has the factorization

$$L(\pi_f \otimes \pi_f \otimes \pi_f, s) = L(\mathrm{sym}^3\pi_f, s)L(\pi_f, s)^2.$$

By [**KiSh**], $L(\mathrm{sym}^3\pi_f, s)$ is the L-function of a cusp form on GL_4, hence the convexity bound holds ,

$$L(\mathrm{sym}^3\pi_f, 1/2) \ll_\varepsilon (1 + |t_f|)^{1+\varepsilon}.$$

Thus one finds that

$$\int_{X_0(1)} \varphi^3(z)\frac{dxdy}{y^2} \ll_\varepsilon (1 + |t_f|)^{-1+\varepsilon} L(\pi_f, 1/2)^2;$$

once again, any subconvex bound for $L(\pi_f, 1/2)$ in the spectral aspect is sufficient to show that $\frac{3}{\pi}\int_{X_0(1)} \phi^3(z)\frac{dxdy}{y^2} \to 0$. The first estimate of this kind was obtained by Iwaniec [**I3**] with the subconvex exponent $1/2 - 1/12$, and it was improved subsequently by Ivic to $1/2 - 1/6$ [**Iv2**].

□

Sarnak and Watson have addressed the problem of the fourth moment for modular curves. Using the triple product identities, subconvex estimates and a Voronoi formula for GL_3-forms [**MiS**], they show (assuming **RPC** at the moment) that the fourth moment conjecture holds for dihedral forms, and that for non-dihedral forms one has

$$|\frac{3}{\pi}\int_{X_0(1)} \frac{|f(z)|^4}{\langle f, f\rangle^2}\frac{dxdy}{y^2}|^2 \ll_\varepsilon (1 + |t_f|)^\varepsilon$$

for any $\varepsilon > 0$.

5.4.4. QUE in the level aspect

Arithmetic QUE has a natural generalization to the level aspect. To fix ideas, consider $f \in S_2^p(q)$ a primitive holomorphic form of level q and fixed weight 2, and the associated probability measure on $X_0(q)$:

$$\mu_f(z) := \frac{|f(z)|^2}{\langle f, f\rangle} y^{k_f}\frac{dxdy}{y^2}.$$

Denote by $\pi_q : X_0(q) \to X_0(1)$ the natural projection induced by the inclusion $\Gamma_0(q) \subset \Gamma_0(1)$.

Conjecture. *For $f \in S_2^p(q)$, the probability measure $\mu_{f,1} := \pi_{q*}\mu_f$ (the direct image of μ_f by π_q) weakly converges to $\mu_P = \frac{1}{\mathrm{vol}(X_0(1))}\frac{dxdy}{y^2}$ as $q \to +\infty$.*

Remark 5.9. One can consider more general versions of **QUE** in the level aspect; for instance, one can pull back μ_f to some $X_0(qq')$ and then push forward the resulting measure to $X_0(q')$. The conjecture is that the image of μ_f weakly converges to the Poincaré measure on $X_0(q')$.

When q is squarefree at least, one can check that this conjecture follows from **ScP** for the L-functions (5.14) in the level aspect.

Analogously to remark 5.8, this conjecture would have the pleasing consequence that the projections of the zeros of f by π_q become equidistributed on $X_0(1)$, relative to μ_P. A suggestive corollary is that the multiplicity of any such zero is $o(q)$ as $q \to +\infty$. this improves over the Riemann/Roch theorem, which bounds the multiplicity by $O(q\prod_{p|q}(1 + 1/p))$. When $f = f_E$ corresponds to an elliptic curve, the zeros are the ramification points of the modular parametrization E, and have multiplicity negligible by comparison with the conductor. Such corollaries would certainly be meaningful in connection with the *abc*-conjecture, as very little is known about the ramification divisor of a strong Weil curve (it is not even clear that the ramification cannot be concentrated on one point).

BIBLIOGRAPHY

[A] Arenas, A. *An arithmetic problem on the sums of three squares.* Acta Arith. 51 (1988), no. 2, 131–140.

[AB] Arenas, A., Bayer, P.,*Arithmetic behavior of the sums of three squares.* J. Number Theory 27 (1987), no. 3, 273–284.

[AC] J. Arthur and L. Clozel, *Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula*, Ann. of Math. Stud. 120, Princeton Univ. Press, Princeton, 1989.

[AL] A. Atkin, J. Lehner: *Hecke operators on* $\Gamma_0(N)$, Math. Ann. 185, (1970) 134-160.

[ALi] A. Atkin, W. Li: *Twists of newforms and pseudo-Eigenvalues of* W*-operators*, Invent. Math. 48 (1978), no. 3, 221–243.

[Ar] S. Arno, M. L. Robinson, and F. S. Wheeler, *Imaginary Quadratic Fields with Small Odd Class Number,* Acta Arith. 83 (1998), no. 4, 295-330.

[AU] A. Abbes, E. Ullmo *Comparaison des métriques d'Arakelov et de Poincaré sur* $X_0(N)$, Duke Math. J. 80 (1995), no. 2, 295–307.

[Ban] Banks, W. D., *Twisted symmetric-square* L*-functions and the nonexistence of Siegel zeros on* $\mathrm{GL}(3)$. Duke Math. J. 87 (1997), no. 2, 343–353.

[Bar] Barban, M.B.*The "Large Sieve" method and its applications in the theory of numbers.* (Russian, English) Russ. Math. Surv. 21, No.1, 49-103 (1966); translation from Usp. Mat. Nauk 21, No.1(127), 51-102 (1966)

[BR] Barthel, Laure; Ramakrishnan, Dinakar *A nonvanishing result for twists of* L*-functions of GL*(n). Duke Math. J. 74, No.3, 681-700 (1994).

[BM] Baruch, E.; Mao, Z., *Central value of automorphic* $L-$*functions*. Preprint http://arXiv.org/abs/math/0301115.

[BD1] M. Bertolini, H. Darmon, *Heegner points of Mumford-Tate curves* Invent. Math. 126 (1996), 413-456.

[BD2] M. Bertolini, H. Darmon, *A rigid analytic Gross-Zagier formula and arithmetic applications* With an appendix by Bas Edixhoven. Ann. of Math. (2) 146 (1997), no. 1, 111–147.

[BD3] M. Bertolini, H. Darmon, *Iwasawa's Main Conjecture for elliptic curves over anticyclotomic* $\mathbf{Z}_p$*-extensions*, preprint 2001.

[BR] Bernstein, J.; Reznikov, A. *Analytic continuation of representations and estimates of automorphic forms.* Ann. Math. (2) 150, No.1, 329-352 (1999).

[BR2] Bernstein, J.; Reznikov, A. *Estimates of automorphic functions* Mosc. Math. J. **4** (2004), no. 1, 19-37, 310.

[BR3] Bernstein, J.; Reznikov, A. *Periods, subconvexity and representation theory*, J. of Diff. Geom. 70 (2005), no. 1, 129-141.

[Bo] Bombieri, E.: *Le grand crible dans la théorie analytique des nombres*, Second edition. Asterisque No. 18 (1987), 103 pp.

[BFI1] Bombieri, E.; Friedlander, J. B.; Iwaniec, H. *Primes in arithmetic progressions to large moduli.* Acta Math. 156 (1986), no. 3-4, 203–251.

[BFI2] Bombieri, E.; Friedlander, J. B.; Iwaniec, H. *Primes in arithmetic progressions to large moduli II.* Math. Ann. 277 (1987), no. 3-4, 361-393.

[BFI3] Bombieri, E.; Friedlander, J. B.; Iwaniec, H. *Primes in arithmetic progressions to large moduli. III.* J. Amer. Math. Soc. 2 (1989), no. 2, 215–224.

[BS] Bocherer, S.; Schulze-Pillot, R. *On the central critical value of the triple product L-function.* David, Sinnou (ed.), Number theory. Seminaire de Theorie des Nombres de Paris 1993–94. Cambridge: Cambridge University Press. Lond. Math. Soc. Lect. Note Ser. 235, 1-46 (1996).

[BSSP] Böcherer, S., Sarnak, P. and Schulze-Pillot, R. *Arithmetic and equidistribution of measures on the sphere*, Comm. Math. Phys. 242 (2003), no. 1-2, 67–80.

[BL] Bourgain, J., Lindenstrauss, E., *Entropy of quantum limits*, Comm. Math. Phys. 233 (2003), no. 1, 153–171.

[BCDT] Breuil, C., Conrad, B, Diamond, F., Taylor, R.: *On the modularity of elliptic curves over* $\mathbf{Q}$, J. Am. Math. Soc. 14, No.4, 843-939 (2001).

[BH] Bushnell, C. J. and Henniart, G.: *An upper bound on conductors for pairs*, J. Number Theory 65 no. 2 (1997), 183–196.

[Bu] Burgess, D. A.*On character sums and L-series.* Proc. London Math. Soc. (3) 12 (1962) 193–206.

[BFG] Bump, D., Friedberg, S. and Goldfeld, D., *Poincare series and Kloosterman sums for* $\mathrm{SL}(3, Z)$. Acta Arith. 50 (1988), no. 1, 31–89.

[CMP] Carletti, E., Monti Bragadin, G., Perelli, A. *On general L-functions.* Acta Arith. 66 (1994), no. 2, 147–179.

[Co1] Cogdell, J.,*On sums of three squares*, Les XXIIèmes Journées Arithmetiques (Lille, 2001). J. Théor. Nombres Bordeaux 15 (2003), no. 1, 33–44.

[Co2] Cogdell, J., *L-functions and Converse Theorems for GL_n*, this volume.

[CoPS] Cogdell, J., I.I. Piatetskii-Shapiro, *The arithmetic and spectral analysis of Poincare series.* Perspectives in Mathematics, 13. Academic Press, Inc., Boston, MA, 1990. vi+182 pp.

[CoPSS] Cogdell, J., I.I. Piatetskii-Shapiro, Sarnak, P. . In preparation.

[ClU] Clozel, L., Ullmo, E.,*Equidistribution des points de Hecke*. Contributions to automorphic forms, geometry, and number theory, 193–254, Johns Hopkins Univ. Press, Baltimore, MD, 2004.

[C-V] Colin de Verdière, Comm. Math. Phys., 102 (1985), 497-502.

[ConI] Conrey, B., Iwaniec, H., *The cubic moment of central values of automorphic L-functions*, Annals of Math. 151 (2000), 1175-1216.

[ConS] J.B. Conrey, K. Soundararajan: *Real zeros of quadratic Dirichlet L-functions* Invent. math. 150 (2002) 1, 1-44.

[Cor] Cornut, C. *Mazur's conjecture on higher Heegner points.* Invent. Math. 148, No.3, 495-523 (2002).

[Da] Daghigh, H. *Modular forms, quaternion algebras, and special values of L-functions*, PhD thesis Mac Gill, Spring 1998.

[Dav] Davenport, H. *Multiplicative Number Theory 2nd. Edition*. Springer.

[De] Deligne, P., *Formes modulaires et representations ℓ-adiques* Semin. Bourbaki 1968/69, No.355, 139-172 (1971).

[De2] P. Deligne.*La conjecture de Weil. I*. Inst. Hautes Études Sci. Publ. Math. No. 43, (1974), 273–307.

[De3] Deligne, P., *La conjecture de Weil II*, Publ.Math.IHES 52 (1981), 313-428.

[De4] Deligne, P., *Cohomologie etale. Seminaire de Geometrie Algebrique du Bois-Marie SGA $4\frac{1}{2}$*. Lecture Notes in Mathematics, Vol. 569. Springer-Verlag, Berlin-New York, 1977. iv+312pp.

[DS] Deligne, P., Serre, J.-P., *Formes modulaires de poids 1.* Ann. Sci. Ec. Norm. Super., IV. Ser. 7, 507-530 (1974).

[DI] Deshouillers, J.-M., Iwaniec, H. *Kloosterman sums and Fourier coefficients of cusp forms.* Invent. Math. 70 (1982/83), no. 2, 219–288.

[DIPS] Deshouillers, J.-M.; Iwaniec, H.; Phillips, R. S.; Sarnak, P. *Maass cusp forms.* Proc. Nat. Acad. Sci. U.S.A. 82 (1985), no. 11, 3533–3534.

[Du] Duke, W. Hyperbolic distribution problems and half-integral weight Maass forms. Invent. Math. 92 (1988), no. 1, 73–90.

[Du2] Duke, W. *The dimension of the space of cusp forms of weight one.* Internat. Math. Res. Notices 1995, no. 2, 99–109.

[DuI] Duke, W.; Iwaniec, H. *Estimates for coefficients of L-functions. II.* Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), 71–82, Univ. Salerno, Salerno, 1992.

[DuI2] Duke, W.; Iwaniec, H. *Estimates for coefficients of L-functions. IV.* Amer. J. Math. 116 (1994), no. 1, 207–217.

[DuI3] Duke, W.; Iwaniec, H. *Convolution L-series.* Compositio Math. 91 (1994), no. 2, 145–158.

[DFI1] Duke, W., Friedlander, J., Iwaniec, H. *Bounds for automorphic L-functions.* Invent. Math. 112 (1993), no. 1, 1–8.

[DFI2] Duke, W., Friedlander, J., Iwaniec, H. *A quadratic divisor problem.* Invent. Math. 115 (1994), no. 2, 209–217.

[DFI3] Duke, W., Friedlander, J. et Iwaniec, H.: *Bounds for automorphic L-functions, II*, Invent. Math. 115 (1994), 219–239.

[DFI4] Duke, W., Friedlander, J., Iwaniec, H. *Class group L-functions.* Duke Math. J. 79 (1995), no. 1, 1–56.

[DFI5] Duke, W., Friedlander, J., Iwaniec, H. *Bilinear forms with Kloosterman fractions.* Invent. Math. 128 (1997), no. 1, 23–43.

[DFI6] Duke, W., Friedlander, J., Iwaniec, H. *Representations by the determinant and mean values of L-functions.* Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), 109–115, London Math. Soc. Lecture Note Ser., 237, Cambridge Univ. Press, Cambridge, 1997.

[DFI7] Duke, W., Friedlander, J. et Iwaniec, H.: *Bounds for automorphic L-functions, III*, Invent. Math. 143, No.2, 221-248 (2001).

[DFI8] Duke, W., Friedlander, J. et Iwaniec, H.: *The subconvexity problem for Artin L-functions*, Invent. math. 149 (2002) 3, 489-577.

[DK] Duke, W. and Kowalski, E.: *A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations. With an appendix* *by Dinakar Ramakrishnan.*, Invent. math. 139 (2000), 1–39.

[DSP] Duke, W.; Schulze-Pillot, R. *Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids.* Invent. Math. 99, No.1, 49-57 (1990).

[El] Ellenberg, J.; *On the average number of octahedral modular forms* Math. Res. Lett. 10 (2003), no. 2-3, 269–273.

[Fo1] Fouvry, É.; *Autour du theoreme de Bombieri-Vinogradov.* Acta Math. 152, 219-244 (1984).

[Fo2] Fouvry, É.; *Autour du theoreme de Bombieri-Vinogradov.* Ann. Scient. Ec. Norm. Sup. (4) (20) (1987), 617-640.

[FoI1] Fouvry, É.; Iwaniec, H. *On a theorem of Bombieri-Vinogradov type.* Mathematika 27 (1980), 135–152.

[FoI2] Fouvry, É.; Iwaniec, H. *Primes in arithmetic progressions.* Acta Arithmetica 42 (1983), 197-218.

[FoI3] Fouvry, É.; Iwaniec, H. *A subconvexity bound for Hecke L-functions.* Ann. Sci. Ecole Norm. Sup. (4) 34 (2001), no. 5, 669–683.

[Fr] Friedlander, J. B., *Bounds for L-functions.* Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), 363–373, Birkhäuser, Basel, 1995.

[Ge] S. Gelbart *Automorphic forms on adele groups.* Annals of Mathematics Studies. No.83. Princeton, N. J.: Princeton University Press and University of Tokyo Press. X, 267 p. $ 9.00 (1975).

[GeJ1] S. Gelbart and H. Jacquet,*A relation between automorphic representations of* $\mathrm{GL}(2)$ *and* $\mathrm{GL}(3)$, Ann. Sci. Ecole Norm. Sup. (4) 11 (1978), 471–552.

[GeSh] S. Gelbart, Shahidi, F., *Boundeness of automorphic L-functions in vertical strips*, J. Amer. Math. Soc. 14 (2001), 79-107.

[GoJ] Godement, R.; Jacquet, H., *Zeta functions of simple algebras.* Lecture Notes in Mathematics. 260. Berlin-Heidelberg-New York: Springer-Verlag. VIII, 188 p. (1972).

[Go1] Goldfeld, D., *A simple proof of Siegel's theorem.* Proc. Nat. Acad. Sci. U.S.A. 71 (1974), 1055.

[Go2] Goldfeld, D., *The class number of quadratic fields and the conjecture of Birch and Swinnerton-Dyer.* Ann. Sc. norm. super. Pisa, Cl. Sci., IV. Ser. 3, 623-663 (1976).

1. Go3 Goldfeld, D.,*The Gauss class number problem for imaginary quadratic fields.* Heegner points and Rankin L-series, 25–36, Math. Sci. Res. Inst. Publ., 49, Cambridge Univ. Press, Cambridge, 2004.

[GH] Goldfeld, D.; Hoffstein, J. *On the number of Fourier coefficients that determine a modular form.* [CA] Knopp, Marvin (ed.) et al., A tribute to Emil Grosswald: number theory and related analysis. Providence, RI: American Mathematical Society. Contemp. Math. 143, 385-393 (1993). [ISBN 0-8218-5155-1/pbk]

[GHL] Goldfeld, D., Hoffstein J., Lieman, D., *Appendix: An effective zero-free region* Ann. Math. (2) 140, No.1, 177-181 (1994).

[GL] Goldfeld, D., Lieman, D., *Effective bounds on the size of the Tate-Shafarevich group.* Math. Res. Lett. 3 (1996), no. 3, 309–318.

[GS] Goldfeld, Dorian; Szpiro, Lucien *Bounds for the order of the Tate-Shafarevich group.* Compos. Math. 97, No.1-2, 71-87 (1995).

[Go] Good, A. *Cusp forms and eigenfunctions of the Laplacian.* Math. Ann. 255 (1981), no. 4, 523–548.

[GR] Graham, S.W.; Ringrose, C.J. *Lower bounds for least quadratic non-residues.* Analytic number theory, Proc. Conf. in Honor of Paul T. Bateman, Urbana/IL (USA) 1989, Prog. Math. 85, 269-309 (1990).

[GrS] Granville, A., Soundararajan K.: *The distribution of value of* $L(1,\chi)$, Geom. Funct. Anal. 13 (2003), no. 5, 992–1028.

[Gr] B. Gross, *Heights and the special values of* L*-series*. Number theory (Montreal, Que., 1985), 115–187, CMS Conf. Proc. 7, Amer. Math. Soc., Providence, RI, 1987.

[GK] Gross, Benedict H.; Kudla, Stephen S. *Heights and the central critical values of triple product* L*-functions.* Compos. Math. 81, No.2, 143-209 (1992).

[GZ] Gross, B. and Zagier, D. *Heegner points and derivatives of* L*-series*, Invent. Math., 84 (1986), 225–320.

[H] Hadamard, J., Hadamard, J. *Sur la distribution des zeros de la fonction* $\zeta(s)$ *et ses consequences arithmetiques.* S. M. F. Bull. 24, 199-220. (1896)

[Ha1] Harcos, G. *Uniform approximate functional equation for principal L-functions.* Int. Math. Res. Not. 2002, no. 18, 923–932.

[Ha2] Harcos, G.,*An additive problem in the Fourier coefficients of Maass forms,* Math. Ann. 326 (2003), no. 2, 347–365.

[HM] Harcos, G., Michel, Ph., *The subconvexity problem for Rankin-Selberg* L*-functions and equidistribution of Heegner points. II.* Invent. Math. 163 (2006), no. 3, 581–655.

[HK] Harris, Michael; Kudla, Stephen S. *The central critical value of a triple product L-function.* Ann. Math. (2) 133, No.3, 605-672 (1991).

[HB] D. R. Heath-Brown. A mean value estimate for real character sums. Acta Arith. 72 (1995), no. 3, 235–275.

[HB-M] D. R. Heath-Brown and Ph. Michel *Exponential in the frequency of analytic ranks of Automorphic* L*-functions*, Duke Mathematical Journal 102, 3, p. 475-484 (2000).

[HL] Hoffstein, J., Lockhart, P., *Coefficients of Maass forms and the Siegel zero. With an appendix by Dorian Goldfield, Hoffstein and Daniel Lieman.* Ann. of Math. (2) 140 (1994), no. 1, 161–181.

[HR] Hoffstein, J.; Ramakrishnan, D. *Siegel zeros and cusp forms.* Internat. Math. Res. Notices 1995, no. 6, 279–308.

[Hu] Huxley, M., Oberwolfach lecture, fall 2001.

[HJ] Huxley, M. N.; Jutila, M. *Large values of Dirichlet polynomials. IV.* Acta Arith. 32 (1977), no. 3, 297–312.

[Iv] Ivić, A. *The Riemann zeta-function. The theory of the Riemann zeta-function with applications.* A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1985. xvi+517 pp. ISBN: 0-471-80634-X

[Iv2] Ivić, A. *On sums of Hecke series in short intervals.* J. Théor. Nombres Bordeaux 13 (2001), no. 2, 453–468.

[I1] H. Iwaniec, *Fourier coefficients of modular forms of half-integral weight.* Invent. Math. 87 (1987), no. 2, 385–401.

[I2] H. Iwaniec, *On the order of vanishing of modular* L*-functions at the critical point.* Sém. Théor. Nombres Bordeaux (2) 2 (1990), no. 2, 365–376.

[I3] H. Iwaniec, *The spectral growth of automorphic L-functions.* J. Reine Angew. Math. 428 (1992), 139–159.

[I4] H. Iwaniec, *Introduction to the spectral theory of automorphic forms.* Biblioteca de la Revista Matemática Iberoamericana. Revista Matemática Iberoamericana, Madrid, 1995. xiv+247 pp.

[I5] H. Iwaniec,*Topics in classical automorphic forms.* Graduate Studies in Mathematics, 17. American Mathematical Society, Providence, RI, 1997. xii+259 pp. ISBN: 0-8218-0777-3

[I6] H. Iwaniec, *Harmonic analysis in number theory.* Prospects in mathematics (Princeton, NJ, 1996), 51–68, Amer. Math. Soc., Providence, RI, 1999.

[ILS] H. Iwaniec, W. Luo, P. Sarnak,*Low Lying zeros of families of L-functions*, Publ. Math. IHES, No. 91, (2000), 55–131 (2001).

[IM] Iwaniec, H.; Michel, Ph. *The second moment of the symmetric square L-functions.* Ann. Acad. Sci. Fenn. Math. 26 (2001), no. 2, 465–482.

[IS1] H. Iwaniec, P. Sarnak, L^∞ norms of eigenfunctions of arithmetic surfaces. Ann. of Math. (2) 141 (1995), no. 2, 301–320.

[IS2] H. Iwaniec, P. Sarnak, *Perspectives in the Analytic Theory of L functions*, GAFA, Special Volume GAFA 2000, 705–741.

[IS3] H. Iwaniec, P. Sarnak, *The non-vanishing of central values of automorphic L-functions and Landau-Siegel zeros.* Israel J. Math. 120 (2000), part A, 155–177.

[J] H. Jacquet, *Automorphic forms of $GL(2)$*, Lect. Notes Math. 278 (1972).

[JS1] H. Jacquet, J. A. Shalika,*A non-vanishing theorem for zeta-functions of GL_n,* invent. Math. 38 (1976), 1–16.

[JS2] H. Jacquet, J. A. Shalika,*On Euler products and the classification of automorphic representations. I& 2,* Amer. J. Math. 103 (1981), no. 3, 499–558& 777-815.

[JPSS1] H. Jacquet, I.I. Piatetskii-Shapiro, J. A. Shalika, *Automorphic forms on $GL_3 i \& II$,* Ann. Math. 109 (1979), 169-258.

[JPSS2] H. Jacquet, I.I. Piatetskii-Shapiro, J. A. Shalika, *Rankin–Selberg convolutions,* Amer. J. Math. 105 (1983), no. 2, 367–464.

[Ju1] Jutila, M. A variant of the circle method. Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), 245–254, London Math. Soc. Lecture Note Ser., 237, Cambridge Univ. Press, Cambridge, 1997.

[Ju2] Jutila, M. *The additive divisor problem and its analogs for Fourier coefficients of cusp forms. I.* Math. Z. 223, No.3, 435-461 (1996).

[Ju3] Jutila, M. *The additive divisor problem and its analogs for Fourier coefficients of cusp forms. II.* Math. Z. 225 (1997), no. 4, 625–637.

[Ju4] Jutila, M. *The fourth moment of central values of Hecke series.* [CA] Jutila, Matti (ed.) et al., Number theory. Proceedings of the Turku symposium on number theory in memory of Kustaa Inkeri, Turku, Finland, May 31-June 4, 1999. Berlin: de Gruyter. 167-177 (2001). [ISBN 3-11-016481-7/hbk]

[JM] Jutila, M.; Motohashi, Y. *Mean value estimates for exponential sums and L-functions: a spectral-theoretic approach.* J. Reine Angew. Math. 459 (1995), 61–87.

[Ki] Kim, H.,*Functoriality for the exterior square of GL_4 and symmetric fourth of GL_2, Appendix 1 by Dinakar Ramakrishnan; Appendix 2 by Henry H. Kim and Peter Sarnak*, J. Amer. Math. Soc. 16 (2003), 139-183.

[KiSa] Kim, H., Sarnak, P.,*Appendix: refined estimates towards the Ramanujan and Selberg Conjectures*, J. Amer. Math. Soc. 16 (2003).

[KiSh] Kim, H., Shahidi. F., *Functorial products for $GL_2 \times GL_3$ and functorial symmetric cube for GL_2. (with an appendix by C. J. Bushnell and G. Henniart)* Annals of Math. Volume 155, No. 3, 837-893 (2002).

[KiSh2] Kim, H., Shahidi. F.,*Cuspidality of symmetric powers with applications,* Duke Math. J. 112 (2002), no. 1, 177–197.

[Ko] Korobov, N. M.*Estimates of trigonometric sums and their applications.* (Russian) Uspehi Mat. Nauk 13 1958 no. 4 (82), 185–192.

[Kow] E. Kowalski, Dependency on the group in automorphic Sobolev inequalities Forum Math. (to appear).

[KrS] B. Kroetz, R. J. Stanton *Holomorphic extension of representations: (I) automorphic functions* . Ann. of Math. (2) 159 (2004), no. 2, 641–724.

[KM1] E. Kowalski, Ph. Michel, *Sur les zéros des fonctions L automorphes de grand niveau*, Preprint de l'Université d' Orsay 54 (1997).

[KM2] E. Kowalski, Ph. Michel, *The analytic rank of $J_0(q)$ and zeros of automorphic L-functions*, Duke Mathematical Journal 100, 3 (1999), p. 503-542 .

[KM3] E. Kowalski, Ph. Michel, *Zeros of families of Automorphic L-functions close to 1*, Pacific Journal of Mathematics 207 (2002), 411-431.

[KMV1] E. Kowalski, Ph. Michel, J. VanderKam, *Mollification of the fourth moment of automorphic L-functions and arithmetic applications*, Inventiones Mathematicae, 142, 1, p.95-151 (2000).

[KMV2] Kowalski, E., Michel, Ph., Vanderkam, J.: *Rankin-Selberg L functions in the level aspect*, Duke Math. J. 114 (2002), no. 1, 123–191.

[La] Landau, E., *Uber die Anzahl der Gitterpunkte in gewissen Bereichen. (Zweite Abhandlung.)* Gott. Nachr. 1915, 209-243. Published: 1915. and Collected works. Ed. by P. T. Bateman, L. Mirsky, H. L. Montgomery, W. Schaal, I. J. Schoenberg, W. Schwarz, H. Wefelscheid. Volume 6. [B] Essen: Thales Verlag. 512. p. DM 224.00 (1986). [ISBN 3-88908-215-7]

[Lin] Lindenstrauss, E., *Invariant measures and arithmetic quantum unique ergodicity.* Ann. of Math. (2) 163 (2006), no. 1, 165–219.

[Li] W. Li, *Newforms and functional equations.*, Math Ann. 212 (1975), 285–315.

[Li2] W. Li, *L-series of Rankin Type and their functional equations*, Math Ann. 244 135-166(1979).

[Lin1] Linnik, Yu.V. *The large sieve.* C. R. (Dokl.) Acad. Sci. URSS, n. Ser. 30, 292-294 (1941).

[Lin2] Linnik, Yu. V.: *On the least prime in an arithmetic progression, I, II*, Rec. Math (Mat. Sbornik) N.S. 15 (57) (1944), 139–178 and 347–368.

[Lin3] Linnik, Y. V., *Ergodic properties of algebraic fields*, ergebnisse Math, 45 Springer (1968).

[LY] Liu, J., Ye, Y.,*Subconvexity for Rankin-Selberg L-functions for Maass forms*, Geom. Funct. Anal. 12 (2002), no. 6, 1296–1323.

[Lu] Luo, W. *On the nonvanishing of Rankin-Selberg L-functions.* Duke Math. J. 69 (1993), no. 2, 411–425.

[Lu2] Luo, W. *Nonvanishing of L-values and the Weyl law.* Ann. Math. (2) 154, No.2, 477-502 (2001).

[Lu3] Luo, W.*Values of symmetric L-functions at 1*, J. Reine Angew. Math. 506, (1999), 215–235.

[LS] W. Luo, P. Sarnak, *Quantum ergodicity of eigenfunctions on $SL_2(\mathbf{Z})\backslash\mathbf{H}$*, Publ. Math. IHES 81 (1995), 207-237.

[LRS] W. Luo, Z. Rudnick, and P. Sarnak, *On Selberg's eigenvalue conjecture.* Geom. Funct. Anal. 5 (1995), no. 2, 387–401.

[LRS2] W. Luo, Z. Rudnick, and P. Sarnak, *On the generalized Ramanujan conjecture for* $\mathrm{GL}(n)$ in Automorphic Forms, Automorphic Representations, and Arithmetic (Fort Worth, Tex., 1996), Proc. Sympos. Pure Math. 66, Part 2, Amer. Math. Soc., Providence, 1999, 301–310.

[Me] T. Meurman, On the order of the Maass L-function on the critical line. [CA] Number theory. Vol. I. Elementary and analytic, Proc. Conf., Budapest/Hung. 1987, Colloq. Math. Soc. Janos Bolyai 51, 325-354 (1990).

[Mi] Ph. Michel, *The subconvexity problem for Rankin-Selberg L functions with nebentypus and equidistribution of Heegner points*, Ann. of Math. (2) 160 (2004), no. 1, 185–236.

[MiV] Ph. Michel, A. Venkatesh *On the dimension of the space of cusp forms associated to 2-dimensional complex Galois representations.* International Math. Research Notices. 2002:38 (2002) 2021-2027.

[MiS] S. Miller, and W. Schmid, *Automorphic Distributions, L-functions, and Voronoi Summation for $GL(3)$*, Annals of Math.(to appear).

[MW1] C. Moeglin, J.-L. Waldspurger,*Le spectre résiduel de $GL(n)$,* Ann. Sci. Ecole Norm. Sup. (4), 22 (1989), 605-674.

[MW2] C. Moeglin, J.-L. Waldspurger,*Spectral Decomposition and Eisenstein Series: Une paraphrase de l'Ecriture,* Cambridge Tracts in Math. 113, Cambridge Univ. Press, Cambridge, 1995.

[Mol] Molteni, G.,*Upper and lower bounds at s = 1 for certain Dirichlet series with Euler product* , Duke Math. J., Volume 111 Issue 1, 133-158, 2001.

[Mon] Montgomery, Hugh L.*Topics in Multiplicative Number Theory.* Lecture Notes in Mathematics, Vol. 227. Springer-Verlag, Berlin-New York, 1971. ix+178 pp.

[MoVa] Montgomery, H.L., Vaughan, R.: *Extreme values of Dirichlet L-functions at 1*, Number Theory in Progress (edts. Gyory, K., Iwaniec, H., Urbanowicz, J.), de Gruyter, Berlin, 1999.

[Mor1] Moreno, Carlos J. The method of Hadamard and de la Vallee-Poussin (according to Pierre Deligne). Enseign. Math. (2) 29 (1983), no. 1-2, 89–128.

[Mor2] Moreno, Carlos J. *Analytic proof of the strong multiplicity one theorem.* Amer. J. Math. 107 (1985), no. 1, 163–206.

[Mor3] Moreno, Carlos J. *Some problems of effectivity in arithmetic, geometry and analysis.* Number theory (New York, 1982), 252–272, Lecture Notes in Math., 1052, Springer, Berlin, 1984.

[Mot] Motohashi, Yōichi. *The binary additive divisor problem.* Ann. Sci. École Norm. Sup. (4) 27 (1994), no. 5, 529–572.

[O] Oesterlé, J.,*Nombres de classes des corps quadratiques imaginaires.* Semin. Bourbaki, 36e annee, Vol. 1983/84, Exp. 631, Asterisque 121- 122, 309-323 (1985).

[Pe] Peng, A., *PhD Thesis*, Princeton University 2001.

[PS] Petridis, Y. N.; Sarnak, P. *Quantum unique ergodicity for* $\mathrm{SL}_2(O)\backslash\mathbf{H}^3$ *and estimates for* L*-functions.* Dedicated to Ralph S. Phillips. J. Evol. Equ. 1 (2001), no. 3, 277–290.

[Pi] Pitt, Nigel J. E.*On shifted convolutions of* $\zeta^3(s)$ *with automorphic* L*-functions.* Duke Math. J. 77 (1995), no. 2, 383–406.

[PS] Phillips, R. S.; Sarnak, P. *On cusp forms for co-finite subgroups of* $\mathrm{PSL}(2,\mathbf{R})$. Invent. Math. 80 (1985), no. 2, 339–364.

[PS2] Phillips, R.; Sarnak, P. *Perturbation theory for the Laplacian on automorphic functions.* J. Amer. Math. Soc. 5 (1992), no. 1, 1–32.

[Pr] Proskurin, N.V. *Summation formulas for general Kloosterman sums. (English)* J. Sov. Math. 18, 925-950 (1982).

[Ra1] Ramakrishnan, D., *Pure motives and automorphic forms.* Motives (Seattle, WA, 1991), 411–446, Proc. Sympos. Pure Math., 55, Part 2, Amer. Math. Soc., Providence, RI, 1994.

[Ra1] Ramakrishnan, D., *Modularity of the Rankin-Selberg* L*-series, and multiplicity one for* $SL(2)$, Annals of Math. 152 (2000), 45-111.

[Ra2] Ramakrishnan, D., *Expository talk at the IAS on Siegel zeros* available at http://math.caltech.edu/papers/ramak.pdf.

[RaW] Ramakrishnan, D., Wang, S., *On the exceptional zeros of Rankin-Selberg L-functions,* Compositio Math. 135 (2003), no. 2, 211–244.

[Ran] Rankin. Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions. I, II, III., Proc. Camb. Philos. Soc. 35, 35 and 36, 351-356 (1939), 357-372 (1939), 150-151 (1940).

[Ri] Ribet, K. A. *On modular representations of* $\mathrm{Gal}(\overline{Q}/Q)$ *arising from modular forms.* Invent. Math. 100 (1990), no. 2, 431–476.

[Ro] Royer, E.: *Statistique de la variable* $L(\mathrm{sym}^2 f, 1)$, Math. Ann. 321, No.3, 667-687, (2001).

[Ro2] Royer, E.: *Interpretation combinatoire des moments negatifs des valeurs de fonctions L au bord de la bande critique* , Ann. Sci. École Norm. Sup. (4) 36 (2003), no. 4, 601–620.

[RoWu] Royer, E., Wu, J.: *Taille des valeurs de fonctions* L *de carrés symétriques au bord de la bande critique.* Rev. Mat. Iberoamericana 21 (2005), no. 1, 263–312.

[Ru] Rudnick Z. *On the asymptotic distribution of zeros of modular forms.* Int. Math. Res. Not. 2005, no. 34, 2059–2074.

[RS] Rudnick Z., Sarnak, P.*The behaviour of eigenstates of arithmetic hyperbolic manifolds.* Comm. Math. Phys. 161 (1994), no. 1, 195–213.

[RS2] Z. Rudnick, P. Sarnak, *Zeros of principal* L*-functions and random matrix theory,* A celebration of John F. Nash, Jr. Duke Math. J. 81 (1996), no. 2, 269–322.

[Sa1] Sarnak, P., *Some Applications of Modular forms* Cambridge Tracts in Mathematics, 99. Cambridge University Press, Cambridge, 1990. x+111 pp.

[Sa2] Sarnak, P.,*Arithmetic Quantum Chaos*. the Schur Lectures (1992), Israel Math. Conf. Proc., 8, Bar-Ilan Univ., Ramat Gan, 1995, 183–236.

[Sa3] Sarnak, P.,*Integrals of products of eigenfunctions*, IMRN 6, 251-260 (1994).

[Sa4] Sarnak, P.,*Estimates for Rankin-Selberg L-functions and quantum unique ergodicity.* J. Funct. Anal. 184 (2001), no. 2, 419–453.

[Sa5] Sarnak, P.,*Nonvanishing of L-functions on* $\Re e(s) = 1$. Contributions to automorphic forms, geometry, and number theory, 719–732, Johns Hopkins Univ. Press, Baltimore, MD, 2004.

[Sa6] Sarnak, P.,*Spectra of hyperbolic surfaces* . Bull. Amer. Math. Soc. (N.S.) 40 (2003), no. 4, 441–478

[SS] Seeger, A.; Sogge, C.D. *Bounds for eigenfunctions of differential operators.* Indiana Univ. Math. J. 38, No.3, 669-682 (1989).

[Se] Selberg, A. *Bemerkungen uber eine Dirichletsche Reihe, die mit der Theorie der Modulformen nahe verbunden ist.* Arch. Math. Naturvid. B 43, No.4, 1-4 (1940).

[Se2] Selberg, A. *On the estimation of Fourier coefficients of modular forms.* 1965 Proc. Sympos. Pure Math., Vol. VIII pp. 1–15 Amer. Math. Soc., Providence, R.I.

[Ser1] Serre, J.-P. *Modular forms of weight one and Galois representations.* Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pp. 193–268. Academic Press, London, 1977.

[Ser2] Serre, J.-P., *Quelques applications du theoreme de densite de Chebotarev.* Publ. Math., Inst. Hautes Etud. Sci. 54, 123-202 (1981).

[Ser3] Serre, J.-P., *lettre à Deshouillers*.

[Sha1] Shahidi, F.,*On certain L-functions*, Amer. Journ. Math. 103 (1981), 297-355.

[Sha2] Shahidi, F.,*Fourier transforms of intertwining operators and Plancherel measures for* $\mathrm{GL}(n)$. Amer. J. Math. 106 (1984), no. 1, 67–111.

[Sha3] Shahidi, F.,*Local coefficients as Artin factors for real groups.* Duke Math. J. 52 (1985), no. 4, 973–1007.

[Sh] Shimura, G., *On modular forms of half integral weight.* Ann. of Math. (2) 97 (1973), 440–481.

[Shn] Shnirelman, A., Uspenski Math. Nauk., 29/6 (1974), 79-98.

[So] Soundararajan, K. *Nonvanishing of quadratic Dirichlet L-functions at* $s = 1/2$. Ann. of Math. (2) 152 (2000), no. 2, 447–488.

[St] Stark, H. M. *On the 'gap' in a theorem of Heegner,* J. Number Theory 1, 16-27 (1969).

[St2] Stark, H. M. *Some effective cases of the Brauer-Siegel theorem*. Invent. Math. 23 (1974), 135–152.

[Szp] Szpiro, L. *Degrés, Discriminants, Hauteurs*. Seminaire sur Les Pinceaux de Courbes Elliptiques. A la recherche de "Mordell effectif". Edited by L. Szpiro. Asterisque No. 183 (1990). Societe Mathematique de France, Paris, 1990. pp. 1–135.

[Ta] Tatuzawa, T., *On a theorem of Siegel.* Jap. J. Math. 21 (1951), 163–178 (1952).

[Te] Tenenbaum, G., *Introduction a la theorie analytique et probabiliste des nombres.* 2eme ed. (French) Cours Specialises. 1. Paris: Societe Mathematique de France. xv, 457 p. FF 450.00 (1995).

[TV] Takhtadzhyan, L.A., Vinogradov, A.I. *The zeta-function of the additive divisor problem and spectral decomposition of the automorphic Laplacian.* J. Sov. Math. 36, 57-78 (1987).

[TW] Taylor, R., Wiles A.: *Ring theoretic properties of certain Hecke algebras*, Annals of Math. 141, No. 3 (1995), 553–572.

[VP] De la Vallee Poussin, Ch. J. *Recherches analytiques sur la theorie des nombres premiers.* Brux. S. sc. 21 B, 183-256, 281-362, 363-397. (1896)

[VP2] De la Vallee-Poussin, Ch. J. *Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inferieurs a une limite donnee.* Belg. Mem. cour. in 8°, 59, 74 S. Published: (1899)

[Vdk] VanderKam, Jeffrey M. *The rank of quotients of $J_0(N)$.* Duke Math. J. 97 (1999), no. 3, 545–577.

[Vdk2] VanderKam, Jeffrey. *Linear independence of Hecke operators in the homology of $X_0(N)$.* J. London Math. Soc. (2) 61 (2000), no. 2, 349–358.

[Va] Vatsal, V., *Uniform distribution of Heegner points*, Invent. Math. 148, No.1, 1-46 (2002).

[Ve] Venkatesh, V., *Sparse equidistribution problems, period bounds, and subconvexity*, Preprint (2005).

[Vi] Vinogradov, I. M.*A new estimate of the function $\zeta(1+it)$.* (Russian) Izv. Akad. Nauk SSSR. Ser. Mat. 22 1958 161–164.

[Wal] Waldspurger, J-L. *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, J. Math. Pures et Appliquées 60 (1981), 374–484.

[Wa] Watkins, M.,*Class numbers of imaginary quadratic fields.* Math. Comp. 73 (2004), no. 246, 907–938

[Wat] Watson, T., *Rankin triple products and quantum chaos*, Annals of Math. (to appear).

[Wats] G. N. Watson, A Treatise on the Theory of Bessel Functions, Reprint of the second (1944) edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1995, viii+804 pp.

[We] Weyl, H., *Zur Abschatzung von $\zeta(1+ti)$.* Math. Zeitschr. 10, 88-101 (1921).

[Wi] Wiles, A.: *Modular elliptic curves and Fermat's last theorem*, Annals of Math. 141, No. 3 (1995), 443–551.

[Wo] Wong, S.:*Automorphic forms on* $\mathrm{GL}(2)$ *and the rank of class groups.* J. Reine Angew. Math. 515 (1999), 125–153.

[Ze] Zelditch, S., Duke Math. Journal, 55 (1987), 919-941.

[Z1] Zhang, S., *Heights of Heegner points on Shimura curves*, Ann. Math. 153, 27-147 (2001)

[Z2] Zhang, S., *Gross-Zagier formula for $GL(2)$*, Asian J. Math. Vol 5, N2, 2001.

[Z3] Zhang, S., *Gross-Zagier formula for* $\mathrm{GL}(2)$. *II.* Heegner points and Rankin L-series, 191–214, Math. Sci. Res. Inst. Publ., 49, Cambridge Univ. Press, Cambridge, 2004.