

TWO REMARKS ON THE LARGE SIEVE

E. KOWALSKI

1. INTRODUCTION

This paper contains two remarks concerning the *large sieve inequality*, which are partly related to and motivated by the general formalism developed in [K1]. The first one, in Section 2, is a new proof of the arithmetic large sieve inequality

$$(1) \quad |\{n \leq N \mid n \pmod{p} \notin \Omega_p \text{ for } p \leq Q\}| \leq \frac{\Delta}{H}$$

where $\Omega_p \subset \mathbf{Z}/p\mathbf{Z}$ are arbitrary sets of residues and

$$H = \sum_{q \leq Q}^{\flat} \prod_{p|q} \frac{|\Omega_p|}{p - |\Omega_p|},$$

while Δ is any constant for which the “harmonic” large sieve inequality holds, i.e., for any complex numbers (a_n) , we have

$$(2) \quad \sum_{q \leq Q} \sum_{a \pmod{q}}^* \left| \sum_{n \leq N} a_n e\left(\frac{an}{q}\right) \right|^2 \leq \Delta \sum_{n \leq N} |a_n|^2,$$

the notation \sum^{\flat} and \sum^* denoting, respectively, a sum over squarefree integers, and one over integers coprime with q , while $e(z) = \exp(2i\pi z)$.

The second remark, in Section 3, shows how one version of the large sieve inequality for Fourier coefficients of modular forms (as introduced by Iwaniec and developed by Deshouillers and Iwaniec) *can* be interpreted in such a way that a similar proof leads to an analogue of (1), where the Hecke eigenvalues $\lambda_f(p)$ play the role of the reductions of integers modulo primes. This provides a kind of “unification” of the various inequalities called, sometimes awkwardly, “large sieve” (see [IK, Ch. 7]). Moreover, this sieve inequality is most easily proved by following the pattern of Section 2, thus providing a link between the two parts of the paper. However, they may be read independently if desired.

In Section 4, we give a first application of the sieve inequality of Section 3: we estimate how many primitive holomorphic modular forms of given weight k and level q have Hecke eigenvalues with given signs ± 1 for p up to some small limit (compare with [DK]; such results have also been obtained by Lau and Wu). In an Appendix, we also prove a variant of the “analytic multiplicity one theorem” where only the signs of the Hecke eigenvalues are considered – this serves as background and motivation for Section 3; note that although it is a fairly elementary computation, it requires in an essential way to understand the behavior of the sixth symmetric power of a modular form [KS].

Notation. As usual, $|X|$ denotes the cardinality of a set. By $f \ll g$ for $x \in X$, or $f = O(g)$ for $x \in X$, where X is an arbitrary set on which f is defined, we mean synonymously that there exists a constant $C \geq 0$ such that $|f(x)| \leq Cg(x)$ for all $x \in X$. The “implied constant” is any admissible value of C . It may depend on the set X which is always specified or clear in context. On the other hand, $f = o(g)$ as $x \rightarrow x_0$ means that $f/g \rightarrow 0$ as $x \rightarrow x_0$.

For any $z \geq 1$, and any integer $q \geq 1$, we write

$$(3) \quad P(z) = \prod_{p \leq z} p, \quad P_q(z) = \prod_{\substack{p \leq z \\ p \nmid q}} p$$

and an integer all of whose prime factors p divide $P(z)$ is called z -friable. If $n \geq 1$ is an integer, the integer

$$\varrho(n) = \prod_{p|n} p$$

is the (*squarefree*) radical of n .

2. A PROOF OF THE ARITHMETIC LARGE SIEVE INEQUALITY

There are a number of derivations of (1) from (2).¹ For one of the earliest, see [M1, Ch. 3], but the most commonly used is probably the argument of Gallagher involving a “submultiplicative” property of some arithmetic function (see, e.g., [K1, §II.2] for a very general version).

We will in fact prove (1) straightforwardly from the *dual* version of (2): Δ is also any constant for which

$$(4) \quad \sum_{n \leq N} \left| \sum_{q \leq Q} \sum_{a \pmod{q}}^* \beta(q, a) e\left(\frac{an}{q}\right) \right|^2 \leq \Delta \sum_{q \leq Q} \sum_{a \pmod{q}}^* |\beta(q, a)|^2,$$

holds for arbitrary complex numbers $(\beta(q, a))$. This is of some interest because, quite often,² the inequality (2) is proved by duality from (4), and because, in recent generalized versions of the large sieve, it often seems that the analogue of (4) is the most natural inequality to prove, or at least the most easily accessible (see, e.g., [K1, Ch. 7, Ch. 8] for examples related to random walks on discrete groups and the distribution of Frobenius conjugacy classes over finite fields). So, in some sense, one could dispense entirely with (2) for many applications! In particular, note that both known proofs of the optimal version with $\Delta = N - 1 + Q^2$ proceed by duality.

There are other proofs of (1) working directly from (4), which often carry explicit connections with the Selberg sieve (see the references to papers of Huxley, Kobayashi, Matthews and Motohashi in [M2, p. 561]), but none of those that the author has seen seems to give an argument which is as well motivated. Moreover, as we will see in Section 3, our proof leads to new sieve inequalities that seem difficult to obtain otherwise.

We now explain the argument. Let

$$S = \{n \leq N \mid n \pmod{p} \notin \Omega_p \text{ for } p \leq Q\},$$

be the sifted set; we wish to estimate from above the cardinality of this finite set. From (4), the idea is to find an “amplifier” of those integers remaining in the sifted set, i.e., an expression of the form

$$A(n) = \sum_{q \leq Q} \sum_{a \pmod{q}}^* \beta(q, a) e\left(\frac{an}{q}\right)$$

which is *large* (in some sense) when $n \in S$. Then an estimate for $|S|$ follows from the usual Chebychev-type manoeuvre.

¹ For applications, it is of course crucial to know a value of Δ , and we recall that it is known that (1) and (2) hold with $\Delta = N - 1 + Q^2$ (see, e.g., [IK, Th. 7.7]), which is due to Montgomery-Vaughan and Selberg, and is essentially best possible.

² But not always – Gallagher’s very short proof, found e.g. in [M2, Th. 1, p. 549], proceeds directly.

To construct the amplifier $A(n)$, we look first at a single prime $p \leq Q$. If $n \in S$, we have $n \pmod{p} \notin \Omega_p$. If we expand the characteristic function of Ω_p in terms of additive characters,³ we have then

$$0 = \mathbb{1}_{\Omega_p}(n) = \sum_{a \pmod{p}} \alpha(p, a) e\left(\frac{an}{p}\right), \quad \alpha(p, a) = \frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} \mathbb{1}_{\Omega_p}(x) e\left(\frac{ax}{p}\right),$$

and the point is that the contribution of the constant function (0-th harmonic) is, indeed, relatively “large”, because it is

$$\alpha(p, 0) = \frac{|\Omega_p|}{p},$$

and exactly reflects the probability of a random element being in Ω_p . Thus for $n \pmod{p} \notin \Omega_p$, we have

$$(5) \quad \sum_{a \pmod{p}}^* \beta(p, a) e\left(\frac{an}{p}\right) = c_p$$

with

$$c_p = \frac{|\Omega_p|}{p}, \quad \beta(p, a) = -\alpha(p, a).$$

If we only use the contribution of the primes in (4), and the amplifier

$$A(n) = \sum_{p \leq Q} \sum_{a \pmod{p}} \beta(p, a) e\left(\frac{an}{p}\right),$$

then by (4), we get

$$\sum_{n \in S} |A(n)|^2 \leq \sum_{n \leq N} |A(n)|^2 \leq \Delta \sum_{p \leq Q} \sum_{a \pmod{p}} |\beta(p, a)|^2.$$

For $n \in S$, the size of the amplifier is

$$|A(n)|^2 = \left| \sum_{p \leq Q} \sum_{a \pmod{p}} \beta(p, a) e\left(\frac{an}{p}\right) \right|^2 = \left| \sum_{p \leq Q} c_p \right|^2 = K^2,$$

by (5), while on the other hand, by applying the Parseval identity in $\mathbf{Z}/p\mathbf{Z}$, we get

$$\begin{aligned} \sum_{p \leq Q} \sum_{a \pmod{p}} |\beta(p, a)|^2 &= \sum_{p \leq Q} \left(\frac{1}{p} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} |\mathbb{1}_{\Omega_p}(x)|^2 - \alpha(p, 0)^2 \right) \\ &= \sum_{p \leq Q} c_p(1 - c_p) \leq K. \end{aligned}$$

So we obtain $K^2|S| \leq \Delta K$, which is a version of (1) originally due to Rényi’s (see [K1, §II.4]).

To go further, we must exploit all the squarefree integers $q \leq Q$ (and not only the primes) to construct the amplifier. This is most easily described using the Chinese Remainder Theorem to write

$$\mathbf{Z}/q\mathbf{Z} \simeq \prod_{p|q} \mathbf{Z}/p\mathbf{Z}, \quad (\mathbf{Z}/q\mathbf{Z})^\times \simeq \prod_{p|q} (\mathbf{Z}/p\mathbf{Z})^\times,$$

³ We use this specific basis to use (4), but any orthonormal basis containing the constant function 1 would do the job, as in [K1].

and putting together the amplifiers modulo primes $p \mid q$: if $n \in S$ then $n \pmod{p} \notin \Omega_p$ for all $p \mid q$, and hence multiplying out (5) over $p \mid q$, we find constants $\beta(q, a) \in \mathbf{C}$, defined for $(a, q) = 1$ (because $\beta(p, a)$ is defined for a coprime with p), such that

$$\sum_{a \pmod{q}}^* \beta(q, a) e\left(\frac{an}{q}\right) = \prod_{p \mid q} c_p.$$

Moreover, because the product decomposition of the Chinese Remainder Theorem is compatible with the Hilbert space structure involved, we have

$$\sum_{a \pmod{q}}^* |\beta(q, a)|^2 = \prod_{p \pmod{q}} \sum_{a \pmod{p}}^* |\beta(p, a)|^2 = \prod_{p \mid q} c_p(1 - c_p).$$

Arguing as before, we obtain from (4), using all squarefree moduli $q \leq Q$ this time, that

$$(6) \quad |S| \leq \Delta \frac{A}{B^2}, \quad \text{with} \quad A = \sum_{q \leq Q}^b \prod_{p \mid q} c_p(1 - c_p), \quad B = \sum_{q \leq Q}^b \prod_{p \mid q} c_p.$$

This is not quite (1), but we have some flexibility to choose another amplifier, namely, notice that this expression is not homogeneous if we multiply the coefficients $\beta(q, a)$ by scalars independent of a , and we can use this to find a better inequality. Precisely, let

$$\gamma(q, a) = \left(\prod_{p \mid q} \xi_p \right) \beta(q, a),$$

where $\xi_p > 0$ are arbitrary positive real coefficients. Then we have the new amplification property

$$\sum_{a \pmod{q}}^* \gamma(q, a) e\left(\frac{an}{q}\right) = \prod_{p \mid q} \xi_p c_p$$

with altered ‘‘cost’’ given by

$$\sum_{a \pmod{q}}^* |\gamma(q, a)|^2 = \prod_{p \mid q} \xi_p^2 c_p(1 - c_p),$$

so that, arguing as before, we get

$$|S| \leq \Delta \frac{A_1(\xi)}{B_1(\xi)^2}$$

with

$$A_1(\xi) = \sum_{q \leq Q}^b \prod_{p \mid q} \xi_p^2 c_p(1 - c_p), \quad B_1(\xi) = \sum_{q \leq Q}^b \prod_{p \mid q} \xi_p c_p.$$

Now we can select

$$\xi_p = \frac{1}{1 - c_p} = \frac{p}{p - |\Omega_p|},$$

which leads to $A_1(\xi) = B_1(\xi) = H$, hence to $|S| \leq \Delta H^{-1}$ (i.e., to (1)).

This choice is in fact optimal, and corresponds to the standard problem of minimizing a quadratic form (namely $A_1(\xi)$) with a linear constraint (given by $B_1(\xi)$): by Cauchy’s inequality, writing

$$c_q = \prod_{p \mid q} c_p, \quad \tilde{c}_q = \prod_{p \mid q} (1 - c_p), \quad \xi_q = \prod_{p \mid q} \xi_p$$

for ease of notation, we have

$$B_1(\xi)^2 = \left(\sum_{q \leq Q}^b \xi_q c_q \right)^2 \leq \left(\sum_{q \leq Q}^b \xi_q^2 c_q \tilde{c}_q \right) \left(\sum_{q \leq Q}^b \frac{c_q}{\tilde{c}_q} \right) = A_1(\xi)H,$$

with equality if ξ_p is as given above.

Remark. (1) The last “optimization” step is reminiscent of the Selberg sieve. Indeed, it is well known that the Selberg sieve is related to the large sieve, and particularly with the dual inequality (4), as explained in [HR, p. 125]. Note however that the coefficients we optimize for, being of an “amplificatory” nature, and different from the coefficients ξ_d typically sought for in Selberg’s sieve, which are akin to the Möbius function and of a “mollificatory” nature.

(2) The argument does not use any particular feature of the classical sieve, and thus extends immediately to provide a proof of the general large sieve inequality of [K1, Prop. 2.3] which is directly based on the dual inequality [K1, Lemma 2.8]; readers interested in the formalism of [K1] are encouraged to check this. The next section will show that it also extends beyond the classical cases.

Example 1. What are the amplifiers above in some simple situations? In the case – maybe the most important – where we try to count primes, we then take $\Omega_p = \{0\}$ to detect integers free of small primes by sieving, and (5) becomes

$$\sum_{a \pmod{p}}^* \left(-\frac{1}{p}\right) \cdot e\left(\frac{an}{p}\right) = \frac{1}{p},$$

if $p \nmid n$. Then, for q squarefree, the associated detector is the identity

$$\sum_{a \pmod{q}}^* \frac{\mu(q)}{q} e\left(\frac{an}{q}\right) = \frac{1}{q},$$

if $(n, q) = 1$, or in other words, it amounts to the well-known formula

$$\sum_{a \pmod{q}}^* e\left(\frac{an}{q}\right) = \mu(q)$$

for the values of a Ramanujan sum with coprime arguments. Note that in this case, the optimization process above replaced $c_p = \frac{1}{p}$ with

$$\xi_p c_p = \frac{1}{p-1},$$

which is not a very big change – and indeed, for small sieves, the bound (6) is not far from (1), and remains of the right order of magnitude.

On the other hand, for an example in a large sieve situation, we can take Ω_p to be the set of squares in $\mathbf{Z}/p\mathbf{Z}$. The characteristic function (for odd p) is

$$\mathbf{1}_{\Omega_p}(x) = \sum_{a \pmod{p}} \tau(p, a) e\left(\frac{ax}{p}\right)$$

with coefficients given – essentially – by Gauss sums

$$\tau(p, a) = \frac{1}{p} \left(1 + \frac{1}{2} \sum_{x \pmod{p}}^* e\left(\frac{ax^2}{p}\right) \right).$$

Then c_p tends to $1/2$ as $p \rightarrow +\infty$, while $\xi_p c_p$ tends to 1. This difference leads to a discrepancy in the order of magnitude of the final estimate: using standard results on bounds for sums of multiplicative functions, (6) and taking $Q = \sqrt{N}$, we get

$$|S| \ll \sqrt{N}(\log N)^{1/4},$$

instead of $|S| \ll \sqrt{N}$ that follows from (1).

3. AN ARITHMETIC SIEVE FOR HECKE EIGENVALUES

For an even integer $k \geq 2$ and positive integer $q \geq 1$, we denote by $S_k(q)^*$ the finite set of primitive holomorphic cusp forms of weight k and level q with trivial nebentypus (see, e.g., [IK, Ch. 14] for basic references). Each $f \in S_k(q)^*$ has the well-known expansion

$$f(z) = \sum_{n \geq 1} \lambda_f(n) n^{(k-1)/2} e(nz)$$

where $\lambda_f(n)$ are the Hecke eigenvalues, which are real numbers here. Since the pioneering work of Iwaniec, and the major developments by Deshouillers and Iwaniec [DI], many “large sieve inequalities” for those coefficients, and their analogues in the context of Maass cusp forms, have been established. However, it has not yet been observed that some at least of these inequalities *can*, when seen in the right way, be interpreted as leading to analogues of (1).

The starting point is the analogy between the equidistribution of integers modulo primes (which is at the root of (2)) and the equidistribution of Hecke eigenvalues $\lambda_f(p)$, for a fixed prime p and f varying, which has been remarked and used in a number of contexts already.⁴ This leads to the idea of considering sifted sets of cusp forms of the type

$$(7) \quad \{f \in S_k(q)^* \mid \lambda_f(p) \notin \Omega_p \text{ for } p \leq Q\}$$

where $\Omega_p \subset \mathbf{R}$ are suitable sets. Although we won’t be able to deal with such sets in full generality here, we will provide some estimates for the size of some of them where Ω_p has some special description.

To be more precise, consider on $S_k(q)^*$ the measure $\mu = \mu_q$ defined by

$$\mu_q(\{f\}) = \frac{\Gamma(k-1)}{(4\pi)^{k-1} \langle f, f \rangle},$$

where $\langle \cdot, \cdot \rangle$ is the Petersson inner product. This is the familiar “harmonic weight” often used in the study of automorphic L -functions, and we denote

$$(8) \quad \mathbf{E}_q(\alpha) = \sum_{f \in S_k(q)^*} \mu_q(\{f\}) \alpha(f), \quad \mathbf{P}_q(\mathcal{P} \text{ is true}) = \sum_{\substack{f \in S_k(q)^* \\ \mathcal{P}(f) \text{ is true}}} \mu_q(\{f\})$$

the corresponding averaging operator and “probability”, for an arbitrary property $\mathcal{P}(f)$ referring to the modular forms $f \in S_k(q)^*$. Note that (as we will recall) this measure is *only asymptotically* a probability measure (as $q \rightarrow +\infty$ for a fixed k , for instance): we have

$$\lim_{q \rightarrow +\infty} \mathbf{E}_q(1) = 1.$$

⁴ There is an implicit version already present in Bruggeman’s work (see [B, §4], where it is shown that, on average, “most” Maass forms with Laplace eigenvalue $\leq T$, satisfy the Ramanujan-Petersson conjecture for a fixed p), and the first explicit result goes back to Sarnak [Sa], still in the case of Maass forms. Serre [S] and (independently) Conrey, Duke and Farmer [CDF] gave similar statements for holomorphic forms, and Royer [R] described quantitative versions in the latter case.

We will measure the size of sifted sets like (7) through their measure for μ_q , which turns out to be easier to handle because the local equidistribution of $\lambda_f(p)$ (for fixed p) holds with a measure independent of p .

Imitating the notation in [K1, Ch. 1], we now denote by

$$\rho_p : \begin{cases} S_k(q)^* \rightarrow \mathbf{R} \\ f \mapsto \lambda_f(p), \end{cases}$$

the p -th Fourier coefficient maps, which we see as random variables on $(S_k(q)^*, \mu_q)$, or as “global-to-local” data. Moreover, if $d \geq 1$ is a squarefree integer coprime with q , we denote

$$\rho_d : \begin{cases} S_k(q)^* \rightarrow \mathbf{R}^{\omega(d)} \\ f \mapsto (\rho_p(f))_{p|d} = (\lambda_f(p))_{p|d}, \end{cases}$$

which we emphasize is a *tuple* of Fourier coefficients, that should not be mistaken with the single number $\lambda_f(d)$.

Our basic tool is the following: provided z is small enough, the $(\rho_p(f))_{p \leq z}$ become equidistributed as $q \rightarrow +\infty$ for the product Sato-Tate measure

$$\nu_z = \bigotimes_{p \leq z} \mu_{ST}, \quad \text{on } \mathbf{R}^{\pi(z)}, \quad \mu_{ST} = \sqrt{1 - \frac{x^2}{4}} \mathbf{1}_{[-2,2]}(x) dx,$$

which means that the image of μ_q under the individual ρ_p converge to μ_{ST} , uniformly and independently. To express this quantitatively is easy if z is bounded (independently of q), but requires some care when $z = z(q)$ grows with q . For our purpose, we choose to express it through large-sieve bounds for Weyl-type sums associated with a suitable orthonormal basis of $L^2(\nu_z)$. The most convenient is that given by the Chebychev polynomials of the second kind.⁵ We recall the definition: for $n \geq 0$, let $X_n \in \mathbf{Z}[x]$ be defined by

$$(9) \quad X_n(2 \cos \theta) = \frac{\sin((n+1)\theta)}{\sin \theta}$$

for $\theta \in [0, \pi]$; then $(X_n)_{n \geq 0}$ is an orthonormal basis of $L^2(\mu_{ST})$.

It follows immediately that if we define

$$\Lambda_m(x) = \prod_p X_{n_p}(x_p), \quad \text{for all } x = (x_p) \in [-2, 2]^{\pi(z)},$$

for any z -friable integer $m \geq 1$, factored as

$$m = \prod_{p \leq z} p^{n_p}, \quad n_p \geq 0,$$

then the functions (Λ_m) , for m z -friable, form an orthonormal basis of $L^2(\nu_z)$. Note in particular that we have $\Lambda_1 = 1$, the constant function 1.

We have now the following fundamental (and easy) fact which gives the link between this orthonormal basis and our local data (ρ_p) :

Fact. *With notation as above, for any z -friable integer $m \geq 1$ coprime with q , and any $f \in S_k(q)^*$, we have*

$$(10) \quad \Lambda_m(\rho_d(f)) = \lambda_f(m), \quad \text{where } d = \varrho(m) = \prod_{p|m} p.$$

⁵ Which are probably best seen as characters of $SL(2)$, though we do not use this.

Remark 2. Our situation is indeed similar to that of classical sieve problems. In the framework of [K1], we have a set X (with a finite measure μ) and surjective maps $X \xrightarrow{\rho_\ell} Y_\ell$ with *finite* target sets Y_ℓ , each equipped with a probability density ν_ℓ , so that the equidistribution can be measured by the size of the remainders $r_\ell(y)$ defined by

$$\mu(\rho_\ell^{-1}(y)) = \mu(X)\nu_\ell(y) + r_\ell(y)$$

and the independence by using finite sets $m = \{\ell_1, \dots, \ell_k\}$, and

$$Y_m = \prod_{\ell \in m} Y_\ell, \quad \rho_m = \prod_{\ell \in m} \rho_\ell : X \rightarrow Y_m, \quad \nu_m(y_1, \dots, y_k) = \nu_{\ell_1}(y_1) \cdots \nu_{\ell_k}(y_k),$$

and looking at

$$\mu(\rho_m^{-1}(y)) = \mu(X)\nu_m(y) + r_m(y).$$

Here the compact set $[-2, 2]$ requires the use of infinitely many functions to describe an orthonormal basis. Another (less striking) difference is that our local information lies in the same set $[-2, 2]$ for *all* primes, whereas classical sieves typically involve reduction modulo primes, which lie in different sets.

The equidistribution statement for an individual prime p is that

$$\lim_{q \rightarrow +\infty} \mathbf{E}_q(X_m(\rho_p(f))) = \lim_{q \rightarrow +\infty} \mathbf{E}_q(\lambda_m(f)) = 0,$$

for any fixed $m \geq 1$ coprime with q ; in this formulation, it is a trivial consequence of the Petersson formula. Here is the large-sieve version:

Proposition 3. *With notation as above, for all integers $N \geq 1$, all complex numbers $\alpha(m)$ defined for integers $m \leq M$ coprime with q , we have*

$$(11) \quad \mathbf{E}_q \left(\left| \sum_{\substack{m \leq M \\ (m, q) = 1}} \alpha(m) \Lambda_m(\rho_d(f)) \right|^2 \right) \ll (1 + q^{-1} N(\log N)) \sum_m |\alpha(m)|^2,$$

where the implied constant is absolute and $d = \varrho(m)$ is the radical of m . If $k > 2$, then the factor $\log N$ may be removed.

Proof of Proposition 3. This is in fact simply a consequence of one of the large sieve inequalities for Fourier coefficients of cusp forms (see [DI]). The point is that because of (10), the left-hand side W of (11) can be rewritten

$$W = \sum_{f \in S_k(q)^*} \mu_q(f) \left| \sum_{\substack{m \leq M \\ (m, q) = 1}} \alpha(m) \lambda_f(m) \right|^2 = \frac{\Gamma(k-1)}{(4\pi)^{k-1}} \sum_{f \in S_k(q)^*} \left| \sum_{\substack{m \leq M \\ (m, q) = 1}} \alpha(m) \frac{\lambda_f(m)}{\|f\|} \right|^2.$$

We can now enlarge this by positivity; remarking that

$$S_k(q)^\sharp = \left\{ \frac{f}{\|f\|} \mid f \in S_k(q)^* \right\}$$

is a subset of an orthonormal basis of the space $S_k(q)$ of cusp forms of weight k and level q . Selecting any such basis $\mathcal{B}_{k,q} \supset S_k(q)^\sharp$, we have therefore

$$W \leq \frac{\Gamma(k-1)}{(4\pi)^{k-1}} \sum_{\varphi \in \mathcal{B}_{k,q}} \left| \sum_{m \leq N} \alpha(m) \lambda_\varphi(m) \right|^2,$$

where we put $\alpha(m) = 0$ if $(m, q) \neq 1$, and where the $\lambda_\varphi(m)$ are again the Fourier coefficients of φ (as earlier for Hecke forms). Now by the large sieve inequality stated in [IK, Theorem 7.26],⁶ we have

$$(12) \quad \frac{\Gamma(k-1)}{(4\pi)^{k-1}} \sum_{\varphi \in \mathcal{B}_{k,q}} \left| \sum_{1 \leq m \leq N} \alpha(m) \lambda_\varphi(m) \right|^2 \ll (1 + q^{-1}N(\log N)) \sum_{m \leq N} |\alpha(m)|^2$$

where the implied constant is absolute (and $\log N$ can be removed for $k > 2$), and this leads to (11). \square

We derive from this two genuinely sieve-like consequences, where the basic description (7) is specialized to sets Ω_p , $p \nmid q$, where a *polynomial* takes values “far” from its mean (computed with respect to the limiting measure μ_{ST}).

Corollary 4. *With notation as above, let*

$$Y(x) = \sum_{i=0}^s \beta_i X_i(x)$$

be a polynomial of degree $\leq s$ with complex coefficients, expressed in the basis of Chebychev polynomials, so that

$$\beta_0 = \int_{-2}^2 Y d\mu_{ST}, \quad \sum_{1 \leq i \leq s} |\beta_i|^2 = \int_{-2}^2 |Y|^2 d\mu_{ST} - \left(\int_{-2}^2 Y d\mu_{ST} \right)^2.$$

(1) *We have*

$$(13) \quad \mathbf{E}_q \left(\left| \sum_{\substack{p \leq z \\ p \nmid q}} (Y(\lambda_f(p)) - \beta_0) \right|^2 \right) \ll (1 + q^{-1}z^s(\log z^s)) \sum_{1 \leq i \leq s} |\beta_i|^2,$$

where the implied constant is absolute and the $\log z^s$ can be removed for $k > 2$.

(2) *Assume $\beta_i \in \mathbf{R}$ for all i . For any $z \geq 2$, any $N \geq 2$ and any δ with $\delta \geq 0$, we have*

$$(14) \quad \mathbf{P}_q \left(Y(\lambda_f(p)) \leq \beta_0 - \delta \text{ for all } p \leq z, p \nmid q \right) \ll (1 + q^{-1}N^s(\log N^s))H^{-1}$$

where

$$H = \sum_{m \in \Psi_q(N,z)}^b \left(\frac{\delta^2}{|\beta_1|^2 + \dots + |\beta_s|^2} \right)^{\omega(m)},$$

with $\Psi_q(N, z)$ the set of integers $m \leq N$ which are coprime with q and z -friable, the implied constant is absolute and $\log N^s$ can be removed if $k > 2$.

Remark 5. Proposition 3 and this Corollary give the first known link between the arithmetic versions of the large sieve inequalities (generalizing (1), as in [K1]) and the “spectral” large sieve inequalities for coefficients of modular forms. The comparison with the dual inequality in ([K1, (2.7)]) is particularly clear when the left-hand side of (11) is written

$$\mathbf{E}_q \left(\left| \sum_{d|P_q(z)} \sum_{\mathbf{n}=(n_p)_{p|d}} \gamma(d, \mathbf{n}) \prod_{p|d} X_{n_p}(\rho_p(f)) \right|^2 \right)$$

⁶ Taking into account the slightly different normalization, and the fact that the result is proved for $k > 2$ there, without the $\log N$ factor; the case $k = 2$ is obtained by separating the constant term of the Bessel function J_1 , which requires the $\log N$ term, from the rest.

where the sum over $\mathbf{n} = (n_p)_{p|d}$ is a sum over those tuples of positive integers $n_p \geq 1$ such that

$$(15) \quad \prod_{p|d} p^{n_p} \leq N,$$

while, in addition, we have

$$\gamma(d, \mathbf{n}) = \alpha \left(\prod_{p|d} p^{n_p} \right).$$

Thus we see that, if we forget about the size restriction (15), we have for given $d \mid P_q(z)$ a sum over all “primitive” elements of the orthonormal basis of

$$\bigotimes_{p|d} L^2(\mu_{ST}),$$

given by products of Chebychev polynomials (where primitivity means those basis elements for which none of the components X_{n_p} at a prime $p \mid d$ is the constant function X_0). This is entirely similar with the general expression in [K1, equ. (2.7)] where the corresponding bases are finite (and there are no restriction similar to (15)).

One interesting point arises: what we have written is the analogue of the *dual* large sieve inequalities (i.e., the sum over the harmonics are in the inner sum; compare it with [K1, equ. (2.3)], where the sum over the harmonics is the outer one), whereas in the spectral interpretation of the underlying inequalities of Deshouillers and Iwaniec, the harmonics are the modular forms, so (12) is the “non-dual” form of their large-sieve inequality!

In Corollary 4, the inequality (13) is clearly an analogue of [K1, Prop. 2.15], e.g., of

$$\sum_{n \leq N} \left(\sum_{\substack{p \leq Q \\ n \pmod{p} \in \Omega_p}} 1 - \sum_{p \leq Q} \frac{|\Omega_p|}{p} \right)^2 \leq (N - 1 + Q^2) \sum_{p \leq Q} \frac{|\Omega_p|}{p},$$

while (14) is (maybe less recognizably) the analogue of (1) (take $z = N$ so that $\Psi_q(N, z)$ is the set of integers $\leq N$ coprime with q) or [K1, Prop. 2.3]. This means that the characteristic functions of the residue classes Ω_p play the role of the polynomial $Y(x)$. The “sieve” content of Corollary 4 is that the values $Y(\lambda_f(p))$ of any polynomial $Y(x)$ are typically “close” to the average value (with respect to the Sato-Tate measure).

Note on the other hand that the *quality* of the upper-bound is not as good as that in (1): the right-hand side $1 + N^s q^{-1}$ involves the maximal “size” of an harmonic Λ_m , not the total number of them (which is about N).

Proof of Corollary 4. The first formulas for β_0 and the sum of $|\beta_i|^2$ are consequence of the orthonormality of the Chebychev polynomials for μ_{ST} .

To prove (13), we apply (11) with $N = z^s$ and $\alpha(m) = 0$ unless $m = p^j$ with $1 \leq j \leq s$ and $p \leq z$, $p \nmid q$, in which case $\alpha(p^j) = \beta_j$. By definition of $Y(x)$ and of Λ_d , we get

$$\sum_{m \leq N} \alpha(m) \Lambda_m(\rho_d(f)) = \sum_{p \leq z} (Y(\rho_p(f)) - \beta_0),$$

showing that (13) is indeed a special case of (11).

To prove (14), we use the same amplification technique used in Section 2 (it is not clear if the standard proof found, e.g., in [K1, Ch. 2] can be adapted here; for instance, Lemma 2.6 of loc. cit. does not make sense in this context). First, to simplify notation, let $\sigma_0 \geq 0$ be defined by

$$\sigma_0^2 = \beta_1^2 + \cdots + \beta_s^2.$$

The basic observation is that if, for some prime $p \nmid q$, we have

$$(16) \quad Y(\lambda_f(p)) \leq \beta_0 - \delta, \quad \text{with } \delta \geq 0,$$

then it follows that

$$\sum_{1 \leq i \leq s} (-\beta_i) X_i(\lambda_f(p)) \geq \delta \geq 0.$$

Now, as in Section 2, let ξ_p , for $p \leq z$ (coprime with q) be arbitrary auxiliary positive real numbers, and let

$$\xi_d = \prod_{p|d} \xi_p$$

for $d \mid P_q(z)$. If (16) holds for all $p \leq z$ coprime with q , then we find by multiplying out that, for any squarefree integer $d \in \Psi_q(N, z)$, i.e., such that

$$(17) \quad d \leq N, \quad d \mid P(z), \quad (d, q) = 1, \quad d = p_1 \cdots p_k, \quad (\text{say}),$$

and for such (ξ_p) , we have

$$\xi_d \sum_{1 \leq n_1, \dots, n_k \leq s} (-1)^k \beta_{n_1} \cdots \beta_{n_k} X_{n_1}(\lambda_f(p_1)) \cdots X_{n_k}(\lambda_f(p_k)) \geq \xi_d \delta^{\omega(d)},$$

which translates to

$$\xi_d \sum_{m \in S_d} \alpha(m) \Lambda_m((\lambda_f(p))_{p \leq z}) \geq \xi_d \delta^{\omega(d)},$$

where m runs over the set S_d of integers of the type

$$m = \prod_{p|d} p^{n_p}, \quad \text{with } 1 \leq n_p \leq s$$

(as described in (17), so $m \leq N^s$), and

$$\alpha(m) = \prod_{p|d} (-\beta_{n_p}).$$

Thus, summing over d subject to (17), squaring, then averaging over f and applying (11), we find that

$$\mathbf{P}_q \left(Y(\lambda_f(p)) \leq \beta_0 - \delta \text{ for all } p \leq z \right) \ll (1 + q^{-1} N^s (\log N^s)) \frac{A(\xi)}{B(\xi)^2},$$

where

$$A(\xi) = \sum_d \xi_d^2 \sum_{m \in S_d} |\alpha(m)|^2 = \sum_d \xi_d^2 \sigma_0^{2\omega(d)}, \quad B(\xi) = \sum_d \xi_d \delta^{\omega(d)}.$$

We now find the optimal choice of the parameters (ξ_p) , as in the previous section by Cauchy's inequality: we have $B(\xi)^2 \leq HA(\xi)$, with equality if

$$\xi_p = \frac{\delta}{\sigma_0^2}, \quad \text{for all } p \leq z,$$

and the inequality above, with this choice, leads to

$$\mathbf{P}_q \left(Y(\lambda_f(p)) \leq \beta_0 - \delta \text{ for all } p \leq z \right) \ll (1 + q^{-1} N^s (\log N^s)) H^{-1},$$

which was the desired conclusion. □

Remark 6. In the case of classical sieves, this last proof is performed with $\delta = \beta_0$, since $n \pmod{p} \notin \Omega_p$ implies that the characteristic function of Ω_p is zero at n , giving (16) with $\delta = \beta_0 = |\Omega_p|p^{-1}$.

Note also that there is no obstruction that prevents us from doing the same thing with different polynomials $Y_p(x)$ and different $\delta_p \geq 0$ for every $p \leq z$; if $\deg(Y_p) \leq s$ for all $p \leq z$, we get

$$(18) \quad \mathbf{P}_q \left(Y_p(\lambda_f(p)) \leq \beta_{0,p} - \delta_p, \text{ for all } p \leq z \right) \ll (1 + q^{-1} N^s (\log N^s)) H^{-1}$$

where

$$(19) \quad H = \sum_{m \in \Psi_q(N, z)}^b \left(\prod_{p|m} \frac{\delta_p}{\sigma_{0,p}} \right)^2,$$

the meaning of the notation $\beta_{0,p}$, $\sigma_{0,p}$ being clear. This looks, of course, even more similar to (1).

4. APPLICATION: MODULAR SIGNS

There are many results in the arithmetic of modular forms which are, more or less, concerned with various ways of characterizing a given (primitive) cusp form f from its siblings, starting from the fact that Fourier coefficients determine uniquely a cusp form f relative to a congruence subgroup Γ of $SL(2, \mathbf{Z})$, going through stronger forms of the multiplicity one theorem for automorphic representations, and then to various explicit forms of these statements, where only finitely many coefficients are required (say at primes $p \leq X$, for some explicit X depending on the parameters defining f), and to “statistic” versions of the latter, where X can be reduced drastically, provided one accepts some possible exceptions.

Recently, Lau and Wu [LW2] have found, for real characters, a precise “threshold” $y(Q)$ for which the upper bound on the number of real characters of conductor $q \leq Q$ with value $+1$ for primes $p \leq y(Q)$ almost coincides with a lower bound for this number. Then they proved in [LW1] a corresponding upper-bound for recognition problems for modular forms. However, as they mention, it is doubtful that the analogue of the lower bound holds, because there are much less restrictions on the values of Fourier coefficients (or Hecke eigenvalues) than on values of real characters (the latter take value in $\{-1, 0, 1\}$, whereas, for instance, there are about $4\sqrt{p}$ values for the p -th Fourier coefficient of the modular form of weight 2 attached to an elliptic curve, and even more for a general one where the coefficients are not necessarily integers).

One may be tempted, at least out of curiosity, to try to remove this discrepancy by considering only the signs of the Fourier coefficients – assuming they are real.⁷ In this spirit, here are the three natural questions:

Question 1. *Let $f \in S_k(q)^*$ for some $k \geq 2$ and $q \geq 1$. Is it true that f is determined uniquely by the sequence of signs of its Hecke eigenvalues $\lambda_f(p)$, for p prime, where the sign is interpreted in a relaxed way so that zero has the same sign as both positive and negative numbers? If yes, how many exceptions can we possibly allow to get the same conclusion, i.e., how large can a set E of primes be so that if another modular form g has Hecke eigenvalues of the same sign as f for $p \notin E$, then $g = f$?*

⁷ In so doing, we cross another ongoing track in the field of modular forms: there have been quite a few works investigating the first sign-change in the sequence of Hecke eigenvalues of a primitive form.

Question 2. Given a sequence of signs $(\varepsilon_p)_p$, $\varepsilon_p = \pm 1$ for all primes p , and a threshold $N = N(k, q) \geq 1$, what is the size of the set

$$\{f \in S_k(q)^* \mid \text{the sign of } \lambda_f(p) \text{ is the same as } \varepsilon_p \text{ for } p \leq N(k, q)\}$$

where, as before, a zero coefficient is meant to be compatible with both signs ± 1 ?

Question 3. In particular, for what type of threshold $N = N(k, q)$ do we have

$$\lim_{k, q} |\{f \in S_k(q)^* \mid \lambda_f(p) \geq 0 \text{ for } p \leq N(k, q)\}| = +\infty,$$

where the limit may be for $k \rightarrow +\infty$, $q \rightarrow +\infty$, or simply $\max(k, q) \rightarrow +\infty$.

Note that here we define the sign function sgn for $x \in \mathbf{R}$ by

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases},$$

and we will write $x \simeq_s y$ to say that the sign of x and that of y are *compatible*, i.e., that either $xy = 0$, or $\text{sgn}(x) = \text{sgn}(y)$. Note that this is equivalent with $xy \geq 0$.

Using Corollary 4 (or rather (18)), we will give an answer to Question 2 which is quite comparable with the results of [DK].

Theorem 7. Let $k \geq 2$ be a fixed even integer and $q \geq 1$ an integer. Let (ε_p) be a sequence of signs for prime numbers p . For any $\alpha > 1$, and any $\varepsilon > 0$, we have

$$|\{f \in S_k(q)^* \mid \lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha\}| \ll q^{1/2+(2\alpha)^{-1+\varepsilon}}$$

for $q \geq 2$, where the implied constant depends only on α , ε and k . We recall the well-known fact that for any fixed $k \geq 2$, we have $|S_k(q)^*| \gg q$ for $q \geq q_0$, the implied constant and q_0 depending only on k , so the estimate above is non-trivial for all $\alpha > 1$.

In the Appendix, we also show that the answer to the first Question 1 is “Yes”, and indeed we can specify the possibility of an exceptional set E of (small) positive analytic density (in the sense of (28) below). One can also obtain partial answers to Question 3 using the Petersson formula to express the uniform equidistribution of Fourier coefficients, but we leave this aside in this note.

Proof. We will apply the variant (18) of Part (2) of Corollary 4 (if ε_p is independent of p , then one can apply directly the result of the corollary) with $z = (\log q)^\alpha$. We assume we have at our disposal a polynomial $Y \in \mathbf{R}[X]$, of degree $s \geq 1$, such that

$$(20) \quad Y(x) \leq \text{sgn}(x), \quad \text{for all } x \in [-2, 2], \quad \text{and} \quad \beta_0 = \int_{-2}^2 Y d\mu_{ST} > -1.$$

Moreover, we let $\tilde{Y}(x) = Y(-x)$; observe that

$$(21) \quad \int_{-2}^2 \tilde{Y} d\mu_{ST} = \int_{-2}^2 Y d\mu_{ST} = \beta_0, \quad \int_{-2}^2 \tilde{Y}^2 d\mu_{ST} - \beta_0^2 = \int_{-2}^2 Y^2 d\mu_{ST} - \beta_0^2,$$

and that (by continuity) $Y(0) = \tilde{Y}(0) \leq -1$.

For each $p \leq (\log q)^\alpha$, coprime with q , we let $Y_p = Y$ if $\varepsilon_p = 1$ and $Y_p = \tilde{Y}$ if $\varepsilon_p = -1$. Then notice, by inspecting both cases, that

$$\lambda_f(p) \simeq_s \varepsilon_p \Rightarrow Y_p(\lambda_f(p)) \leq -1 = \beta_0 - \delta,$$

where $\delta > 0$ is fixed. Therefore,⁸ by (18) with $N = q^{1/s}(\log q)^{-1/s}$, we get

$$\mathbf{P}_q(\lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha) \ll H^{-1}$$

where

$$H = \sum_{m \in \Psi_q(N, (\log q)^\alpha)}^b \gamma^{\omega(m)}, \quad \text{with} \quad \gamma = \frac{(\beta_0 + 1)^2}{|\beta_1|^2 + \cdots + |\beta_s|^2},$$

and the implied constant is absolute (see (19), noting that by (21), both δ_p and $\sigma_{0,p}$ are independent of p). By assumption, we have $\gamma > 0$, and a lower bound for H in the range of interest to us is given by the following lemma:

Lemma 8. *Let $\gamma > 0$ be a real number and $q \geq 1$ an integer. Then for $\alpha > 1$ and for any $\varepsilon > 0$, we have*

$$\sum_{m \in \Psi_q(y, (\log q)^\alpha)}^b \gamma^{\omega(m)} \gg y^{\kappa - \varepsilon}, \quad \kappa = 1 - \frac{1}{\alpha},$$

for $y \geq 2$, the implied constant depending only on α , γ and ε .

Proof. We use a simplistic argument, though stronger versions are certainly available in the literature on sums over friable integers. First, assume $\gamma < 1$, and notice that if $1 \leq n \leq y$ is a squarefree integer, we have

$$\omega(n) \ll \frac{\log y}{\log \log y}$$

with an absolute implied constant. Then if S is the sum to estimate, we have

$$S \geq \exp\left(\log \gamma \cdot \frac{\log y}{\log \log y}\right) |\Psi_q^b(y, (\log q)^\alpha)| \gg |\Psi_q^b(y, (\log q)^\alpha)| y^{-\varepsilon},$$

where $\Psi_q^b(y, z)$ is the set of squarefree integers $n \in \Psi_q(y, z)$.

If $\gamma \geq 1$, on the other hand, we have trivially $S \geq |\Psi_q^b(y, (\log q)^\alpha)|$, so it is enough to prove

$$|\Psi_q^b(y, (\log q)^\alpha)| \gg y^{\kappa - \varepsilon},$$

the implied constant depending on α , γ and ε . But this is a familiar fact (see, e.g., [DK, p. 10], [K3, Lemma 2.2]). \square

Continuing the proof of Theorem 7, this lemma gives

$$H \gg q^{s^{-1}(1 - \alpha^{-1} - \varepsilon)}$$

for arbitrary $\varepsilon > 0$, the implied constant depending on ε , $\alpha > 1$ and the choice of Y , so

$$(22) \quad \mathbf{P}_q(\lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha) \ll q^{-s^{-1} + (s\alpha)^{-1} + \varepsilon},$$

for any $\varepsilon > 0$, the implied constant depending on ε , α and Y .

To conclude the proof of Theorem 7, we must go back from the measure μ_q to the natural counting measure on $S_k(q)^*$. This is easily done here because of the very simple⁹ estimate

$$(23) \quad \langle f, f \rangle \ll q(\log q)^3,$$

⁸ We ignore the conditions at primes $p \mid q$.

⁹ Better results are known.

where the implied constant depends only on k , which follows from the occurrence of the left-hand side in the residue of $L(f \times f, s)$ at $s = 1$ (see, e.g., [IK, p. 138]). We thus obtain from Proposition 22, (8) and (23) that

$$(24) \quad |\{f \in S_k(q)^* \mid \lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha\}| \ll q(\log q) \times \mathbf{P}_q(\lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha) \ll q^{1-s^{-1}+(s\alpha)^{-1}+\varepsilon}$$

where the implied constant depends only on k , α , ε and the choice of the polynomial.

Now to conclude we describe a suitable polynomial Y of degree 2. For this we recall the first few Chebychev polynomials:

$$(25) \quad X_0 = 1, \quad X_1 = x, \quad X_2 = x^2 - 1, \quad X_3 = x^3 - 2x, \\ X_4 = x^4 - 3x^2 + 1, \quad X_5 = x^5 - 4x^3 + 3x, \quad X_6 = x^6 - 5x^4 + 6x^2 - 1.$$

The graph of Y appears in Figure 1 (together with the graph of the sign function). It is given precisely by

$$Y = -\frac{3}{4}X_0 + \frac{1}{2}X_1 + \frac{1}{4}X_2 = -1 + \frac{x}{2} + \frac{x^2}{4},$$

so the numerical values of β_0 , δ and γ are given by

$$\beta_0 = -\frac{3}{4}, \quad \delta = \frac{1}{4}, \quad \beta_1^2 + \beta_2^2 = \frac{5}{16}, \quad \gamma = \frac{4}{5}.$$

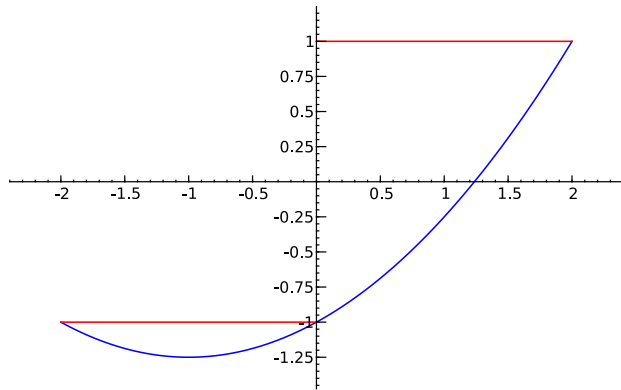


FIGURE 1

So (24) translates to

$$(26) \quad |\{f \in S_k(q)^* \mid \lambda_f(p) \simeq_s -1 \text{ for } p \leq (\log q)^\alpha\}| \ll q^{1/2+(2\alpha)^{-1}+\varepsilon},$$

where the implied constant depends only on k and α , proving Theorem 7. \square

Remark 9. The polynomial above was found by looking for a polynomial with $Y(0) = -1$, $Y \leq \text{sgn}$, having smallest possible degree s , and such that $\beta_0 + 1 > 0$ (and is as large as possible). It is easily checked that this is not possible for $s = 1$; for $s = 2$, one quickly derives the following conditions on β_0, β_2 :

$$\begin{cases} \beta_0 - \beta_2 \leq -1, \\ \beta_0 + 2\beta_2 \leq 0, \end{cases}$$

and the intersection of those two half-spaces has a unique point with maximal value of β_0 , namely $(\beta_0, \beta_2) = (-3/4, 1/4)$. Then it is easy to see that the only possible value of β_1 to have

$$-\frac{3}{4}X_0 + \beta_1 X_1 + \frac{1}{4}X_2 \leq \text{sgn} \quad \text{on } [-2, 2]$$

is $\beta_1 = 1/2$ (evaluate at ± 2); then a final check shows that the resulting polynomial does, indeed, satisfy the required condition.

Of course, by standard approximation arguments (in $L^2(\mu_{ST})$), one can find (higher degree) polynomials $Y \leq \text{sgn}$ such that

$$\int Y d\mu_{ST} \quad \text{and} \quad \int (Y - \beta_0)^2 d\mu_{ST}$$

are arbitrarily close to

$$\int \text{sgn} d\mu_{ST} = 0 \quad \text{and} \quad \int (\text{sgn} - 0)^2 d\mu_{ST} = 1,$$

respectively (and $\beta_0 < 0$ since the condition implies $\beta_0 \leq \int \text{sgn} d\mu_{ST} = 0$). But this does not improve the upper bound for

$$\mathbf{P}_q(\lambda_f(p) \simeq_s \varepsilon_p \text{ for } p \leq (\log q)^\alpha),$$

because the degree increase leads to a loss of a power of q after choosing $N = q^{1/s}$ (this is related to the last part of Remark 5 on the strength of the large sieve inequality we have proved).

APPENDIX: INDIVIDUAL RECOGNITION

We show here that if two primitive cusp forms (non CM for simplicity) have Fourier coefficients with identical (or compatible) signs for all but a smallish set of exceptions, then they are identical.

Theorem 10. *Let $q_1, q_2 \geq 1$ be integers, let $k_1, k_2 \geq 2$ be even integers and $f_1 \in S_{k_1}(q_1)^*$, $f_2 \in S_{k_2}(q_2)^*$ be primitive holomorphic cusp forms of level q_1, q_2 and weight k_1, k_2 , respectively. Assume neither of f_1 or f_2 is of CM type, for instance assume that q_1 and q_2 are squarefree.*

If we have

$$(27) \quad \lambda_{f_1}(p) \simeq_s \lambda_{f_2}(p)$$

for every prime p , except those in a set E of analytic density κ , with $\kappa < 0.0011116875$, then in fact $f_1 = f_2$, and of course $q_1 = q_2, k_1 = k_2$.

The definition of the analytic density we use is as follows: a set E of primes has density $\kappa > 0$ if and only if

$$(28) \quad \sum_{p \in E} \frac{1}{p^\sigma} \sim \kappa \sum_p \frac{1}{p^\sigma} \sim -\kappa \log(\sigma - 1), \quad \text{as } \sigma \rightarrow 1.$$

The theorem is also valid for the natural density, since the existence of the latter implies that of the analytic density, and that they are equal.

The idea is that the assumption translates to

$$\lambda_{f_1}(p)\lambda_{f_2}(p) \geq 0$$

for all primes p (with few exceptions), but it is well-known from Rankin-Selberg theory that if $f_1 \neq f_2$, we have

$$(29) \quad \sum_p \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} = O(1)$$

as $\sigma \rightarrow 1$ (see, e.g. [IK, §5.12] for a survey and references; the underlying fact about automorphic forms is due to Mœglin and Waldspurger). Thus we only need to find a lower bound for the left-hand side (which is a sum of non-negative terms) which is unbounded as σ tends to 1. Since Rankin-Selberg theory also gives

$$(30) \quad \sum_p \frac{\lambda_{f_1}(p)^2}{p^\sigma} \sim -\log(\sigma - 1), \quad \text{as } \sigma \rightarrow 1,$$

the only difficulty is that one might fear that the coefficients of f_1 and f_2 are such that whenever $\lambda_{f_1}(p)$ is not small, the value of $\lambda_{f_2}(p)$ is very small. In other words, what we must show is that the smaller order of magnitude of (29) compared with (30) is not due to the small size of the summands, but to sign compensations. This would be immediate if we knew the Sato-Tate Conjecture, but we can work around the fact that we do not yet know it (in full generality, see Mazur's survey [Ma]) by using small symmetric power L -functions. Here's the lemma we will use:

Lemma 11. *Let $q \geq 1$ be an integer, $k \geq 2$ be an even integer and $f \in S_k(q)^*$ a primitive cusp form of level q and weight k which is not of CM type. Then there exists a constant $\alpha > 0$ and $\delta > 1/2$ such that*

$$\sum_{|\lambda_f(p)| > \alpha} \frac{1}{p^\sigma} \geq \delta \sum_p \frac{1}{p^\sigma} + O(1),$$

for $\sigma > 1$. In fact, one can take $\alpha = 0.231$, $\delta = 1/2 + 1/24$.

The lemma does not hold for CM forms, because we then have $\lambda_f(p) = 0$ for a set of primes of density $1/2$. However, one could easily prove a version of Theorem 10 when f_1 or f_2 is of CM-type using the (better known) distribution of the eigenvalues in those cases.

Proof. We recall first that for any prime $p \nmid q$, we have the relation $\lambda_f(p^n) = X_n(\lambda_f(p))$, where X_n is defined by (9). Now we claim that there exists a polynomial

$$Y = \beta_0 + \beta_2 X_2 + \beta_4 X_4 + \beta_6 X_6 \in \mathbf{R}[x]$$

with the following properties: (1) $\beta_0 > 1/2$; (2) for some $\alpha > 0$, and $x \in [-2, 2]$, we have

$$(31) \quad Y(x) \leq \chi_A(x), \quad \text{where } A = \{x \in [-2, 2] \mid |x| > \alpha\}.$$

Assuming this, we conclude as follows: by (2), we have

$$\sum_{|\lambda_f(p)| > \alpha} \frac{1}{p^\sigma} \geq \sum_{p \nmid q} \frac{Y(\lambda_f(p))}{p^\sigma} = \beta_0 \sum_{p \nmid q} \frac{1}{p^\sigma} + \sum_{i=1}^3 \beta_{2i} \sum_{p \nmid q} \frac{X_{2i}(\lambda_f(p))}{p^\sigma}.$$

By the holomorphy and non-vanishing at $s = 1$ of the second, fourth and sixth symmetric power L -functions (see [KS, Th. 3.3.7, Prop. 4.3] for the last two, and [Sh] for a survey concerning those L -functions), since $X_n(\lambda_f(p))$ is exactly the p -th coefficient of the n -th symmetric power for $p \nmid q$, standard analytic arguments show that

$$\sum_{p \nmid q} \frac{X_{2i}(\lambda_f(p))}{p^\sigma} = O(1)$$

for $\sigma \geq 1$ and $i = 1, 2, 3$. Hence the result follows with $\delta = \beta_0 > 1/2$.

Now to check the claim, and verify the values of α and δ , we just exhibit a suitable polynomial, namely

$$Y = \frac{1}{2} + \frac{1}{24} + \frac{1}{4}X_2 - \frac{1}{4}X_4 + \frac{136}{1000}X_6 = \frac{17}{125}x^6 - \frac{93}{100}x^4 + \frac{227}{125}x^2 - \frac{283}{3000},$$

(see (25)) which is even, and the graph of which on $[-2, 2]$ is in Figure 2 (see Remark 12 for an explanation of the origin of this polynomial).

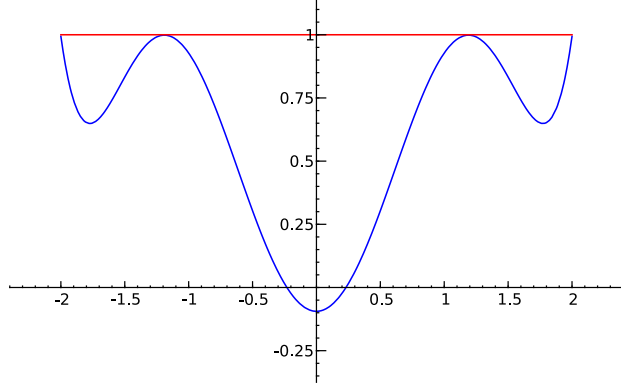


FIGURE 2

The value of α is an approximation (from below) to the real root

$$\alpha_0 = 0.23107202470801418176315245050693402580 \dots$$

of Y in $[0, 2]$; the maximum value of Y on $[0, 2]$ is very close to 1. \square

Proof of Theorem 10. As already noticed, the assumption (27) implies that for all primes $p \notin E$, we have $\lambda_{f_1}(p)\lambda_{f_2}(p) \geq 0$. Moreover, we have $|\lambda_{f_1}(p)\lambda_{f_2}(p)| \leq 4$ for all p , by the Deligne bound (the Ramanujan-Petersson conjecture). Hence, with $\alpha > 0$ and $\delta > 1/2$ as in Lemma 11, we have

$$\begin{aligned} \sum_p \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} &= \sum_{p \notin E} \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} + \sum_{p \in E} \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} \\ &\geq \alpha^2 \sum_{\substack{|\lambda_{f_1}(p)| > \alpha, |\lambda_{f_2}(p)| > \alpha \\ p \notin E}} \frac{1}{p^\sigma} - 4\kappa |\log(\sigma - 1)| + O(1) \\ &\geq \alpha^2 \left\{ \sum_p \frac{1}{p^\sigma} - \sum_{p \in E} \frac{1}{p^\sigma} - \sum_{|\lambda_{f_1}(p)| \leq \alpha} \frac{1}{p^\sigma} \right. \\ &\quad \left. - \sum_{|\lambda_{f_1}(p)| \leq \alpha} \frac{1}{p^\sigma} \right\} - 4\kappa |\log(\sigma - 1)| + O(1) \\ &\geq \alpha^2 \left\{ (1 - 2(1 - \delta)) \sum_p \frac{1}{p^\sigma} + O(1) \right\} - 4\kappa |\log(\sigma - 1)| + O(1) \\ &= (\alpha^2(2\delta - 1) - 4\kappa) |\log(\sigma - 1)| + O(1) \end{aligned}$$

for any $\sigma > 1$.

Since $2\delta > 1$, we find that the left-hand side goes to $+\infty$ as $\sigma \rightarrow 1$ under the condition

$$\kappa < \frac{\alpha^2(2\delta - 1)}{4} = \frac{17787}{16000000} = 0.0011116875$$

(with the values of Lemma 11). However, as already mentioned, the theory of Rankin-Selberg L -functions shows that if $f_1 \neq f_2$, we have

$$\sum_p \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} = O(1)$$

since there is no pole (or zero) of $L(f_1 \times f_2, s)$ at $s = 1$ (recall that f_1 and f_2 have real coefficients so $f_2 = \bar{f}_2$). So we must indeed have $f_1 = f_2$. \square

Remark 12. We now explain how the polynomial of Lemma 11 was found, and why it must necessarily involve the sixth symmetric power (or higher). Indeed, there is no polynomial

$$Y = \beta_0 + \beta_2 X_2 + \beta_4 X_4 \in \mathbf{R}[X], \quad \text{with } \beta_0 > \frac{1}{2},$$

such that $Y \leq \chi_I$ for some interval $I = [\alpha, 2] \subset]0, 2]$, because such a polynomial would have to satisfy

$$(1) Y(0) \leq 0; (2) Y \leq 1 \text{ on } [0, 2]; (3) \beta_0 > 1/2.$$

But then, expressing $Y \in \mathbf{R}[x]$ in the basis of powers of x , we check that

$$Y(0) + Y(\sqrt{2}) = (\beta_0 - \beta_2 + \beta_4) + (\beta_0 - \beta_2 + \beta_4 + 2\beta_2 - 6\beta_4 + 4\beta_4) = 2\beta_0$$

so condition (3) leads to $Y(0) + Y(\sqrt{2}) > 1$, and if $Y(0) \leq 0$, this means that $Y(\sqrt{2}) > 1$, showing that (1) and (2) are then incompatible. (This formula is a very special case of ‘‘Gaussian quadrature’’ using zeros of orthogonal polynomials.)

Looking at this argument, however, reveals that it ‘‘almost’’ works: precisely, this suggests looking at the value $Y(\sqrt{2})$ for a general polynomial as above, and then one notices quickly that the polynomial

$$Y_0 = \frac{1}{2} + \frac{1}{4}X_2 - \frac{1}{4}X_4 = x^2 - \frac{x^4}{4} = x^2(1 - x/2)(1 + x/2)$$

is ‘‘borderline’’: we have $0 \leq Y_0 \leq 1$ on $[0, 2]$ and $\beta_0 = 1/2$; in fact

$$Y_0(0) = 0, \quad \beta_0 = 1/2, \quad \max_{x \in [0, 2]} Y_0(x) = Y_0(\sqrt{2}) = 1.$$

So we looked for, and found, our polynomial by simply ‘‘deforming’’ slightly this example, increasing the constant coefficient to make it $> 1/2$ and compensating with a small multiple of X_6 . We did spend some time trying to adjust the parameters to maximize $\alpha^2(2\delta - 1)/4$, but we do not know what is the best possible result with polynomials of degree 6.

In the same direction of showing that the sixth symmetric power (and thus the deep results of Kim and Shahidi) are necessary, note that the sequences (x_p) and (y_p) defined by $x_2 = y_2 = 0$ and for primes $p \geq 3$ by

$$x_p = \begin{cases} 0, & \text{if } p \equiv 3 \pmod{4}, \\ (-1)^{(p-1)/4}\sqrt{2} & \text{if } p \equiv 1 \pmod{4}, \end{cases} \quad y_p = \begin{cases} (-1)^{(p-3)/4}\sqrt{2}, & \text{if } p \equiv 3 \pmod{4}, \\ 0 & \text{if } p \equiv 1 \pmod{4}, \end{cases}$$

have the ‘‘right’’ moments of order 1 to 5 for being Sato-Tate distributed,¹⁰ i.e., we have

$$\sum_p X_k(x_p)p^{-\sigma} = O(1), \quad \sum_p X_k(y_p)p^{-\sigma} = O(1), \quad \text{for } \sigma > 1, \text{ and } 1 \leq k \leq 5,$$

¹⁰ The sixth moment fails: it is 4 instead of 5 for the Sato-Tate distribution.

and yet $x_p y_p \geq 0$ for all p , in fact $x_p y_p = 0$, so that we most certainly have

$$\sum_p x_p y_p p^{-\sigma} = O(1) \text{ for } \sigma > 1.$$

Remark 13. Lemma 11 is somewhat “dual” to well-known investigations of consequences of holomorphy of the first symmetric power L -functions towards the Sato-Tate conjecture, explained in particular in Serre’s letter to Shahidi included as an Appendix to [Sh] (and refined most recently by Kim and Shahidi [KS, §4]). In those works, one is interested in finding $c \in [-2, 2]$, as large as possible, such that $\lambda_f(p) > c$ for infinitely many p and $\lambda_f(p) < -c$ for infinitely many p . In our lemma, the value of c (i.e., α) is not the most important, but the density of the set of primes has to be quite large.

Remark 14. Here are some natural further questions concerning Theorem 10:

– What is the optimal density κ one can obtain in Theorem 10? If one assumes that f_1 and f_2 obey the pair-Sato-Tate conjecture, namely that for $a_1 < b_1$, $a_2 < b_2$ the set of primes

$$\{p \mid \lambda_{f_1}(p) \in [a_1, b_1] \text{ and } \lambda_{f_2}(p) \in [a_2, b_2]\}$$

has density $\mu_{ST}([a_1, b_1])\mu_{ST}([a_2, b_2])$ (in other words, the Fourier coefficients are independently Sato-Tate distributed), one may easily get the result for any $\kappa < 1/2$ (corresponding to the probability for $\mu_{ST} \otimes \mu_{ST}$ of having the same sign). But this can only hold if f_1 and f_2 are not related by quadratic twists, of course (and in that case, for elliptic curves, Harris is currently making progress on the problem, according to [Ma, Footnote 12]). If, on the other hand, $f_2 = f_1 \otimes \chi$ for a real character χ , the coefficients are of the same sign for a set of primes of density exactly $1/2$.

– The limit of the argument we used is fairly easy to determine: if we have all even symmetric power L -functions at our disposal, we can use polynomials Y approximating arbitrarily closely to $\chi_A(x)$ (see (31)) for any $\alpha \in]0, 2]$. The value $\delta = \beta_0$ is then the probability under the Sato-Tate distribution of $A = \{x \mid |x| \geq \alpha\}$, namely

$$\beta_0 = \frac{2}{\pi} \arccos\left(\frac{\alpha}{2}\right),$$

and we are led to maximize over $[0, 2]$ the quantity

$$\kappa = \frac{\alpha^2}{4} \left(\frac{4}{\pi} \arccos\left(\frac{\alpha}{2}\right) - 1 \right),$$

and we find numerically that the best value is around $\alpha \simeq 0.971$, allowing to take $\kappa < 0.083571\dots$ (i.e., knowing only the individual Sato-Tate conjecture, we can allow about 8% of the primes to be in the exceptional set).

– What happens with Maass forms? The problem there is that the exceptional set might correspond to primes where the Fourier coefficients are large, since we do not know the Ramanujan-Petersson bound in this case. This means a direct adaptation must assume much stronger conditions on the set E ; for instance, using $|\lambda(p)| \leq 2p^{7/64}$ (the Kim-Sarnak bound) for eigenvalues of Hecke operators for a Maass form, Theorem 10 holds if

$$|\{p \in E \mid p \leq X\}| \ll X^{25/32 - \delta_1}$$

for $X \geq 2$ and some $\delta_1 > 0$ (where the implied constant may depend on f_1 , f_2 and δ_1).

– What is the size, as a function of the weight and conductor, of the smallest prime for which the sign of $\lambda_{f_1}(p)$ and $\lambda_{f_2}(p)$ are different? If f_2 is taken (in effect) to be an Eisenstein series with positive Fourier coefficients (e.g., $d(p)$, the divisor function, for a non-holomorphic Eisenstein series, or $\theta(z)^4$, where θ is the theta function of weight $1/2$,

for $\Gamma_0(4)$), then the question is to find the first negative Hecke eigenvalue for f_1 , and there have been a number of works recently on this issue (see for instance [IKS] where Iwaniec, Kohnen and Sengupta show that smallest sign-change is $\ll (k_1 q_1^2)^{29/60}$; the main point is that the exponent is $< 1/2$, which can be obtained quite easily from the convexity bound for L -functions).

REFERENCES

- [B] R. Bruggeman: *Fourier coefficients of cusp forms*, Invent. math. 45 (1978), 1–18.
- [CDF] B. Conrey, W. Duke and D. Farmer: *The distribution of the eigenvalues of Hecke operators*, Acta Arith. 78 (1997), 405–409.
- [DI] J-M. Deshouillers and H. Iwaniec: *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. math. 70 (1982), 220–288.
- [DK] W. Duke and E. Kowalski: *A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations*, with an Appendix by D. Ramakrishnan, Invent math. 139 (2000), 1–39.
- [HR] H. Halberstam and H.E. Richert: *Sieve methods*, London Math. Soc. Monograph, Academic Press (London), 1974
- [IKS] H. Iwaniec, W. Kohnen and J. Sengupta: *The first negative Hecke eigenvalue*, International J. Number Theory 3 (2007), 355–363.
- [IK] H. Iwaniec and E. Kowalski: *Analytic number theory*, Colloquium Publ. 53, American Math. Soc. 2004.
- [KS] H. Kim and F. Shahidi: *Cuspidality of symmetric powers with applications*, Duke Math. J. 112 (2002), 177–197.
- [K1] E. Kowalski: *The large sieve and its applications*, Cambridge Tracts Math. 175, Cambridge Univ. Press 2008.
- [K3] E. Kowalski: *Variants of recognition problems for modular forms*, Arch. Math. 84 (2005), 57–70.
- [LW1] Y.K. Lau and J. Wu: *A Large Sieve inequality of Elliott–Montgomery–Vaughan type for automorphic forms and two applications*, International Mathematics Research Notices, Vol. 2008, doi:10.1093/imrn/rmn162.
- [LW2] Y.K. Lau and J. Wu: *On the least quadratic non-residue*, Intl. J. Number Theory (to appear).
- [Ma] B. Mazur: *Finding meaning in error terms*, Bull. A.M.S 45 (2008), 185–228.
- [M1] H-L. Montgomery: *Topics in multiplicative number theory*, Lecture Notes Math. 227, Springer-Verlag 1971.
- [M2] H-L. Montgomery: *The analytic principle of the large sieve*, Bull. A.M.S. 84 (1978), 547–567.
- [R] E. Royer: *Facteurs \mathbf{Q} -simples de $J_0(N)$ de grande dimension et de grand rang*, Bull. Soc. Math. France 128 (2000), 219–248.
- [Sa] P. Sarnak: *Statistical properties of eigenvalues of the Hecke operators*, in “Analytic Number Theory and Diophantine Problems” (Stillwater, OK, 1984), Progr. Math. 70, Birkhäuser, 1987, 321–331.
- [S] J-P. Serre: *Répartition asymptotique des valeurs propres de l’opérateur de Hecke T_p* , J. American Math. Soc. 10 (1997), 75–102.
- [Sh] F. Shahidi: *Symmetric power L -functions for $GL(2)$* , in “Elliptic curves and related topics”, edited by E. Kishilevsky and M. Ram Murty, CRM Proc. and Lecture Notes 4, 1994, 159–182.

ETH ZÜRICH – D-MATH, RÄMISTRASSE 101, 8092 ZÜRICH, SWITZERLAND
E-mail address: kowalski@math.ethz.ch