EXPONENTIAL SUMS, TWISTED MULTIPLICATIVITY AND MOMENTS

E. KOWALSKI AND K. SOUNDARARAJAN

Abstract. We study averages over squarefree moduli of the size of exponential sums with polynomial phases. We prove upper bounds on various moments of such sums, and obtain evidence of un-correlation of exponential sums associated to different suitably unrelated and generic polynomials. The proofs combine analytic arguments with the algebraic interpretation of exponential sums and their monodromy groups.

Dedicated to the memory of Jean Bourgain

1. Introduction

Some of Jean Bourgain's many interactions with number theory involved exponential sums in different ways. Among these, one can mention his ground-breaking use of ideas from the circle method to solve Bellow's problems concerning pointwise ergodic theorems at times $f(n)$, where $f \in \mathbf{Z}[X]$ is a polynomial (see in particular [5, 6, 8]) or its combination with bilinear forms in joint works with A. Kontorovich to study some aspects of the sieve in orbits beyond a simple appeal to expansion and spectral gaps (see for instance [7]). We respectfully dedicate this paper to his memory.

1.1. **Exponential sums with polynomials.** This paper is primarily concerned with exponential sums with polynomial phases. Let $f \in \mathbf{Z}[X]$ be a non-constant polynomial with degree d. For $q \geqslant 1$ squarefree and a coprime to q, we define

$$\mathrm{W}(a;q) = \mathrm{W}_f(a;q) = \frac{1}{\sqrt{q}} \sum_{x \,(\mathrm{mod}\, q)} e\Big(\frac{af(x)}{q}\Big),$$

where the sum is over residue classes modulo q. For simplicity we restrict attention to square-free q, and set $\mathrm{W}(a;q) = 0$ if q is not square-free or if $(a,q) > 1$.

An application of the Chinese Remainder Theorem shows that the exponential sums $\mathrm{W}(a;q)$ satisfy the following "twisted multiplicativity": if $(q_1,q_2) = 1$ then

$$\mathrm{W}(a;q_1q_2) = \mathrm{W}(a\bar{q_1};q_2)\mathrm{W}(a\bar{q_2};q_1),$$

where $q_1\bar{q_1} \equiv 1 \mod q_2$ and $q_2\bar{q_2} \equiv 1 \mod q_1$. Apart from finitely many primes, the Weil bound gives $|\mathrm{W}(a;p)| \leqslant (d-1)$, so that $|\mathrm{W}(a;q)| \ll (d-1)^{\omega(q)}$ where $\omega(q)$ denotes the number of (distinct) prime factors of q. It follows that

$$\sum_{q \leqslant x} |\mathrm{W}(a;q)| \ll \sum_{q \leqslant x} (d-1)^{\omega(q)} \ll x(\log x)^{d-2},$$

Date: June 26, 2022, 15:31.

and we seek an improvement over this "trivial" bound, as well as bounds for related mean values such as $\sum_{q\leqslant x}|\mathrm{W}(a;q)|^2$. The possibility of obtaining such improvements was first recognized by Hooley, and explored further in the work of Fouvry and Michel [12].

One of our main theorems gives a refinement of these earlier results. Given a field K, we say that a polynomial $f\in \mathrm{K}[\mathrm{X}]$ is *decomposable* if there are polynomials g and h in $\mathrm{K}[\mathrm{X}]$, both with degree $\geqslant 2$, such that $f=g\circ h$. If f cannot be expressed as such a composition, we call f *indecomposable*.

Theorem 1.1. *Let $f\in \mathbf{Q}[\mathrm{X}]$ be an indecomposable polynomial with $\deg(f)=d\geqslant 3$.*

(1) *For any $a\geqslant 1$,*

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|^2\ll x(\log\log x)^{(d-1)^2}.$$

(2) *There exists $\gamma>0$, depending only on d, such that for any $a\geqslant 1$,*

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|\ll \frac{x}{(\log x)^{\gamma}}.$$

Remark 1.2. The implied constants above (and in what follows) are allowed to depend on f. Throughout we ignore linear polynomials where $\mathrm{W}(a;q)$ is usually 0, and quadratic polynomials where $|\mathrm{W}(a;q)|$ is usually 1 (since these are quadratic Gauss sums).

One can compute effectively a possible value of the constant γ (see Remark 5.4).

The possibility of obtaining non-trivial bounds for

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|$$

(with f allowed to be a rational function) was first pointed out by Hooley in [18] in the case of Kloosterman sums. Introducing ideas from algebraic geometry (notably from the work of Katz [22]), Fouvry and Michel [12] refined and extended Hooley's work to more general exponential sums. Under a hypothesis that the polynomial f is generic (in a sense to be made precise below, see [12, H.1, H.2, H.3, H.3']; note that in fact Fouvry and Michel consider rational functions and not only polynomials), Fouvry and Michel proved in [12, Th. 1.5] that

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|\ll x(\log\log x)^{k_f-1} \tag{1}$$

for some explicit integer $k_f\geqslant 1$. Theorem 1.1 refines this in two ways. Firstly it applies to a larger class of polynomials f, with the much simpler criterion of being indecomposable (for instance, if the degree $d\geqslant 3$ of f is prime, then f is automatically indecomposable, so that our result applies, but any polynomial f such that f' has a multiple root, say $f=\mathrm{X}^3g$ for some g of degree $d-3$, fails to satisfy the condition H.1 of [12], since the zeros of f' are not simple; on the other hand, it is elementary to check that a polynomial f satisfying the conditions H.1, H.2, H.3 (or H.3') is indecomposable (see Lemma 6.1 below, combined with Remark 1.10, (4) for the terminology). Secondly, part (2) of the theorem improves on (1) qualitatively by showing that the average of $|\mathrm{W}(a;q)|$ over $q\leqslant x$ tends to 0, which does not follow from the method of Fouvry–Michel.

The proof of the second part of Theorem 1.1 relies on the following result, which may be of independent interest.

Theorem 1.3. *Let $f \in \mathbf{Z}[X]$ of degree $d \geqslant 3$. Then, one of the following two possibilities holds:*

(1) *The limit*

$$\lim_{p\to+\infty} \frac{1}{p} \sum_{a\in\mathbf{F}_p^\times} |W(a;p)|^4 \qquad \text{exists and equals } 2.$$

(2) *There exists $\delta > 0$ (depending only on d) and a subset of primes with positive density $\geqslant \delta$ on which*

$$\frac{1}{p} \sum_{a\in\mathbf{F}_p^\times} |W(a;p)|^4 \geqslant 3 + O(p^{-1/2}).$$

For a generic (again in a sense to be made precise later) polynomial f, the first case of the theorem holds.

Remark 1.4. (1) The work of Katz [21] contains material from which it is likely that one can deduce Theorem 1.3. However, in view of the different focus and the generality of [21], our independent and slightly more elementary proof seems worth including.

(2) Using the method of [12, § 4] one can show that

$$\sum_{q\leqslant x} |W(a;q)| \gg \frac{x}{\log x}$$

(or even a slightly better lower bound), and it is a natural question to ask whether there exists a constant $\delta > 0$ such that

$$\frac{x}{(\log x)^{\delta+\varepsilon}} \ll \sum_{q\leqslant x} |W(a;q)| \ll \frac{x}{(\log x)^{\delta-\varepsilon}} \tag{2}$$

for any $\varepsilon > 0$. This is an open problem; in Remark 3.6, we will mention a potential candidate value of δ, at least for the upper bound for generic polynomials.

(3) It might be possible to extend Theorem 1.1 to certain rational functions, but some additional work is required (e.g., to properly understand the analogue of indecomposability for rational functions, and to extend [22, Lemma 7.7.5]).

1.2. **Sums of twisted-multiplicative functions.** A key feature of the exponential sums considered above is their twisted multiplicativity. In this section we formulate, following Hooley [18], Fouvry and Michel [12], and our own recent paper [27], a general result on bounding averages of twisted multiplicative functions.

Suppose we are given a function V that associates to each prime p and each reduced residue class $a \pmod p$ a complex number $V(a;p)$. Extend this to a function $V(a;q)$ where q is square-free and $a \pmod q$ is a reduced residue class by "twisted multiplicativity": that is, if $q = q_1q_2$ with $(q_1, q_2) = 1$ then

$$V(a;q_1q_2) = V(a\bar{q_1};q_2)V(a\bar{q_2};q_1). \tag{3}$$

Set $V(a;q) = 0$ if q is not square-free, or if a is not coprime to q. For each prime p, let $G(p) \geqslant g(p) \geqslant 0$ be such that

$$\max_{(a,p)=1} |V(a;p)| \leqslant G(p), \qquad \text{and} \qquad \frac{1}{p} \sum_{(a,p)=1} |V(a;p)| \leqslant g(p). \tag{4}$$

Extend g and G to all square-free integers using multiplicativity, so that (4) remains valid for all q.

The question then is to obtain, under suitable conditions, a bound for

$$\sum_{q\leqslant x}|\mathrm{V}(a;q)|$$

that improves upon the trivial bound

$$\sum_{q\leqslant x}|\mathrm{V}(a;q)|\leqslant\sum_{q\leqslant x}\mathrm{G}(q).$$

Theorem 1.5. *Let* $\mathrm{M}>0$ *be such that* $\mathrm{G}(p)\leqslant\mathrm{M}$ *for all primes* p. *Then, for any fixed integer* $a\geqslant 1$ *and for all large* x, *we have*

$$\sum_{q\leqslant x}|\mathrm{V}(a;q)|\ll\frac{x}{\log x}\prod_{p\leqslant x}\Big(1+\frac{g(p)}{p}\Big)(\log\log x)^{\mathrm{M}},$$

where the implied constant may depend on M.

Remark 1.6. (1) The twisted multiplicativity (3) is naturally connected to the Chinese Remainder Theorem via the Fourier transform. Suppose that for each prime p and any residue class $a \pmod p$, we are given a complex number $v(a;p)$. We extend v to square-free moduli q and any residue class $a \pmod q$ by means of the Chinese Remainder Theorem: that is we set

$$v(a;q)=\prod_{p|q}v(a;p).$$

Consider now the Fourier transform of v:

$$\mathrm{V}(a;q)=\sum_{b\,(\mathrm{mod}\,q)}v(b;q)e(ab/q).$$

Then $\mathrm{V}(a;q)$ satisfies the twisted multiplicative relation (3).

If $v(a;p)$ corresponds to a probability measure (thus all $v(a;p)$ are non-negative and $\sum_a v(a;p)=1$) then $|\mathrm{V}(a;p)|\leqslant 1$ for all $a \pmod p$, so that we may use $\mathrm{G}(p)=1$. Bounding the L^1-norm by the L^2-norm, we may take

$$g(p)=\Big(\frac{1}{p}\sum_{a=1}^{p-1}|\mathrm{V}(a;p)|^2\Big)^{\frac12}=\Big(\sum_{a=1}^{p}|v(a;p)|^2-\frac{1}{p}\Big)^{\frac12},$$

upon using Parseval.

(2) In the applications to equidistribution in [27], the functions that occur are Weyl sums of the form

$$\mathrm{V}(a;q)=\frac{1}{\varrho(q)}\sum_{x\in\mathrm{A}_q}e\Big(\frac{a\cdot x}{q}\Big)$$

for some $h\in\mathbf{Z}^n-\{0\}$, where $\mathrm{A}_q\subset(\mathbf{Z}/q\mathbf{Z})^n$ are non-empty sets "defined by the Chinese Remainder Theorem", and $\varrho(q)=|\mathrm{A}_q|$.

1.3. **Non-correlation of exponential sums for different polynomials.** Our next results are attempts to establish that the exponential sums associated to two different polynomials f and g are uncorrelated. Here we use the notation $\mathrm{W}_f(a;q)$ instead of $\mathrm{W}(a;q)$ to keep track of the dependency on the polynomial. The results here will depend on polynomials being suitably generic (as in the work of Fouvry and Michel [12] mentioned earlier), and we begin by making this notion precise.

Definition 1.7 (Morse polynomial)**.** Let K be a field. A polynomial $f \in \mathrm{K}[\mathrm{X}]$ of degree $d \geqslant 1$ is called *Morse* if it has no repeated roots, its derivative f' is squarefree of degree $d-1$, and the values of f at the zeros of f' (in an algebraic closure of K) are distinct.

Remark 1.8. The values of f at the zeros of the derivative of f are known as *critical values* of f. Note that when f' is even, the critical values appear in pairs $a+f(0)$, $-a+f(0)$ where a is a critical value of $f(x)-f(0)$.

If d is smaller than the characteristic of K, then the condition that $\deg(f') = d-1$ is automatically fullfilled.

If f is a Morse polynomial, then 0 is not a critical value of f (since there would then be a double zero).

We recall that in an abelian group A, a subset $\mathrm{S} \subset \mathrm{A}$ is called *Sidon* if the equation $a+b=c+d$ with $(a,b,c,d) \in \mathrm{S}^4$ has only the obvious solutions where $a \in \{c,d\}$.

We will say that $\mathrm{S} \subset \mathrm{A}$ is a *symmetric Sidon set* if there exists $\alpha \in \mathrm{A}$ such that $\mathrm{S} = \alpha - \mathrm{S}$, and the equation $a+b=c+d$ with $(a,b,c,d) \in \mathrm{S}^4$ has only the obvious solutions where $a \in \{c,d\}$ or $b = \alpha - a$.

We require one last item of terminology. For any field K, two polynomials f and g in $\mathrm{K}[\mathrm{X}]$ are *linearly equivalent over* K if there exist a, b, c, d in K, with a and c non-zero, such that

$$g(\mathrm{X}) = af(c\mathrm{X}+d)+b.$$

Note that the sets V_f and V_g of critical values of f and g are then related by

$$\mathrm{V}_g = a\mathrm{V}_f + d.$$

In particular, if V_f is a Sidon set (resp. a symmetric Sidon set) then so is V_g.

Definition 1.9 (Sidon–Morse polynomial)**.** Let K be a field. A polynomial $f \in \mathrm{K}[\mathrm{X}]$ of degree $d \geqslant 2$ is called *Sidon–Morse* if it is Morse and one of the following holds:

(1) The set of critical values of f is a Sidon set in the additive group of K.
(2) The polynomial f is linearly equivalent to an odd polynomial g and the set of critical values of g is a symmetric Sidon set in K.

For a polynomial $f \in \mathrm{A}[\mathrm{X}]$, with A an integral domain, we say that f is Morse (or Sidon–Morse) if the definition is satisfied for the field of fractions of A.

Remark 1.10. (1) To distinguish between the two alternatives above, we will say that f is a *symmetric Sidon–Morse polynomial* in the second case.

(2) It would seem to be more natural to define a symmetric Sidon polynomial to be one where the set of critical values of f is a symmetric Sidon set. This condition is implied by our definition, and it may in fact be that this is an equivalent definition (at least over $\mathbf{Q}$), but we do not know if this is the case. We will see how, at some crucial point in the proof of Theorem 6.3 below, this alternative definition is not sufficient to proceed.

(3) Any polynomial f of degree $d \geqslant 3$ in $\mathbf{Z}[X]$ whose derivative has Galois group $\mathfrak{S}_{d-1}$ is a (non-symmetric) Sidon–Morse polynomial over $\mathbf{Q}$ (see [22, proof of Th. 7.10.6]). It is then a Sidon–Morse polynomial over $\mathbf{F}_p$ for all but finitely many p. In particular, a "generic" polynomial in $\mathbf{Z}[X]$, in a natural sense, is Sidon–Morse over $\mathbf{Q}$.

(4) The genericity conditions H.1, H.2, H.3 for $f \in \mathbf{Z}[X]$ used by Fouvry and Michel are equivalent to asking that f is a Sidon–Morse polynomial; if f satisfies H.1, H.2, H.3', then it is a symmetric Sidon–Morse polynomial (but the converse is not always true, since H.3' requires f to be odd, not merely linearly equivalent to an odd polynomial).

Theorem 1.11. (1) *Let f and g be polynomials in $\mathbf{Z}[X]$ with degree $d_f \geqslant 3$ and d_g respectively. Assume that f is Sidon–Morse over $\mathbf{Q}$ and that $d_f > d_g$. Then*

$$\sum_{q\leqslant x} |\mathrm{W}_f(a;q)\overline{\mathrm{W}_g(a;q)}|^2 \ll x(\log\log x)^{\mathrm{A}}$$

for some A depending only on d_f and d_g, where the implied constant depends on f and g.

(2) *Let $m \geqslant 1$ be an integer and let $f_1, \dots, f_m$ be polynomials of degrees $d_i = \deg(f_i) \geqslant 3$. Assume that all f_i are Sidon–Morse polynomials over $\mathbf{Q}$ and moreover that for any $i \neq j$, the polynomials f_i and f_j are not linearly equivalent over $\bar{\mathbf{Q}}$.*

Let s be the number of polynomials f_i such that f_i is a symmetric Sidon–Morse polynomial of odd degree $\geqslant 5$. Then for $x \geqslant 2$, we have

$$\sum_{q\leqslant x} |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)| \ll \frac{x}{(\log x)^{\gamma}}$$

$$\sum_{q\leqslant x} |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)|^2 \ll x(\log\log x)^{\mathrm{A}}$$

$$\sum_{q\leqslant x} |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)|^4 \ll x(\log x)^{2^{m-s}3^s-1}(\log\log x)^{\mathrm{A}}$$

for some $\gamma > 0$ and some $\mathrm{A} \geqslant 0$ depending only on m and $(d_1, \dots, d_m)$, where $\mathrm{W}_i(a;q) = \mathrm{W}_{f_i}(a;q)$. The implied constants depend on the polynomials.

Remark 1.12. (1) Since the upper-bounds for two polynomials essentially match those in Theorem 1.1, this result suggests that the exponential sums are uncorrelated. However, we cannot prove it rigorously, since we would need to prove some matching lower-bound, such as

$$\sum_{q\leqslant x} |\mathrm{W}_f(a;q)|^4 \gg x(\log\log x)^{\mathrm{B}}$$

for any $\mathrm{B} \geqslant 1$, for instance. The best current lower-bound that we can achieve in general (by adapting the method of Fouvry and Michel [12, §4]) is

$$\sum_{q\leqslant x} |\mathrm{W}_f(a;q)|^4 \gg \frac{x}{\log x}(\log\log x)^{\mathrm{B}}$$

for any $\mathrm{B} \geqslant 1$ (and the best upper-bound that we can give for the last sum is

$$\sum_{q\leqslant x} |\mathrm{W}_f(a;q)|^4 \ll x(\log x)(\log\log x)^{\mathrm{A}}$$

for some A).

(2) The genericity assumptions that we impose are not the best possible. We will investigate related issues in the paper [28], where we will describe in particular other classes of polynomials for which Theorem 1.11 will apply.

(3) In another paper, Fouvry and Michel [13, Th. 1.2, 1.3] proved that if f is a Sidon–Morse polynomial, then there are infinitely many squarefree integers q with two prime factors such that

$$|\mathrm{W}_f(a;q)| \leqslant q^{-\beta}$$

where $\beta > 0$ depends only on the degree of f. It would be interesting to extend this property to all indecomposable polynomials.

1.4. **Previous work.** Fouvry and Michel also consider rational functions and lower-bounds. In the case of the Kloosterman sums

$$\mathrm{Kl}_2(a;q) = \frac{1}{\sqrt{q}} \sum_{(x,q)=1} e\Big(\frac{ax+\bar{x}}{q}\Big)$$

(i.e., $f(x) = x + 1/x$), they obtain

$$\frac{x}{\log x} \exp((\log\log x)^{5/12}) \ll \sum_{q\leqslant x} |\mathrm{Kl}_2(a;q)| \ll \frac{x}{(\log x)^{\delta}} \tag{5}$$

for any $\delta < 1 - \frac{8}{3\pi}$ (see [12, Th. 1.2, 1.3]).

In this particular case, it is known that if we sum the Kloosterman sums without taking absolute values, one can prove much stronger estimates using the spectral theory of automorphic forms, like

$$\sum_{q\leqslant x} \mathrm{Kl}_2(1;q) \ll x^{2/3+\varepsilon}$$

for any $\varepsilon > 0$ (see, e.g., [19, §16.6]). Patterson [29] has also proved a strong result for certain cubic sums, namely for any non-zero integer a, the asymptotic formula

$$\sum_{q\leqslant x} \sum_{0\leqslant n<q} e\Big(\frac{an^3}{q}\Big) \sim c(a)\mathrm{X}^{4/3}$$

holds for some explicit constant $c(a) > 0$, and Patterson [30, Conj. 2.2] has conjectured similar asymptotic formulas for all cubic polynomials.

It would be of considerable interest to obtain general conditions on a twisted-multiplicative function $\mathrm{V}(a;q)$, bounded at primes, that ensure a power saving in the sums

$$\sum_{q\leqslant x} \mathrm{V}(a;q).$$

Outline of the paper. We prove Theorem 1.5 in the next section. Section 3 gathers a number of properties of exponential sums with polynomials, and Section 4 uses these results to prove Theorem 1.1, assuming Theorem 1.3. The latter is proved in Section 5, and Section 6 discusses generic polynomials. In both of these, we rely heavily on the foundational studies of Katz. Section 7 concludes with the proof of Theorem 1.11, and Section 8 contains some hopefully enlightening comments concerning parts of the results of Katz that we use.

Acknowledgments. E.K. was partially supported by a DFG-SNF lead agency program grant (grant number 200020L_175755). K.S. is partially supported through a grant from the National Science Foundation, and a Simons Investigator Grant from the Simons Foundation. This work was started when K.S. was a senior Fellow at the ETH Institute for Theoretical Studies, whom he thanks for their warm and generous hospitality.

We thank W. Sawin for his comments concerning Section 5.

2. Sums of twisted-multiplicative functions

Since the proof of Theorem 1.5 follows the broad plan of our earlier work (and is not far from that of Fouvry and Michel [12, §3]), we shall be brief.

Put $z = x^{1/(\alpha \log\log x)}$ with $\alpha = 3(\mathrm{M}^2+1)$. We factor any integer $q \leqslant x$ as $q = rs$ where all prime factors of s are $\leqslant z$, and all prime factors of r are $> z$. We then have

$$\mathrm{V}(a;q) = \mathrm{V}(a;rs) = \mathrm{V}(\bar{r}a;s)\mathrm{V}(\bar{s}a;r)$$

by twisted multiplicativity, hence

$$|\mathrm{V}(a;q)| \leqslant \mathrm{G}(r)|\mathrm{V}(\bar{r}a;s)|.$$

We handle first the terms where $s \leqslant x^{1/3}$. We split the sum over $q \leqslant x$ according to the residue class of r modulo s, getting

$$\sum_{\substack{q\leqslant x\\ s\leqslant x^{1/3}}} |\mathrm{V}(a;q)| \leqslant \sum_{s\leqslant x^{1/3}} \sum_{r\leqslant x/s} \mathrm{G}(r)|\mathrm{V}(\bar{r}a;s)| \leqslant \sum_{s\leqslant x^{1/3}} \sum_{t \,(\mathrm{mod}\, s)} |\mathrm{V}(\bar{t}a;s)| \sum_{\substack{r\leqslant x/s\\ r\equiv t \,(\mathrm{mod}\, s)}} \mathrm{G}(r).$$

By Shiu's work on the Brun–Titchmarsh Theorem for multiplicative functions (see [33, Th. 1]) we may bound the sum over r above by

$$\ll \frac{x/s}{\varphi(s)\log(x/s)} \exp\Big(\sum_{z<p\leqslant x} \frac{\mathrm{G}(p)}{p}\Big) \ll \frac{x}{s\varphi(s)\log x}\Big(\frac{\log x}{\log z}\Big)^{\mathrm{M}} \ll \frac{x}{s\varphi(s)\log x}(\log\log x)^{\mathrm{M}}.$$

Therefore

$$\begin{aligned}\sum_{\substack{q\leqslant x\\ s\leqslant x^{1/3}}} |\mathrm{V}(a;q)| &\ll \frac{x}{\log x}(\log\log x)^{\mathrm{M}} \sum_{s\leqslant x^{1/3}} \frac{1}{s\varphi(s)} \sum_{t \,(\mathrm{mod}\, s)} |\mathrm{V}(\bar{t}a;s)| \\ &\ll \frac{x}{\log x}(\log\log x)^{\mathrm{M}} \sum_{s\leqslant x^{1/3}} \frac{g(s)}{\varphi(s)} \ll \frac{x}{\log x}(\log\log x)^{\mathrm{M}} \prod_{p\leqslant x}\Big(1+\frac{g(p)}{p}\Big).\end{aligned}$$

We now consider the contribution of the terms with $s > x^{1/3}$. Since $\mathrm{G}(p) \leqslant \mathrm{M}$ for all p,

$$\sum_{\substack{q\leqslant x\\ s>x^{1/3}}} |\mathrm{V}(a;q)| \leqslant \sum_{r\leqslant x^{2/3}} \mathrm{M}^{\omega(r)} \sum_{x^{1/3}<s\leqslant x/r} \mathrm{M}^{\omega(s)}.$$

Applying the Cauchy–Schwarz inequality and [27, Lemma 3.2] to the inner sum, we find that

$$\begin{aligned}\sum_{x^{1/3}<s\leqslant x/r} \mathrm{M}^{\omega(s)} &\ll \Big(\sum_{s\leqslant x/r} \mathrm{M}^{2\omega(s)}\Big)^{1/2}\Big(\sum_{x^{1/3}<s\leqslant x/r} 1\Big)^{1/2} \\ &\ll \frac{x}{r}(\log x)^{(\mathrm{M}^2-1)/2}\exp\Big(-\frac{\log(x/r)}{2\log z}\Big) \ll \frac{x}{r}(\log x)^{(\mathrm{M}^2-1)/2-\alpha/6} \ll \frac{x}{r\log x}.\end{aligned}$$

Therefore

$$\sum_{\substack{q\leqslant x\\ s>x^{1/3}}}|\mathrm{V}(a;q)|\ll\frac{x}{\log x}\sum_{r\leqslant x^{2/3}}\frac{\mathrm{M}^{\omega(r)}}{r}\ll\frac{x}{\log x}\exp\Big(\sum_{z\leqslant p\leqslant x}\frac{\mathrm{M}}{p}\Big)\ll\frac{x}{\log x}(\log\log x)^{\mathrm{M}}.$$

The proof of Theorem 1.5 is now complete.

3. Exponential sums of polynomials, preliminary results

In this section we collect together some results on the exponential sums $\mathrm{W}_f(a;p)$. We shall use and expand on some of these results in later sections. First we recall the Weil bound: if $f\in\mathbf{Z}[\mathrm{X}]$ has degree $d\geqslant 1$ and $(a,p)=1$ then

$$|\mathrm{W}_f(a;p)|\leqslant(d-1). \tag{6}$$

Next we quote a result from Shao [32, Th. 2.1].

Lemma 3.1. *Let $f\in\mathbf{Z}[\mathrm{X}]$ be a polynomial of degree d. Let κ denote the number of irreducible factors of $f(\mathrm{X})-f(\mathrm{Y})\in\mathbf{Q}[\mathrm{X},\mathrm{Y}]$. Then $\kappa\leqslant\tau(d)$ (the number of divisors of d) and for large x we have*

$$\sum_{p\leqslant x}\frac{1}{p}\Big(\frac{1}{p}\sum_{(a,p)=1}|\mathrm{W}(a;p)|^2\Big)=(\kappa-1)\log\log x+\mathrm{O}(1).$$

Proof. The asymptotic for the sum over primes is given in Theorem 2.1 of Shao [32], and the bound on κ is described in the remark after Theorem 2.1 there. □

While Lemma 3.1 involves the factorization of $\mathrm{F}(\mathrm{X},\mathrm{Y})=(f(\mathrm{X})-f(\mathrm{Y}))/(\mathrm{X}-\mathrm{Y})$ in $\mathbf{Q}[\mathrm{X},\mathrm{Y}]$, it is of greater significance to understand the factorization of $\mathrm{F}(\mathrm{X},\mathrm{Y})$ over $\bar{\mathbf{Q}}[\mathrm{X},\mathrm{Y}]$ (or equivalently over $\mathbf{C}[\mathrm{X},\mathrm{Y}]$).

Lemma 3.2. *Let $f\in\mathbf{Z}[\mathrm{X}]$ be a polynomial of degree d, and suppose that the polynomial $\mathrm{F}(\mathrm{X},\mathrm{Y})=(f(\mathrm{X})-f(\mathrm{Y}))/(\mathrm{X}-\mathrm{Y})$ factors into m irreducible factors over $\bar{\mathbf{Q}}[\mathrm{X},\mathrm{Y}]$. If $m=1$ then for all p we have*

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2=1+\mathrm{O}(p^{-1/2}).$$

If $m>1$, then there is a set of primes $\mathscr{P}$ of density $\geqslant\delta>0$ (with δ depending only on the degree d) such that for $p\in\mathscr{P}$

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2=m+\mathrm{O}(p^{-1/2}).$$

Proof. If $m=1$ then the affine curve with equation $\mathrm{F}(\mathrm{X},\mathrm{Y})=0$ is geometrically irreducible over $\mathbf{Q}$, so that for all large p it is geometrically irreducible over $\mathbf{F}_p$. Orthogonality of characters and the Riemann Hypothesis for curves over finite fields then show that

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2=\frac{1}{p}\Big|\Big\{(x,y)\in\mathbf{F}_p^2:\mathrm{F}(x,y)=0\Big\}\Big|=1+\mathrm{O}(p^{-1/2}).$$

Now suppose $m>1$, and let K be a finite Galois extension of $\mathbf{Q}$ such that $\mathrm{F}(\mathrm{X},\mathrm{Y})$ factors in $\mathrm{K}[\mathrm{X},\mathrm{Y}]$ into m different factors, each of which is irreducible in $\bar{\mathbf{Q}}[\mathrm{X},\mathrm{Y}]$. Thus the affine curve defined by $\mathrm{F}(\mathrm{X},\mathrm{Y})$ is the union of m geometrically irreducible curves over K. Note

that the degree of the field K may be bounded in terms of d. We take $\mathscr{P}$ to be the set of primes splitting completely in K. By the Chebotarev density theorem $\mathscr{P}$ has density $1/[\mathrm{K}:\mathbf{Q}]$, which is bounded away from 0 by an amount depending only on d. For $p \in \mathscr{P}$, the m geometrically irreducible components of the curve $\mathrm{F}(\mathrm{X},\mathrm{Y}) = 0$ are defined over $\mathbf{F}_p$, and the Riemann Hypothesis gives here

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}(a;p)|^2 = m + \mathrm{O}(p^{-1/2}).$$

□

Our next result is due to Fried [16, Th. 1] (see also the more elementary account by Turnwald in [34, Th. 1]). It describes when the polynomial $\mathrm{F}(\mathrm{X},\mathrm{Y}) = (f(\mathrm{X})-f(\mathrm{Y}))/(\mathrm{X}-\mathrm{Y})$ is absolutely irreducible, i.e., when $m = 1$ in the notation of the previous lemma, and therefore $\kappa = 2$ in the notation of Lemma 3.1.

We recall that for any integer $d \geqslant 0$, the Dickson polynomial $\mathrm{D}_d \in \mathbf{Z}[\mathrm{X}, a]$ is defined to be the unique polynomial such that

$$\mathrm{D}_d(\mathrm{X} + a\mathrm{X}^{-1}, a) = \mathrm{X}^d + (a/\mathrm{X})^d$$

(see, e.g., [34, §1]); in particular, $\mathrm{D}_d(\mathrm{X}, 0) = \mathrm{X}^d$.

Proposition 3.3 (Fried)**.** *Let $f \in \mathbf{Z}[\mathrm{X}]$ with degree $d \geqslant 1$ and let*

$$\mathrm{F} = (f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y}) \in \mathbf{Q}[\mathrm{X}, \mathrm{Y}].$$

(1) *If $\deg(f)$ is not an odd prime, then F is absolutely irreducible if and only if f is indecomposable in $\mathbf{Q}[\mathrm{X}]$.*

(2) *If d is an odd prime $\geqslant 5$, then F is absolutely irreducible if it is not linearly equivalent in $\mathbf{Q}[\mathrm{X}]$ to a Dickson polynomial $\mathrm{D}_d(\mathrm{X}, a)$.*

(3) *If $d = 3$, then F is absolutely irreducible if and only if f is not linearly equivalent in $\mathbf{Q}[\mathrm{X}]$ to a Dickson polynomial $\mathrm{D}_3(\mathrm{X}, 0)$.*

Putting Lemmas 3.1, 3.2 and Proposition 3.3 together, we arrive at the following corollary.

Corollary 3.4. *Let $f \in \mathbf{Z}[\mathrm{X}]$ be a polynomial of degree $d \geqslant 1$. If f is indecomposable then for large x we have*

$$\sum_{p\leqslant x} \frac{1}{p}\Big(\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}(a;p)|^2\Big) = \log\log x + \mathrm{O}(1),$$

whereas if f is decomposable then for large x we have

$$\sum_{p\leqslant x} \frac{1}{p}\Big(\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}(a;p)|^2\Big) \geqslant 2\log\log x + \mathrm{O}(1).$$

Proof. If d is prime, then κ must be $2 = \tau(d)$ in Lemma 3.1. Moreover, f is automatically indecomposable, and so the stated result holds in this case. If $f = g \circ h$ is decomposable, then $f(\mathrm{X})-f(\mathrm{Y})$ has $(\mathrm{X}-\mathrm{Y})$, $(h(\mathrm{X})-h(\mathrm{Y}))/(\mathrm{X}-\mathrm{Y})$ and $(g(h(\mathrm{X}))-g(h(\mathrm{Y})))/(h(\mathrm{X})-h(\mathrm{Y}))$ as factors, so that $\kappa \geqslant 3$ in Lemma 3.1 and the stated result holds. Finally if the degree d is composite and f is indecomposable, then the first part of Proposition 3.3 shows that $(f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y})$ is irreducible in $\bar{\mathbf{Q}}[\mathrm{X}, \mathrm{Y}]$ and therefore in $\mathbf{Q}[\mathrm{X}, \mathrm{Y}]$. Either Lemma 3.1 or Lemma 3.2 now gives the stated result. □

Lastly we consider the behavior of $\mathrm{W}(a;p)$ when f is assumed to be Sidon–Morse over $\mathbf{Q}$. Here the work of Katz permits a very precise understanding of such exponential sums.

Proposition 3.5. *Let $f \in \mathbf{Z}[\mathrm{X}]$ be a polynomial of degree d, and suppose that f is Sidon–Morse over $\mathbf{Q}$. Let K_d denote the compact group $\mathrm{USp}_{d-1}(\mathbf{C})$ if f is symmetric Sidon–Morse, and the compact group $\mathrm{SU}_{d-1}(\mathbf{C})$ if f is Sidon–Morse but not symmetric. For any integer $k \geqslant 0$ we have*

$$\lim_{p\to+\infty} \frac{1}{p} \sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}(a;p)|^{2k} = \int_{\mathrm{K}_d} |\operatorname{tr}(g)|^{2k} d\mu(g),$$

where μ is the Haar measure on K_d normalized to have total volume 1. Furthermore

$$\int_{\mathrm{USp}_{d-1}(\mathbf{C})} |\operatorname{tr}(g)|^{2k} d\mu(g) \begin{cases} = (2k-1)!! & \text{for } 1 \leqslant k \leqslant (d-1)/2 \\ \leqslant (2k-1)!! & \text{for all } k \geqslant 1, \end{cases}$$

and

$$\int_{\mathrm{SU}_{d-1}(\mathbf{C})} |\operatorname{tr}(g)|^{2k} d\mu(g) \begin{cases} = k! & \text{for } 0 \leqslant k \leqslant (d-1) \\ \leqslant k! & \text{for all } k \geqslant 0, \end{cases}$$

Proof. This is largely a consequence of the work of Katz [22]. We recall the relevant result of Katz in Theorem 6.3 below, and explain the link to the moments over K_d in Remark 6.10. Further discussion of Katz's theorem may be found in Section 8.

The moments over K_d for small k (which match the moments of a standard complex Gaussian for $\mathrm{K}_d = \mathrm{SU}_{d-1}(\mathbf{C})$, and the moments of a standard real Gaussian for $\mathrm{K}_d = \mathrm{USp}_{d-1}(\mathbf{C})$) were computed by Diaconis and Shahshahani, and the upper bounds for all k may be found in the work of Perret-Gentil [31, Prop. 2.2]. □

Remark 3.6. Katz's Theorem also leads to a possible guess for the optimal value of the upper-bound in (2), in the case of Sidon–Morse polynomials, namely

$$\delta = 1 - \int_{\mathrm{K}_d} |\operatorname{tr}(g)| d\mu(g)$$

(for instance, if $\mathrm{K}_d = \mathrm{SU}_2(\mathbf{C})$, this leads to $\delta = 1 - 8/(3\pi)$, as in (5)).

Asymptotically, for large d, we have the Gaussian approximations

$$\int_{\mathrm{K}_d} |\operatorname{tr}(g)| d\mu(g) \approx \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} |x| e^{-x^2/2} dx = \sqrt{\frac{2}{\pi}} = 0.79788456\ldots,$$

$$\int_{\mathrm{K}_d} |\operatorname{tr}(g)| d\mu(g) \approx \frac{1}{\pi} \int_{\mathbf{C}} |z| e^{-|z|^2} dz = \frac{\sqrt{\pi}}{2} = 0.8662269\ldots$$

in the $\mathrm{USp}_{d-1}(\mathbf{C})$ and $\mathrm{SU}_{d-1}(\mathbf{C})$ cases, respectively.

4. Proof of Theorem 1.1

We begin with the first part of the theorem, which seeks a bound for $\sum_{q\leqslant x} |\mathrm{W}(a;q)|^2$. We apply Theorem 1.5 to the function $q \mapsto \mathrm{W}(a;q)^2$, which is twisted-multiplicative. The Weil bound (6) allows us to take $\mathrm{G}(p) = (d-1)^2$ for all but finitely many primes. Writing

$$g(p) = \frac{1}{p} \sum_{(a,p)=1} |\mathrm{W}(a;p)|^2,$$

and recalling that f is indecomposable, Corollary 3.4 gives

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

Theorem 1.5 yields

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|^2\ll\frac{x}{\log x}\exp\Big(\sum_{p\leqslant x}\frac{g(p)}{p}\Big)(\log\log x)^{(d-1)^2}\ll x(\log\log x)^{(d-1)^2}.$$

Now we turn to the proof of the second part of the theorem, which we will deduce from Theorem 1.5 and Theorem 1.3 (to be proved in Section 5). Applying Theorem 1.5 to the twisted multiplicative function $|\mathrm{W}(a;q)|$ and using the Weil bound (which permits $\mathrm{M}=d-1$ here) we obtain

$$\sum_{q\leqslant x}|\mathrm{W}(a;q)|\ll\frac{x}{\log x}(\log\log x)^{d-1}\exp\Big(\sum_{p\leqslant x}\frac{1}{p}\Big(\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|\Big)\Big). \tag{7}$$

Let ϵ be a small positive number, and let $\mathscr{P}$ denote the set of primes p for which

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^4\geqslant 2-\epsilon.$$

By Theorem 1.3 we know that the set $\mathscr{P}$ has density $\geqslant\delta=\delta(d)>0$ with δ depending only on d. For any real number y with $|y|\leqslant d-1$ we claim that

$$|y|\leqslant\frac{1+y^2}{2},\qquad\text{and}\qquad |y|\leqslant\frac{1+y^2}{2}+\frac{3/2-y^4}{200(d-1)^4}.$$

The first inequality is clear, and so is the second inequality in the range $y^4\leqslant 3/2$. In the range $3/2<y^4\leqslant(d-1)^4$, note that $(1+y^2)/2-|y|\geqslant(1+\sqrt{3/2})/2-(3/2)^{1/4}>1/200$, so that the desired inequality holds in this case also.

Applying the first inequality above for primes $p\notin\mathscr{P}$, we find

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|\leqslant\frac{1}{2}+\frac{1}{2p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2,$$

while applying the second inequality above for primes $p\in\mathscr{P}$ we find

$$\begin{aligned}\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|&\leqslant\frac{1}{2}+\frac{1}{2p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2+\frac{1}{200(d-1)^4}\Big(\frac{3}{2}-\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^4\Big)\\&\leqslant\frac{1}{2}+\frac{1}{2p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^2-\frac{1}{400(d-1)^4}.\end{aligned}$$

Combining both inequalities, and using the first part of Corollary 3.4, we conclude that

$$\sum_{p\leqslant x}\frac{1}{p^2}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|\leqslant\Big(\frac{1}{2}+\frac{1}{2}-\frac{\delta}{400(d-1)^4}+o(1)\Big)\log\log x.$$

Inserting this bound in (7), the second part of the theorem follows.

5. The fourth moment: Proof of Theorem 1.3

As we shall see, for Sidon–Morse polynomials, the work of Katz [22] can be used to show that Case (1) of Theorem 1.3 holds. The main challenge is to handle all polynomials of degree $\geqslant 3$, and not just the generic ones.

Let $f \in \mathbf{Z}[\mathrm{X}]$ be a polynomial with $d = \deg(f) \geqslant 3$. If $(f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y})$ is not absolutely irreducible, then Lemma 3.2 shows that there is a positive density of primes on which the second moment of $\mathrm{W}(a; p)$ is at least $2 + \mathrm{O}(p^{-1/2})$, so that by Cauchy–Schwarz a stronger form of the second case of Theorem 1.3 holds (with the fourth moment being $\geqslant 4 + \mathrm{O}(p^{-1/2})$).

From now on, we will therefore assume that the polynomial

$$\mathrm{F}(\mathrm{X}, \mathrm{Y}) = (f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y})$$

is absolutely irreducible. The remaining part of the proof will use in an essential way the algebraic interpretation of the exponential sums $\mathrm{W}(a; p)$, which goes back to Weil, and it seems difficult to prove the lower bound for the fourth moment with a direct elementary argument.

Fix a prime ℓ (for instance $\ell = 2$); all primes p below will be assumed to be different from ℓ and to be larger than d. Let ι be a fixed isomorphism $\bar{\mathbf{Q}}_\ell \to \mathbf{C}$; we use it to identify ℓ-adic numbers and complex numbers.

Let $p \neq \ell$, $p > d$, be a prime number. We denote by ψ_p the ℓ-adic additive character of $\mathbf{F}_p$ such that

$$\iota(\psi_p(a)) = e\Big(\frac{a}{p}\Big)$$

for $a \in \mathbf{F}_p$.

Let $\mathscr{G}_p$ be the ℓ-adic sheaf $f_*\bar{\mathbf{Q}}_\ell/\bar{\mathbf{Q}}_\ell$ on the affine line $\mathbf{A}^1_{\mathbf{F}_p}$; it has rank $d-1$ and is everywhere tamely ramified (since $p > d$). The sheaf $\mathscr{G}_p$ is a Fourier sheaf in the sense of Katz ([22, 7.3.5]), and we denote by $\mathscr{F}_p$ its (unitarily normalized) Fourier transform with respect to ψ_p (defined in [22, 7.3.3], up to the normalization). The trace function of $\mathscr{F}_p$ takes value 0 for $a = 0$ and takes value (after applying ι)

$$\frac{1}{\sqrt{p}} \sum_{x \in \mathbf{F}_p} e\Big(\frac{af(x)}{p}\Big) = \mathrm{W}(a; p)$$

for $a \in \mathbf{F}_p^\times$ (see [22, Th. 7.3.8, (4)], where again the Fourier transform is not normalized). The rank of $\mathscr{F}_p$ is also equal to $d-1$, and $\mathscr{F}_p$ is lisse and pure of weight 0 outside 0 and ∞ (see [22, Lemma 7.3.9]).

Lemma 5.1. *If the polynomial $(f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y})$ is absolutely irreducible over $\mathbf{Q}$, then for all p large enough, the sheaf $\mathscr{F}_p$ is geometrically ireducible.*

Proof. This is a Fourier-side variant of Lemma 3.2. If the polynomial

$$\mathrm{F}(\mathrm{X}, \mathrm{Y}) = (f(\mathrm{X}) - f(\mathrm{Y}))/(\mathrm{X} - \mathrm{Y})$$

is absolutely irreducible, then the curve $\mathrm{C}_{f,p}$ over $\mathbf{F}_p$ with equation

$$(f(x) - f(y))/(x - y) = 0$$

is geometrically irreducible, which by the Riemann Hypothesis for curves implies that as $\nu \to +\infty$, we have

$$|\mathrm{C}_{f,p}(\mathbf{F}_{p^\nu})| \sim p^\nu.$$

But the discrete Parseval formula implies that

$$\frac{1}{p^\nu}|\mathrm{C}_{f,p}(\mathbf{F}_{p^\nu})| = \frac{1}{p^\nu}\sum_{a\in\mathbf{F}_{p^\nu}^\times}\Big|\frac{1}{p^{\nu/2}}\sum_{x\in\mathbf{F}_{p^\nu}} e\Big(\frac{\mathrm{tr}(af(x))}{p}\Big)\Big|^2$$

(with the trace from $\mathbf{F}_{p^\nu}$ to $\mathbf{F}_p$) so we obtain

$$\lim_{\nu\to+\infty}\frac{1}{p^\nu}\sum_{a\in\mathbf{F}_{p^\nu}^\times}\Big|\frac{1}{p^{\nu/2}}\sum_{x\in\mathbf{F}_{p^\nu}} e\Big(\frac{\mathrm{tr}(af(x))}{p}\Big)\Big|^2 = 1,$$

and this implies that $\mathscr{F}_p$ is geometrically irreducible by Katz's diophantine criterion for irreducibility (see e.g. [26, Lemma 4.14]). □

We now consider only primes p such that the sheaf $\mathscr{F}_p$ is geometrically irreducible.

Let G_p be the arithmetic monodromy group of $\mathscr{F}_p$ and G_p^g the geometric monodromy subgroup; we can view these as algebraic subgroups of $\mathrm{GL}_{d-1}(\bar{\mathbf{Q}}_\ell)$. The irreducibility property of $\mathscr{F}_p$ means that G_p^g acts irreducibly on $\bar{\mathbf{Q}}_\ell^{d-1}$.

By a deep theorem of Deligne (see [11, Th. 3.4.1 (iii) and Cor. 1.3.9]), the connected component of the identity $\mathrm{G}_{p,0}^g$ of the group G_p^g is semisimple. It is invariant under all automorphisms of G_p^g, hence it is a normal subgroup of G_p (since inner automorphisms of G_p induce automorphisms of its normal subgroup G_p^g). Let f_p denote a fixed element of the conjugacy class of the Frobenius automorphism at p.

Let $\mathscr{E}_p$ be the sheaf $\mathrm{End}(\mathrm{End}(\mathscr{F}_p))$. Its trace function for $a \in \mathbf{F}_p^\times$ is $|\mathrm{W}(a;p)|^4$.

Let V_p be the subspace $\mathrm{End}(\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1}))^{\mathrm{G}_p^g}$ of vectors invariant under G_p^g, the action of G_p on the space $\mathrm{End}(\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1}))$ being "the obvious one" induced by the action on $\bar{\mathbf{Q}}_\ell^{d-1}$ (if a group G acts on a vector space E, it acts on End(E) by $g\cdot u = g\circ u\circ g^{-1}$).

Applying the Grothendieck–Lefschetz trace formula and Deligne's version of the Riemann Hypothesis, we get a formula

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}|\mathrm{W}(a;p)|^4 = \iota(\mathrm{tr}(f_p|\mathrm{V}_p)) + \mathrm{O}(p^{-1/2}) \tag{8}$$

where the implied constant depends only on d (e.g. by conductor estimates, much as in [14, Th. 9.1]).

Proposition 5.2. *There exists a finite Galois extension* K *of* $\mathbf{Q}$ *of degree bounded in terms of* d *only such that for all but finitely many primes* p *that are totally split in* K*, the action of* f_p *on* $\mathrm{V}_p = \mathrm{End}(\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1}))^{\mathrm{G}_p^g}$ *is trivial.*

Let us admit this proposition and conclude the proof of Theorem 1.3. For primes totally split in the number field K, we have $\iota(\mathrm{tr}(f_p|\mathrm{V}_p)) = \dim(\mathrm{V}_p)$. On the other hand, the definition of the action of G_p^g on $\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1})$ shows that the space V_p is the space of all linear maps $\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1}) \to \mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1})$ which commute with the G_p^g-action. The identity is an element of this space, so its dimension is $\geqslant 1$. Since the action on $\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1})$ is semisimple (e.g. by Deligne's Theorem [11, Th. 3.4.1] because it is still pure of weight 0), Schur's Lemma

in representation theory (see, e.g., [25, Prop. 2.7.15 (3)]) implies that the dimension of V_p is exactly 1 if and only if the action of G_p^g on $\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1})$ is irreducible. But V_p contains both the multiples of the identity and the space $\mathrm{End}^0(\bar{\mathbf{Q}}_\ell^{d-1})$ of matrices of trace zero as stable subspaces, so this irreducibility can only hold if $\mathrm{End}^0(\bar{\mathbf{Q}}_\ell^{d-1})$ is zero, i.e., if $d = 2$. So for primes totally split in K, we have $\dim(V_p) \geqslant 2$ hence

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times} |W(a;p)|^4 \geqslant 2 + O(p^{-1/2})$$

by (8).

To improve on this unless the limit is equal to 2, we use very deep work of Katz [22, Th. 14.3.4] that implies that $G_{p,0}^g$ is independent of p for all p large enough. Take a prime p large enough so that $G_{p,0}^g$ has stabilized and suppose that $\dim(V_p) = 2$ for some p split in K. Then the group $G_{p,0}^g$ must act irreducibly on matrices of trace zero. But the Lie algebra of G_p^g is a stable subspace, so that we must have $\mathrm{Lie}(G_{p,0}^g) = \mathrm{End}^0(\bar{\mathbf{Q}}_\ell^{d-1})$. That means that $G_{p,0}^g$ is equal to $\mathrm{SL}_{d-1}(\bar{\mathbf{Q}}_\ell)$. Then for all primes p large enough we have $ZG_{p,0}^g = \mathrm{GL}_{d-1}(\bar{\mathbf{Q}}_\ell)$, where Z is the group of scalar matrices in $\mathrm{GL}_{d-1}(\bar{\mathbf{Q}}_\ell)$, which implies that f_p acts trivially for all p large enough, and then that the limit of the fourth moments exists and is equal to 2.

To finally show that the constant 2 is best possible, we recall that Katz has proved that if f is a Sidon–Morse polynomial (e.g., the derivative f' has Galois group S_{d-1}), then G_p^g contains $\mathrm{SL}_{d-1}(\bar{\mathbf{Q}}_\ell)$ for all p large enough (see Theorem 6.3), in which case it is well-known that the action of G_p^g on the space of matrices of trace zero is irreducible, so that the dimension of V_p is then equal to 2 for all p large enough.

Remark 5.3. The arguments above are related to the easiest part of the Larsen Alternative [24].

Proof of Proposition 5.2. We will begin by proving the statement without the information that the degree of K can be bounded in terms of d only, since the latter requires extra ingredients.

Step 1. We first prove that, for all primes p large enough, the action of f_p on V_p is of finite order. Since we are assuming that G_p^g acts irreducibly on $\bar{\mathbf{Q}}_\ell^{d-1}$, a result of Katz shows that the connected component of the identity $G_{p,0}^g$ of G_p^g acts irreducibly on $\bar{\mathbf{Q}}_\ell^{d-1}$, provided p is large enough (see [22, 7.7.3, Lemma 7.7.5]).

Recall that the group of outer automorphisms of $G_{p,0}^g$ is the group $\mathrm{Out}(G_{p,0}^g)$ of automorphisms modulo inner automorphisms. For $g \in G_p$, let $\alpha_p(g) \in \mathrm{Out}(G_{p,0}^g)$ be the class modulo inner automorphisms of the automorphism $x \mapsto xgx^{-1}$ of $G_{p,0}^g$ (it is an automorphism since $G_{p,0}^g$ is normal in G_p). This defines a group homomorphism

$$G_p \xrightarrow{\alpha_p} \mathrm{Out}(G_{p,0}^g).$$

We claim that the kernel of α_p is $G_{p,0}^g Z \cap G_p$ where Z is again the group of scalar matrices in $\mathrm{GL}_{d-1}(\bar{\mathbf{Q}}_\ell)$. Indeed, the condition $\alpha_p(g) = 1$ means that there exists $h \in G_{p,0}^g$ such that $gxg^{-1} = hxh^{-1}$ for all $x \in G_{p,0}^g$, which is equivalent to $h^{-1}g$ belonging to the centralizer of $G_{p,0}^g$ in $\mathrm{GL}_{d-1}(\bar{\mathbf{Q}}_\ell)$, or in other words, to $h^{-1}g$ commuting with the action of $G_{p,0}^g$ on $\bar{\mathbf{Q}}_\ell^{d-1}$. By Schur's Lemma (see, e.g., [25, Prop. 2.7.15 (2)]), the irreducibility of the action of $G_{p,0}^g$ implies that this centralizer is equal to Z. Thus $g \in \ker(\alpha_p)$ is equivalent to $g \in G_{p,0}^g Z \cap G_p$.

We deduce therefore that we have an injective group homomorphism

$$\mathrm{G}_p/(\mathrm{G}_{p,0}^g\mathrm{Z}\cap\mathrm{G}_p)\xrightarrow{\alpha_p}\mathrm{Out}(\mathrm{G}_{p,0}^g).$$

Because $\mathrm{G}_{p,0}^g$ is a connected semisimple group, its outer automorphism group is finite (see, e.g., [4, p. 42, prop. 18] in the case of compact groups). Hence α_p injects $\mathrm{G}_p/(\mathrm{G}_{p,0}^g\mathrm{Z}\cap\mathrm{G}_p)$ in a finite group. Since Z acts trivially on End(W) for any representation W, and since G_p^g acts trivially on V_p, this shows that the order of the action of f_p on V_p is a divisor of the order of the outer automorphism group.

Step 2. We next prove that there exists a finite-dimensional continuous ℓ-adic Galois representation

$$\varrho\colon\ \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})\to\mathrm{GL}(\mathrm{E})$$

for some $\bar{\mathbf{Q}}_\ell$-vector space E, such that for all but finitely many primes, the action of Frobenius at p on E "is" is the same as the action of f_p on V_p. It is enough to define a constructible ℓ-adic sheaf $\mathscr{V}$ on $\mathrm{Spec}(\mathbf{Z}[1/\ell\mathrm{N}])$ for some integer $\mathrm{N}\geqslant 1$ such that the stalk over all but finitely many primes p "is" the space V_p, and such that the action of f_p coincides with the action of the Frobenius at p. Indeed, this sheaf $\mathscr{V}$ will be lisse outside of a finite set S of primes, hence will correspond to a Galois representation of the Galois group of the maximal extension unramified outside S, and this is a quotient of the Galois group of $\mathbf{Q}$.

To construct $\mathscr{V}$, we use [26, Lemma 4.23] (see also [26, Lemma 4.27] for a more difficult application), applied to the data

$$(\mathrm{X},\mathrm{Y},f,g)=(\mathbf{A}^4,\mathrm{Spec}(\mathbf{Z}[1/\ell]),\text{the structure morphism},$$
$$g(x,y,z,w)=f(x)+f(y)-f(z)-f(w))$$

and take the second cohomology sheaf of the complex resulting from this application of [26, Lemma 4.23].

That this "works" results from the expression

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times}\Big|\frac{1}{\sqrt{p}}\sum_{x\in\mathbf{F}_p}e\Big(\frac{af(x)}{p}\Big)\Big|^4=\frac{1}{p^3}\sum_{x,y,z,w\in\mathbf{F}_p}\ \sum_{a\in\mathbf{F}_p^\times}e\Big(\frac{ag(x,y,z,w)}{p}\Big),$$

combined with the cohomological expression

$$\mathrm{V}_p\simeq\mathrm{H}_c^2(\mathbf{G}_m\times\bar{\mathbf{F}}_p,\mathrm{End}(\mathrm{End}(\mathscr{F}_p)))(1). \tag{9}$$

Step 3. By the compatibility with Frobenius of the isomorphism (9) in Step 2, and by Step 1, the action of Frobenius at p under ϱ is of finite order for all but finitely many primes p. The image H of ϱ is a compact ℓ-adic Lie group (identifying GL(E) with $\mathrm{GL}_m(\bar{\mathbf{Q}}_\ell)$ for some $m\geqslant 1$, we first note that H is contained in $\mathrm{GL}_m(\mathrm{L})$ for some finite extension L of $\mathbf{Q}_\ell$, by an oft-rediscovered lemma – see for instance [23, Lemma 9.0.8] – and then it is a closed subgroup of an ℓ-adic Lie group, hence itself an ℓ-adic Lie group by, e.g., [1, p. 227, th. 2]). It follows from [1, Cor. 1, p. 169] that there is a neighborhood U of $1\in\mathrm{H}$ which contains no non-trivial finite subgroup; there is then a number field K such that the finite-index subgroup $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathrm{K})$ maps to U. All the Frobenius elements in this subgroup (which exist outside any given finite set of primes because Frobenius elements are dense, by a form of Chebotarev's density theorem) must map to the identity, which means that $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathrm{K})$ is in the kernel of ϱ. This implies that for a prime p that is totally split in K, the action

of f_p, which "is" the action of Frobenius under ϱ, is trivial. This proves the result, up to the bound on the degree of K.

Step 4. Now we explain how to bound the degree of K in terms of d only.

The first ingredient is a fact from the theory of finite groups: for given positive integers k and m, if Γ is a finite subgroup of $\mathrm{GL}_k(\bar{\mathbf{Q}}_\ell)$ such that all elements of Γ have order dividing m, then the order of Γ is bounded in terms of k and m only. Indeed, by a well-known theorem of Jordan (see, e.g., [10, Th. 36.13]), there exists a normal abelian subgroup Γ_0 of Γ of index bounded in terms of k and m. This reduces the problem to the abelian case; but Γ_0 can be diagonalized, and the bound on the order of its elements show that Γ_0 is isomorphic to a subgroup of $(\mathbf{Z}/m\mathbf{Z})^k$, hence the result.

We want to apply this to the image $\Gamma \subset \mathrm{GL}(\mathrm{E})$ of the Galois representation ϱ. We have $\dim(\mathrm{E}) \leqslant (d-1)^4$. By the Chebotarev Density Theorem, it is then enough to prove that the order of the action of f_p on V_p is uniformly bounded in terms of d only. For this we use the fact that there are, up to isomorphism, only finitely many possibilities for $\mathrm{G}_{p,0}^g$, since it is a connected and semisimple subgroup of GL_{d-1} (this follows, in the equivalent case of compact Lie groups, from the discussion in [4, §4, n^{0}9, Scholie], which shows that such subgroups are classified by their root system R, which here has rank $\leqslant d-1$, which gives only finitely many possibilities, and for each root system R by a subgroup of the quotient $\mathrm{Q}(\mathrm{R})/\mathrm{P}(\mathrm{R})$ discussed in loc. cit.; since this quotient is finite by [2, §1, n^{o}10], there are again only finitely many possibilities). So the order of f_p is a divisor of the order of one of finitely many finite groups (depending only on d). □

Remark 5.4. The argument in Step 4 shows that it is possible to give an effective value for the constant γ in Theorem 1.1. Indeed, the index of Γ_0 in Jordan's Theorem can be bounded effectively (for instance, one gets from [10, Th. 36.14] that

$$|\Gamma| \leqslant m^k|\Gamma_0| \leqslant m^k(\sqrt{8k}+1)^{2k^2},$$

and better bounds are known), and the order of the groups $\mathrm{G}_{p,0}/\mathrm{G}_{0,0}^g$ can also be bounded effectively from the classification of roots systems.

6. Generic polynomials

In this section, we will prove the kind of non-correlation estimates modulo primes that are needed in the proof of Theorem 1.11. We also explain Proposition 3.5 at the end.

We first make some remarks concerning Sidon–Morse polynomials:

Lemma 6.1. *Let* K *be any field and let* $f \in \mathrm{K}[\mathrm{X}]$ *be a Morse polynomial of degree* $d \geqslant 2$.

(1) *The polynomial* f *is indecomposable over* K.

(2) *For any* $c \in \mathrm{K}$, *the polynomials* $f+c$ *and* $-f+c$ *are Morse polynomials. If* f *is a Sidon–Morse polynomial, then* $f+c$ *and* $-f+c$ *are Sidon–Morse polynomials.*

Proof. (1) We show that if f is decomposable, then it is not a Morse polynomial. Let $f = g \circ h$ where $\deg(g) \geqslant 2$ and $\deg(h) \geqslant 2$ be a decomposable polynomial. Note that p does not divide either $\deg(g)$ or $\deg(f)$ since $p \nmid d$.

For any critical point α of g, the critical values of f contain, with multiplicity, the values $g(h(\beta))$ where $h(\beta) = \alpha$. This will give rise to a critical value with multiplicity at least 2 *unless* $h - \alpha = \gamma(\mathrm{X}-\beta)^{\deg(h)}$ for some $\gamma \in \mathrm{K}^\times$. Since p does not divide $\deg(h)$, this can

only occur for a single value of α, so that g is of the form

$$g = \delta(\mathrm{X} - \alpha)^{\deg(g)} + \eta$$

for some $\delta \in \mathrm{K}^\times$ and $\eta \in \mathrm{K}$. Then we get

$$g \circ h = \eta + \delta\gamma^{\deg(g)}(\mathrm{X} - \beta)^d,$$

which has a single critical value, and is therefore not a Morse polynomial.

(2) This is straightforward from the definition, since the critical points of $g = f + c$ (resp. $g = -f + c$) are the same as those of f, so the critical values of g are those of f translated by c (resp. the negative of those of f, translated by c). □

Let p be a prime number and $f \in \mathbf{F}_p[\mathrm{X}]$ a Sidon–Morse polynomial. We define the ℓ-adic sheaf $\mathscr{F}_f$ associated to f as in the previous section. We will now normalize it in a specific way. We denote by $\mathscr{L}_2$ the Kummer sheaf associated to the Legendre character, with trace function $a \mapsto (a/p)$.

Definition 6.2 (Normalized sheaf)**.** Let p be a prime and $f \in \mathbf{F}_p[\mathrm{X}]$ a Sidon–Morse polynomial with $p \nmid \deg(f) - 1$.

(1) If f is not symmetric Sidon, then there is a unique $c \in \mathbf{F}_p$ such that the sum of the critical values of $f + c$ is equal to 0, and the *normalized* sheaf $\widetilde{\mathscr{F}_f}$ of f is defined to be

$$\widetilde{\mathscr{F}_f} = \mathscr{F}_{f+c} \otimes \mathscr{L}_2^{d-1}.$$

We then say that $c = c_f$ is the *critical shift* of f.

(2) If f is symmetric Sidon polynomial, and

$$f = g(\beta\mathrm{X} + \gamma) + \delta$$

where g is odd, then we put

$$\widetilde{\mathscr{F}_f} = \mathscr{F}_g.$$

We note that the sum of critical values of g is then equal to 0.

The trace function of $\widetilde{\mathscr{F}_f}$ is 0 for $a = 0$ and for $a \in \mathbf{F}_p^\times$ is given either by

$$\widetilde{\mathrm{W}}_f(a;p) = \frac{1}{\sqrt{p}}\Big(\frac{a}{p}\Big)^{d-1} \sum_{x \in \mathbf{F}_p} e\Big(\frac{a(f(x)+c)}{p}\Big) = \Big(\frac{a}{p}\Big)^{d-1} e\Big(\frac{ac}{p}\Big) \sum_{x \in \mathbf{F}_p} e\Big(\frac{af(x)}{p}\Big) \tag{10}$$

or by

$$\begin{aligned}\widetilde{\mathrm{W}}_f(a;p) &= \frac{1}{\sqrt{p}} \sum_{x \in \mathbf{F}_p} e\Big(\frac{ag(x)}{p}\Big) \\ &= \frac{1}{\sqrt{p}} e\Big(-\frac{a\delta}{p}\Big) \sum_{x \in \mathbf{F}_p} e\Big(\frac{af((x-\gamma)/\beta)}{p}\Big) = e\Big(-\frac{a\delta}{p}\Big)\mathrm{W}_f(a;p).\end{aligned} \tag{11}$$

in the symmetric case. In particular, we see that in all cases, the formula

$$|\widetilde{\mathrm{W}}_f(a;p)| = |\mathrm{W}_f(a;p)|$$

is valid all a modulo p.

The point of this normalization is the following theorem of Katz:

Theorem 6.3 (Katz)**.** *Let p be a prime number. Let $f \in \mathbf{F}_p[\mathrm{X}]$ be a Sidon–Morse polynomial of degree $d \geqslant 3$. Assume that $p > 2d-1$ and that $p \nmid d-1$.*

(1) If f is not a symmetric Sidon–Morse polynomial, then the geometric monodromy group of $\widetilde{\mathscr{F}_f}$ is equal to $\mathrm{SL}_{d-1}(\bar{\mathbf{Q}}_\ell)$.

(2) If f is a symmetric Sidon–Morse polynomial, which implies that d is odd, then the geometric monodromy group of $\widetilde{\mathscr{F}_f}$ is isomorphic to $\mathrm{Sp}_{d-1}(\bar{\mathbf{Q}}_\ell)$.

Proof. (1) If f is not of symmetric type, then the geometric monodromy group contains SL_{d-1} under the assumption on p, by [22, Th. 7.9.6], and has trivial determinant by [22, Lemma 7.10.4, (2)], so it must be SL_{d-1}.

(2) If f is of symmetric type, then under the assumption on p, a conjugate of the geometric monodromy group of $\widetilde{\mathscr{F}_f}$ is contained in Sp_{d-1} by [22, Lemma 7.10.4, (3)] (since the associated polynomial g is odd). By [22, Th. 7.9.7], it contains either SL_{d-1} or Sp_{d-1} or SO_{d-1}; the only possibility that is compatible with both these facts is that it is Sp_{d-1}. □

Remark 6.4. If we consider a Morse polynomial f such that the set of critical values is a symmetric Sidon set, we might hope that (2) still holds. However, although one can still deduce from the work of Katz that the geometric monodromy group of $\widetilde{\mathscr{F}_f}$ contains a symplectic group, we currently do not know if this condition is sufficient to ensure that $\widetilde{\mathscr{F}}$ has conversely a symplectic symmetry.

We will also need a result that is essentially a consequence of the ideas of Fried.

Proposition 6.5. *Let p be a prime number. Let f and g in $\mathbf{F}_p[\mathrm{X}]$ be Sidon–Morse polynomials of respective degree $d_f \geqslant 3$ and $d_g \geqslant 3$. Assume that $d_f < p$ and $d_g < p$.*

If f and g are not linearly equivalent over $\bar{\mathbf{F}}_p$, then $f(\mathrm{X}) - g(\mathrm{Y}) + c$ and $f(\mathrm{X}) + g(\mathrm{Y}) + c$ are absolutely irreducible for any c.

Proof. Since $f + c$ is a Sidon–Morse polynomial (Lemma 6.1), and linearly equivalent to g if and only f is, we can assume that $c = 0$. Since $-g$ is a Sidon–Morse polynomial, and linearly equivalent to f if and only if so is g, we need only consider the case of $f(\mathrm{X}) - g(\mathrm{Y})$.

Let G be the Galois group of the equation $f(\mathrm{X}) - \mathrm{Y} = 0$ over the field $\bar{\mathbf{F}}_p(\mathrm{Y})$ (so X is the variable). If $f(\mathrm{X}) - g(\mathrm{Y})$ is not absolutely irreducible then G is also isomorphic to the one for the equation $g(\mathrm{X}) - \mathrm{Y} = 0$ by [9, §2.1.1].[1] By [9, §2.1.4], if f and g are not linearly equivalent over $\bar{\mathbf{F}}_p$, then the faithful permutation representations of G on the roots of these two equations are not equivalent as permutation representations, but have the same character (i.e., are equivalent as linear representations). However, for Sidon–Morse polynomials f and g, the group G and its permutation representation are isomorphic to $\mathfrak{S}_d$ with the standard permutation representation on d letters (see [22, Proof of Lemma 7.10.2.3]). But this is a contradiction, since this faithful permutation representation of $\mathfrak{S}_d$ is characterized by its character (the only non-obvious case is when $d = 6$ and we consider the standard permutation representation and that given by a non-trivial outer automorphism of $\mathfrak{S}_6$, but these have different characters, e.g. because a transposition is mapped to, respectively, a transposition, with 4 fixed points, or a product of three disjoint transpositions, without fixed points). □

[1] This is written for the base field $\mathbf{C}$, but the argument extends to any algebraically closed field when the polynomials involved have degree less than the characteristic of the field.

Proposition 6.6. *Let p be a prime. Let $m \geqslant 1$ be an integer and let $f_1, \ldots, f_m$ be Sidon–Morse polynomials in $\mathbf{F}_p[\mathrm{X}]$. Assume that $p > 2\deg(f_i) - 1$ and $p \nmid (\deg(f_i) - 1)$ for all i. Assume also that for all $i \neq j$, the polynomials f_i and f_j are not linearly equivalent over $\bar{\mathbf{F}}_p$.*

Then the geometric monodromy group of the sheaf

$$\bigoplus_{1 \leqslant i \leqslant m} \widetilde{\mathscr{F}_{f_i}}$$

is the direct product of the geometric monodromy groups of the sheaves $\widetilde{\mathscr{F}_{f_i}}$.

Proof. We write $d_i = \deg(f_i)$ and $\widetilde{\mathscr{F}_i} = \widetilde{\mathscr{F}_{f_i}}$. We also denote by $\widetilde{\mathscr{F}_i}^\vee$ the dual of $\widetilde{\mathscr{F}_i}$.

We will apply the Goursat–Kolchin–Ribet Criterion, as developed by Katz [22, Prop. 1.8.2], and expounded by Fouvry, Kowalski and Michel [15, Lemma 2.4]. In the language of loc. cit., it suffices to check that the family $(\widetilde{\mathscr{F}_i})$ is $\mathbf{G}_m$-generous ([15, Def. 2.1]), since the individual geometric monodromy groups of $\widetilde{\mathscr{F}_i}$ are connected by Theorem 6.3.

This desired property is the combination of four conditions. Condition (1) holds because the sheaves $\widetilde{\mathscr{F}_i}$ are pure of weight 0 on $\mathbf{G}_m$, and have a geometric monodromy group (namely SL_{d_i-1} or Sp_{d_i-1} by Theorem 6.3) that acts irreducibly on $\bar{\mathbf{Q}}_\ell^{d_i-1}$. Conditions (2) and (3) are then known properties of SL_{d_i-1} and Sp_{d_i-1} (see [15, §3.1]).

To prove the most important Condition (4), it is enough to check that if $i \neq j$, there is no geometric isomorphism

$$\widetilde{\mathscr{F}_i} \simeq \widetilde{\mathscr{F}_j} \otimes \mathscr{L}, \quad \text{or} \quad \widetilde{\mathscr{F}_i}^\vee \simeq \widetilde{\mathscr{F}_j} \otimes \mathscr{L} \tag{12}$$

where $\mathscr{L}$ is a rank one sheaf lisse on $\mathbf{G}_m$ (see [15, Remark 2.2]). This is impossible unless $d_i = d_j$ and unless either none or both of f_i and f_j are symmetric Sidon–Morse. We now assume that $d_i = d_j$ and we denote by d this common value.

Case (1). Assume first that neither f_i nor f_j is symmetric, and that we have the isomorphism $\widetilde{\mathscr{F}_i} \simeq \widetilde{\mathscr{F}_j} \otimes \mathscr{L}$ in (12). We denote by c_i and c_j the critical shifts of f_i and f_j.

We recall that since $p > 2d-1$, the sheaf $\widetilde{\mathscr{F}_i}$ is, for all i, tamely ramified at 0 ([22, Lemma 7.10.4, (1)]), with local monodromy isomorphic to the sum of the non-trivial characters of order d ([22, Lemma 7.10.4, (1)]). These must be permuted by multiplication by the monodromy character χ_0 of $\mathscr{L}$ at 0, which is only possible if $\chi_0 = 1$, i.e., if $\mathscr{L}$ is lisse at 0.

Next, by [22, Th. 7.8.4, (2)] and the construction of $\widetilde{\mathscr{F}_i}$, the wild monodromy representation of $\widetilde{\mathscr{F}_i}$ at ∞ is the direct sum

$$\bigoplus_{v \in \mathrm{V}_i} \mathscr{L}_{\psi(v\mathrm{X})} \tag{13}$$

where V_i is the set of critical values of $f_i + c_i$, and $\mathscr{L}_{\psi(v\mathrm{X})}$ denotes the Artin–Schreier sheaf modulo p with trace function $a \mapsto e(av/p)$. Let $v \in \mathrm{V}_i$. The putative isomorphism $\widetilde{\mathscr{F}_i} \simeq \widetilde{\mathscr{F}_j} \otimes \mathscr{L}$ implies that there exists $w \in \mathrm{V}_j$ such that

$$\mathscr{L}_{\psi(v\mathrm{X})} = \mathscr{L} \otimes \mathscr{L}_{\psi(w\mathrm{X})},$$

as representations of the wild inertia group at ∞. In particular, $\mathscr{L}$ is an Artin–Schreier sheaf at infinity, say $\mathscr{L} \simeq \mathscr{L}_{\psi(c\mathrm{X})}$ for some c, as representations of the wild inertia group.

The local isomorphism becomes

$$\bigoplus_{v\in \mathrm{V}_i} \mathscr{L}_{\psi(v\mathrm{X})} \simeq \bigoplus_{w\in \mathrm{V}_j} \mathscr{L}_{\psi((c+w)\mathrm{X})},$$

so that $\mathrm{V}_i = \mathrm{V}_j + c$ as subsets of $\bar{\mathbf{F}}_p$. But taking the sum of the values on both sides, and using the definition of the normalized sheaf, we deduce that $c = 0$. Thus the sheaf $\mathscr{L}$ is trivial on the wild monodromy group, and therefore is also tamely ramified at ∞.

Since $\mathscr{L}$ is lisse on $\mathbf{G}_m$ and tame, it is a Kummer sheaf attached to some multiplicative character χ of $\mathbf{F}_p^\times$ (which is its trace function). Since it is lisse at 0, this character must be trivial. Hence we deduce that $\widetilde{\mathscr{F}_i}$ and $\widetilde{\mathscr{F}_j}$ are in fact geometrically isomorphic.

By the Diophantine Criterion for Irreducibility (see e.g. [26, Lemma 4.14]), this implies that

$$\limsup_{\nu\to+\infty} \frac{1}{p^\nu}\Big|\sum_{a\in\mathbf{F}_{p^\nu}^\times} \widetilde{\mathrm{W}}_i(a;p^\nu)\overline{\widetilde{\mathrm{W}}_j(a;p^\nu)}\Big| = \limsup_{\nu\to+\infty} \frac{1}{p^\nu}\sum_{a\in\mathbf{F}_{p^\nu}^\times} |\widetilde{\mathrm{W}}_i(a;p^\nu)|^2 = 1, \tag{14}$$

where $\widetilde{\mathrm{W}}_i(a;p^\nu)$ is the trace function of $\widetilde{\mathscr{F}_i}$ over the extension of degree ν of $\mathbf{F}_p$. By (10) and orthogonality of characters, the sum on the left-hand side is equal to

$$\frac{1}{p^\nu}|\{(x,y)\in\mathbf{F}_{p^\nu}^2 \mid f_i(x)+c_i = f_j(y)+c_j\}| - 1$$

(noting that if the trace function of f_i has the Legendre factor, then so does f_j, and they cancel out). If the polynomial $f_i(\mathrm{X}) - f_j(\mathrm{Y}) + c_i - c_j$ is absolutely irreducible, then we get

$$\frac{1}{p^\nu}|\{(x,y)\in\mathbf{F}_{p^\nu}^2 \mid f_i(x)+c_i = f_j(y)+c_j\}| - 1 \ll p^{-\nu/2}$$

by the Riemann Hypothesis for curves, which contradicts (14). Thus the polynomial

$$f_i(\mathrm{X}) - f_j(\mathrm{Y}) + c_i - c_j$$

is not absolutely irreducible, which can only happen if f_i and f_j are linearly equivalent over $\bar{\mathbf{F}}_p$ (Proposition 6.5).

Case 2. We continue assuming that neither f_i nor f_j is symmetric, and consider the second case of an hypothetical isomorphism (12). It is elementary that the dual $\widetilde{\mathscr{F}_i}^\vee$ is the normalized sheaf associated to $-f_i$ (because $\mathscr{F}_i$ is the Fourier transform of a sheaf that is self-dual, being the direct image of the self-dual constant sheaf; see [22, Th. 7.3.8, (2)]). Thus we are reduced to the previous case.

Case 3. Now we assume that f_i and f_j are symmetric. Since $\widetilde{\mathscr{F}_i}$ and $\widetilde{\mathscr{F}_j}$ are then self-dual by Theorem 6.3 (2), we need only exclude the possibility of a geometric isomorphism of the form

$$\widetilde{\mathscr{F}_i} \simeq \widetilde{\mathscr{F}_j} \otimes \mathscr{L}.$$

Assume there is such an isomorphism. We denote by g_i and g_j the odd polynomials associated to f_i and f_j so that $\widetilde{\mathscr{F}_i} = \mathscr{F}_{g_i}$ and $\widetilde{\mathscr{F}_j} = \mathscr{F}_{g_j}$. Arguing exactly as in Case 1, we see that the sheaf $\mathscr{L}$ is trivial. Then continuing again as in Case 1 using (11) we find that there are δ_i and δ_j such that

$$f_i(\mathrm{X}) - f_j(\mathrm{Y}) - \delta_i + \delta_j$$

is not absolutely irreducible, and Proposition 6.5 allows us to conclude that f_i and f_j would have to be linearly dependent. □

Lemma 6.7. *Let f and g in $\mathbf{Z}[\mathrm{X}]$ be polynomials of common degree $d \geqslant 3$. The polynomials f and g are linearly equivalent over $\bar{\mathbf{Q}}$ if and only if $f \pmod p$ and $g \pmod p$ are linearly equivalent over an algebraic closure $\bar{\mathbf{F}}_p$ of $\mathbf{F}_p$ for infinitely many primes.*

Proof. The set $\mathrm{X}_{f,g}$ of tuples (a,b,c,d) in $\bar{\mathbf{Q}}$ such that

$$g = af(c\mathrm{X}+d)+b$$

is defined by polynomial equations with rational coefficients. The polynomials f and g are linearly equivalent over $\bar{\mathbf{Q}}$ if and only if $\mathrm{X}_{f,g}(\bar{\mathbf{Q}})$ is not empty. Since $\mathrm{X}_{f,g}$ is an algebraic variety, this is true if and only if $\mathrm{X}_{f,g}(\bar{\mathbf{F}}_p)$ is not empty for all p large enough (e.g, by the Nullstellensatz: if $\mathrm{X}_{f,g}(\bar{\mathbf{Q}})$ is empty, then there is a representation of 1 as belonging to the ideal generated by the equations of $\mathrm{X}_{f,g}$, and this leads to a representation of 1 over $\bar{\mathbf{F}}_p$ for all primes large enough), which proves the assertion. □

Corollary 6.8. *Let $m \geqslant 1$ be an integer and let $f_1, \dots, f_m$ be Sidon–Morse polynomials in $\mathbf{Z}[\mathrm{X}]$ that are pairwise not linearly equivalent over $\bar{\mathbf{Q}}$. Let $s \leqslant m$ be the number of f_i such that f_i is symmetric Sidon–Morse of degree $\geqslant 5$.*

We have

$$\frac{1}{p}\sum_{(a,p)=1} |\mathrm{W}_{f_1}(a;p)\cdots \mathrm{W}_{f_m}(a;p)|^2 = 1+\mathrm{O}(p^{-1/2}) \tag{15}$$

$$\frac{1}{p}\sum_{(a,p)=1} |\mathrm{W}_{f_1}(a;p)\cdots \mathrm{W}_{f_m}(a;p)|^4 = 2^{m-s}3^s+\mathrm{O}(p^{-1/2}) \tag{16}$$

where the implied constant depends only on m and on the degrees of the polynomials f_i.

Proof. Applying Lemma 6.7, we see that for p large enough, the assumptions of Proposition 6.6 hold modulo p. Let p be such a prime. Using the same notation as in (8), the left-hand side of (15) is equal to

$$\iota(\mathrm{tr}(f_p|\,\mathrm{End}(\mathrm{W}_p)^{\mathrm{G}}))+\mathrm{O}(p^{-1/2})$$

where W_p is the tensor product space

$$\bigotimes_i \bar{\mathbf{Q}}_\ell^{d_i-1}$$

as a representation of the geometric monodromy group G of

$$\bigoplus_i \widetilde{\mathscr{F}_i}.$$

By Proposition 6.6, this representation can be identified with the external tensor product of the representations of the individual geometric monodromy groups; since this external tensor produt is an irreducible representation (see, e.g., [25, Prop. 2.3.23]), the invariant space has dimension one, spanned by the scalar matrices, on which f_p acts trivially, and the first result follows.

For the second result, we get in the same way the main term of (16) equal to

$$\prod_{i=1}^{m} \dim(\mathrm{End}(\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d_i-1}))^{\mathrm{G}_i}$$

where G_i is the geometric monodromy group of $\widetilde{\mathscr{F}_i}$. By the simplest case of the Larsen Alternative (see [24, Th. 1.1.6]), each factor is equal to 3 if f_i is a symmetric Sidon–Morse polynomial of degree $\geqslant 5$ (with symplectic monodromy) and to 2 for the others. □

We conclude this section with the following proposition will is used in the proof of the first part of Theorem 1.11, where only one polynomial is assumed to be a Sidon–Morse polynomial.

Proposition 6.9. *Let f and g be non-constant polynomials in $\mathbf{Z}[\mathrm{X}]$ of degrees d_f and d_g respectively. Suppose that f is a Sidon–Morse polynomial, that $d_f < d_g$ and that g is absolutely irreducible. Then*

$$\frac{1}{p}\sum_{(a,p)=1} |\mathrm{W}_f(a;p)\overline{\mathrm{W}_g(a;p)}|^2 = 1 + \mathrm{O}(p^{-1/2})$$

where the implied constant depends only on d_f and d_g.

Proof. This is a variant of the Goursat–Kolchin–Ribet argument, but where we only fully control one of the sheaves.

Let $p > d_f - 1$ be a prime such that f is a Sidon–Morse polynomial modulo p. We denote by $\widetilde{\mathscr{F}_f}$ the normalized sheaf associated to f modulo p, and by G_f (resp. G_g) the geometric monodromy group of $\widetilde{\mathscr{F}}$ (resp. of $\mathscr{F}_g$). Since f is a Sidon–Morse polynomial, we have $\mathrm{G}_f = \mathrm{SL}_{d_f-1}$ or $\mathrm{G}_f = \mathrm{Sp}_{d_f-1}$ (the latter when f is symmetric Sidon–Morse) by Theorem 6.3.

Let further H be the geometric monodromy group of $\widetilde{\mathscr{F}_f} \oplus \mathscr{F}_g$. We have a natural inclusion $\mathrm{H} \to \mathrm{G}_f \times \mathrm{G}_g$, and the composition of this inclusion with either projection is surjective.

We denote by W_p the space

$$\mathrm{End}(\widetilde{\mathscr{F}_f} \otimes \mathscr{F}_g)^{\mathrm{H}},$$

and by f_p a representative of the Frobenius automorphism in H. The analogue of (8) in this case is the formula

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}_f(a;p)\overline{\mathrm{W}_g(a;p)}|^2 = \iota(\mathrm{tr}(f_p|\mathrm{W}_p)) + \mathrm{O}(p^{-1/2})$$

where the implied constant depends only on d_f and d_g (and we used the fact that the trace function of $\widetilde{\mathscr{F}_f}$ has the same modulus as that of $\mathscr{F}_f$). By Schur's Lemma, it then suffices to prove that the representation of H on $\widetilde{\mathscr{F}_f} \otimes \mathscr{F}_g$ is irreducible, and in turn it is enough to prove that $\mathrm{H} = \mathrm{G}_f \times \mathrm{G}_g$ (using again the irreducibility of external tensor product of irreducible representations, see [25, Prop. 2.3.23]).

We denote by L the kernel of the composition homomorphism

$$\mathrm{G}_f \to \mathrm{H} \subset \mathrm{G}_f \times \mathrm{G}_g \to \mathrm{G}_g.$$

This is a normal subgroup of G_f, hence L is either finite or equal to G_f. If the latter holds, then H contains $\mathrm{G}_f \times \{1\}$, and it follows easily that $\mathrm{H} = \mathrm{G}_f \times \mathrm{G}_f$.

Thus we need to exclude the possibility that L is finite. However, if that is the case, then G_f/L is isomorphic to a subgroup of G_g, hence the Lie algebra of G_f has a faithful representation of dimension $\leqslant d_g - 1$. Since we assumed that $d_f > d_g$, this is impossible in view of the minimal dimensions of faithful representations of the Lie algebras of SL_{d_f-1} or Sp_{d_f-1} (which are equal to $d_f - 1$, see e.g. [3, p. 249, Exercice 2 et p. 214, Table 2]). □

Remark 6.10. Theorem 6.3 also implies Proposition 3.5. Indeed, using the same notation as in (8), the Riemann Hypothesis and conductor estimates imply that for k fixed and p large, we have

$$\sum_{a\in\mathbf{F}_p^\times} |\mathrm{W}(a;p)|^{2k} = \nu_k + \mathrm{O}(p^{-1/2}),$$

where ν_k is the multiplicity of the trivial representation of the geometric monodromy group in the representation $\mathrm{End}(\bar{\mathbf{Q}}_\ell^{d-1})^{\otimes k}$. By character theory for compact groups, we have

$$\nu_k = \int_{\mathrm{K}_d} |\operatorname{tr}(g)|^{2k} d\mu(g)$$

for a maximal compact subgroup K_d of the geometric monodromy group, where μ is the Haar measure on K_d normalized to have total volume 1. We can take $\mathrm{K}_d = \mathrm{SU}_{d-1}(\mathbf{C})$ if the geometric monodromy group is SL_{d-1}, and $\mathrm{K}_d = \mathrm{USp}_{d-1}(\mathbf{C})$ if it is Sp_{d-1}.

7. Multiple correlations

We now come to Theorem 1.11. For the first part, we apply Theorem 1.5 to the function $a \mapsto |\mathrm{W}_f(a;q)\overline{\mathrm{W}_g(a;q)}|^2$. We can take $\mathrm{M} = (d_f-1)^2(d_g-1)^2$. By Proposition 6.9, we have

$$\frac{1}{p}\sum_{(a,p)=1} |\mathrm{W}_f(a;p)\overline{\mathrm{W}_g(a;p)}|^2 = 1 + \mathrm{O}(p^{-1/2})$$

so we can take $g(p) = 1 + \mathrm{O}(p^{-1/2})$. Thus Theorem 1.5 gives, for some constant $\mathrm{C} \geqslant 0$, the bound

$$\begin{aligned}\sum_{q\leqslant x} |\mathrm{W}_f(a;q)\mathrm{W}_g(a;q)|^2 &\ll \frac{x}{\log x}\prod_{p\leqslant x}\Big(1+\frac{1}{p}+\frac{\mathrm{C}}{p^{3/2}}\Big)(\log\log x)^{(d_f-1)^2(d_g-1)^2}\\ &\ll x(\log\log x)^{(d_f-1)^2(d_g-1)^2}.\end{aligned}$$

For the second part, we apply Theorem 1.5 to the functions

$$\begin{aligned} a &\mapsto |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)|,\\ a &\mapsto |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)|^2,\\ a &\mapsto |\mathrm{W}_1(a;q)\cdots\mathrm{W}_m(a;q)|^4\end{aligned}$$

and argue as in the proof of Theorem 1.1 using Corollary 6.8.

8. Remarks on Katz's Theorem

We want to observe that Katz's Theorem (Theorem 6.3) can be explained, in the case of monodromy SL_{d-1}, as the combination of two facts:

(1) the local monodromy computation (13), which has an intuitive meaning as the algebraic analogue of the stationary phase expansion for oscillatory integrals

$$g(t) = \int e^{itf(x)} dx,$$

(2) a result of Gabber (see [22, Th. 1.0]) which (essentially) deduces the nature of the monodromy group from the Sidon property of the critical values.

Since the proof of Gabber's result, in this special case, is relatively accessible and (in our opinion) quite enlightening with respect to the relevance of the Sidon condition, we include the precise statement and its proof.

Proposition 8.1. *Let* V *be a finite-dimensional complex vector space of dimension* $r \geqslant 1$*, and let* G *be a connected semisimple compact subgroup of* $\mathrm{GL}(\mathrm{V})$ *which acts irreducibly on* V*. Let* D *be the subgroup of elements of* $\mathrm{GL}(\mathrm{V})$ *which are diagonal with respect to some basis, and let* χ_i*, for* $1 \leqslant i \leqslant r$*, be the characters* $\mathrm{D} \to \mathbf{C}^\times$ *giving the coefficients of the elements of* D*.*

Let $\mathrm{A} \subset \mathrm{D}$ *be a subgroup of the normalizer of* G *in* $\mathrm{GL}(\mathrm{V})$*. Let* $\mathrm{S} \subset \widehat{\mathrm{A}}$ *be the subset of the group of characters of* A *given by the restrictions to* A *of the diagonal characters* χ_i*. If* $|\mathrm{S}| = r$ *and* S *is a Sidon set in* $\widehat{\mathrm{A}}$*, then* $\mathrm{G} = \mathrm{SU}(\mathrm{V})$*.*

Proof. We denote by $\mathrm{Z} \subset \mathrm{D}$ the subgroup of scalar matrices. We may assume that $\mathrm{G} \subset \mathrm{U}(\mathrm{V})$.

The group G is a compact real Lie group. We consider the representation of A on $\mathrm{End}(\mathrm{V})$ by conjugation. It acts on the elementary matrices $\mathrm{E}_{i,j}$ by $\chi_i\chi_j^{-1}$. The assumption that S has r elements and is a Sidon set means then that

$$\mathrm{End}(\mathrm{V}) = \bigoplus_{i,j} \mathbf{C}\mathrm{E}_{i,j}$$

is a decomposition of the representation as a sum of characters where, for $i \neq j$, the line $\mathbf{C}\mathrm{E}_{i,j}$ is a non-trivial character of multiplicity one.

Since $\mathrm{A} \subset \mathrm{N}_{\mathrm{GL}(\mathrm{V})}(\mathrm{G})$, the complexified Lie algebra $\mathrm{L} \subset \mathrm{End}(\mathrm{V})$ of G is a subrepresentation of the representation of A on $\mathrm{End}(\mathrm{V})$. Thus there exists a subspace H of the diagonal matrices, and a subset X of pairs (i, j) of distinct integers such that

$$\mathrm{L} = \mathrm{H} \oplus \bigoplus_{(i,j)\in \mathrm{X}} \mathbf{C}\mathrm{E}_{i,j}.$$

This implies that L is in fact stable under conjugation by all of D. We have therefore an induced morphism

$$\mathrm{D} \to \mathrm{Aut}(\mathrm{L}),$$

which induces an injective morphism $\mathrm{D}/\mathrm{Z} \to \mathrm{Aut}(\mathrm{L})$. Its image is contained in the neutral component of $\mathrm{Aut}(\mathrm{L})$. Since L is semisimple, the latter is equal to the adjoint group of G (see, e.g., [2, p. 244, Prop. 30, (ii)]). It follows that the connected semisimple group $\mathrm{G} \subset \mathrm{SU}(\mathrm{V})$ has rank $r - 1$; it follows that $\mathrm{G} = \mathrm{SU}(\mathrm{V})$ (e.g., by the Borel–de Siebenthal Theorem: the group G coincides with the connected component of the identity of the centralizer in $\mathrm{SU}(\mathrm{V})$ of the center of G, for instance by [4, p. 36, prop. 13], and the center is contained in the group of scalar matrices by Schur's Lemma, so its centralizer is $\mathrm{SU}(\mathrm{V})$). □

This proposition is applied to a conjugate of the *finite* subgroup A of elements of the form

$$\mathrm{diag}(e(xv_1/p), \dots, e(xv_{d-1}/p))$$

where $(v_1, \dots, v_{d-1})$ are the critical values of f; indeed, the local monodromy computation implies that such a subgroup is contained in a maximal compact subgroup of the monodromy group.

References

[1] N. Bourbaki: *Groupes et algèbres de Lie*, chapitre III, Springer.
[2] N. Bourbaki: *Groupes et algèbres de Lie*, chapitre VI, Springer.
[3] N. Bourbaki: *Groupes et algèbres de Lie*, chapitre VIII, Springer.
[4] N. Bourbaki: *Groupes et algèbres de Lie*, chapitre IX, Springer.
[5] J. Bourgain: *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. of Math. 61 (1988), 39–72.
[6] J. Bourgain: *An approach to pointwise ergodic theorems*, in "Geometric aspects of functional analysis (1986/87)", Lecture Notes in Math. 1317, 204–223, Springer,1988.
[7] J. Bourgain and A. Kontorovich: *On the local-global conjecture for integral Apollonian gaskets*, with an appendix by Péter Varjú, Invent. math. 196 (2014), 589–650.
[8] J. Bourgain: *Pointwise ergodic theorems for arithmetic sets*, Publications Mathématiques de l'IHÉS 69 (1989), 5–41.
[9] P. Cassou-Noguès and J.M. Couveignes: *Factorisations explicits de $g(y) - h(z)$*, Acta Arith. 87 (1999), 291–317.
[10] C. Curtis and I. Reiner: *Representation theory of finite groups and associative algebras*, AMS Chelsea Publishing 356 (1962).
[11] P. Deligne: *La conjecture de Weil, II*, Publ. Math. IHÉS 52 (1980), 137–252.
[12] É. Fouvry and Ph. Michel: *Sommes de modules de sommes d'exponentielles*, Pacific J. of Math. 209 (2003), 261–288; erratum, Pacific J. of Math. 225 (2006), 199–200.
[13] É. Fouvry and Ph. Michel: *À la recherche de petites sommes d'exponentielles*, Annales de l'Institut Fourier 52 (2002), 47–80.
[14] É. Fouvry, E. Kowalski and Ph. Michel: *Algebraic twists of modular forms and Hecke orbits*, Geom. Funct. Anal. 25 (2015), 580–657; doi:10.1007/s00039-015-0310-2.
[15] É. Fouvry, E. Kowalski and Ph. Michel: *A study in sums of products*, Phil. Trans. R. Soc. A 373:20140309.
[16] M. Fried: *On a conjecture of Schur*, Michigan Math. J. 17 (1970), 41–55.
[17] M. Fried and M. Jarden: *Field arithmetic*, Ergebnisse der Math. 11, Springer, 2008.
[18] C. Hooley: *On the distribution of the roots of polynomial congruences*, Mathematika 11 (1964), 39–49.
[19] H. Iwaniec and E. Kowalski: *Analytic number theory*, AMS Colloquium Publ. 53, 2004.
[20] N.M. Katz: *Gauss sums, Kloosterman sums and monodromy groups*, Annals of Math. Studies 116, Princeton Univ. Press (1988).
[21] N.M. Katz: *Perversity and exponential sums*, in "Algebraic Number Theory in honor of K. Iwasawa", Adv. Studies Pure Math. 17, 1989, p. 209–259.
[22] N.M. Katz: *Exponential sums and differential equations*, Annals of Math. Studies 124, Princeton Univ. Press, 1990.
[23] N.M. Katz and P. Sarnak: *Random matrices, Frobenius eigenvalues and monodromy*, Colloquium Publ. 45, A.M.S, 1999.
[24] N.M. Katz: *Larsen's alternative, moments and the monodromy of Lefschetz pencils*, in "Contributions to automorphic forms, geometry, and number theory (collection in honor of J. Shalika's 60th birthday)", J. Hopkins Univ. Press (2004), 521–560.
[25] E. Kowalski: *An introduction to the representation theory of groups*, Grad. Studies in Math. 155, A.M.S, 2014.

[26] E. Kowalski, Ph. Michel and W. Sawin: *Bilinear forms with Kloosterman sums and applicatins*, Annals of Math. 186 (2017), 413–500.
[27] E. Kowalski and K. Soundararajan: *Equidistribution from the Chinese Remainder Theorem*, Advances in Math. 385 (2021) 107776.
[28] E. Kowalski and K. Soundararajan: *Remembrances of polynomial values: du côté de chez Fourier*, in progress.
[29] S.J. Patterson: *On the distribution of certain Hua sums, II*, Asian J. Math. 6 (2002), 719–730.
[30] S. J. Patterson: *The asymptotic distribution of exponential sums, II*, Experiment. Math. 14 (2005), 87–98.
[31] C. Perret-Gentil: *Gaussian distribution of short sums of trace functions over finite fields*, Math. Proc. Camb. Phil. Soc. 163 (2017), 385–422.
[32] X. Shao: *Polynomial values modulo primes on average and sharpness of the larger sieve*, Algebra Number Theory 9 (2015), 2325–2346.
[33] P. Shiu: *A Brun-Titchmarsh theorem for multiplicative functions*, J. reine angew. Math. 313 (1980), 161–170.
[34] G. Turnwald: *On Schur's conjecture*, J.. Austral. Math. Soc. 58 (1995), 312–357.

ETH Zürich – D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
Email address: kowalski@math.ethz.ch

Department of Mathematics, Stanford University, Stanford, CA 94305
Email address: ksound@stanford.edu