Exponential Sums

A tour through number theory

Gabriel Ribeiro

May 2022

1 How exponential sums appear in nature

Ever since Gauss, *exponential sums* of the form

$$S(f,p) = \sum_{x \in \mathbb{F}_p^n} \exp\left(\frac{2\pi i f(x)}{p}\right),$$

where p is a prime number and f is some function, play a key role in number theory. The simplest example probably being the case $f(x) = x^2$, which appeared in Gauss' fourth proof of quadratic reciprocity.

A large part of twentieth-century analytic number theory was devoted to the study of these sums. For example, they can be used to estimate $\zeta(s)$ on vertical lines. Indeed, the approximation

$$\zeta(s) = \sum_{n=1}^{N} n^{-s} + \frac{N^{1-s}}{s-1} + O(N^{-\sigma}),$$

reduces the problem to sums of the form $\sum_n n^{-it}$, which are of the form considered above for $f(x) = -t\log(x)/2\pi$. The reader interested in more alike examples may find a plethora thereof in the book *Analytic Number Theory* by H. Iwaniec and E. Kowalski.

Whenever the function f is well-approximated by another function g, the sums $S(f,p)$ and $S(g,p)$ are very close. This allows us to focus our attention on the case where f is a polynomial. As we shall see, this case already encodes much of number theory.

Most foundational questions in number theory aim to describe the set of integer or rational solutions of an equation like $f(x) = 0$ for some $f \in \mathbb{Z}[x_1, \dots, x_n]$. Is this set finite or infinite? If it's finite, what's its cardinality? If it's infinite, can we describe some numbers which "generate" all the solutions in some sense?

Very often these questions are way out of reach for our methods. This leads us to consider solutions mod p of the desired equations. Let's then define a function $\mathrm{Sol}(f,p,t)$

which counts the number of solutions to

$$f(x) \equiv t \pmod{p}.$$

Now, we lose no information if we consider $t \mapsto \mathrm{Sol}(f, p, t)$ as being complex-valued and if we take its Fourier transform. This Fourier transform is given by

$$\psi \mapsto \sum_{t \in \mathbb{F}_p} \psi(t)\, \mathrm{Sol}(f, p, t) = \sum_{x \in \mathbb{F}_p^n} \psi(f(x)).$$

The Pontryagin dual of $\mathbb{F}_p$ is itself, since every additive character is of the form $x \mapsto \psi_a(x) := \exp(2\pi i a x/p)$ for a unique $a \in \mathbb{F}_p$. Via this identification, the function above is none other than

$$a \mapsto \sum_{x \in \mathbb{F}_p^n} \exp\left(\frac{2\pi i a f(x)}{p}\right);$$

an exponential sum!

Another omnipresent example of exponential sums first appeared on Poincaré's posthumous paper on modular forms. Those are the *Kloosterman sums* given by

$$\begin{aligned} \mathrm{Kl}_n(a, q) &:= \sum_{\substack{x_1, \ldots, x_n \in \mathbb{F}_q^\times \\ x_1 \cdots x_n = a}} \psi_q(x_1 + \cdots + x_n) \\ &= \sum_{x_1, \ldots, x_{n-1} \in \mathbb{F}_q^\times} \psi_q\left(x_1 + \cdots + x_{n-1} + \frac{a}{x_1 \cdots x_{n-1}}\right), \end{aligned}$$

where $\psi_q := \psi_1 \circ \mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p}$. Such sums appear, among a myriad of places, in the Fourier expansion of the Poincaré series. Bounds on Kloosterman sums form an important part of Y. Zhang's celebrated Annals of Mathematics paper about the twin-prime conjecture.

In order to deal systematically with exponential sums, let's give a proper definition which encompasses all our polynomial examples and many interesting others.

Definition 1.1 — Exponential sum. Let k be a finite field and X be a finite-type scheme over k. An *exponential sum* is an expression of the form

$$S(f, \varphi) := \sum_{x \in X(k)} \varphi(f(x)),$$

where G is a commutative algebraic group over k, φ is a character of $G(k)$, and $f : X \to G$ is a morphism of schemes over k.

As before, we remark that

$$\begin{aligned}\widehat{G(k)} &\to \mathbb{C}\\ \varphi &\mapsto \sum_{x\in X(k)} \varphi(f(x))\end{aligned}$$

is the Fourier transform of

$$\begin{aligned}G(k) &\to \mathbb{C}\\ t &\mapsto \#\{x \in X(k) \mid f(x) = t\}.\end{aligned}$$

This definition allows for more natural descriptions of some of the exponential sums we encountered. For example, sums of the form (called Gauss sums)

$$\sum_{x\in\mathbb{F}_p^\times} \chi(x)\psi(x),$$

where ψ is an additive and χ is a multiplicative character of $\mathbb{F}_p$, appear in this way by taking $X = \mathbb{A}^1 \setminus \{0\}$, $G = \mathbb{G}_m \times \mathbb{G}_a$, and f as the product of the identity and the inclusion. When $\psi = \psi_1$ and χ is Legendre's symbol we recover the sum used in Gauss' fourth proof of quadratic reciprocity in a possibly more natural way than before.

This point of view also allows us to put numerous number-theoretic questions under the umbrella of exponential sums. The case where φ is the trivial character is already interesting and highly non-trivial. Indeed, the exponential sum becomes the number of k-points of X (independently of G and f).

Let's consider an explicit example. Take $A = \mathbb{Z}[1/26]$ and X as the elliptic curve defined by $y^2 = 4x^3 - x - 1$. We denote by $N(X, q)$ the number of $\mathbb{F}_q$-points of X and wonder how the numbers $N(X, q)$ varies as a function of q.

In analytic number theory, we usually divide the analysis into two cases: either we consider only the cases where q varies between the prime numbers (which are not 2 or 13), or we fix one such prime number p and make q vary among the numbers of the form p^n, for some n.

We begin with the latter. Ever since Artin's thesis in the 1920's, it is known that there exist two complex numbers α_p and β_p, satisfying $\alpha_p\beta_p = p$, such that

$$N(X, p^n) = p^n + 1 - \alpha_p^n - \beta_p^n$$

for all $n \geqslant 1$. In particular, in order to determine $N(X, p^n)$ for all n, it suffices to know $N(X, p)$.

The former case is much harder. By the Hasse bound, we know that

$$|N(X,p) - (p+1)| \leqslant 2\sqrt{p}$$

and so there exists a unique "angle" $\theta_p \in [0, \pi]$ such that

$$N(X,p) - (p+1) = 2\sqrt{p}\cos(\theta_p).$$

Our question, then, is about how the angles θ_p vary as a function of p. If X is an elliptic curve with complex multiplication, it's known since Deuring's 1955 paper *Die Zetafunktion einer algebraischen Kurve von Geschlechte Eins* that the θ_p are uniformly distributed in $[0, \pi]$. Our elliptic curve, however, doesn't have complex multiplication (its j-invariant is not an algebraic integer, for example).

The distribution of the angles θ_p for elliptic curves without complex multiplication was the subject of a famous conjecture of Sato and Tate, which says that the sequence (θ_p) is equidistributed in $[0, \pi]$ for the *Sato-Tate measure* $\mu_{ST} := (2/\pi)\sin^2\theta\, d\theta$. This conjecture very recently became a theorem by Clozel, Barnet-Lamb, Geraghty, Harris, Sheperd-Barron and Taylor, whose proof builds from all the arithmetic geometry used on the modularity theorem. Several natural variants and generalizations remain wide-open.

2 Cohomology to the rescue!

The hero of our story is the theory of étale cohomology and, more precisely, Deligne's groundbreaking paper *La conjecture de Weil II*, which we'll henceforth call "Weil II". Since this is a huge machinery, we'll begin by explaining its main features.

For now, let k be a finite field with q elements, and ℓ a prime different from the characteristic of k. Moreover, let X be a (smooth geometrically connected) variety over k of dimension d.

Since the Zariski topology is so coarse, lots of spaces of interest have trivial fundamental group. This happens, for example, for every variety as above. In other words, the usual tools from algebraic topology are not very adapted to the study of algebraic varieties over fields which have no natural topology. This led Grothendieck to define the *étale fundamental group* $\pi_1(X)$, a profinite group which classifies the finite étale covers of X.

Given the assumption that X is connected, the étale fundamental group is independent of a base point up to inner automorphism. As in topology, a morphism of schemes $f : X \to Y$ induces a morphism

$$f_* : \pi_1(X) \to \pi_1(Y).$$

Moreover, if $X = \operatorname{Spec} k$, its fundamental group is nothing but the absolute Galois group of k. In our case, where k is $\mathbb{F}_q$, this is the free profinite group $\widehat{\mathbb{Z}}$ on one canonical generator given by

$$\overline{\mathbb{F}}_q \to \overline{\mathbb{F}}_q, \qquad x \mapsto x^q.$$

This is the so-called *arithmetic Frobenius*. As we'll see, its inverse, denoted by Frob_k and called *geometric Frobenius*, will play a key role on the theory.

In topology, the category of local systems is equivalent to the category of finite-dimensional representations of the fundamental group by taking the fiber of a local system on a fixed point. This suggests the following definition.

Definition 2.1 A ℓ-adic local system $\mathscr{L}$ of rank r over X is a continuous representation $\rho : \pi_1(X) \to \mathrm{GL}_r(\overline{\mathbb{Q}}_\ell)$.

Given a finite extension E of k, we may define a *trace function* $t_{\mathscr{L}} : X(E) \to \overline{\mathbb{Q}}_\ell$ in the following way: a point $x \in X(E)$ is a morphism $\operatorname{Spec} E \to X$, and so it induces a map

$$\mathrm{Gal}(E) \to \pi_1(X).$$

We denote by $\mathrm{Frob}_{E,x}$ the image of Frob_E via this morphism. Since all of this is only canonical up to a choice of base point, $\mathrm{Frob}_{E,x}$ is a conjugation class in $\pi_1(X)$. It follows that $\rho(\mathrm{Frob}_{E,x})$ is a conjugation class in $\mathrm{GL}_r(\overline{\mathbb{Q}}_\ell)$, and we may take its trace. This number, often denoted $\mathrm{tr}(\mathrm{Frob}_{E,x} \mid \mathscr{L})$, is the image of x by $t_{\mathscr{L}}$.

There's an interesting way to obtain local systems. Let G be a (smooth connected) commutative group scheme over k. We consider the *absolute Frobenius* $F_G : G \to G$; the morphism of k-schemes which is the identity on the underlying topological space and acts as $x \mapsto x^q$ on the structure sheaf $\mathcal{O}_G$. The *Lang isogeny*

$$\mathrm{id}_G - F_G : G \to G$$

is a finite étale cover, which is also Galois with group $G(k)$.

Since $\pi_1(G)$ is the limit of the Galois groups of all finite étale Galois covers, we obtain a natural surjection $\pi_1(G) \to G(k)$. Now, if $\varphi : G(k) \to \overline{\mathbb{Q}}_\ell^\times$ is a character, we may compose those morphisms to obtain a representation

$$\pi_1(G) \to G(k) \to \overline{\mathbb{Q}}_\ell^\times,$$

corresponding to a rank one local system over G; denoted $\mathscr{L}_\varphi$. More generally, if we're also given a morphism $f : X \to G$ of k-schemes, we compose the morphism above with f_* to obtain a rank one local system $f^*\mathscr{L}_\varphi$, commonly denoted $\mathscr{L}_{\varphi(f)}$. Its trace in a point $x \in X(E)$ is the image of Frob_E through the composition

$$\mathrm{Gal}(E) \xrightarrow{x} \pi_1(X) \xrightarrow{f_*} \pi_1(G) \longrightarrow G(k) \xrightarrow{\varphi} \overline{\mathbb{Q}}_\ell^\times.$$

As one may verify, this is simply $\varphi(\mathrm{tr}^G_{E/k} f(x))$, where the $\mathrm{tr}^G_{E/k} : G(E) \to G(k)$ function sends $g \in G(E)$ to $g + \mathrm{Frob}_E(g) + \ldots + \mathrm{Frob}_E^{n-1}(g)$ for $n = [E : k]$. (This coincides with the usual trace when $G = \mathbb{G}_a$ and with the usual norm when $G = \mathbb{G}_m$.) In particular, up to identifying $\overline{\mathbb{Q}}_\ell$ with $\mathbb{C}$, we may write our exponential sum $S(f, \varphi)$ as

$$S(f,\varphi) = \sum_{x \in X(k)} \operatorname{tr}(\mathrm{Frob}_{k,x} \mid \mathscr{L}_{\varphi(f)}).$$

Believe it or not, this is a tremendous achievement!

In order to go further in the étale cohomology world, we need to enlarge our category of ℓ-adic local systems to the so-called *constructible sheaves*, which behave much better functorially. In topology, the constructible sheaves are those which restrict to local systems on a given stratification. Up to some minor technical details, the same definition works in the ℓ-adic setting.

Since constructible sheaves are "locally" local systems, given a constructible sheaf $\mathscr{F}$ and a geometric point $\overline{x}$ over $x \in X(E)$, the fiber $\mathscr{F}_{\overline{x}}$ is a local system. As before, we may make the Frobenius automorphism act on this local system, extending the trace function to constructible sheaves.

Given a constructible sheaf $\mathscr{F}$ on X, Grothendieck defined the cohomology groups $H^i(X_{\bar{k}}, \mathscr{F})$ and the compactly supported cohomology groups $H^i_c(X_{\bar{k}}, \mathscr{F})$. These are finite-dimensional $\overline{\mathbb{Q}}_\ell$-vector spaces, endowed with actions of $\mathrm{Gal}(k)$, which vanish for $i < 0$ or $i > 2d$. They satisfy the Grothendieck trace formula

$$\sum_{x \in X(E)} \operatorname{tr}(\mathrm{Frob}_{E,x} \mid \mathscr{F}) = \sum_{i=0}^{2d} (-1)^i \operatorname{tr}(\mathrm{Frob}_E \mid H^i_c(X_{\bar{k}}, \mathscr{F})).$$

Our approach then becomes clear. We'll write exponential sums as the left-hand side of the equation above, and we'll estimate the eigenvalues of Frob_E acting on $H^i_c(X_{\bar{k}}, \mathscr{F})$.

Definition 2.2 Let $\mathscr{F}$ be a constructible sheaf and let $\iota : \overline{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$ be an embedding. We say that $\mathscr{F}$ is ι-pure of weight w if, for all finite extensions E/k and for all $x \in X(E)$, the eigenvalues α_i of Frob_E acting on $\mathscr{F}_{\overline{x}}$ satisfy $|\iota(\alpha_i)| = |E|^{w/2}$. It is ι-mixed of weight $\leqslant w$ (resp. $\geqslant w$) if we have $\leqslant$ (resp. $\geqslant$) on the equation above. We say that $\mathscr{F}$ is pure / mixed of some weight if it is ι-pure / ι-mixed of the same weight for all ι.

The relation between the definition above and our desired estimates is given by (a particular case of) the main theorem in Weil II.

Theorem 2.1 — Deligne. Let $\mathscr{F}$ be a constructible sheaf over X and let $\iota : \overline{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$ be an embedding. If $\mathscr{F}$ is ι-mixed of weight $\leqslant w$, then $H^i_c(X_{\bar{k}}, \mathscr{F})$ is ι-mixed of weight $\leqslant w + i$.

We remark that, in this case, Poincaré duality implies that $H^i(X_{\bar{k}}, \mathscr{F})$ is ι-mixed of weight $\geqslant w + i$. If the natural morphism $H^i_c(X_{\bar{k}}, \mathscr{F}) \to H^i(X_{\bar{k}}, \mathscr{F})$ is an isomorphism (which happens if X is proper over k), then $H^i(X_{\bar{k}}, \mathscr{F}) = H^i_c(X_{\bar{k}}, \mathscr{F})$ is ι-pure of weight $w + i$.

While this won't really be needed for us, it would be a shame to not remark at this point that this theorem finishes the proof of the Weil conjectures. Let us recall how. We define the *zeta function* of X as the formal power series

$$Z(X, t) := \exp\left(\sum_{n=1}^{\infty} |X(\mathbb{F}_{q^n})| \frac{t^n}{n}\right) \in \mathbb{Q}[\![t]\!].$$

If X is supposed to be projective, the Weil conjectures say, among other things, that $Z(X, t)$ may be written as

$$\frac{P_1(t)P_3(t)\cdots P_{2d-1}(t)}{P_0(t)P_2(t)\cdots P_{2d}(t)},$$

where each P_i is a polynomial in $\mathbb{Z}[t]$, which factors over $\mathbb{C}$ as $\prod_j (1 - \alpha_{ij} t)$ for some complex numbers α_{ij} satisfying $|\alpha_{ij}| = q^{i/2}$ for all i, j.

These conjectures shaped the development of algebraic geometry for over twenty years. All of it now falls under the umbrella of the formalism above. Indeed, we may define P_i to be the (image under some ι of the) determinant of $1 - t\,\mathrm{Frob}_k$, acting on $H^i_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell)$. A simple calculation using the Grothendieck trace formula then implies that $Z(X, t)$ is indeed the desired rational function on the P_i.

The hardest part of these conjectures was the *Riemann Hypothesis;* the calculation that the α_{ij} satisfy $|\alpha_{ij}| = q^{i/2}$ for all i, j. This is now a simple consequence of Deligne's theorem, for the α_{ij} are precisely the (image under the same ι as before of the) eigenvalues of Frob_k acting on $H^i_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell)$, which is ι-pure of weight i. (Since $\overline{\mathbb{Q}}_\ell$ is pure of weight 0.)

Another magnificent example of the applications of Weil II is given by the *Lang-Weil* bound. (Here we don't suppose X to be projective anymore.) By taking $\mathscr{F} = \overline{\mathbb{Q}}_\ell$ on the Grothendieck trace formula we obtain

$$|X(E)| = \sum_{i=0}^{2d} (-1)^i \operatorname{tr}(\mathrm{Frob}_E \mid H^i_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell)).$$

(Since $\mathrm{Frob}_E = \mathrm{Frob}_k^n$, this equation sheds light into the simple case where X is an elliptic curve. Indeed, the complex numbers α_p and β_p that we encountered long ago are nothing but the eigenvalues of the Frobenius acting on $H^1_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell)$!) Consider the numbers

$$b^i_c(X) := \dim_{\overline{\mathbb{Q}}_\ell} H^i_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell) \qquad \text{and} \qquad A(X) := \sum_{i=0}^{2d} b^i_c(X).$$

Since $H^{2d}_c(X_{\bar{k}}, \overline{\mathbb{Q}}_\ell)$ is a one-dimensional vector space endowed with an action of Frob_E

given by multiplication by $|E|^d$, and $\overline{\mathbb{Q}}_\ell$ is pure of weight 0, we obtain

$$\left|X(E) - |E|^d\right| \leqslant \sum_{i=0}^{2d-1} b_c^i(X)|E|^{i/2} \leqslant A(X)|E|^{(2d-1)/2}.$$

In particular, as soon as $|E| > A(X)^2$, the variety X has a E-point.

3 Let's work out the case of Gauss' sums

Let's recall an ancient friend that we encountered in our tour; the Gauss sum $g(\psi, \chi)$, defined as

$$g(\psi, \chi) := \sum_{x \in \mathbb{F}_q^\times} \psi(x)\chi(x),$$

where ψ is an additive and χ is a multiplicative character of $\mathbb{F}_q$. Consider, for each prime p, a non-trivial additive character ψ_p of $\mathbb{F}_p$ and denote by ψ_q the character of $\mathbb{F}_q$ obtained by composing with the trace. If χ is trivial, $g(\psi_q, \chi)$ is simply -1. Else, its absolute value is $\sqrt{q}$ and we find $q-2$ points

$$\theta_{q,\chi} := \frac{g(\psi_q, \chi)}{\sqrt{q}} \in S^1,$$

one for each non-trivial multiplicative character.

As in Sato-Tate's conjecture, we may wonder how do these "angles" are distributed on the unit circle as q tends to infinity.

Theorem 3.1 — Deligne. As q tends to infinity, the angles $\{\theta_{q,\chi}\}_{\chi \neq 1}$ become equidistributed on S^1 with respect to its normalized Haar measure. In other words, the equation

$$\frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\, d\theta = \lim_{q \to \infty} \frac{1}{q-2} \sum_{\chi \neq 1} f(\theta_{q,\chi})$$

is satisfied for all continuous functions $f : S^1 \to \mathbb{C}$.

As the Laurent polynomials are dense in $\mathscr{C}(S^1)$, it suffices to consider functions of the form $f(z) = z^n$, for $n \in \mathbb{Z}$. The case $n = 0$ is trivial and the relation $g(\psi_q, \chi)^{-1} = g(\overline{\psi_q}, \overline{\chi})q^{-1}$ allows us to only consider $n \geqslant 1$. In this case the integral always vanishes, so we must prove that the sequence

$$\frac{1}{q-2}\sum_{\chi \neq 1} f(\theta_{q,\chi}) = \frac{1}{q^{n/2}(q-2)} \sum_{\chi \neq 1} g(\psi_q, \chi)^n$$

tends to zero as q goes to infinity. Then, we remark that

$$\begin{aligned} g(\psi_q,\chi)^n &= \sum_{x_1,\dots,x_n\in\mathbb{F}_q^\times} \psi_q(x_1+\ldots+x_n)\chi(x_1\cdots x_n) \\ &= \sum_{a\in\mathbb{F}_q^\times}\chi(a)\sum_{\substack{x_1,\dots,x_n\in\mathbb{F}_q^\times\\ x_1\cdots x_n=a}}\psi_q(x_1+\cdots+x_n) \\ &= \sum_{a\in\mathbb{F}_q^\times}\chi(a)\,\mathrm{Kl}_n(a,q). \end{aligned}$$

That is, $\chi\mapsto g(\psi_q,\chi)^n$ is the Fourier transform of the Kloosterman sums that we encountered before!

As we do now, Kloosterman himself needed to bound the sums $\mathrm{Kl}_n(a,q)$, but only for $n=2$. By calculating the fourth moment,

$$\sum_{a\in\mathbb{F}_q^\times}\mathrm{Kl}_2(a,q)^4 = 2q^3-3q^2-3q-1,$$

he concluded that $|\mathrm{Kl}_2(a,q)|<2q^{3/4}$. The estimation of the sixth moment allowed Salié and Davenport to upgrade the exponent from 3/4 to 2/3. Finally, Hasse observed that the optimal bound $|\mathrm{Kl}_2(a,q)|<2\sqrt{q}$ would follow from the Riemann Hypothesis for curves over finite fields.

The optimal bound for $\mathrm{Kl}_n(a,q)$ with $n>2$ was only proved, by Deligne, almost 40 years after Weil proved the Riemann Hypothesis for curves over finite fields and established the $n=2$ case. Now, in great Grothendieckian style, it is a somewhat straighforward application of all the breathtaking machinery of the previous section.

Let $k=\mathbb{F}_q$, X be the vanishing set of $x_1\cdots x_n-a$ inside $\mathbb{G}_m^n$, and take $f:X\to\mathbb{G}_a$ be the "sum" function. As we explained, we have that

$$\mathrm{Kl}_n(a,q)=\sum_{i=0}^{2n}(-1)^i\,\mathrm{tr}(\mathrm{Frob}_k\,|\,H_c^i(X_{\bar k},\mathscr{L}_{\psi_q(f)})).$$

In the SGA4$\frac{1}{2}$, Deligne calculated these cohomology groups and concluded that $H_c^i=0$ for all $i\neq n-1$, and that $H^{n-1}=H_c^{n-1}$ is n-dimensional. Moreover, since $\psi_q(f(x))$ is always a p-th root of unity, $\mathscr{L}_{\psi_q(f)}$ is pure of weight 0. All these facts, along with Weil II, implies that

$$|\mathrm{Kl}_n(a,q)|=|\,\mathrm{tr}(\mathrm{Frob}_k\,|\,H_c^{n-1}(X_{\bar k},\mathscr{L}_{\psi_q(f)}))|\leqslant nq^{(n-1)/2},$$

the optimal bound.

This allows us to finish our proof of the equidistribution of the angles of Gauss sums. By summing over the non-trivial χ, we obtain

$$\sum_{\chi\neq 1} g(\psi_q,\chi)^n=-g(\psi_q,1)^n+\sum_{a\in\mathbb{F}_q^\times}\mathrm{Kl}_n(a,q)\sum_{\chi}\chi(a)=(-1)^{n+1}+(q-1)\,\mathrm{Kl}_n(1,q).$$

Finally, using Deligne's bound, we conclude that

$$\left| \frac{1}{q^{n/2}(q-2)} \sum_{\chi \neq 1} g(\psi_q, \chi)^n \right| \leqslant \frac{2n+1}{\sqrt{q}},$$

which goes to zero as q tends to infinity. This finishes the proof.

References

[Bar+11] Tom Barnet-Lamb et al. "A family of Calabi–Yau varieties and potential automorphy II". In: *Publications of the Research Institute for Mathematical Sciences* 47.1 (2011), pp. 29–98.

[Del77] Pierre Deligne. "SGA 4 1/2–Cohomologie étale". In: *Lecture Notes in Mathematics* 569 (1977).

[Del80] Pierre Deligne. "La conjecture de Weil: II". In: *Publications Mathématiques de l'IHÉS* 52 (1980), pp. 137–252.

[Deu53] Max Deuring. *Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins, von Max Deuring*. Vandenhoeck und Ruprecht, 1953.

[Fre19] Javier Fresán. "Équirépartition de sommes exponentielles (travaux de Katz)". In: *Astérisque* 414 (2019), pp. 205–250. DOI: `10.24033/ast.1085`. URL: `https://doi.org/10.24033%2Fast.1085`.

[Gau08] Carl Friedrich Gauss. *Summatio quarundam serierum singularium*. Vol. 1. 1808.

[Has35] Helmut Hasse. "Theorie der relativ-zyklischen algebraischen Funktionenkörper, insbesondere bei endlichem Konstantenkörper." In: (1935).

[IK21] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*. Vol. 53. American Mathematical Soc., 2021.

[Kat80] Nicholas M Katz. *Sommes exponentielles*. Vol. 79. Soc. Math. De France, 1980.

[LW54] Serge Lang and André Weil. "Number of points of varieties in finite fields". In: *American Journal of Mathematics* 76.4 (1954), pp. 819–827.

[Poi11] Henri Poincaré. "Fonctions modulaires et fonctions fuchsiennes". In: *Annales de la Faculté des sciences de Toulouse: Mathématiques*. Vol. 3. 1911, pp. 125–149.

[Sal32] Hans Salié. "Über die Kloostermanschen Summen $S(u, v; q)$". In: *Mathematische Zeitschrift* 34.1 (1932), pp. 91–109.

[Zha14] Yitang Zhang. "Bounded gaps between primes". In: *Annals of Mathematics* (2014), pp. 1121–1174.