arXiv:1712.03173v1 [math.NT] 8 Dec 2017

LECTURES ON APPLIED ℓ-ADIC COHOMOLOGY

PHILIPPE MICHEL

Abstract. We describe how a systematic use the deep methods from ℓ-adic cohomology pioneered by Grothendieck and Deligne and further developed by Katz, Laumon allow to make progress on various classical questions from analytic number theory. This text is an extended version of a series of lectures given during the 2016 Arizona Winter School and is based first and foremost on the works of Deligne, Katz and Laumon and on our ongoing joint work with Fouvry, Kowalski, Sawin and others.

Contents

1. Introduction

One of the most basic question in number theory is to understand how several subsets of integers behave when restricted (intersected with) to *congruence classes*, a notion that goes back at least to Euclid and was exposed systematically by Gauss in his 1801 *Disquisitiones Arithmeticae* (following works of Fermat, Euler, wilson, Lagrange, Legendre and their predecessors from the middle ages and antiquity), and which is fundamental to number theory.

Let us recall that given an integer $q \in \mathbf{Z} - \{0\}$, a *congruence class* (a.k.a. an *arithmetic progression*) modulo q is a subset of $\mathbf{Z}$ of the shape

$$a \,(\mathrm{mod}\, q) = a + q\mathbf{Z} \subset \mathbf{Z}$$

Mon 11th Dec, 2017.

for some integer a. The set of congruence classes modulo q is denoted $\mathbf{Z}/q\mathbf{Z}$; it is a finite ring of cardinality q (with addition and multiplication induced by that of $\mathbf{Z}$).

In number theory, especially analytic number theory one is interested in studying the behaviour of some given arithmetic function along congruence classes for instance to determine whether a set of integers has finite or infinite intersection with some congruence class. The analysis of such problem, which may involve quite sophisticated manipulations often makes certain specific classes of functions on $\mathbf{Z}/q\mathbf{Z}$.

When studying such function it is natural to invoque the *Chinese Remainder Theorem*

$$\mathbf{Z}/q\mathbf{Z} \simeq \prod_{p^\alpha \| q} \mathbf{Z}/p^\alpha\mathbf{Z}$$

which largely reduces the study to the case of prime power moduli; then, in many instances the deepest case is when q is a prime; the ring $\mathbf{Z}/q\mathbf{Z}$ is then a finite field, de noted $\mathbf{F}_q$, and the functions that occur are called *trace functions.*

The tone of these lectures is utilitarian: our aim is to describe these trace functions, many examples, their theory and most importantly how they are handled when they occur in analytic number theory. Indeed the mention of "étale" or "ℓ-adic cohomology", "sheaves", "purity", "functors", "local systems" or "vanishing cycles" sounds forbidding to the working analytic number theorist and often prevents him/her to embrasse the subject and make full use of the powerful methods that Deligne, Katz, Laumon have developed for us. It is our hope that after these introductory lectures, any of the remaining reader will feel ready for and at ease with more serious activities such as the reading of wonderful series of orange books by Nick Katz and eventually will be able to tackle by him/herself any trace function that nature has laid in front of him/her.

Acknowledgements. These expository notes are an expanded version of a series of lectures given together with Will Sawin during the 2016 Arizona Winter School. Many thanks to Will for helping me shaping the course and for the evening sessions as well as for our (ongoing) collaboration during which I probably learn much more from him than he does from me. I would also like to thank the general audience for its attention and its numerous questions during the daily lecture as well as the teams of student who had engaged in research activities with us during the evening sessions for their enthusiasm. Big thanks are also due to Alina Bucur, Bryden Cais and David Zureick-Brown for the perfect organisation making this edition of the AWS a memorable experience. Last but not least I would like express my deep gratitude to my collaborators of the first hour, Etienne Fouvry and Emmanuel Kowalski; without them, none of this would have existed.

2. Examples of trace functions

Unless stated otherwise, we now assume that q is a prime number.

2.1. Characters.

Trace functions modulo q are special classes of $\mathbf{C}$-valued functions on $\mathbf{F}_q$ of geometric origin. Perhaps the first significant example is the *Legendre symbol*

$$(\frac{\cdot}{q}) : x \in \mathbf{F}_q \mapsto \begin{cases} 0 & \text{if } x = 0 \\ +1 & \text{if } x \in (\mathbf{F}_q^\times)^2 \\ -1 & \text{if } x \in \mathbf{F}_q^\times - (\mathbf{F}_q^\times)^2. \end{cases}$$

which detects the squares modulo q, and whose arithmetic properties (especially the *quadratic reciprocity law*) where studied by Gauss in the *Disquisitiones.*

The class of trace function was further enriched by C. L. Dirichlet : on his way to proving his famous theorem on primes in arithmetic progressions, he introduced what are now called *Dirichlet*

characters, i.e. the homomorphisms of the multiplicative group

$$\chi : (\mathbf{Z}/q\mathbf{Z})^\times \to \mathbf{C}^\times$$

which are extended by 0 to the whole of $\mathbf{Z}/q\mathbf{Z}$.

Another significant class of trace functions are the additive characters

$$\psi : (\mathbf{Z}/q\mathbf{Z}, +) \to \mathbf{C}^\times.$$

These are all of the shape

$$x \in \mathbf{Z}/q\mathbf{Z} \mapsto \mathrm{e}_q(ax) := \exp(2\pi i \frac{\tilde{a}\tilde{x}}{q})$$

(say) for some $a \in \mathbf{Z}/q\mathbf{Z}$, where $\tilde{a}$ and $\tilde{x}$ denote elements (lifts) of the congruence classes $a \,(\mathrm{mod}\, q)$ and $x \,(\mathrm{mod}\, q)$. Both additive and multiplicative characters satisfy the important *orthogonality relations*

$$\frac{1}{q}\sum_{x\in\mathbf{F}_q} \psi(x)\overline{\psi'(x)} = \delta_{\psi=\psi'},\ \frac{1}{q-1}\sum_{x\in\mathbf{F}_q} \chi(x)\overline{\chi'(x)} = \delta_{\chi=\chi'};$$

we will see later a generalization of these relations to arbitrary trace functions.

Additive and multiplicative characters can be combined together (by means of a Fourier transform) to form the (normalized) *Gauss sums*

$$g_\chi(a) = \frac{1}{q^{1/2}}\sum_{x\in\mathbf{F}_q} \chi(x)\mathrm{e}_q(ax),$$

but these are not really new functions of a: by a simple change of variable, one has

$$g_\chi(a) = \overline{\chi}(a) g_\chi(1)$$

for $a \in \mathbf{F}_q^\times$. For χ non-trivial, Gauss proved that

$$|g_\chi(1)| = 1.$$

2.2. **Algebraic exponential sums.** Another big source of trace function comes from the study of the diophantine equations

$$Q(\boldsymbol{x}) = 0,\ \boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbf{Z}^n,\ Q(X_1, \cdots, X_n) \in \mathbf{Z}[X_1, \cdots, X_n]. \tag{2.1}$$

For instance the analysis of the *major arcs* in the Hardy-Littlewood circle method leads to the following algebraic exponential sums on $(\mathbf{Z}/q\mathbf{Z})^n$ obtained by Fourier transform

$$(a, \boldsymbol{x}) \in (\mathbf{Z}/q\mathbf{Z})^{n+1} \mapsto \frac{1}{q^{n/2}}\sum_{\boldsymbol{y}\in(\mathbf{Z}/q\mathbf{Z})^n} \mathrm{e}_q(aQ(\boldsymbol{y}) + \boldsymbol{x}.\boldsymbol{y}).$$

In the 1926, while studying the case of a positive definite homogeneous polynomial Q of degree 2 in four variables (a positive definite integral quaternary quadratic form), and introducing a new variant of the circle method, Kloosterman defined the so-called (normalized) *Kloosterman sums*

$$\mathrm{Kl}(a; q) = \frac{1}{q^{1/2}}\sum_{\substack{x,y\in\mathbf{F}_q^\times \\ xy=a}} \mathrm{e}_q(x + y).$$

This is another example of trace function, and indeed one that is defined by a Fourier transform.

By computing the fourth moment of these sums (see [Iwa97, (4.26)]), Kloosterman was able to obtain the first non-trivial bound for these sums, namely

$$|\,\mathrm{Kl}(a; q)| \leqslant 2q^{1/4}$$

. This proved crucial for the study of equation (2.1) in the case of quaternary positive definite quadratic forms. In the 1940's, this bound was improved by A. Weil, who as a consequence of

his proof of the Riemann hypothesis for curves over finite fields ([IK04, §11.7]) proved the best individual upper bound:

$$|\operatorname{Kl}(a;q)| \leqslant 2.$$

In 1939, Kloosterman sums appeared again in the work of Petersson who related them to Fourier coefficients of modular forms.[1] Since then, via the works of Selberg, Kuznetsov, Deshouillers-Iwaniec and many others, Kloosterman sums play a fundamental role in the analytic theory of automorphic forms[2].

A further important example of trace functions are the (normalized) *hyper-Kloosterman sums*. These are higher dimensional generalisations of Kloosterman sums, and are given by

$$\operatorname{Kl}_k(a;q) = \frac{1}{q^{(k-1)/2}} \sum_{\substack{x_1,\cdots,x_k \in \mathbf{F}_q^\times \\ x_1.x_2.\cdots.x_k = a}} \mathrm{e}_q(x_1 + x_2 + \cdots x_k).$$

Hyper-Kloosterman sums were introduced by P. Deligne, who also established the following generalization of the Weil bound:

$$|\operatorname{Kl}_k(a;q)| \leqslant k.$$

Hyper-Kloosterman sums can be interpreted as inverse (discrete) Mellin transforms of powers of Gauss sums, and therefore can be used to study the distribution of Gauss sums. As was noted by Katz in [Kat80], this fact and Deligne's bound, imply the following[3]

Theorem 2.1. *As $q \to \infty$, the set of (normalized) Gauss sums*

$$\{g_\chi(1),\ \chi \text{ non trivial}\} \subset \mathbf{C}^1$$

become equidistributed on the unit circle $\mathbf{C}^1 \subset \mathbf{C}^\times$ with respect to the uniform measure on the circle.

Hyper-Kloosterman sums also occur in the theory of automorphic forms; for instance, Luo, Rudnick and Sarnak used the fact that powers of Gauss sums occur in the root number of the functional equation of certain automorphic L-function, the inverse Mellin transform property and Deligne's bound, to obtain non-trivial estimates for the Langlands parameters of automorphic representations on GL_n (giving in particular the first improvement of Selberg's famous 3/16 bound for the Laplace eigenvalues of Maass cusp forms).

In addition, just as for the classical Kloosterman sums, hyper-Kloosterman sums also occur in the spectral thoery of GL_k automorphic forms.

There are many more examples of trace functions, and we will describe below some ways to obtain new trace functions from older ones. For the moment, we will just say that trace functions are functions on the set of $\mathbf{F}_q$-points of the affine line $\mathbf{A}^1_{\mathbf{F}_q}$ coming from geometry, or alternatively/equivalently as we will see below, from Galois representation of the Galois group $\operatorname{Gal}(\mathbf{F}_q(X)^{sep}/\mathbf{F}_q(X))$ of the field of functions of the affine line $\mathbf{A}^1_{\mathbf{F}_q}$.

3. Trace functions and Galois representations

Let $\mathbf{P}^1_{\mathbf{F}_q}$ be the projective line and $\mathbf{A}^1_{\mathbf{F}_q} \subset \mathbf{P}^1_{\mathbf{F}_q}$ be the affine line. Trace functions are functions defined on the set of $\mathbf{F}_q$-points $\mathbf{A}^1(\mathbf{F}_q) \simeq \mathbf{F}_q$ which are constructed from *constructible* ℓ-adic sheaves relative to the étale topology on $\mathbf{P}^1_{\mathbf{F}_q}$, where ℓ denotes a prime number coprime to q.

The category of *constructible* ℓ-adic sheaves on a curve is an abelian category which can be described rather explicitly in terms of Galois representations (see [Kat80, §4.4] and [Kat88, Chap.

[1] In fact, Poincaré had already written them down in one of list papers, published posthumously.

[2] The double occurence of Kloosterman sums in the context of quadratic forms and of modular forms is explained by the theta correspondence

[3] See [Kat12] for a considerable generalisation of this theorem.

2]). For the moment, it will be sufficent to discuss the case of sheaves which are "lisse" on a dense open subset of the affine line.

Let $K = \mathbf{F}_q(X)$ be the field of rational functions in one variable over $\mathbf{F}_q$ (ie. the function field of $\mathbf{P}^1_{\mathbf{F}_q}$), let $K^{\mathrm{sep}} \supset K$ be a separable closure of K, and $\overline{\eta}$ the associated geometric generic point (so that $\mathrm{Spec}(\overline{\eta}) = K^{\mathrm{sep}}$). Let $\overline{\mathbf{F}_q} \subset K^{\mathrm{sep}}$ denote the separable (or algebraic) closure of $\mathbf{F}_q$ in K^{sep}. We denote

$$G^{\mathrm{geom}} := \mathrm{Gal}(K^{\mathrm{sep}}/\overline{\mathbf{F}_q}.K) \subset G^{\mathrm{arith}} = \mathrm{Gal}(K^{\mathrm{sep}}/K),$$

the *geometric*, resp. *arithmetic*, Galois group. We have the following exact sequence of Galois groups:

$$1 \to G^{\mathrm{geom}} \to G^{\mathrm{arith}} \to \mathrm{Gal}(\overline{\mathbf{F}_q}/\mathbf{F}_q) \to 1.$$

Let x be a closed point of $\mathbf{P}^1_{\mathbf{F}_q}$. We denote by $\mathcal{O}_x$ its associated local ring, by k_x its residue field, by $q_x = |k_x| = q^{\deg x}$ the size of the latter, by K_x the field of fractions of the henselization of $\mathcal{O}_x$ and by $K_x^{\mathrm{sep}} \hookleftarrow K^{\mathrm{sep}}$ a separable closure of K_x with an embedding of K^{sep}. To these data are associated decomposition and inertia subgroups

$$I_x \subset D_x \subset G^{\mathrm{arith}}$$

fitting in the exact sequence

$$1 \to I_x \to D_x \to \mathrm{Gal}(\overline{\mathbf{F}_q}/k_x) = \overline{\langle \mathrm{Fr}_{k_x}^{\mathrm{geom}} \rangle} \to 1. \tag{3.1}$$

Here $\mathrm{Fr}_{k_x}^{\mathrm{geom}}$ (which is denoted Fr_x in the sequel) denotes the *geometric Frobenius element*, namely the topological generator of $\mathrm{Gal}(\overline{\mathbf{F}_q}/k_x)$ given by the inverse of the usual Frobenius automorphism

$$\mathrm{Fr}_{k_x}^{\mathrm{arith}} : u \mapsto u^{|k_x|}.$$

Let ℓ be a prime coprime with q and $\iota : \overline{\mathbf{Q}_\ell} \hookrightarrow \mathbf{C}$ be an algebraic closure of the field of ℓ-adic numbers $\mathbf{Q}_\ell$ with a fixed embedding into the complex numbers. With these definitions, we can define the notion of a lisse ℓ-adic sheaf.[4]

Definition 3.1. Let $U \subset \mathbf{A}^1_{\mathbf{F}_q}$ be a non-empty open subset of $\mathbf{A}^1_{\mathbf{F}_q}$ that is defined over $\mathbf{F}_q$. An ℓ-adic sheaf lisse on U, say $\mathcal{F}$, is a continuous finite- dimensional Galois representation

$$\varrho_{\mathcal{F}} : G^{\mathrm{arith}} \to \mathrm{GL}(V_{\mathcal{F}})$$

where $V_{\mathcal{F}}$ is a finite dimensional $\overline{\mathbf{Q}_\ell}$-vector space, which is unramified at every closed point x of U, in the sense that the inertia subgroup $I_x \subset G^{\mathrm{geom}}$ acts trivially on $V_{\mathcal{F}}$ for all closed points x of U.

The dimension $\dim V_{\mathcal{F}}$ is called the rank of $\mathcal{F}$ and is denoted $\mathrm{rk}(\mathcal{F})$. The vector space $V_{\mathcal{F}}$ is also denoted $\mathcal{F}_{\overline{\eta}}$.

From this definition, one can import the vocabulary and constructions from representation theory.

An ℓ-adic sheaf will be said to be *arithmetically irreducible, isotypic, semisimple, trivial etc...* if it is so as a representation of G^{arith}. It will be said to be *geometrically irreducible, isotypic, semisimple, trivial etc...* if its restriction to G^{geom} is so.

Also one can easily form new sheaves from old ones:

- The dual sheaf $D(\mathcal{F})$ is the contragredient representation $D(\varrho_{\mathcal{F}})$ acting on the dual space $\mathrm{Hom}(V_{\mathcal{F}}, \overline{\mathbf{Q}_\ell})$. This sheaf is also lisse on U.
- Let $H \subset \mathrm{GL}(V_{\mathcal{F}})$ be an algebraic group containing $\varrho_{\mathcal{F}}(G^{\mathrm{arith}})$ and let $r : H \to \mathrm{GL}(V')$ be a finite-dimensional continuous ℓ-adic representation; the composite representation $r \circ \varrho_{\mathcal{F}}$ defines an ℓ-adic sheaf, denoted $r \circ \mathcal{F}$, which lisse on U and has rank $\dim V'$.

[4]This is NOt the original definition, but this one is particularly well adapted to our utilitarian purposes, although it wouldn't be suitable for higher-dimensional generalizations.

– Given another sheaf $\mathcal{G}$ lisse on some U', one can form the tensor product representation $\varrho_{\mathcal{F}} \otimes \varrho_{\mathcal{G}}$; the corresponding sheaf $\mathcal{F} \otimes \mathcal{G}$ is lisse (at least) on $U \cap U'$.
– As a special case, one obtains the endomorphism sheaf $\mathrm{End}(\mathcal{F})$, which is (isomorphic to) the tensor product $\mathcal{F} \otimes D(\mathcal{F})$.
– Let $f \in \mathbf{F}_q(X)$ be non-constant; we can view f as a non-constant morphism $\mathbf{P}^1_{\mathbf{F}_q} \to \mathbf{P}^1_{\mathbf{F}_q}$ (ramified) covering; the subgroup

$$G'^{\mathrm{arith}} = \mathrm{Gal}(K^{\mathrm{sep}}/\mathbf{F}_q(f(X)))$$

of G^{arith} corresponding to this covering is isomorphic to G^{arith}, and therefore the Galois representation $\varrho_{G'^{\mathrm{arith}}}$ restricted to G'^{arith} corresponds to an ℓ-adic sheaf, which is lisse on $f^{-1}(U)$. It is denoted $f^*\mathcal{F}$ and is called the pull-back of $\mathcal{F}$ by f.

Remark 3.2. There is, a priori, no reason to limit ourselves to the affine line: if $\mathcal{C}_{\mathbf{F}_q}$ is any geometrically connected curve over $\mathbf{F}_q$ with function field $K_{\mathcal{C}}$ (which is a finite extension of $\mathbf{F}_q(X)$) and any dense open subset $U \subset \mathcal{C}$ defined over $\mathbf{F}_q$, an ℓ-adic sheaf $\mathcal{F}$ on $\mathcal{C}$ lisse on U is a continuous representation

$$\varrho_{\mathcal{F}} : \mathrm{Gal}(K_{\mathcal{C}}^{sep}/K_{\mathcal{C}}) \to \mathrm{GL}(V_{\mathcal{F}})$$

which is unramified at every closed point of U. More generally, we may also consider ℓ-adic sheaves over a curve $\mathcal{C}$ defined over a finite field extension $\mathbf{F}_{q^n}$, or even over $\overline{\mathbf{F}_q}$, as representations of the Galois groups $\mathrm{Gal}(K^{sep}/\mathbf{F}_{q^n}(\mathcal{C}))$ (or of $\mathrm{Gal}(K^{sep}/\overline{\mathbf{F}_q}(\mathcal{C}))$).

3.1. **The trace function attached to a lisse sheaf.** Given $\mathcal{F}$ as above and $x \in U(\mathbf{F}_q)$ some closed point of degree 1; since the inertia subgroup I_x acts trivially , there is a well-defined action of the Frobenius element $\mathrm{Fr}_{k_x}^{\mathrm{geom}} \in D_x/I_x$ on $V_{\mathcal{F}}$ via $\varrho_{\mathcal{F}}$. We denote by

$$\varrho_{\mathcal{F}}(\mathrm{Fr}_x) \text{ or } (\mathrm{Fr}_x\,|V_{\mathcal{F}})$$

the conjugacy class on the corresponding automorphism; that class does not depend on the choice of the embedding $K_x \hookrightarrow K_x^{\mathrm{sep}}$.

Definition 3.3. Given $\mathcal{F}$ and U as above; the trace function $K_{\mathcal{F}}$ associated to this situation is the function on $U(\mathbf{F}_q)$ given by

$$x \in U(\mathbf{F}_q) \mapsto K_{\mathcal{F}}(x) = \mathrm{tr}(\mathrm{Fr}_x\,|V_{\mathcal{F}}).$$

This is a priori a $\overline{\mathbf{Q}_\ell}$-valued function which can be considered complex valued by means of the chosen embedding $\overline{\mathbf{Q}_\ell} \hookrightarrow \mathbf{C}$.

Remark 3.4. There are several ways by which one could extend $K_{\mathcal{F}}$ to the whole of $\mathbf{A}^1(\mathbf{F}_q)$. The simplest way is the extension by zero outside $U(\mathbf{F}_q)$; another possible extension (called the *middle extension*) would be to set for any $x \in \mathbf{A}^1(\mathbf{F}_q)$,

$$K_{\mathcal{F}}(x) := \mathrm{tr}(\mathrm{Fr}_x\,|V_{\mathcal{F}}^{I_x})$$

where $V_{\mathcal{F}}^{I_x} \subset V_{\mathcal{F}}$ is the subspace of I_x-invariant vectors: the action of the Frobenius element $\mathrm{Fr}_{k_x}^{\mathrm{geom}}$ is well defined. For our purpose any of the two extensions would work (cf. Remark 3.7).

Example 3.5. One has

– $K_{D(\mathcal{F})}(x) = \mathrm{tr}(\mathrm{Fr}_x^{-1}\,|V_{\mathcal{F}})$,
– $K_{\mathcal{F}\otimes\mathcal{G}}(x) = K_{\mathcal{F}}(x)K_{\mathcal{G}}(x)$,
– $K_{f^*\mathcal{F}}(x) = K_{\mathcal{F}}(f(x))$,
– $K_{r\circ\mathcal{F}}(x) = \mathrm{tr}(r(\mathrm{Fr}_x\,|V_{\mathcal{F}})|V')$.

Regarding the third examples we will write

$$[+a] : x \mapsto x + a,\ [\times a] : x \mapsto ax$$

the additive and multiplicative translate maps by $a \in \mathbf{F}_q$. ; more generally for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PGL}_2(\mathbf{F}_q)$ (the group of automorphisms of $\mathbf{P}^1_{\mathbf{F}_q}$) we will write

$$[\gamma] : x \mapsto \frac{ax+b}{cx+d}$$

for the corresponding fractional linear transformation.

3.2. **Trace functions over $\mathbf{A}^1(\mathbf{F}_{q^n})$.** In fact an ℓ-adic sheaf lisse on $U_{\mathbf{F}_q}$ give rise to a whole family of trace functions on $U(\mathbf{F}_{q^n})$ for $n \geqslant 1$. For this we repeat the exact same construction replacing $\mathbf{P}^1_{\mathbf{F}_q}$ by $\mathbf{P}^1_{\mathbf{F}_{q^n}}$, $K = \mathbf{F}_q(X)$ by $K_n = \mathbf{F}_{q^n}(X)$, G^{arith} by $G_n^{\mathrm{arith}} = \mathrm{Gal}(K^{\mathrm{sep}}/K_n)$ (of index n.) We obtain in that way

$$K_{\mathcal{F},n} : \begin{matrix} U(\mathbf{F}_{q^n}) & \mapsto & \mathbf{C} \\ x & \mapsto & \mathrm{tr}(\mathrm{Fr}_x \,|V_{\mathcal{F}}) \end{matrix}$$

where Fr_x denote the conjugacy class of $\mathrm{Fr}_{k_x}^{\mathrm{geom}}$ acting on $V_{\mathcal{F}}$. As we will see below the existence of this sequence of auxilliary functions is very important: by the Chebotareff density theorem, the full sequence $(K_{\mathcal{F},n})_{n\geqslant 1}$ suffice to to characterize the representation $\varrho_{\mathcal{F}}$ up to semi-simplification.

Remark. As the reader has noticed the Frobenius element at x is relative to the local field at x, k_x of the curve $\mathbf{P}^1_{\mathbf{F}_{q^n}}$ (defined over $\mathbf{F}_{q^n}$). In particular, given $x \in U(\mathbf{F}_q)$ aclosed point of degree 1 with Frobenius element Fr_x; the point x give rise to a closed point x_n of degree 1 in $U_{\mathbf{F}_{q^n}}$ and its associated Frobenius conjugacy class $(\mathrm{Fr}_{x_n} \,|V_{\mathcal{F}})$ is related to the previous one by the relation

$$(\mathrm{Fr}_{x_n} \,|V_{\mathcal{F}}) = (\mathrm{Fr}_x^n \,|V_{\mathcal{F}}).$$

In particular, if $n > 1$ the obvious injective map $U(\mathbf{F}_q) \hookrightarrow U(\mathbf{F}_{q^n})$ does not translate to the trace function $K_{\mathcal{F}}$ on $U(\mathbf{F}_q)$ being the restriction of $K_{\mathcal{F},n}$.

3.3. **Purity.** We will be interested in the size of these function. For this the notion of purity is particularly relevant.

Definition 3.6. Given $w \in \mathbf{Z}$; an ℓ-adic sheaf lisse on U as above is punctually pure of weight w if for any $x \in U_{\mathbf{F}_q}$, the various eigenvalues of $(\mathrm{Fr}_x \,|V_{\mathcal{F}})$ are complex numbers[5] of modulus $q_x^{w/2}$. An ℓ-adic sheaf is said mixed of weight $\leqslant w$ if (as a representation) it is a successive extension of sheaves punctually pure of weights $\leqslant w$.

In particular, if $\mathcal{F}$ is mixed of weight $\leqslant w$, one has for any $x \in U(\mathbf{F}_q)$

$$|K_{\mathcal{F}}(x)| \leqslant \mathrm{rk}(\mathcal{F}) q^{w/2}. \tag{3.2}$$

Remark 3.7. It is a deep result of Deligne that for a sheaf punctually pure of weight w,for any x closed point $x \in \mathbf{P}^1_{\mathbf{F}_q}$, the eigenvalues of $(\mathrm{Fr}_x \,|V_{\mathcal{F}}^{I_x})$ has modulus $\leqslant q_x^{w/2}$. In particular

$$|\,\mathrm{tr}(\mathrm{Fr}_x \,|V_{\mathcal{F}}^{I_x})| \leqslant \mathrm{rk}(\mathcal{F}) q_x^{w/2}.$$

Remark 3.8. It is always possible to reduce to the case of ℓ-adic sheaves of weight $w = 0$. For any $w \in \mathbf{Z}$ there exist an ℓ-adic sheaf noted $\overline{\mathbf{Q}_\ell}(w/2)$ of rank 1, lisse on $\mathbf{P}^1_{\mathbf{F}_q}$, whose restriction to G^{geom} is trivial and such that whose Frobeniuses act by multiplication by $q^{-w/2}$ (in particular $\overline{\mathbf{Q}_\ell}(w/2)$ is pure of weight $-w$). Given $\mathcal{F}$ of some weight w', the tensor product

$$\mathcal{F}(w/2) := \mathcal{F} \otimes \overline{\mathbf{Q}_\ell}(w/2)$$

[5] via the fixed embedding $\overline{\mathbf{Q}_\ell} \hookrightarrow \mathbf{C}$

has weight $w' - w$ and trace function given by

$$x \mapsto q^{-w/2} K_{\mathcal{F}}(x).$$

In the sequel, unless stated otherwise, we will always assume that trace function are associated to punctually pure sheaves of weight 0.

3.4. **Other functions.** There are other q-periodic functions of great interest which do not qualify under our current definition of trace function; for instance the Dirac function at some point[6] $a \in \mathbf{F}_q$

$$\delta_a(n) = \begin{cases} 1 & \text{if } n \equiv a \,(\mathrm{mod}\, q) \\ 0 & \text{otherwise .} \end{cases}$$

which extended to $\mathbf{Z}$ is the characteristic function of the arithmetic progression $a + q\mathbf{Z}$ (obviously of considerable interest for analytic number theory.) It turn out that such functions can be related to trace functions in our sense by vary natural transformation and this what will allow to make progresses on problems from "classical" analytic number theory.

3.5. **Local monodromy representations.** Given $\mathcal{F}$ some ℓ-adic sheaf, let $D_{\mathcal{F}}^{ram} \subset \mathbf{P}^1(\overline{\mathbf{F}_q})$ be the set of points where the representation $\varrho_{\mathcal{F}}$ is ramified, that is the inertia I_x acts non-trivially. The restricted representation

$$\varrho_{\mathcal{F},|I_x} = \varrho_{\mathcal{F},x}$$

is call the local monodromy representation of $\mathcal{F}$ at x. Although $D_{\mathcal{F}}^{ram}$ is disjoint from $U(\overline{\mathbf{F}_q})$, the knowledge of these finite set of representations is fundamental to study $\mathcal{F}$ and it trace function. Let us recall [Kat88, Chap. 1] thatone has an exact sequence

$$1 \to P_x \to I_x \to I_x^{tame} \to 1$$

where I_x^{tame} is the *tame inertia quotient* and is isomorphic to $\prod_{p \neq q} \mathbf{Z}_p$ while P_x is the q-Sylow of I_x and is called the wild inertia subgroup.

Definition 3.9. The monodromy representation at x is called tamely ramified if P_x acts trivially on $V_{\mathcal{F}}$ (so that $\varrho_{\mathcal{F},x}$ factors through I_x^{tame}) and is called wildly ramified otherwise.

3.5.1. *The Swan conductor.* If the representation is wildly ramified one can measure how deep is it by means of a numerical invariant: the Swan conductor. The wild inertia subgroup I_x is equipped with a decreasing *the upper numbering filtration* indexed by the real numbers $I_x^{(\lambda)},\ \lambda \geqslant 0$ such that $P_x = I_x^{(>0)}$. Given $V = V_{\mathcal{F}}$ as above there is a P-stable direct sum decomposition

$$V = \bigoplus_{\lambda \in \mathrm{Break}(V)} V(\lambda)$$

indexed by some finite set of rational numbers $\mathrm{Break}(V) \subset \mathbf{Q}_{\geqslant 0}$ (the set of breaks of the I-module V) such that

$$V(0) = V^{P_x},\ V(\lambda)^{I_x^{(\lambda)}} = 0,\ V(\lambda)^{I_x^{(\lambda')}} = V(\lambda),\ \lambda' > \lambda$$

(see [Kat88, Chap. 1]). The swan conductor is defined as

$$\mathrm{Swan}_x(\mathcal{F}) = \sum_{\lambda \in \mathrm{Break}(V)} \lambda \dim V(\lambda)$$

and turns out to be an non-negative integer.

[6] in fact this function could be interpreted as the trace function of a sheaf with punctual support at a but we will not do this here

In the decomposition

$$V = V(0) \oplus \bigoplus_{\substack{\lambda \in \mathrm{Break}(V) \\ \lambda > 0}} V(\lambda) = V(0) \oplus V(>0) := V^{tame} \oplus V^{wild}$$

the first summand is called the tame part and the remaining one the wild part.

4. Summing trace function over $\mathbf{F}_q$

Let $K_{\mathcal{F}}$ be the trace function associated to a sheaf $\mathcal{F}$ lisse on $U_{\mathbf{F}_q}$. We have a function on $U(\mathbf{F}_q)$ which we may extend by zero to $\mathbf{A}^1(\mathbf{F}_q) = \mathbf{F}_q = \mathbf{Z}/q\mathbf{Z}$.

The Grothendieck-Lefschetz trace formula provides an alternative expression for the sum of $K_{\mathcal{F}}$ over the whole $\mathbf{A}^1(\mathbf{F}_q)$.

Theorem 4.1 (Grothendieck-Lefschetz trace formula)**.** *Let $\mathcal{F}$ be lisse on U; there exists three finite dimensional ℓ-adic representations of* $\mathrm{Gal}(\overline{\mathbf{F}_q}/\mathbf{F}_q)$*, $H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})$ such that*

$$\sum_{x \in U(\mathbf{F}_q)} K_{\mathcal{F}}(x) = \sum_{x \in U(\mathbf{F}_q)} \mathrm{tr}(\mathrm{Fr}_x \,|\mathcal{F}) = \sum_{i=0}^{2} (-1)^i \,\mathrm{tr}(\mathrm{Fr}_q \,| H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})). \tag{4.1}$$

More generally, for any $n \geqslant 1$,

$$\sum_{x \in U(\mathbf{F}_{q^n})} K_{\mathcal{F},n}(x) = \sum_{x \in U(\mathbf{F}_{q^n})} \mathrm{tr}(\mathrm{Fr}_x \,|\mathcal{F}) = \sum_{i=0}^{2} (-1)^i \,\mathrm{tr}(\mathrm{Fr}_q^n \,| H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})).$$

The $\overline{\mathbf{Q}_\ell}$-vector spaces $H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})$ are the so-called compactly supported étale cohomology groups of $\mathcal{F}$ and can also be considered as ℓ-adic sheave over the point $\mathrm{Spec}(\mathbf{F}_q)$.

The above formula reduce the evaluation of average of trace functions to that of the three summands

$$\mathrm{tr}(\mathrm{Fr}_q \,| H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})), \ i = 0, 1, 2.$$

we need netherfore to controle the size of these spaces as well as the size of the eigenvalues. We start with the former.

4.1. **Bounding the dimension of the cohomology groups.** The extremal cohomology groups have a simple interpretation

$$H^0_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) = \begin{cases} 0 & \text{if } U \neq \mathbf{P}^1_{\mathbf{F}_q} \\ V_{\mathcal{F}}^{G^{\mathrm{geom}}} & \text{if } U = \mathbf{P}^1_{\mathbf{F}_q}. \end{cases}$$

As a $\mathrm{Gal}(\overline{\mathbf{F}_q}/\mathbf{F}_q)$-representation, one has the isomorphism

$$H^2_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) \simeq V_{\mathcal{F}, G^{\mathrm{geom}}}(-1) \tag{4.2}$$

(ie $H^2_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F})$ is isomorphic to the subquotient of G^{geom}-coinvariants of $V_{\mathcal{F}}$ twisted by $\overline{\mathbf{Q}_\ell}(-1)$). In particular if $\mathcal{F}$ is geometrically irreducible (non geometrically trivial) or more generally geometrically isotypic (the underlying geometric irreducible representation being non trivial) one has

$$H^2_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) = 0.$$

In any cases, one has

$$\dim H^0_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}), \ \dim H^2_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) \leqslant \mathrm{rk}(\mathcal{F}).$$

The dimension of the middle cohomology group is now controlled by the

Theorem 4.2 (The Grothendieck-Ogg-Shafarevich formula)**.**

$$\chi(u|\mathcal{F}) = \sum_{i=0}^{2}(-1)^i \dim H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) = \mathrm{rk}(\mathcal{F})(2 - |\mathbf{P}^1(\overline{\mathbf{F}_q}) - U(\overline{\mathbf{F}_q})|) - \sum_{x \in \mathbf{P}^1(\overline{\mathbf{F}_q}) - U(\overline{\mathbf{F}_q})} \mathrm{Swan}_x(\mathcal{F}).$$

Observe that the quantities that occurs are local geometric data associated to the sheaf yet this collection of local data provides global informations.

We then define the following ad-hoc numerical invariant which serves as a measure of the complexity of the sheaf $\mathcal{F}$:

Definition 4.1. The conductor of $\mathcal{F}$ is defined via the following formula

$$C(\mathcal{F}) = \mathrm{rk}(\mathcal{F}) + |\mathbf{P}^1(\overline{\mathbf{F}_q}) - U(\overline{\mathbf{F}_q})| + \sum_{x \notin U(\overline{\mathbf{F}_q})} \mathrm{Swan}_x(\mathcal{F})$$

In view of this definition we have

$$\sum_{i=0}^{2} \dim H^i_c(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) \ll C(\mathcal{F})^2. \tag{4.3}$$

4.2. **Examples.**

4.2.1. *The trivial sheaf.* The trivial representation $\overline{\mathbf{Q}_\ell}$ is everywhere lisse, pure of weight 0, of rank 1 and conductor 1 and

$$K_{\overline{\mathbf{Q}_\ell}}(x) = 1.$$

4.2.2. *Kummer sheaf* [SGA4$\frac{1}{2}$]. For any non-trivial Dirichlet character $\chi : (\mathbf{F}_q^\times, \times) \to \mathbf{C}^\times$ there exists an ℓ-adic sheaf (the Kummer sheaf) noted $\mathcal{L}_\chi$ which is of rank 1, pure of weight 0, lisse on $\mathbf{G}_{m,\mathbf{F}_q} = \mathbf{P}^1_{\mathbf{F}_q} - \{0, \infty\}$ with trace functions

$$K_{\mathcal{L}_\chi}(x) = \chi(x),\ K_{\mathcal{L}_\chi, n}(x) = \chi(\mathrm{Nr}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(x)) =: \chi_n(x)$$

and conductor

$$C(\mathcal{L}_\chi) = 3;$$

indeed $\mathrm{Swan}_0(\mathcal{L}_\chi) = \mathrm{Swan}_\infty(\mathcal{L}_\chi) = 0$.

4.2.3. *Artin-Schreier sheaf* [SGA4$\frac{1}{2}$]. For any additive character $\psi : (\mathbf{F}_q, +) \to \mathbf{C}^\times$ there exists an ℓ-adic sheaf (the Kummer sheaf) noted $\mathcal{L}_\psi$ which is of rank 1, pure of weight 0, lisse on $\mathbf{A}^1_{\mathbf{F}_q} = \mathbf{P}^1_{\mathbf{F}_q} - \{\infty\}$ with trace function

$$K_{\mathcal{L}_\psi}(x) = \psi(x)\ K_{\mathcal{L}_\psi, n}(x) = \psi(\mathrm{tr}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(x)) =: \psi_n(x)$$

and conductor (if ψ is non-trivial)

$$C(\mathcal{L}_\psi) = 3.$$

(indeed $\mathrm{Swan}_\infty(\mathcal{L}_\psi) = 1$.) If $f \in \mathbf{F}_q(X) - \mathbf{F}_q$, the pull-back sheaf $\mathcal{L}_{\psi(f)}$ is geometrically irreducible and has conductor

$$1 + \text{ number of poles } + \text{ sum of multiplicities of the poles.}$$

More generally a character ψ of $(\mathbf{F}_{q^n}, +)$ is of the shape

$$x \mapsto \psi_1(\mathrm{tr}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(ax))$$

for ψ_1 a character of $(\mathbf{F}_q, +)$ and $a \in \mathbf{F}_{q^n}$ and associated to each such character is an Artin-Schreier sheaf $\mathcal{L}_\psi$.

4.2.4. *(hyper)-Kloosterman sheaves* [Kat88]. Hyper-Kloosterman sums are formed by multiplicative convolution out of additive characters.

Given $K_1, K_2 : \mathbf{F}_q^\times \to \mathbf{C}$ one define their (normalized) multiplicative convolution:

$$K_1 \star K_2 : x \in \mathbf{F}_q^\times \mapsto \frac{1}{q^{1/2}} \sum_{\substack{x_1,x_2 \in \mathbf{F}_q^\times \\ x_1 x_= x}} K_1(x_1)K_2(x_2) = \frac{1}{q^{1/2}} \sum_{x_1 \in \mathbf{F}_q^\times} K_1(x_1)K_2(x/x_1).$$

Similarly for any $n \geqslant 1$ one defines the multiplicative convolution of $K_{1,n}, K_{2,n} : \mathbf{F}_{q^n}^\times \to \mathbf{C}$ as

$$K_{1,n} \star K_{2,n} : x \in \mathbf{F}_{q^n}^\times \mapsto \frac{1}{q^{n/2}} \sum_{\substack{x_1,x_2 \in \mathbf{F}_{q^n}^\times \\ x_1 x_= x}} K_{1,n}(x_1)K_{2,n}(x_2).$$

Now, given ψ a non-trivial additive character and $k \geqslant 2$, the hyper-Klooerman sums are defined by k-times multiplicative convolutions of ψ:

$$\mathrm{Kl}_{k,\psi}(x;q) = \star_k \text{ times} \psi(x) = \frac{1}{q^{\frac{k-1}{2}}} \sum_{\substack{x_1,\cdots,x_k \in \mathbf{F}_q^\times \\ x_1 \cdots x_k = x}} \psi(x_1 + \cdots + x_k)$$

and more generally, one defines hyper-Kloosterman sums over $\mathbf{F}_{q^n}^\times$

$$\mathrm{Kl}_{k,\psi}(x;q^n) = \star_k \text{ times} \psi_n(x) = \frac{1}{q^{n\frac{k-1}{2}}} \sum_{\substack{x_1,\cdots,x_k \in \mathbf{F}_{q^n}^\times \\ x_1 \cdots x_k = x}} \psi_n(x_1 + \cdots + x_k).$$

That these are trace functions is the following important theorem of Katz [Kat88]:

Theorem 4.3. *For any $k \geqslant 2$, there exists an ℓ-adic sheaf (the Kloosterman sheaf) noted $\mathcal{K}\ell_{k,\psi}$, of rank k, pure of weight 0, geometrically irreducible, lisse on $\mathbf{G}_{m,\mathbf{F}_q}$ with trace function*

$$K_{\mathcal{K}\ell_{k,\psi}}(x) = \mathrm{Kl}_{k,\psi}(x;q)$$

and more generally, for any $n \geqslant 1$

$$K_{\mathcal{K}\ell_{k,\psi},n}(x) = \mathrm{Kl}_{k,\psi}(x;q^n).$$

One has $\mathrm{Swan}_0(\mathcal{K}\ell_{k,\psi}) = 0$ *and* $\mathrm{Swan}_\infty(\mathcal{K}\ell_{k,\psi}) = 1$ *so that the conductor of that sheaf equals*

$$C(\mathcal{K}\ell_{k,\psi}) = k + 2 + 1$$

The Kloosterman sheaves have trivial determinant

$$\det \mathcal{K}\ell_k = \overline{\mathbf{Q}_\ell}$$

and if (and only if) k is even, the Kloosterman sheaf $\mathcal{K}\ell_k$ is self-dual:

$$D(\mathcal{K}\ell_k) \simeq \mathcal{K}\ell_k.$$

Remark. When $\psi(\cdot) = \mathrm{e}_q(\cdot)$ we will not mention the additive character e_q in the notations.

4.3. **Deligne's Theorem on the weight.** Now that we control the dimension it remains to control the size of the Frobenius eigenvalues; suppose that $\mathcal{F}$ is pure of some weight 0 so that

$$|K_{\mathcal{F}}(x)| \leqslant \mathrm{rk}(\mathcal{F}).$$

As we have seen as long as $U \neq \mathbf{P}^1$, $H_c^0(U_{\overline{\mathbf{F}_q}}, \mathcal{F}) = 0$.

By (4.2), the eigenvalues of Fr_q acting on $H_c^2(U_{\overline{\mathbf{F}_q}}, \mathcal{F})$ are of the form

$$q\alpha_i, \ i = 1, \cdots, (V_{\mathcal{F}})_{G^{\mathrm{geom}}} \text{ with } |\alpha_i| = 1.$$

The trace of the middle cohomology group $\mathrm{tr}(\mathrm{Fr}_q\,|H^1_c(U_{\overline{\mathbf{F}_q}},\mathcal{F}))$ is much more mysterious but fortunately we have the following deep result[7] of Deligne [Del80].

Theorem 4.4 (Deligne's theorem on the weight)**.** *The eigenvalues of* Fr_q *acting on* $H^1_c(U_{\overline{\mathbf{F}_q}},\mathcal{F})$ *are complex numbers of modulus* $\leqslant q^{1/2}$.

We deduce from this

Corollary 4.2. *Let* $\mathcal{F}$ *be an* ℓ*-adic sheaf lisse on some* U *pure of weight* 0*; one has*

$$\sum_{x\in\mathbf{F}_q} K_{\mathcal{F}}(x) - \mathrm{tr}(\mathrm{Fr}_q\,|H^2_c(U_{\overline{\mathbf{F}_q}},\mathcal{F})) \ll C(\mathcal{F})^2 q^{1/2}.$$

More generally for any $n\geqslant 1$

$$\sum_{x\in\mathbf{F}_{q^n}} K_{\mathcal{F},n}(x) - \mathrm{tr}(\mathrm{Fr}_q^n\,|H^2_c(U_{\overline{\mathbf{F}_q}},\mathcal{F})) \ll C(\mathcal{F})^2 q^{n/2}.$$

In particular if $\mathcal{F}$ *is geometrically irreducible or isotypic with no trivial components, one has*

$$\sum_{x\in\mathbf{F}_q} K_{\mathcal{F}}(x) \ll C(\mathcal{F})^2 q^{1/2}.$$

In practical application we will be faced with situations where we dispose of a sequence of sheaves $(\mathcal{F}_q)_q$ indexed by and infinite set of primes (with $\mathcal{F}_q$ a sheaf over over the field $\mathbf{F}_q$) such that the sequence of conductors $(C(\mathcal{F}_q))_q$ remains uniformly bounded (by C say). In such situation the above formula represents an asymptotic formula as $q\to\infty$ for the sum of $q-O(1)$ terms

$$\sum_{x\in U(\mathbf{F}_q)} K_{\mathcal{F}}(x)$$

with main term $\mathrm{tr}(\mathrm{Fr}_q\,|H^2_c(U_{\overline{\mathbf{F}_q}},\mathcal{F}))$ (possibly 0) and up to an error term of size $\ll C^2 q^{1/2}$.

5. Quasi-orthogonality relations

We will often apply the trace formula and Deligne's theorem to the following sheave: given $\mathcal{F}$ and $\mathcal{G}$ two ℓ-adic sheaves both lisse on some non-empty open set $U\subset\mathbf{A}^1_{\mathbf{F}_q}$ and both pure of weight 0; consider the tensor product $\mathcal{F}\otimes D(\mathcal{G})$: this sheave is also lisse on U and pure of weight 0; moreover from the definition of the conductor (see [Kat88, Chap. 1]) that

$$C(\mathcal{F}\otimes D(\mathcal{G}))\leqslant C(\mathcal{F})C(\mathcal{G}). \tag{5.1}$$

Its associated trace functions are given by (for x in U)

$$x\to K_{\mathcal{F}\otimes D(\mathcal{G}),n}(x) = K_{\mathcal{F},n}(x)\overline{K_{\mathcal{G},n}(x)}.$$

Therefore the trace formula can be used to evaluate the correlation sums between the trace function of $\mathcal{F}$ and $\mathcal{G}$,

$$\mathcal{C}(\mathcal{F},\mathcal{G}) := \frac{1}{q}\sum_{x\in\mathbf{F}_q} K_{\mathcal{F}}(x)\overline{K_{\mathcal{G}}(x)};$$

more generally for any $n\geqslant 1$ we set

$$\mathcal{C}_n(\mathcal{F},\mathcal{G}) := \frac{1}{q^n}\sum_{x\in\mathbf{F}_{q^n}} K_{\mathcal{F},n}(x)\overline{K_{\mathcal{G},n}(x)}.$$

[7] which implies the Riemann hypothesis for higher dimensional varieties over finite fields

Indeed, by Corollary 4.2 one has

$$\mathcal{C}_n(\mathcal{F},\mathcal{G}) = \mathrm{tr}(\mathrm{Fr}_q^n \,|V_{\mathcal{F}\otimes D(\mathcal{G}),G^{\mathrm{geom}}}) + O(\frac{C(\mathcal{F})C(\mathcal{G})}{q^{n/2}}). \tag{5.2}$$

In particular if $C(\mathcal{F})C(\mathcal{G})$ are bounded while $q^n \to \infty$ one obtained as assymptotic formula whose main terms is given by the traces of the powers of Frobenius acting on the coinvariants of $\mathcal{F}\otimes D(\mathcal{G}) = \mathrm{End}(\mathcal{F},\mathcal{G})$.

5.1. **Decomposition of sheaves and trace functions.** Using first a weaker version of the formula (with an error term converging to 0 as $n \to \infty$) Deligne, on his way to the proof of Theorem 4.4 established that any ℓ-adic sheaf as above pure of some weight 0 is geometrically semisimple (the representation $\varrho_{\mathcal{F}|G^{\mathrm{geom}}}$ decomposes into a direct sum of irreducible representations (of G^{geom})); the irreducible components occuring in the decomposition of $\varrho_{\mathcal{F}|G^{\mathrm{geom}}}$ are called the geometric irreducible components of $\mathcal{F}$.

This is not exactly valid for the arithmetic representation but considering its semi-simplification[8] one obtains a decomposition

$$\varrho_{\mathcal{F}}^{ss} = \bigoplus_{i\in I} \varrho_{\mathcal{F}_i}$$

where the $\varrho_{\mathcal{F}_i}$ are arithmetically irreducible (and pure) and lisse on U. Regarding geometric reducibility, each $\varrho_{\mathcal{F}_i}$ is either geometrically isotypic or is induced from a representation of $\mathrm{Gal}(K^{\mathrm{sep}}/k.K)$ for k some finite extension of $\mathbf{F}_q$. Regarding the associated trace function $K_{\mathcal{F}}$ on $U(\mathbf{F}_q)$, since semi-simplification does not change the trace function we obtain a decomposition

$$K_{\mathcal{F}} = \sum_i K_{\mathcal{F}_i}.$$

Moreover a computation shows that whenever $\mathcal{F}_i$ is induced one has $K_{\mathcal{F}_i} \equiv 0$ on $U(\mathbf{F}_q)$. Therefore we obtain

Proposition 5.1. *The trace function associated to some punctually pure sheaf $\mathcal{F}$ lisse on U can be decomposed into the sum of $\leqslant C(\mathcal{F})$ of trace functions whose associated sheaves $\mathcal{F}_i$ are lisse on U, punctually pure, geometrically isotypic with conductors $C(\mathcal{F}_i) \leqslant C(\mathcal{F})$.*

This proposition enable to reduce the study of trace functions to trace functions associated to geometrically isotypic or (most of the time) geometrically irreducible sheaves. From now (unless stated otherwise) we will assume that the trace functions are associated to punctually pure of weight 0, geometrically isotypic sheaves. To ease notations we say that such sheaves are "isotypic" or "irreducible" omitting the mention "geometrically" and likewise will speak of isotypic or irreducible trace function. In such situation, using Schur lemma, the formula for (5.2) specialize to the

Theorem 5.1 (Quasi-orthogonality relations)**.** *Supppose that $\mathcal{F}$ and $\mathcal{G}$ are both geometrically isotypic with $n_{\mathcal{F}}$ copies of the irreducible component $\overline{\mathcal{F}}_{irr}$ for $\mathcal{F}$ and $n_{\mathcal{G}}$ copies of the irreducible component $\overline{\mathcal{G}}_{irr}$ for $\mathcal{G}$. There exists $n_{\mathcal{F}}.n_{\mathcal{G}}$ complex numbers $\alpha_{i,\mathcal{F},\mathcal{G}}$ of modulus 1 such that*

$$\mathcal{C}_n(\mathcal{F},\mathcal{G}) = (\sum_{i=1}^{n_{\mathcal{F}}n_{\mathcal{G}}} \alpha_{i,\mathcal{F},\mathcal{G}}^n)\delta_{\overline{\mathcal{F}}\sim_{geom}\mathcal{G}} + O(C(\mathcal{F})^2C(\mathcal{G})^2q^{-n/2}). \tag{5.3}$$

In particular if $\mathcal{F}$ and $\mathcal{G}$ are both geometrically irreducible there exist $\alpha_{\mathcal{F},\mathcal{G}} \in \mathbf{C}^1$ such that

$$\mathcal{C}_n(\mathcal{F},\mathcal{G}) = \alpha_{\mathcal{F},\mathcal{G}}^n\delta_{\overline{\mathcal{F}}\sim_{geom}\mathcal{G}} + O(C(\mathcal{F})^2C(\mathcal{G})^2q^{-n/2}). \tag{5.4}$$

Remark 5.2. Observe that for $\mathcal{F}$ and $\mathcal{G}$ either the Kummer or Artin-Schreier sheaves these correspond to the orthogonality relations of characters.

[8] which does not change the trace function

Remark 5.3. If two geometrically irreducible sheaves $\mathcal{F}, \mathcal{G}$ are geometrically isomorphic then their trace functions are proportional: more precisely one has for any n

$$K_{\mathcal{F},n} = \alpha_{\mathcal{F},\mathcal{G}}^n K_{\mathcal{G},n}$$

where $\alpha_{\mathcal{F},\mathcal{G}}$ is the complex number of modulus 1 introduced in the previous statement.

Granted that q^n is large compared to $C(\mathcal{F})^2 C(\mathcal{G})^2$ the above formula give a useful criterion to detect whether $\mathcal{F}$ and $\mathcal{G}$ have geometric irreducible components in common. While our main focus is for $n = 1$ and $q \to \infty$ (while $C(\mathcal{F})^2 C(\mathcal{G})^2$ remaining bounded) the case $n \to \infty$ may also be useful. We start we the following easy lemma

Lemma 5.4. *Given $\alpha_1, \cdots, \alpha_d \in \mathbf{C}^1$, d arbitrary complex numbers of modulus 1, one has*

$$\limsup_{n\to\infty} \alpha_1^n + \cdots + \alpha_d^n = d.$$

Using this lemma together with the decomposition into irreducible one obtains the following

Corollary 5.5 (Katz's Diophantine criterion for irreducibility)**.** *Let $\mathcal{F}$ be an ℓ-adic sheaf lisse on U pure of weight 0 with decomposition into geometrically irreducible subsheaves noted*

$$\mathcal{F}^{geom} = \bigoplus_i \overline{\mathcal{F}}_i^{\oplus n_i}$$

then

$$\limsup_{n\to\infty} \mathcal{C}_n(\mathcal{F}, \mathcal{F}) = \sum_{\overline{\mathcal{F}}_i} n_i^2.$$

In particular, $\mathcal{F}$ is geometrically irreducible if and only if

$$\limsup_{n\to\infty} \mathcal{C}_n(\mathcal{F}, \mathcal{F}) = 1.$$

5.2. **Counting trace functions.** These relations enable to obtain upperbounds for the number of geometric isomorphism classes of ℓ-adic sheaves of bounded conductor (see [FKM13] for the proof)

Theorem 5.2. *Given $C \geqslant 1$, The number of geometric isomorphism classes of ℓ-adic sheaves of conductor $\leqslant C$ is finite and bounded by*

$$q^{O(C^6)}$$

where the implied constant is absolute.

Proof. The principle of the proof is as follows: the sheaf-to-trace-function map $\mathcal{F} \mapsto t_{\mathcal{F}}$ associate to the geometric isomorphism class of some sheaf a line in the q-dimensional space $\mathbf{C}^{\mathbf{F}_q}$ of complex valued functions on $\mathbf{F}_q$ which is hermitian under the inner product

$$\langle K, K' \rangle = \frac{1}{q} \sum_{x \in \mathbf{F}_q} K(x)\overline{K'}(x).$$

The quasi-orthogonality relations show that these different lines are almost orthogonal to one another and so one obtains a number of almost orthogonal (circles of) unit vector in the corresponding unit sphere but then a sphere-packing arguments for high-dimensional hermitian spaces (see [KL78]) show that the number of such vectors cannot be too large. □

6. Trace functions over short intervals

In this section and the next ones we discuss the correlations between trace functions and other classical arithmetic functions. Indeed given a trace function

$$K_{\mathcal{F}} : \mathbf{A}^1(\mathbf{F}_q) = \mathbf{F}_q \to \mathbf{C}$$

(extended from $U(\mathbf{F}_q)$ to $\mathbf{A}^1(\mathbf{F}_q)$ either by zero or by the middle-extension we obtain a q-periodic function on $\mathbf{Z}$ via the $(\mathrm{mod}\, q)$-map (which we also denote by $K_{\mathcal{F}}$)

$$K = K_{\mathcal{F}} : \mathbf{Z} \to \mathbf{Z}/q\mathbf{Z} = \mathbf{A}^1(\mathbf{F}_q) \to \mathbf{C}.$$

Given some other arithmetic function $g : \mathbf{N} \to \mathbf{C}$ it is natural to compare them by evaluating their correlation sums

$$\sum_{n \leqslant N} K(n)\overline{g(n)}$$

as $N \to \infty$ (in suitable ranges depending on $C(\mathcal{F})$ and g.)

6.1. **The Polya-Vinogradov method.** We start with the basic case of $g = 1_I$ is the characteristic function of an interval I of $\mathbf{Z}$ (which we may assume is contained in $[0, q-1]$). We want to evaluate non-trivially the sum

$$S(K; I) := \sum_{n \in I} K(n).$$

We may assume that $\mathcal{F}$ is geometrically isotypic and if $I = [0, q-1]$ such sum can be dealed with by Deligne's theorem.

By Parseval, one has

$$S(K; I) = \sum_{y \in \mathbf{F}_q} \widehat{K}(y)\overline{\widehat{1_I}}(y)$$

where

$$\widehat{K}(y) = \frac{1}{q^{1/2}} \sum_{x \in \mathbf{F}_q} K(x)\mathrm{e}_q(xy) \tag{6.1}$$

and

$$\widehat{1_I}(y) = \frac{1}{q^{1/2}} \sum_{x \in I} \mathrm{e}_q(xy)$$

are the (normalized) Fourier transform of K and 1_I (for the abelian group $(\mathbf{F}_q, +)$). One has

$$|\widehat{1_I}(y)| \ll \frac{1}{q^{1/2}} \min(|I|, \|\frac{x}{q}\|^{-1}) \ll \frac{1}{q^{1/2}} \min(|I|, \frac{q}{|x|})$$

which implies that

$$\|\widehat{1_I}\|_1 \ll \frac{|I|}{q^{1/2}} + q^{1/2} \log q.$$

Therefore one has

$$\sum_{n \in I} K(n) \ll \|\widehat{K}\|_\infty q^{1/2} \log q.$$

We need therefore to look as the size of the Fourier transform $y \mapsto \widehat{K}(y)$. As is well know if K is of the shape $\mathrm{e}_q(ax)$ for some $a \in \mathbf{F}_q$ its Fourier transform is a Dirac type function

$$\widehat{K}(y) = q^{1/2}\delta_{y=a\,(\mathrm{mod}\, q)}$$

Definition 6.1. An isotypic sheaf $\mathcal{F}$ is Fourier if its geometric irreducible component is not (geometrically) isomorphic to any Artin-Schreier sheaf $\mathcal{L}_\psi$.

In particular if K is Fourier of conductor $C(\mathcal{F})$, it follows from Theorem 5.1 that for any $y \in \mathbf{F}_q$

$$\widehat{K}(y) \ll C(\mathcal{F})^2.$$

In that way we obtain the

Theorem 6.1 (Polya-Vinogradov bound)**.** *Let K be an isotypic Fourier trace function of conductor $C(\mathcal{F})$, for any interval I of length $\leqslant q$, one has*

$$\sum_{x\in I} K(x) \ll C(\mathcal{F})^2 q^{1/2}\log q.$$

Remark. This statement was obtained for the first time by Polya and Vinogradov in the case of Dirichlet characters χ. In that case the Fourier transform is the normalized Gauss sum

$$\widehat{\chi}(y) = g(\chi, y) = \frac{1}{q^{1/2}}\sum_{x\in\mathbf{F}_q}\chi(x)\mathrm{e}_q(xy)$$

which is bounded in absolute value by 1.

Observe that this bound is better than the trivial bound

$$|\sum_{x\in I} K(x)| \leqslant C(\mathcal{F})|I|$$

as long as

$$|I| \gg_{C(\mathcal{F})} q^{1/2}\log q.$$

Such range is called the *Polya-Vinogradov range* and the question of bounding non-trivially trace functions non over shorter intervals is an fundamental problem in analytic number theory which would have many striking applications. For now the sproblem is solved in a very limited number of cases starting from the celebrated work of Burgess on Dirichlet characters [Bur62] which we will describe in §16.1. A lot of the forthcoming lectures will indeed be concerned with breaking this barrier in specific cases or in different contexts and to describe applications. For now we will content ourselves with

6.1.1. *Bringing the Polya-Vinogradov range.* The following argument improves slightly the Polya-Vinogradov range:

Theorem 6.2. [FKM$^+$17] *Let K be some Fourier trace function; and $I \subset \mathbf{Z}$ be an interval of length $\sqrt{q} < |I| \leqslant q$; one has*

$$\sum_{x\in I} K(x) \ll C(\mathcal{F})^2 q^{1/2}(1+\log(|I|/q^{1/2})).$$

Proof. Given $r \in \mathbf{Z}$ let $I_r = r + I$; this is again an interval and $S(K; I)$ and $S(K; I_r)$ differ only by $O(\|K\|_\infty r)$ which is will be useful for r not too large; moreover

$$\widehat{1_{I_r}}(y) = \mathrm{e}_q(ry)\widehat{I}(y).$$

We have therefore

$$S(K; I) = \sum_{|y|\leqslant q/2}\widehat{K}(y)\overline{1_I(y)}\frac{1}{R}\sum_{0\leqslant r\leqslant R-1}\mathrm{e}_q(-ry).$$

We choose $R = [q^{1/2}] + 1$; using the bounds

$$|1_I(y)| \ll q^{-1/2}\min(|I|, q/|y|),\quad \sum_{0\leqslant r\leqslant R-1}\mathrm{e}_q(-ry) \ll \min(R, q/|r|)$$

and

$$\|K\|_\infty + \|\widehat{K}\|_\infty \ll C(\mathcal{F})^2$$

we obtain the result. □

6.2. **A smoothed version of the Polya-Vinogradov method.** Often in analytic number theory one is not faced with summing a trace function over an interval but instead against some smooth compactly supported function, for instance one has to evaluate sums of the shape

$$\sum_{n\in\mathbf{Z}} K(n)V(\frac{n}{N}),\ V\in C_c^\infty(\mathbf{R}) \text{ fixed.}$$

By the Poisson summation formula one has the identity

$$\sum_{n\in\mathbf{Z}} K(n)V(\frac{n}{N}) = \frac{N}{q^{1/2}}\sum_{n\in\mathbf{Z}}\widehat{K}(n)\widehat{V}(\frac{nN}{q}) \tag{6.2}$$

where

$$\widehat{V}(y) = \int_{\mathbf{R}} V(x)e(xy)dx$$

is the Fourier transform of $V(x)$ (over $\mathbf{R}$).

Observe that $\widehat{V}(y)$ is not compactly supported but at least is of rapid decay:

$$\forall A\geqslant 0,\ \widehat{V}(y)\ll_{V,A}(1+|y|)^{-A};$$

therefore the dual sum decays rapidly for $n\gg q/N$ and we obtain

Proposition 6.2.

$$\sum_{n\in\mathbf{Z}} K(n)V(\frac{n}{N}) \ll_V q^{1/2}\|\widehat{K}\| \ll_{V,C(\mathcal{F})} q^{1/2}. \tag{6.3}$$

6.3. **The Deligne-Laumon Fourier transform.** The Fourier transform

$$K\mapsto\widehat{K} : y\mapsto \frac{1}{q^{1/2}}\sum_{x\in\mathbf{F}_q} K(x)\mathrm{e}_q(-xy)$$

is a well known and very useful operation on the space of function on $(\mathbf{Z}/q\mathbf{Z},+)$. It serve to realize the spectral decomposition of the functions on $\mathbf{Z}/q\mathbf{Z}$ in terms of eigenvectors of the irreductible representations (characters) of $\mathbf{Z}/q\mathbf{Z}$. Let us recall that the Fourier transform is

- Essentially involutive:
$$\widehat{\widehat{K}}(x) = K(-x);$$
stated otherwise , one has the Fourier decomposition:
$$K(x) = \sum_{y\in\mathbf{F}_q}\widehat{K}(y)\mathrm{e}_q(yx).$$
- The Fourier transform is an isometry on $L^2(\mathbf{Z}/q\mathbf{Z})$; stated otherwise , one has the Plancherel formula
$$\sum_{x\in\mathbf{F}_q} K(x)\overline{K'(x)} = \sum_{y\in\mathbf{F}_q}\widehat{K}(y)\overline{\widehat{K'}(y)}.$$
- The Fourier transform behaves well wrt to additive and multiplicative shifts: for $a\in\mathbf{F}_q,\ z\in\mathbf{F}_q^\times$,
$$\widehat{[+a]K}(y) = \mathrm{e}_q(ay)\widehat{K}(y),\ \widehat{[\times z]K}(y) = [\times z^{-1}]\widehat{K}(y) = \widehat{K}(y/z).$$

A remarkable fact due to Deligne is that, to the Fourier transform at the level of trace function correspond a geometric version of Fourier transform at the level of ℓ-adic sheaves. The following theorem is due to G. Laumon [Lau87]:

Theorem 6.3. *Let $\mathcal{F}$ be a Fourier sheaf lisse on U pure of weight 0, there exists a Fourier sheaf $\widehat{\mathcal{F}}$ lisse on some open set $\widehat{U}$, pure of weight 0 such that if $K_{\mathcal{F},n}$ denote the (middle-extension of the) trace function of $\mathcal{F}$, the (middle extension of the) trace function of $\widehat{\mathcal{F}}$ is given by the Fourier transform $\widehat{K_{\mathcal{F},n}}$ where*

$$\widehat{K_{\mathcal{F},n}}(x) = q^{-n/2} \sum_y K_{\mathcal{F},n}(y) \mathrm{e}_q(\mathrm{tr}_{\mathbf{F}_{q^n}/\mathbf{F}_q}(xy)).$$

The map[9] $\mathcal{F} \mapsto \widehat{\mathcal{F}}$ *is called the geometric Fourier transform. The geometric Fourier transform satisfies (for $a \in \mathbf{F}_q$, $z \in \mathbf{F}_q^\times$)*

$$\widehat{\widehat{\mathcal{F}}} = [\times -1]^*\mathcal{F},\ \widehat{[+a]^*\mathcal{F}} = \mathcal{L}_{\mathrm{e}_q(a.)} \otimes \widehat{\mathcal{F}},\ \widehat{[\times z]^*\mathcal{F}} = [\times z^{-1}]^*\widehat{\mathcal{F}}.$$

In addition Laumon defined local version of the geometric Fourier transform making it possible to compute the local monodromy representation of $\widehat{\mathcal{F}}$ in terms of that of $\mathcal{F}$; using these result one deduce

Proposition 6.3. *Given $\mathcal{F}$ as above, one has*

$$C(\widehat{\mathcal{F}}) \leqslant 10 C(\mathcal{F})^2.$$

Also the Fourier transform preserve irreducibility:

Proposition 6.4. *The Fourier transform maps irreducible (resp. isotypic) sheaves to irreducible (resp. isotypic) sheaves.*

Proof. Given $\mathcal{F}$f a geometrically irreducible sheaf, to prove that $\widehat{\mathcal{F}}$ is irreducible it is sufficient to show that

$$\limsup_n \mathcal{C}_n(\widehat{\mathcal{F}}, \widehat{\mathcal{F}}) = \limsup_n \frac{1}{q^n} \sum_{x \in \mathbf{F}_{q^n}} |\widehat{K_{\mathcal{F},n}}(x)|^2 = 1$$

but by Plancherel formula

$$\frac{1}{q^n} \sum_{x \in \mathbf{F}_{q^n}} |\widehat{K_{\mathcal{F},n}}(x)|^2 = \frac{1}{q^n} \sum_{y \in \mathbf{F}_{q^n}} |K_{\mathcal{F},n}(y)|^2$$

and

$$\limsup_n \frac{1}{q^n} \sum_{y \in \mathbf{F}_{q^n}} |K_{\mathcal{F},n}(y)|^2 = 1$$

by Katz irreducibility criterion applied in the reverse direction.

□

Exercise 6.5. *Prove that the hyper-Kloosterman sheaves are geometrically irreducible (hint: observe that the hyper-Kloosterman sums Kl_{k+1} can be expressed in terms of the Fourier transform of Kl_k.)*

7. Autocorrelation of trace functions; the automorphism group of a sheaf

The next couple of appplications we are going to discuss involve a special type of correlation sums between a trace function and its transform by an automorphism of the projective line.

[9] This is in fact a functor on the derived category of constructible ℓ-adic sheaves

Let $\mathcal{F}$ be an ℓ-adic sheaf lisse on $U \subset \mathbf{P}^1_{\mathbf{F}_q}$, pure of weight 0, geometrically irreducible but non trivial, with conductor $C(\mathcal{F})$. Let γ be an automorphism of $\mathbf{P}^1_{\mathbf{F}_q}$: γ is a fractional linear transformation:

$$\gamma : z \mapsto \gamma.z = \frac{az+b}{cz+d}, \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PGL}_2(\mathbf{F}_q).$$

Let $\gamma^*\mathcal{F}$ be the associated pull-back sheaf; it is lisse on $\gamma^{-1}.U$ and its trace function is

$$\gamma^* K(z) = K(\gamma.z) = K(\frac{az+b}{cz+d}).$$

Moreover since γ is an automorphism of $\mathbf{P}^1_{\mathbf{F}_q}$, one has $C(\gamma^*\mathcal{F}) = C(\mathcal{F})$.

The correlations sums we will consider are the one of K and $\gamma^* K(z)$

$$\mathcal{C}(\mathcal{F}, \gamma) := \mathcal{C}(K, \gamma^* K) = \frac{1}{q} \sum_z K(z)\overline{K(\gamma.z)}$$

and

$$\mathcal{C}_n(\mathcal{F}, \gamma) := \mathcal{C}_n(K, \gamma^* K) = \frac{1}{q^n} \sum_{z \in \mathbf{F}_{q^n}} K_n(z)\overline{K_n(\gamma.z)}$$

which are associated to the tensor product sheaf

$$\mathcal{F} \otimes \gamma^* D(\mathcal{F})$$

which is lisse on $U_\gamma = U \cap \gamma.U$.

7.1. **The automorphism group.** The question of the size of these sums is largely determined by the following invariant of $\mathcal{F}$ (see [FKM15, FKM14])

Definition 7.1. Given $\mathcal{F}$ as above the group of automorphisms of $\mathcal{F}$, $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q) \subset \mathrm{PGL}_2(\mathbf{F}_q)$ is the group of $\gamma \in \mathrm{PGL}_2(\mathbf{F}_q)$ such that

$$\gamma^*\mathcal{F} \simeq_{\mathrm{geom}} \mathcal{F}.$$

The group $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)$ is the group of $\mathbf{F}_q$-points of an algebraic subgroup, $\mathrm{Aut}_{\mathcal{F}} \hookrightarrow \mathrm{PGL}_2$ defined over $\mathbf{F}_q$. Let $B \subset \mathrm{PGL}_2$ the sub-group generated by upper-triangular matrices; we define

$$B_{\mathcal{F}} := \mathrm{Aut}_{\mathcal{F}} \cap B$$

the subgroup of $\mathrm{Aut}_{\mathcal{F}}$ generated by upper-triangular matrices of that group and $B_{\mathcal{F}}(\mathbf{F}_q)$ the group of $\mathbf{F}_q$-points.

The relevance of this notion for the above correlations sums is the following

Proposition 7.2. *For $\gamma \notin \mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)$, one has*

$$\mathcal{C}(K, \gamma) = O_{C(\mathcal{F})}(q^{-1/2}).$$

In view of this proposition it is important to determine how large $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)$ and $B_{\mathcal{F}}(\mathbf{F}_q)$ could be.

Example 7.3. Obviously any element of $\mathrm{Aut}_{\mathcal{F}}$ has to leave $\mathbf{P}^1(\overline{\mathbf{F}_q}) - U(\overline{\mathbf{F}_q})$ invariant and all the points in the same orbit have isomorphic local monodromies. This may impose rather strong constraints on $\mathrm{Aut}_{\mathcal{F}}$.

- If $\mathcal{F}$ is geometrically trivial then $\mathrm{Aut}_{\mathcal{F}} = \mathrm{PGL}_2$.
- If $\psi : (\mathbf{F}_q, +) \to \mathbf{C}^1$ is non trivial then $G_{\mathcal{L}_\psi} = N = \{\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \subset \mathrm{PGL}_2\}$

– If $\chi : (\mathbf{F}_q, +) \to \mathbf{C}^1$ is non trivial then

$$G_{\mathcal{L}_\chi} = T^{0,\infty} = \{\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \subset \mathrm{PGL}_2\}$$

is the diagonal torus unless χ is quadratic in which case $G_{\mathcal{L}_\chi} = NT^{0,\infty}$ is the normalizer of the diagonal torus.

– For the Kloosterman sheaves one can show that $\mathcal{G}_{\mathcal{K}\ell_k}$ is trivial: since $\mathcal{K}\ell_k$ is not lisse at 0 and ∞, with Swan conductor 0 and 0 and 1 at ∞ one has $\mathcal{G}_{\mathcal{K}\ell_k} \subset T^{0,\infty}$. One can then show (see [Mic98]) that $[\times a]^*\mathcal{K}\ell_k \simeq_{\mathrm{geom}} \mathcal{K}\ell_k$ iff $a = 1$.

Given $x \neq y \in \mathbf{P}^1(\overline{\mathbf{F}_q})$ we denote by $T^{x,y}$ the pointwise stabilizer stabilizer of the pair (x, y) (this is a maximal torus defined over some finite extension of $\mathbf{F}_q$) and $N(T^{x,y})$ its normalizer. The torus $T^{x,y}$ is defined over $\mathbf{F}_q$ is x, y below to $\mathbf{P}^1(\mathbf{F}_q)$ or to $\mathbf{P}^1(\mathbf{F}_{q^n})$ and are Galois conjugates.

Proposition 7.4. *Suppose $q \geqslant 7$. Given $\mathcal{F}$ as above. One of the following holds*

– *$C(\mathcal{F}) > q$.*

– *q does not divide $|\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)|$ and either $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)$ is of order $\leqslant 60$ or is a subgroup of the normalizer of some maximal torus $N(T^{x_i,y_i})$ defined over $\mathbf{F}_q$.*

– *q divides $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q)|$ and then $\mathcal{F} \simeq \sigma^*\mathcal{L}_\psi$ for some ψ and $K(x) = \alpha\psi(\sigma.x)$ for for some $\sigma \in \mathrm{PGL}_2(\mathbf{F}_q)$ and $\mathrm{Aut}_{\mathcal{F}}(\mathbf{F}_q) = \sigma N \sigma^{-1}$*

Remark 7.5. Observe that in the later case

$$\mathcal{C}(K, \gamma) = |K(0)|^2\mathcal{C}(\psi(\sigma.x), \gamma)$$

Concerning the size of the group $B_{\mathcal{F}}(\mathbf{F}_q)$ one can show that

Theorem 7.1. *Let $\mathcal{F}$ be an isotypic sheaf whose geometric components are not isomorphic to $[+x]^*\mathcal{L}_\chi$ for some $x \in \mathbf{F}_q$ and some multiplicative character χ and such that*

$$C(\mathcal{F}) < q$$

then

$$|B_{\mathcal{F}}(\mathbf{F}_q)| \leqslant C(\mathcal{F}).$$

The proof of this theorem involves the following rigidity theorems (proven in [Kat96])

Proposition 7.6. *Les $\mathcal{L}$ be irreducible.*

– *If for some $x \in \mathbf{F}_q^\times$, $[+x]^*\mathcal{L} \simeq \mathcal{L}$ then either*

$$C(\mathcal{L}) > q \text{ or } \mathcal{L} \simeq \mathcal{L}_\psi \text{ for some} \psi.$$

– *If $\mathrm{Aut}_{\mathcal{L}}(\mathbf{F}_q)$ contains a subgroup of order m of $\mathrm{Diag}_2(\mathbf{F}_q)$ then either*

$$c(\mathcal{L}) > m \text{ or } \mathcal{L} \simeq \mathcal{L}_\chi \text{ for some} \chi.$$

8. Trace functions vs. primes

After the consideration of short intervals, another possible question to look at (natural from the viewpoint of analytic number theory at least) is how trace functions interact with the primes. In this section, we discuss the structure of the proof of the following result:

Theorem 8.1 (Trace function vs. primes, [FKM14])**.** *Let K be a trace function associated to an isotypic sheaf $\mathcal{F}$, pure of weight 0 and whose geometric components are not of the shape $\mathcal{L}_\psi \otimes \mathcal{L}_\chi$. $V \in C_c^\infty(\mathbf{R}_{>0})$, one has for $X \ll q$ and any $\eta < 1/24$*

$$\sum_{\substack{p \text{ prime} \\ p \leqslant X}} K(p) \ll X(1+q/X)^{1/12} p^{-\eta/2}, \tag{8.1}$$

$$\sum_{p \text{ prime}} K(p) V\Big(\frac{p}{X}\Big) \ll X(1+q/X)^{1/6} q^{-\eta}, \tag{8.2}$$

for any $\eta < 1/24$. The implicit constants depend only on η, $C(\mathcal{F})$ and V. Moreover, the dependency on $C(\mathcal{F})$ is at most polynomial.

Remark. This result exhibit cancellations in summing trace functions along the primes in intervals of length larger than $q^{3/4}$. It is really a pity that Dirichlet characters are excluded by our hypotheses: such a bound in that case would amount to a quasi generalized Riemann hypothesis for the corresponding Dirichlet character L-function !

We discuss the proof for $X = q$.

8.1. **Combinatorial decomposition of the characteristic function of the primes.** As is well know the problem is equivalent to bounding the sum

$$\sum_n \Lambda(n) K(n) V\Big(\frac{n}{q}\Big)$$

where

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^\alpha \\ 0 & \text{otherwise,} \end{cases}$$

is the von Mangolt function. A standard method in analytic number theory is a combinatorial decomposition of this function as a sum of Dirichlet convolution functions; one way to achieve this is to use the celebrated Heath-Brown identity:

Lemma 8.2 (Heath-Brown)**.** *For any integer $J \geqslant 1$ and $n < 2X$, we have*

$$\Lambda(n) = -\sum_{j=1}^{J} (-1)^j \binom{J}{j} \sum_{m_1, \cdots, m_j \leqslant Z} \mu(m_1) \cdots \mu(m_j) \sum_{m_1 \cdots m_j n_1 \cdots n_j = n} \log n_1,$$

where $Z = X^{1/J}$.

Hence splitting the range of summation of the various variables appearing (using partition of unity) and separating these variables, our preferred sum decomposes (essentially) into $O(J)$ sums of the shape

$$\Sigma(M_1, \cdots, M_{2j}) = \sum_{m_1, \cdots m_{2j}} \sum \mu(m_1) \cdots \mu(m_j) K(m_1. \cdots .m_{2j}) V_1\Big(\frac{m_1}{M_1}\Big) \cdots V_{2j}\Big(\frac{m_{2j}}{M_{2j}}\Big)$$

for $j \leqslant J$; here $V_i,\ i = 1, \cdots 2j$ are smooth functions compactly supported in $]1, 2[$, $(M_1, \cdots, M_{2j})$ is a tuple satisfying

$$M_i =: q^{\mu_i},\ \forall i \leqslant j,\ \mu_i \leqslant 1/J,\ \sum_{i \leqslant 2j} \mu_i = 1 + o(1);$$

the objective is to show that

$$\Sigma(M_1, \cdots, M_{2j}) \ll q^{1-\eta}$$

for some fixed $\eta > 0$. We will take $J = 3$ so that $Z = q^{1/3}$. Wlog wma

$$\mu_1 \leqslant \cdots \leqslant \mu_j \leqslant 1/3,\ \mu_{j+1} \leqslant \cdots \leqslant \mu_{2j}.$$

We will bound these sums differently depending on the vector $(\mu_1, \cdots, \mu_{2j})$.

Let $\delta > 0$ be some small fixed parameter to be choosen optimally later.

1) Suppose that for some $\delta \in]0, 1/6[$, $\mu_{2j} \geqslant 1/2 + \delta$, then m_{2j} is a long an "smooth variable" (because the weight attached to it is smooth); therefore using 6.3 we obtain summing over m_{2j} and fixing the other variables

$$\Sigma(M_1, \cdots, M_{2j}) \ll q^{\mu_1 + \cdots \mu_{2j-1}} q^{1/2+o(1)} = q^{1-\delta+o(1)}.$$

In the litterature, sum of that shape are called "type I" sums.

2) We may therefore assume that

$$m_{j+1} \leqslant \cdots \leqslant \mu_{2j} \leqslant 1/2 + \delta;$$

in other terms, there is no very long smooth variable. What one can do is group variables together to form long ones: for this one partition, the indexing set into two blocks

$$\{1, \cdots, 2j\} = \mathcal{I} \sqcup \mathcal{I}',$$

form the variables

$$m = \prod_{i \in \mathcal{I}} m_i, \ n = \prod_{i' \in \mathcal{I}'} m_{i'}$$

so that denoting by α_m the Dirichlet convolutions of either $\mu(\cdot)V(\frac{\cdot}{M_i})$ or $V(\frac{\cdot}{M_i})$ for $i \in \mathcal{I}$ and similarly for β_n for $i' \in \mathcal{I}'$ we are led to bound bilinear sums of the shape

$$B(K; \alpha, \beta) = \sum_{m \ll M, n \ll N} \sum \alpha_m \beta_n K(mn). \tag{8.3}$$

where

$$M = q^{\mu}, \ \mu = \sum_{i \in \mathcal{I}} \mu_i, \ N = q^{\nu}, \ \nu = \sum_{i' \in \mathcal{I}'} \mu_{i'}.$$

The weights α_m, β_n are rather irregular and it is difficult to exploit their structure. Such sums are called "type II".

Assuming that the irreducible component of $\mathcal{F}$ is not of the shape $\mathcal{L}_\chi \otimes \mathcal{L}_\psi$, we will prove in Theorem 9.1 below the following bound

$$\Sigma(M_1, \cdots, M_{2j}) = B(K; \alpha, \beta) \ll_{C(\mathcal{F})} \|\alpha_M\|_2 \|\beta_N\|_2 (MN)^{1/2} (\frac{1}{M} + \frac{q^{1/2} \log q}{N})^{1/2}.$$

Assuming that

$$\mu \geqslant \delta \text{ and } \nu \geqslant 1/2 + \delta$$

we obtain that

$$B(K; \alpha, \beta) \ll q^{1-\delta/2+o(1)}.$$

3) It remains to treat the sums for which neither $\mu_{2j} \leqslant 1/2 + \delta$ nor a decomposition as in 2) exist. This necessarily implies that $\sum_{i \leqslant j} \mu_i \leqslant 1/3$, $j \geqslant 2$ and $\mu_{2j-1} + \mu_{2j} \geqslant 1 - \delta$. Setting $M = M_{2j-1}$ and $N = M_{2j}$

$$a = m_1 \cdots m_{2j-2},$$

it will be sufficient to obtain a bound of the shape

$$\sum_{m,n \geqslant 1} K(amn) V(\frac{m}{M}) W(\frac{n}{N}) \ll_{V,W} (MN)^{1-\eta}$$

for some $\eta > 0$ whenever MN is sufficiently close to q. What we have are is a sum involving two smooth variables who are too short for the Polya-Vinogradov method to work but whose product is rather long. We call these sums "type I1/2". In Section 10 we discuss the proof of

Theorem 8.1. *Let K be a trace function associated to an isotypic Fourier sheaf and $V, W \in C_c^\infty(\mathbf{R}_{>0})$, one has for $M, N \geqslant 1$ and any $\eta < 1/8$*

$$\sum_{m,n\geqslant 1} K(mn)V(\frac{m}{M})W(\frac{n}{N}) \ll_{V,W} MN(1+\frac{q}{MN})^{1/2}q^{-\eta/2}.$$

Observe that this bound is non trivial as long as $MN \geqslant q^{3/4}$; as we will see this result will be a special case of a more general one on the correlation between trace functions and Fourier coefficients of modular forms.

Optimizing parameters in these three approaches one obtains Theorem 8.1.

9. Bilinear sums of trace functions

Let K be a trace function associated to some sheaf isotypic $\mathcal{F}$, pure of weight 0 and let $(\alpha_m)_{m\leqslant M}$, $(\beta_n)_{n\leqslant N}$ be arbitrary complex numbers; in this section we bound for the "type II" bilinear sum we encountered in the previous section (of course such estimates are application beyond the correlation problem for trace function vs. the primes):

$$B(K;\alpha,\beta) = \sum_{m\leqslant M}\sum_{n\leqslant N} \alpha_m\beta_n K(mn).$$

Using the Cauchy-Schwarz inequality, the trivial bound is

$$|B(K;\alpha,\beta)| \leqslant \|\alpha_M\|_2\|\beta_N\|_2(MN)^{1/2}.$$

We which to improve over this bound.

Theorem 9.1 (Bilinear sums of trace functions)**.** *Notations as above; assume that $1 \leqslant M, N < q$ and that the irreducible component of $\mathcal{F}$ is not of the shape $\mathcal{L}_\chi \otimes \mathcal{L}_\psi$, then*

$$B(K;\alpha,\beta) \ll_{C(\mathcal{F})} \|\alpha_M\|_2\|\beta_N\|_2(MN)^{1/2}(\frac{1}{M}+\frac{q^{1/2}\log q}{N})^{1/2}.$$

Remark 9.1. This bound is non-trivial as soon as $M \gg 1$ and $N \gg q^{1/2}\log q$.

We now give an idea of the

Proof. By Cauchy-Schwarz we have

$$|B(K;\alpha,\beta)|^2 \leqslant \|\beta_N\|_2^2 \sum_{m_1,m_2\leqslant M} \alpha_{m_1}\overline{\alpha_{m_2}} \sum_{n\leqslant N} K(mn_1)\overline{K}(mn_2) \tag{9.1}$$

We do not expect to gain anything from the diagonal terms $m_1 \equiv m_2 \,(\mathrm{mod}\, q)$ (equivalent to $m_1 = m_2$ since $M < q$) and the contribution of such terms is bounded trivially by

$$\ll_{C(\mathcal{F})} \|\alpha_M\|^2\|\beta_N\|_2 N. \tag{9.2}$$

As for the non-diagonal terms their contribution is

$$\|\beta_N\|_2^2 \sum_{m_1\neq m_2 \,(\mathrm{mod}\, q)} \alpha_{m_1}\overline{\alpha_{m_2}} \sum_{n\leqslant N} K(mn_1)\overline{K}(mn_2).$$

Using the Polya-Vinogradov method, we are led to evaluate the Fourier transform of

$$n \mapsto K(mn_1)\overline{K}(mn_2).$$

By the Plancherel formula, this Fourier transform equals

$$\begin{aligned} y \mapsto \frac{1}{q^{1/2}} \sum_{x \in \mathbf{F}_q} K(m_1 x)\overline{K}(m_2 x)\mathrm{e}_q(-yx) &= \frac{1}{q^{1/2}} \sum_{z \in \mathbf{F}_q} \widehat{K}((z-y)/m_1)\overline{\widehat{K}}(z/m_2) \\ &= \frac{1}{q^{1/2}} \sum_{z \in \mathbf{F}_q} \widehat{K}((m_2 z - y)/m_1)\overline{\widehat{K}}(z) \\ &= \frac{1}{q^{1/2}} \sum_{z \in \mathbf{F}_q} \widehat{K}(\gamma z)\overline{\widehat{K}}(z) \end{aligned}$$

with

$$\gamma = \begin{pmatrix} m_2/m_1 & -y/m_1 \\ 0 & 1 \end{pmatrix} \in B(\mathbf{F}_q).$$

This sum is the correlation sum is q times $\mathcal{C}(\widehat{\mathcal{F}}, \gamma)$ the correlation sum associated to the isotypic sheaves $\widehat{\mathcal{F}}$ and $\gamma^*\widehat{\mathcal{F}}$ whose conductors are controlled in terms of $C(\mathcal{F})$.

If $\gamma \notin B_{\mathcal{F}}(\mathbf{F}_q)$ we have

$$\mathcal{C}(\widehat{\mathcal{F}}, \gamma) \ll_{C(\mathcal{F})} \frac{1}{q^{1/2}}. \tag{9.3}$$

The condition that the irreducible component of $\mathcal{F}$ is not of the shape $\mathcal{L}_\chi \otimes \mathcal{L}_\psi$ translate into the irreducible component of $\widehat{\mathcal{F}}$ not being of the shape $[+x]^*\mathcal{L}_{\overline{\chi}}$. In that case by Theorem 7.1 there is a set $S_{\mathcal{F}} \subset \mathbf{F}_q^\times$ such that for any $(m_1, m_2, y) \in \mathbf{F}_q^\times \times \mathbf{F}_q^\times \times \mathbf{F}_q$ for which $m_2/m_1 \notin S_{\mathcal{F}}$ one has

$$\mathcal{C}(\widehat{\mathcal{F}}, \gamma)_{C(\mathcal{F})} q^{-1/2}$$

Returning to (9.1), we bound trivially (by (9.2)) the contribution of the pairs $O_{\mathcal{F}}(M)$ pairs (m_1, m_2) such that the ratio $m_2/m_1 \,(\mathrm{mod}\, q)$ is in $S_{\mathcal{F}}$. For the other terms we may use the Polya-Vinogradov method and bound these terms by

$$\ll_{C(\mathcal{F})} \|\alpha_M\|_2^2 \|\beta_N\|_2^2 M q^{1/2} \log q.$$

Combining these bounds leads to the final result. □

10. Trace functions vs. modular forms

In this section we discuss the proof of Theorem 8.1. This theorem is a special case of the resolution of another problem: the question of the correlation between trace functions and the Fourier coefficients $(\varrho_f(n))_n$ of some modular Hecke eigenform. Given some trace function we may then consider the correlation sum

$$\sum_{n \leqslant X} \varrho_f(n) K(n)$$

or its smooth version

$$\sum_n \varrho_f(n) K(n) V(\frac{n}{X}).$$

These sums are bounded trivially (using the ranking Selberg method) by

$$O_{C(\mathcal{F}),f}(X \log^3 X).$$

It turns out that the problem of bounding such sums non-trivially start being interesting for N of size q (or smaller).

In this section, we sketch the proof of the following

Theorem 10.1 (Trace function vs. modular forms, [FKM15])**.** *Let K be a trace function associated to an irreducible Fourier sheaf (of weight 0); let $(\varrho_f(n))_{n\geqslant 1}$ be the sequence of Fourier coefficients of some modular form f and let $V \in C_c^\infty(\mathbf{R}_{>0})$, one has for $X \geqslant 1$ and any $\eta < 1/8$*

$$\mathcal{S}(K, f; X) := \sum_{n\leqslant X} \varrho_f(n)K(n) \ll X(1 + \frac{q}{X})^{1/2}q^{-\eta/2},$$

and

$$\mathcal{S}_V(K, f; X) \sum_{n\geqslant 1} \varrho_f(n)K(n)V(\frac{n}{X}) \ll X(1 + \frac{q}{X})^{1/2}q^{-\eta}.$$

The implicit constants depend only on η, f $C(\mathcal{F})$ and V. Moreover, the dependency on $C(\mathcal{F})$ is at most polynomial.

This result show the absence of correlation for range $X \gg q^{1-1/8}$. The proof which use the amplification method, the Petersson-Kuznetzov trace formula, will ultimately be a consequence of Theorem 7.4.

We give below an idea of the proof. To simplify matters we will assume that $X = q$ and we wish to bound non-trivially the sum

$$\mathcal{S}_V(K, f) := \sum_{n\geqslant 1} \varrho_f(n)K(n)V(\frac{n}{q})$$

for V a fixed smooth function; moreover to simplify things further we will assume that f has level 1 and is cuspidal and holomorphic of very large (but fixed) weight.

10.1. **Trace functions vs. the divisor function.** An important special case of Theorem 10.1 is when f is an Eisenstein series: for instance when

$$f(z) = \frac{\partial}{\partial s}E(z, s)_{|s=1/2} \text{ for } E(z, s) = \frac{1}{2}\sum_{(c,d)=1} \frac{y^s}{|cz + d|^{2s}}$$

is the non-holomorphic Eisenstein series at the central point. In that case

$$\varrho_f(n) = d(n)$$

is the divisor function and so one has

$$\sum_{m,n\geqslant 1} K(mn)V(\frac{mn}{X}) \ll_V X(1 + \frac{q}{X})^{1/2}q^{-\eta} \tag{10.1}$$

whenever K is the trace function of a Fourier sheaf. This bound holds similarly for the unitary Eisenstein series $E(z, s)$ at any $s = \frac{1}{2} + it$ where the divisor function is replaced by

$$d_{it}(n) = \sum_{ab=n} (a/b)^{it}.$$

Such general bounds make it possible to separate the variables m, n in (10.1) and eventually to prove Theorem 8.1.

Remark 10.1. As we will see below the proof of Theorem 10.1 is not a "modular form by modular form" analysis; instead the proof is global involving the full automorphic spectrum and establishes the required bound "for all modular forms f at once".

10.2. **Functional equations.** Our first objective is to understand why the range $X = q$ is interesting; this come from the functional equations satisfied by modular forms as a consequence of their automorphic properties. These equations present themselves in various forms. One is the Voronoi summation formula which in its simplest forme is the following

Proposition 10.2 (Voronoi summation formula)**.** *Let f be an holomorphic modular form of weight k and level 1 with Fourier coefficients $(\varrho_f(n))_n$; let V be a smooth compactly supported function, $q \geqslant 1$ and $(a, q) = 1$, one has we have for $X > 0$*

$$\sum_{n\geqslant 1} \varrho_f(n) V\Big(\frac{n}{X}\Big) e\Big(\frac{an}{q}\Big) = \varepsilon(f)\frac{X}{q}\sum_{n\geqslant 1}\varrho_f(n) e\Big(-\frac{\overline{a}n}{q}\Big)\widetilde{V}\Big(\frac{Xn}{q^2}\Big)$$

where $\varepsilon(f) = \pm 1$ denotes the sign of the functional equation of $L(f, s)$,

$$\widetilde{V}(y) = \int_0^\infty V(u)\mathcal{J}_k(4\pi\sqrt{uy})du,$$

with

$$\mathcal{J}_k(u) = 2\pi i^k J_{k-1}(u).$$

There are several possible proofs of this proposition: one can proceed classically from the Fourier expansion of the modular form f using automorphy relations (see [KMV02, Theorem A.4]). Another more conceptual approach is to working with the Whittaker model of the underlying automorphic representation; this approach offer natural extension to higher rank automorphic forms (see [IT13]). One could also point out other related works like [MS06] as well as the recent paper [KZ16]. We can extend this formula to general functions modulo q. Given $K : \mathbf{Z} \to \mathbf{C}$ a q-periodic function, combining the above formula with the Fourier decomposition

$$K(n) = \frac{1}{q^{1/2}}\sum_{a \,(\mathrm{mod}\, q)} \widehat{K}(a)\mathrm{e}_q(-an).$$

We define the *Voronoi transform* $\breve{K}$ of K as

$$\breve{K}(n) = \frac{1}{\sqrt{q}}\sum_{\substack{h \bmod q\\(h,q)=1}} \widehat{K}(h)\mathrm{e}_q(\overline{h}n) = \frac{1}{\sqrt{q}}\sum_{\substack{h \bmod q\\(h,q)=1}} \widehat{K}(h^{-1})\mathrm{e}_q(hn).$$

We obtain

Corollary 10.3. *Notations are above, given K a q-periodic arithmetic function we have for $N > 0$*

$$\begin{aligned}\sum_{n\geqslant 1}\varrho_f(n)K(n)V\Big(\frac{n}{X}\Big) &= \frac{\widehat{K}(0)}{q^{1/2}}\sum_{n\geqslant 1}\varrho_f(n)V\Big(\frac{n}{X}\Big) + \\ &\quad \varepsilon(f)\frac{X}{q}\sum_{n\geqslant 1}\varrho_f(n)\breve{K}(-n)\widetilde{V}\Big(\frac{nX}{q^2}\Big).\end{aligned}$$

Remark. Another way to obtain such result is to consider the Mellin transform of (the restriction to $\mathbf{F}_q^\times$ of) K:

$$\tilde{K}(\chi) = \frac{1}{(q-1)^{1/2}}\sum_{x\in\mathbf{F}_q^\times} K(x)\chi(x)$$

so that for $x \in \mathbf{F}_q^\times$

$$K(x) = \frac{1}{(q-1)^{1/2}}\sum_{\chi}\tilde{K}(\chi)\chi^{-1}(x).$$

One can then use (archimedean) inverse-Mellin transformation and the functional equation satisfied by the Hecke L-function

$$L(f\otimes\chi,s)=\sum_{n\geqslant 1}\frac{\varrho_f(n)\chi(n)}{n^s}$$

to obtain the formula: for this one observe that the Mellin transform of $\breve{K}_{|\mathbf{F}_q^\times}$ is proportional to

$$\chi\mapsto\varepsilon(\chi)\tilde{K}(\chi^{-1})$$

where $\varepsilon(\chi)$ is the normalized Gauss sum. This method extends easily to automorphic forms of higher rank but uses the fact that q is prime (so that $\mathbf{F}_q^\times$ is not much smaller that $\mathbf{F}_q$).

This identity is formal and has nothing to whether K is a trace function or not. In particular applying it to the Dirac function $\delta_a(n)=\delta_{n\equiv a\,(\mathrm{mod}\,q)}$ for some $a\in\mathbf{F}_q^\times$ we obtain

$$\widehat{\delta_a}(h)=\frac{1}{q^{1/2}}\mathrm{e}_q(ah),\ \breve{\delta}_a(n)=\frac{1}{q^{1/2}}\,\mathrm{Kl}_2(an)$$

so that

$$q^{1/2}\sum_{n\equiv a\,(\mathrm{mod}\,q)}\varrho_f(n)K(n)V\Big(\frac{n}{X}\Big) = \frac{1}{q^{1/2}}\sum_{n\geqslant 1}\varrho_f(n)V\Big(\frac{n}{X}\Big)+ \tag{10.2}$$

$$\varepsilon(f)\frac{X}{q}\sum_{n\geqslant 1}\varrho_f(n)\,\mathrm{Kl}_2(-an)\widetilde{V}\Big(\frac{nX}{q^2}\Big). \tag{10.3}$$

This is an example of a natural transformation which starting from the elementary function δ_a produces a genuine trace function (Kl_2).

Besides this case we would like to use the formula for K a trace function: we observe that the Voronoi transform $\breve{K}$ is "essentially" the Fourier transform of the function $h\in\mathbf{F}_q^\times\mapsto\widehat{K}(h^{-1})=\widehat{K}(wh)$ with $w=\begin{pmatrix}0&1\\0&1\end{pmatrix}$; it is therefore essentially involutive. It would be useful to know that $\breve{K}$ is a trace function. Suppose that K is associated to some isotypic fourier sheaf $\mathcal{F}$, then $\breve{K}$ is a (isotypic) trace function as long as long as $[w]^*\widehat{\mathcal{F}}$ is a Fourier trace function. This mean that $\widehat{\mathcal{F}}$ has not irreducible constituent of the shape $w^*\mathcal{L}_\psi$ which (by involutivity of the Fourier transform mean that $\mathcal{F}$ has no irreducible constituent isomorphic to some Kloosterman sheaf $\mathcal{K}\ell_2$. This reasoning[10] is essentially the reverse of the one leading to (10.2).

Let us assume that $\breve{K}$ is also a trace function; integration by parts show that for V smooth compactly supported, $\widetilde{V}$ has rapid decay for $x\gg 1$, Corollary 10.3 is an equality between a sum of length X and a sum of length about q^2/X (up to the term $\frac{\widehat{K}(0)}{q^{1/2}}\sum_{n\geqslant 1}\varrho_f(n)V\Big(\frac{n}{X}\Big)$ which is easy to understand). The two length are the same when $X=q$.

10.3. **The amplification method.** As mentionned above Theorem 10.1 is proven "for all modular forms at one" as a consequence of the amplification method.

The principle of the amplification method (invented by H. Iwaniec and which in the special case $K=\chi$ was used first by Bykovskii) consist in evaluating the following moments: for $L\geqslant 1$ and $(x_l)_{l\leqslant L}$ real numbers we consider the following average over orthogonal bases of modular forms (holomorphic or general) of level q:

$$M_k(K):=\sum_{g\in\mathcal{B}_k(q)}|A(g)|^2|\mathcal{S}_V(g,K)|^2 \tag{10.4}$$

[10] by involutivity of the Voromoi transform

and

$$M(K) := \sum_{k\equiv 0 \,(\mathrm{mod}\, 2),\ k>0} \dot{\phi}(k)(k-1) \sum_{g\in \mathcal{B}_k(q)} |A(g)|^2 |\mathcal{S}_V(g,K)|^2 + \sum_{g\in \mathcal{B}(q)} \tilde{\phi}(t_g) \frac{4\pi}{\cosh(\pi t_g)} |A(g)|^2 |\mathcal{S}_V(g,K)|^2 + \sum_{\chi} \sum_{g\in \mathcal{B}(\chi)} \int_{-\infty}^{\infty} \tilde{\phi}(t) \frac{1}{\cosh(\pi t)} |A(g,t)|^2 |\mathcal{S}_V(E_{\chi,g}(t),K,p)|^2 \, dt, \tag{10.5}$$

where $\mathcal{B}_k(q)$, $\mathcal{B}(q)$, $\mathcal{B}(\chi)$ denote orthonormal bases of Hecke-eigen modular forms of level q (either holomorphic of weight k or Maass or Eisenstein series), $\dot{\phi}$, $\tilde{\phi}$ are weights constructed from some smooth function, ϕ, rapidly decreasing at 0 and ∞, which depend only on the spectral parameters of the forms and for each form g, $A(g)$ ("A" is for amplifier) denote a suitable linear form in the Hecke eigenvalues $(\lambda_g(n))_{(n,q)=1}$

$$A(g) = \sum_{l\leqslant L} x_l \lambda_g(l)$$

with suitable coefficients x_l and of length some parameter L. The weights $\tilde{\phi}$ are positive while the weight $\dot{\phi}(k)$ is positive at least for k large enough; one can then has to this moment a finite linear combination of the $M(K)$ from which one can bounds

$$|M|(K) := \sum_{k\equiv 0 \,(\mathrm{mod}\, 2),\ k>0} |\dot{\phi}(k)|(k-1) \sum_{g\in \mathcal{B}_k(q)} |A(g)|^2 |\mathcal{S}_V(g,K)|^2 + \sum_{g\in \mathcal{B}(q)} \tilde{\phi}(t_g) \frac{4\pi}{\cosh(\pi t_g)} |A(g)|^2 |\mathcal{S}_V(g,K)|^2 + \sum_{\chi} \sum_{g\in \mathcal{B}(\chi)} \int_{-\infty}^{\infty} \tilde{\phi}(t) \frac{1}{\cosh(\pi t)} |A(g,t)|^2 |\mathcal{S}_V(E_{\chi,g}(t),K,p)|^2 \, dt, \tag{10.6}$$

As we explain below one will be able to prove the following bound

$$M(K), M_k(K) \ll_{C(\mathcal{F})} q^{o(1)} (q \sum_{l\leqslant L} |x_l|^2 + q^{1/2} L (\sum_{l\leqslant L} |x_l|)^2).$$

Now if f is a Hecke-eigenform of level 1 (of L^2 norm 1 for the usual inner product on the level one modular curve) then $f/(q+1)^{1/2}$ embeds in an orthonormal basis of forms of level q.

Since all the terms in $|M|(K)$ are non-negative, $|M|(K)$ is a bound for any single term occuring discretely in the above sum (ie. when f is a cusp form); therefore we obtain

$$\frac{1}{q+1} |A(f)|^2 |\mathcal{S}_V(f,K)|^2 \ll_{C(\mathcal{F}),f} q^{o(1)} (q \sum_{l\leqslant L} |x_l|^2 + q^{1/2} L (\sum_{l\leqslant L} |x_l|)^2).$$

Now we perform amplification by choosing some absolutely bounded sequence $(x_l)_{l\leqslant L}$ taylor made for f such that $A(f)$ is large

$$|A(f)| \gg L^{1+o(1)};$$

specifically choosing

$$x_l = \mathrm{sign}(\lambda_f(l))$$

we obtain

$$|A(f)| \gg L^{1+o(1)}.$$

Dividing by L we obtain

$$|\mathcal{S}_V(f,K)|^2 \ll q^{o(1)}(q^2/L + q^{3/2}L^2)$$

and the optimal choice is $L = q^{1/6}$ giving us

$$\mathcal{S}_V(f,K) \ll q^{1-1/12+o(1)}.$$

10.4. **Computing the moments.** We now bound $M(K)$. Opening squares and using the multiplicative properties of Hecke eigenvalue we are essentially reduced to bounding sums of the shape

$$\sum_{m,n}\sum V(\frac{m}{q})V(\frac{n}{q})K(m)\overline{K(n)}\Delta_{q,\phi}(lm,n) \tag{10.7}$$

and

$$\sum_{m,n}\sum V(\frac{m}{q})V(\frac{n}{q})K(m)\overline{K(n)}\Delta_{q,k}(lm,n) \tag{10.8}$$

where $1 \leqslant l \leqslant L^2$ and

$$\Delta_{q,k}(lm,n) = \sum_{g\in\mathcal{B}_k(q)} \varrho_g(lm)\overline{\varrho_g(n)}$$

and

$$\begin{aligned}\Delta_{q,\phi}(lm,n) \quad &= \sum_{k\equiv 0\,(\mathrm{mod}\,2),\ k>0} \dot{\phi}(k)(k-1)\sum_{g\in\mathcal{B}_k(q)} \varrho_g(lm)\overline{\varrho_g(n)} \\ &+ \sum_{g\in\mathcal{B}(q)} \tilde{\phi}(t_g)\frac{4\pi}{\cosh(\pi t_g)}\varrho_g(lm)\overline{\varrho_g(n)} \\ &+ \sum_{\chi}\sum_{g\in\mathcal{B}(\chi)} \int_{-\infty}^{\infty} \tilde{\phi}(t)\frac{1}{\cosh(\pi t)}\varrho_g(lm,t)\overline{\varrho_g(n,t)}\,dt.\end{aligned}$$

The Petterson-Kuznetzov formula express $\Delta_{q,k}(m,n)$ $\Delta_{q,\phi}(m,n)$ as a sum of Kloosterman sums:

$$\Delta_{q,k}(m,n) = \delta_{m=n} + 2\pi i^{-k}\sum_c \frac{1}{cq}S(m,n;cq)J_{k-1}\left(\frac{4\pi\sqrt{mn}}{cq}\right). \tag{10.9}$$

$$\Delta_{q,\phi}(m,n) = \sum_c \frac{1}{cq}S(m,n;cq)\phi\left(\frac{4\pi\sqrt{mn}}{cq}\right) \tag{10.10}$$

where

$$S(m,n;cq) = \sum_{(x,cq)=1} e_{cq}(mx + n\overline{x})$$

is the non-normalized Kloosterman sum of modulus cq ($x.\overline{x} \equiv 1\,(\mathrm{mod}\,cq)$). In (10.8), Because m and n are of size q and ϕ is rapidly decreasing at 0 the contribution of the $c \gg l^{1/2}$ is small we will simplify further by evalauting only the contribution of $c = 1$, that is

$$\frac{1}{q}\sum_{m,n}\sum V(\frac{m}{q})V(\frac{n}{q})K(m)\overline{K(n)}S(lm,n;q)\phi(\frac{4\pi\sqrt{lmn}}{q}).$$

Our next step will be to open the Kloosterman sum and apply the Poisson summation formula on the m and n variables: we obtain

$$\frac{1}{q}\frac{q^2}{(q^{1/2})^2}\sum_{m^*,n^*}\sum \widehat{W}(m^*,n^*)\sum_{x\in\mathbf{F}_q^\times} \widehat{K}(lx+m^*)\overline{\widehat{K}(x^{-1}+n^*)}$$

where

$$W(x,y) = V(x)V(y)\phi(4\pi\sqrt{lxy}).$$

In particular the Fourier transform $\widehat{W}(m^*, n^*)$ is very small unless $m^* + n^* \ll l$ so the above sum is over $m^*, n^* \ll l$. Setting

$$\gamma_1 = \begin{pmatrix} l & m^* \\ & 1 \end{pmatrix},\ \gamma_2 = \begin{pmatrix} n^* & 1 \\ 1 & 0 \end{pmatrix}$$

we see that the x-sum is the correlation sum $q\mathcal{C}(K, \gamma_2.\gamma_1^{-1})$ which is $\ll q^{1/2}$ iff $\gamma_2.\gamma_1^{-1} \notin G_{\widehat{\mathcal{F}}}$. Now by Theorem 7.4 (which says that $G_{\widehat{F}}$ is constrained) shows if l is a sufficiently small fixed (positive) power of q that $\gamma_2.\gamma_1^{-1}$ belong to $G_{\widehat{\mathcal{F}}}$ for $\ll l^{o(1)}$ pairs (m^*, n^*). To obtain the main result once has to combine such an argument with an averaging over the l parameter.

11. The ternary divisor function in large arithmetic progression

Given $\lambda = (\lambda(n))_{n\geqslant 1}$ some arithmetic function a natural question in analytic number theory is to understand how well λ is distributed in arithmetic progressions, ie. how it correlates with the characteristic function: given $q \geqslant 1$ and $(a,q) = 1$ one would like to evaluate the sum

$$\sum_{\substack{n\leqslant X \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda(n)$$

as $X \to \infty$ and for q as large as possible with respect to X. It is natural to evaluate the difference

$$E(\lambda; q, a) := \sum_{\substack{n\leqslant X \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda(n) - \frac{1}{\varphi(q)} \sum_{\substack{n\leqslant X \\ (n,q)=1}} \lambda(n)$$

and assuming that λ is "essentially" bounded the target would be to obtain a bound of the shape

$$E(\lambda; q, a) \ll_A \frac{X}{q}(\log X)^{-A} \tag{11.1}$$

for any $A \geqslant 0$, as $X \to +\infty$ and for q as large as possible compared to X.

The emblematic case is when $\lambda = 1_{\mathcal{P}}$ is the characteristic function of the primes. In that case the problem can be approached through the analytic properties of Dirichlet L-functions and in particular the localization of their zeros. The Hadamard-de la Vallee-Poussin method (adapted to this setting by Landau) and the Laudau-Siegel theorem show that (11.1) is satisfied for $q \leqslant (\log X)^B$ for any given B while the validity of the generalized Riemann hypothesis would give (11.1) for $q \ll X^{1/2-\delta}$ for any fixed $\delta > 0$. Considering averages over q it is possible to reach the GRH range and this is the content of the Bombieri-Vinogradov theorem

Theorem 11.1 (Bombieri-Vinogradov)**.** *For any $A \geqslant 0$ there exist $B = B(A)$ such that for $Q \leqslant X^{1/2}/\log^B X$*

$$\sum_{q\leqslant Q} \max_{(a,q)=1} |E(1_{\mathcal{P}}; q, a)| \ll X/\log^A X.$$

Passing the GRH/Bombieri-Vinogradov range and reaching the inequality $Q \leqslant x^{1/2+\eta}$ for some $\eta > 0$ is a fundamental problem in analytic number theory with many major applications. For instance, Y. Zhang breakthrough on the existence of bounded gaps between primes went by establishing a version of the BV theorem going beyond the $Q = X^{1/2}$ range[11] [Zha14]; we will discuss some of the techniques entering his proof below.

[11] on average over smooth moduli.

Several arithmetic functions are of interest besides the characteristic fonctions. One the the simplest are the divisor functions

$$d_k(n) = \sum_{n_1 \cdots n_k = n} 1.$$

For $k = 2$, Selberg established the following (still unsurpassed) result

Theorem 11.2 (The divisor function in large arithmetic progressions)**.** *For every non-zero integer a, every $\varepsilon, A > 0$, every $X \geqslant 2$ and every prime q, coprime with a, satisfying*

$$q \leqslant X^{2/3-\varepsilon},$$

we have

$$E(d_2; q, a) \ll \frac{X}{q}(\log X)^{-A},$$

where the implied constant only depends on ε and A (and not on a).

Proof. (Sketch) To simplify matter we replace the problem by evaluating the model sum

$$\sum_{n_1 n_2 \equiv a \,(\mathrm{mod}\, q)} V(\frac{n_1}{N_1})V(\frac{n_2}{N_2})$$

for $N_1 N_2 = X$ and $V \in \mathcal{C}_c^\infty(]1, 2[)$. We apply the Poisson summation formula to the n_1 variable and again on the n_2 variable. The $n_1 n_2 \equiv a \,(\mathrm{mod}\, q)$ condition get transformed into

$$\delta_{n_1 n_2 \equiv a \,(\mathrm{mod}\, q)} \to q^{-1/2} \mathrm{e}_q(a n_1 / n_2) \to q^{-1/2}\, \mathrm{Kl}_2(a n_1 n_2).$$

Regarding ranges the ranges N_1, N_2 are transformed into

$$N_1^* = q/N_1, N_2^* = q/N_2$$

and the whole model sum is transformed into a sum of the shape

$$MT(a; q) + ET(a; q)$$

where $MT(a; q)$ is a main term which we will not specify but is of the right order of magnitude, $ET(a; q)$ is an error term of the shape

$$ET(a; q) = \frac{1}{q^{1/2}} \frac{N_1}{q^{1/2}} \frac{N_2}{q^{1/2}} \sum_{n_1, n_2} \mathrm{Kl}_2(a n_1 n_2) \tilde{V}(\frac{n_1}{N_1^*}) \tilde{V}(\frac{n_2}{N_2^*})$$

where $\tilde{V}$ is a rapidly decreasing function. By Weil bound for Kloosterman sums the error term is bounded by $q^{1/2+\epsilon}$ which smaller that $X(\log X)^{-A}/q$ as long as $X \leqslant q^{2/3-2\varepsilon}$. □

Remark 11.1. Improving the exponent $2/3$ is tantamount to detect cancellation in the sum of Kloosterman sums above. We have given such an improvment in (10.1); unfortunately in the present case the range of the variable $n_1 n_2$ is $N_1^* N_2^* = q^2/X \leqslant q^{1/2}$ which is too short with current technology. See however the [FI92] for an improvement beyond the $q = x^{2/3}$ limit on average over a family of moduli q admitting a specific factorisation.

We now show how to pass the Bombieri-Vinogradov range for the ternary divisor function

$$d_3(n) = \sum_{n_1 n_2 n_3 = n} 1$$

when q is a prime. The very first result of that kind is due to Friedlander-Iwaniec [FI85] (with $\frac{1}{2} + \eta = \frac{1}{2} + \frac{1}{231}$) and was later improved by Heath-Brown (with $\frac{1}{2} + \eta = \frac{1}{2} + \frac{1}{81}$) [HB86]. The best result to date is to be found in [FKM15]

Theorem 11.3 (The ternary divisor function in large arithmetic progressions)**.** *For every non-zero integer a, every $A > 0$, every $X \geqslant 2$ and every prime q, coprime with a, satisfying*

$$q \leqslant x^{\frac{1}{2}+\frac{1}{47}},$$

we have

$$E(d_3; q, a) \ll \frac{x}{q}(\log x)^{-A},$$

where the implied constant only depends on A (and not on a).

Remark. One may wonder why these higher order divisor functions are that interesting: one reason is that these problems can be considered as approximations for the case of the von Mangolt function. Indeed, the Heath-Brown identity (Lemma 8.2) express the von Mangolt function as a linear combination of arithmetic functions involving higher divisor functions, therefore studying higher divisor functions in large arithmetic progressions will enable to progress on the von Mangolt function.[12]

Proof. We consider again a model sum of the shape

$$\sum_{n_1 n_2 n_3 \equiv a \,(\mathrm{mod}\, q)} V(\frac{n_1}{N_1})V(\frac{n_2}{N_2})V(\frac{n_3}{N_3})$$

for $N_1N_2N_3 = X$ and $V \in \mathcal{C}_c^\infty(]1, 2[)$. We apply the Poisson summation formula to the n_1 n_2 and n_3 variable. The $n_1n_2n_3 \equiv a \,(\mathrm{mod}\, q)$ condition is this time transformed into the hyper-Kloosterman sum

$$\frac{1}{q^{1/2}} \mathrm{Kl}_3(an_1n_2n_3).$$

The model sum is transformed into a main term (of the correct order of magnitude) and an error term

$$ET_3(a; q) = \frac{1}{q^{1/2}}\frac{N_1}{q^{1/2}}\frac{N_2}{q^{1/2}}\frac{N_3}{q^{1/2}} \sum_{n_1,n_2,n_3} \mathrm{Kl}_2(an_1n_2n_3)\tilde{V}(\frac{n_1}{N_1^*})\tilde{V}(\frac{n_2}{N_2^*})\tilde{V}(\frac{n_3}{N_3^*})$$

with

$$N_i^* = q/N_i,\ i = 1, 2, 3.$$

The objective is obtain a bound of the shape

$$\Sigma_3 := \sum_{n_1,n_2,n_3} \mathrm{Kl}_3(an_1n_2n_3)\tilde{V}(\frac{n_1}{N_1^*})\tilde{V}(\frac{n_2}{N_2^*})\tilde{V}(\frac{n_3}{N_3^*}) \ll \frac{q}{\log^A q} \tag{11.2}$$

for x of the shape $x = q^{2-\eta}$ for some fixed $\eta > 0$ (small) or equivalently for

$$N_1^*N_2^*N_3^* = q^{1+\eta}.$$

We will show that when $\eta = 0$ (11.2) holds with the stronger bound $\ll q^{1-\delta}$ for some $\delta > 0$. A variation of this argument will show (11.2) for some positive η. Write

$$N_i^* = q^{\nu_i},\ i = 1, 2, 3,\ \nu_1 + \nu_2 + \nu_3 = 1;$$

we assume that

$$0 \leqslant \nu_1 \leqslant \nu_2 \leqslant \nu_3.$$

Suppose that $\nu_3 \geqslant 1/2 + \delta$ then the Polya-Vinogradov method applied to the n_3 variable yield to a bound of the shape

$$\Sigma_3 \ll q^{1-\nu_3+1/2} \log q \ll q^{1-\delta} \log q.$$

[12]This was formalised by Fouvry [Fou85].

Otherwise we have $\nu_3 \leqslant 1/2+\delta$. We assume now that $\nu_1 \geqslant 2\delta$ then $\nu_1 \leqslant 1/3$ so that grouping the variable n_2n_3 into a single variable n of size $\geqslant q^{2/3}$ (weighted by a divisor like function) and applying Theorem 9.1 we obtain the bound

$$\Sigma_3 \ll \log q^{1-\delta}\log^3 q.$$

We may therefore assume that

$$\nu_1 \leqslant 2\delta,\ \ \nu_2+\nu_3 \geqslant 1-2\delta.$$

The n_2n_3-sum is similar to the sum in (10.1) (for $K(n)=\mathrm{Kl}_3(an_1n)$) and indeed the same bound holds so that for any $\varepsilon>0$

$$\Sigma_3 \ll_\varepsilon q^{\nu_1+\frac{\nu_2+\nu_3}{2}+\frac{1}{2}-\frac{1}{8}+\epsilon} \ll_\varepsilon q^{2\delta+1-\frac{1}{8}+\epsilon}$$

which gives the required bounds if δ is chosen $< 1/24$. □

12. The geometric monodromy group and Sato-Tate laws

In this section we discuss an important invariant attached an ℓ-adic sheaf: its geometric monodromy group. This will be crucial in the next section to study more advanced sums of trace functions (multicorrelation sums).

Definition 12.1 ([Kat88][Chap. 3)**.**] Let $\mathcal{F}$ be a sheaf pure of weight 0 and let $\varrho_{\mathcal{F}}$ the associated Galois representation. The geometric (resp. arithmetic) monodromy group $G_{\mathcal{F},\mathrm{geom}}$ (resp. $G_{\mathcal{F},\mathrm{arith}}$) be the Zariski closure of $\varrho_{\mathcal{F}}(G^{\mathrm{geom}})$ (resp. $\varrho_{\mathcal{F}}(G^{\mathrm{arith}})$) inside $\mathrm{GL}(V_{\mathcal{F}})$; in particular

$$G_{\mathcal{F},\mathrm{geom}} \subset G_{\mathcal{F},\mathrm{arith}}.$$

It follows from the work of Deligne that its connected component $G^0_{\mathcal{F},\mathrm{geom}}$ is semisimple.

In the works [Kat88, Kat90a, Kat90b, Kat05a, Kat05b, Kat12] Katz computed the monodromy groups of various classes of sheaves: for instance, he proved in [Kat88] that for Kloosterman sheaves one has (for q large enough)

$$G_{\mathcal{K}\ell_k,\mathrm{geom}} = G_{\mathcal{K}\ell_k,\mathrm{arith}} = \begin{cases} \mathrm{SL}_k & \text{if } k \text{ is odd} \\ \mathrm{Sp}_k & \text{if } k \text{ is even.} \end{cases}$$

12.1. **Sato-Tate laws.** This group is a fundamental invariant of the sheaf. One of the most appealing consequence of its determinantion is the *Sato-Tate law* which describe the distribution of the set

$$\{K(x),\ x\in\mathbf{F}_{q^n}\} \text{ in the disk } D(0,\mathrm{rk}(\mathcal{F}))\subset\mathbf{C}$$

as $q^n\to\infty$. Let us make the simplifying hypothesis that

$$G_{\mathcal{F},\mathrm{geom}} = G_{\mathcal{F},\mathrm{arith}}. \tag{12.1}$$

Before presenting the Sato-Tate law in general let us see how the knowledge of the geometric monodromy group allows to evaluate at least one multiple correlation sum, that is the case $\gamma_i=\gamma_i'=\mathrm{Id}$ and $y=0$. This sum is the average of the trace function associated to the sheaf

$$\mathcal{F}^{\otimes l}\otimes D(\mathcal{F})^{\otimes l}.$$

Consider the representation $\varrho_{l,l}=(\mathrm{Std}\otimes\mathrm{Std}^*)^{\otimes l}$ of the group $G_{\mathcal{F},\mathrm{geom}}$ where $\mathrm{Std}: G_{\mathcal{F},\mathrm{geom}}\hookrightarrow \mathrm{GL}(V_{\mathcal{F}})$ denote the standard representation, then the compositum $\varrho_{l,l}(\mathcal{F})" = "\varrho_{l,l}\circ\varrho_{\mathcal{F}}$ defines an ℓ-adic sheaf pure of weight 0 whose trace function is

$$x\mapsto |K(x)|^{2l}.$$

The decomposition of this representation into irreducible

$$\varrho_{l,l} = m_1(\varrho_{l,l}) \oplus \bigoplus_{1 \neq r \in \mathrm{Irr}(G_{\mathcal{F},\mathrm{geom}})} m_r(\varrho_{l,l}) r$$

yields a decomposition of $\varrho_{l,l}(\mathcal{F})$ into a sum of geometrically irreducible sheaves

$$\varrho_{l,l} \circ \mathcal{F} = m_1(\varrho_{l,l}) \overline{\mathbf{Q}_\ell} \oplus \bigoplus_{1 \neq r \in \mathrm{Irr}(G_{\mathcal{F},\mathrm{geom}})} m_r(\varrho_{l,l}) r \circ \mathcal{F}$$

and a decomposition of $|K(x)|^{2l}$ a sum of irreducible trace functions

$$|K(x)|^{2l} = m_1(\varrho_{l,l}) + \sum_{1 \neq r} m_r(\varrho_{l,l}) K_{r \circ \mathcal{F}}(x).$$

From Deligne's theorem one deduce

$$\frac{1}{q} \sum_x |K(x)|^{2l} = m_1(\varrho_{l,l}) + O_{C(\mathcal{F}),l}(q^{-1/2})$$

where $m_1(\varrho_{l,l})$ is the multiplicity of the trivial representation in the representation $(\mathrm{Std} \otimes \mathrm{Std}^*)^{\otimes l}$ of $G_{\mathcal{F},\mathrm{geom}}$.

The values of the trace function $K_{r \circ \mathcal{F}}$ are given as follows: given $x \in U(\mathbf{F}_q)$ let

$$\theta_{x,\mathcal{F}} = (\mathrm{Fr}_x \,|V_{\mathcal{F}})^{ss} \subset G_{\mathcal{F},\mathrm{arith}} = G_{\mathcal{F},\mathrm{geom}};$$

be the (semisimplifications of the) Frobenius conjugacy class of x acting on $V_{\mathcal{F}}$, then

$$K_{r \circ \mathcal{F}}(x) = \mathrm{tr}(r(\theta_{x,\mathcal{F}})).$$

Let K be a maximal compact subgroup of $G_{\mathcal{F},\mathrm{geom}}(\mathbf{C})$, as explained in [Kat88][Chap. 3], the conjugacy class $\theta_{x,\mathcal{F}} \in G_{\mathcal{F},\mathrm{geom}}(\mathbf{C})^\natural$ defines a unique conjugacy class of K also noted $\theta_{x,\mathcal{F}} \in K^\natural$. The Sato-tate laws describe the distribution of the set $\{\theta_{x,\mathcal{F}},\ x \in U(\mathbf{F}_q)\}$ inside $K^\natural$.

More precisely, let G be some connected semisimple algebraic groupover $\overline{\mathbf{Q}_\ell}$; suppose we are given a sequence of primes $q \to \infty$ and for each such prime some ℓ-adic sheaf $\mathcal{F}$, satisfying (12.1), whose conductor $C(\mathcal{F})$ is bounded independently of q, such that

$$G_{\mathcal{F},\mathrm{geom}} = G.$$

For every non-trivial irreducible representation r, it follows from Deligne's theorem (and the irreducibility of $r \circ \mathcal{F}$) that

$$\frac{1}{U(\mathbf{F}_q)} \sum_{x \in \mathbf{F}_q} \mathrm{tr}(r(\theta_{x,\mathcal{F}})) = O_{C(\mathcal{F}),\dim r}(q^{-1/2}) \to 0,\ q \to \infty;$$

By Peter-Weyl, the characters

$$\theta \in K^\natural \mapsto \mathrm{tr}(r(\theta)),\ r \in \mathrm{Irr}(G)$$

generate a dense subspace in the space of continuous functions on $K^\natural$. Let μ_{ST} be the image of the Haar measure μ_{Haar} on K by the projection $K \to K^\natural$; this measure is the *Sato-tate measure*. Since the characters form an orthonormal family wrt μ_{ST} one deduce that for any $f \in \mathcal{C}(K^\natural)$

$$\frac{1}{U(\mathbf{F}_q)} \sum_{x \in \mathbf{F}_q} f(\theta_{x,\mathcal{F}}) \to \int_{K^\natural} f(\theta) d\mu_{ST}(\theta),\ q \to \infty. \tag{12.2}$$

In other terms one has that

Theorem 12.1 (Sato-Tate law). *As $q \to \infty$ the sets of conjugacy classes*

$$\{\theta_{x,\mathcal{F}},\ x \in U(\mathbf{F}_q)\}$$

become equidistributed wrt to the Sato-Tate measure: the probability measure

$$\frac{1}{U(\mathbf{F}_q)} \sum_{x \in \mathbf{F}_q} \delta_{\theta_{x,\mathcal{F}}}$$

weak-$\star$ converge to the Sato-Tate measure.

12.1.1. *The case of Kloosterman sums.* In [Kat88], Katz computed the monodromy groups of the Kloosterman sheave $(\mathcal{K}\ell_{k,q})_{q \gg 1}$: one has

$$G_{\mathcal{K}\ell_k,\mathrm{geom}} = G_{\mathcal{K}\ell_k,\mathrm{arith}} = \mathrm{SL}_k \text{ or } \mathrm{Sp}_k$$

depending on the parity of k. In particular for the case of Kloosterman sums, $k = 2$, $G = \mathrm{Sp}_2 = \mathrm{SL}_2$, $K = \mathrm{SU}_2(\mathbf{C})$ and on has the identification $K^\natural \simeq [0, \pi]$ given by

$$\theta \in [0, \pi] \mapsto \text{ the conjugacy class of the matrix } \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

and the Sato-Tate measure is the probability measure with density

$$d\mu_{ST} = \frac{2}{\pi} \sin^2(\theta) d\theta.$$

Regarding Kloosterman sums one has

$$\mathrm{Kl}_2(x; q) = \begin{pmatrix} e^{i\theta_{q,x}} & 0 \\ 0 & e^{-i\theta_{q,x}} \end{pmatrix} = 2\cos(\theta_{q,x}),\ \theta_{q,x} \in [0, \pi]$$

The Sato-Tate law become the following explicit statement (due to Katz):

Theorem 12.2 (Sato-Tate law for Kloosterman sums). *For any interval $[a, b] \subset [0, \pi]$*

$$\frac{1}{q-1} |\{x \in \mathbf{F}_q^\times,\ \theta_{q,x} \in [a, b]\}| \to \frac{2}{\pi} \int_a^b \sin^2(\theta) d\theta,\ q \to \infty.$$

The above Sato-Tate law is called "Vertical" as is describe the distribution of Kloosterman angles with varying parameters $x \in \mathbf{F}_q^\times$ as $q \to \infty$. Another possibility discussed by Katz is to consider the Kloosterman angles for a fixed value of the parameter (say $x = 1$) and by varying the modulus q over the primes. The distribution of the angle is a conjecture again due to Katz called the *horizontal Sato-Tate law*

Conjecture 12.2 (Horizontal Sato-Tate conjecture for Kloosterman sums). *As $X \to \infty$, the multiset of Kloosterman angles $\{\theta_{q,1},\ q \leqslant X,\ prime\}$ become equidistributed wrt to the Sato-Tate measure: for any $[a, b] \subset [0, \pi]$*

$$\frac{1}{\pi(X)} |\{q \leqslant X,\ q\ prime,\ \theta_{q,1} \in [a, b]\}| \to \frac{2}{\pi} \int_a^b \sin^2(\theta) d\theta$$

as $X \to \infty$.

This conjecture is very similar to *feu* Sato-Tate conjecture[13] for a given elliptic curve established recently by Taylor and his collaborators, it is also very similar to the equidistribution problem for cubic Gauss sums solved by Heath-Brown and Patterson[HBP79], however prospects for proving this one look much more distant. In particular this conjecture implies the following two simple statements which look as out of reach as the full conjecture.

– There exist infinitely many primes q such that $|\mathrm{Kl}_2(1; q)| \geqslant 2017^{-2017}$,

[13] it also admits a vertical version much simpler to prove

– There exist infinitely many primes q such that $\mathrm{Kl}_2(1;q) > 0$ (resp. $\mathrm{Kl}_2(1;q) < 0$).

In the next section we will explain how some of the results discussed so far enable to say something non-trivial as the constant of replacing the prime moduli q by *almost prime* moduli[14]

12.2. **Towards the horizontal Sato-Tate conjecture for almost prime moduli.** We now describe an application going back to [Mic95] combining Sato-Tate laws together with the results of Section 8 and another ingredient. For $c \geqslant 1$ a squarefree integer and $(a,c) = 1$ the Kloosterman sum of modulus c and parameter a is defined as

$$\mathrm{Kl}_2(a;c) = \frac{1}{c^{1/2}} \sum_{x \in (\mathbf{Z}/c\mathbf{Z})^\times} e(\frac{\overline{x} + ax}{c}).$$

By the Chinese remainder theorem, Kloosterman sums satisfy the *twisted multiplicativity* relation: for $c = c_1c_2$, $(c_1,c_2) = 1$ one has

$$\mathrm{Kl}_2(a;c) = \mathrm{Kl}_2(a\overline{c_2}^2;c_1)\,\mathrm{Kl}_2(a\overline{c_1}^2;c_2) \tag{12.3}$$

so that by Weil's bound one has

$$|\,\mathrm{Kl}_2(a;c)| \leqslant 2^{\omega(c)}$$

where $\omega(c)$ is the number of prime factors of c. One has the following result

Theorem 12.3. *There exist $k \geqslant 2$ such that*

(1) *for infinitely many square-free integers c with at most k prime factors,*

$$|\,\mathrm{Kl}_2(1;c)| \geqslant 2017^{-2017}.$$

(2) *for infinitely many square-free integers c with at most k prime factors,*

$$\mathrm{Kl}_2(1;c) > 0.$$

(3) *for infinitely many square-free integers c with at most k prime factors,*

$$\mathrm{Kl}_2(1;c) < 0.$$

The first statement above was proven in [Mic95] for $k = 2$ (and 2017^{-2017} replaced by $4/25$; the second and the third were first proven in [FM07] for $k = 23$; this value was subsequently improved by Sivak, Matomaki and Ping who holds the current record with $k = 10$[SF09, Mat11, Xi15].

12.2.1. *Kloosterman sums can be large.* We start with the first statement which we prove for $c = pq$ a product of two distinct primes: the main idea is to use the twisted multiplicatity relation

$$\mathrm{Kl}_2(1;pq) = \mathrm{Kl}_2(\overline{p}^2;q)\,\mathrm{Kl}_2(\overline{q}^2;p)$$

and to establish the existence of some κ for which there exist infinitely many pairs of distinct primes (p,q) such that

$$|\,\mathrm{Kl}_2(\overline{p}^2;q)| \& |\,\mathrm{Kl}_2(\overline{q}^2;p)| \geqslant \kappa;$$

for such pairs we have

$$|\,\mathrm{Kl}_2(1;pq)| \geqslant \kappa^2.$$

Given X large, we will consider pairs (p,q) such that $p,q \in [X^{1/2}, 2X^{1/2}[$ and will show that for κ small enough the two sets

$$\{(p,q),\ p \neq q \in [X^{1/2}, 2X^{1/2}[,\ p,q \text{ primes } |\,\mathrm{Kl}_2(\overline{p}^2;q)| \geqslant \kappa\}$$

$$\{(p,q),\ p \neq q \in [X^{1/2}, 2X^{1/2}[,\ p,q \text{ primes } |\,\mathrm{Kl}_2(\overline{q}^2;p)| \geqslant \kappa\}$$

are large enough to have a non-empty (and in fact large) intersection as $X \to \infty$. This is a consequence of the following equidistribution statement

[14] squarefree-integers with an absolutely bounded number of prime factors.

Proposition 12.3. *Given $X \geqslant 1$, $q \in [X, 2X]$ some prime then the (multi)-set of Kloosterman angles*

$$\{\theta_{q,p^{-2}},\ p \in [X^{1/2}, 2X^{1/2}[,\ p \text{ prime},\ p \neq q\}$$

is equidistributed wrt the Sato-Tate measure: for any interval $[a, b] \subset [0, \pi]$

$$\frac{|\{p \in [X^{1/2}, 2X^{1/2}[,\ p \neq q \text{ prime}, \theta_{q,\overline{p}^2} \in [a, b]\}|}{|\{p \in [X^{1/2}, 2X^{1/2}[,\ p \neq q \text{ prime}\}|} \to \frac{2}{\pi}\int_a^b \sin^2(\theta)d\theta$$

as $X \to \infty$.

Proof. We consider the pull-back sheaf $\mathcal{K} := [x \to x^{-2}]^*\mathcal{K}\ell_2$ whose trace function is given by $x \to \mathrm{Kl}_2(\overline{x}^2; q)$: as a representation of the geometric Galois group it corresponds to restricting the representation $\mathcal{K}\ell_2$ to an subgroup of index 2. Since the geometric monodromy group of $\mathcal{K}\ell_2$ is SL_2 the same is true for the pull-back; therefore

$$G_{\mathcal{K},\mathrm{geom}} = G_{\mathcal{K},\mathrm{arith}} = \mathrm{SL}_2\,.$$

The non-trivial irreducible representations of SL_2 are the symetric powers of the standard representation, $\mathrm{Sym}_k(\mathrm{Std})$, $k \geqslant 1$. Given $k \geqslant 1$ the composed sheaf

$$\mathcal{K}_k = \mathrm{Sym}_k \circ \mathcal{K}$$

is by construction geometrically irreducible, has rank $k+1$ with conductor is bounded by a multiple of k and its trace function equals

$$K_k(x) = \mathrm{tr}(\mathrm{Sym}_k \begin{pmatrix} e^{i\theta_{q,\overline{x}^2}} & 0 \\ 0 & e^{-i\theta_{q,\overline{x}^2}} \end{pmatrix}) = \sum_{j=0}^{k} e^{i(k-j)\theta_{q,\overline{x}^2}} e^{-ij\theta_{q,\overline{x}^2}} = \frac{\sin((k+1)\theta_{q,\overline{x}^2})}{\sin(\theta_{q,\overline{x}^2})}.$$

In particular $\mathcal{K}_k$ cannot be geometrically isomorphic to any tensor product of an Artin-Schreier sheaf and a Kummer sheaf (as they have rank 1). Hence by a simple variant of Thm. 8.1 we obtain that

$$\frac{1}{\pi(2X^{1/2}) - \pi(X^{1/2})} \sum_{\substack{p \neq q \\ p \sim X^{1/2}}} K_k(p) \to 0 = \frac{2}{\pi}\int_0^\pi \frac{\sin((k+1)\theta)}{\sin(\theta)} \sin^2(\theta)d\theta$$

□

Averaging over q, we deduce the existence of some $\kappa > 0$ ($\kappa = 0, 4$) such that for X large enough

$$\frac{|\{(p, q),\ p \neq q \in [X^{1/2}, 2X^{1/2}[,\ p, q \text{ primes},\ |\mathrm{Kl}_2(\overline{p}^2; q)| \geqslant \kappa\}|}{|\{(p, q),\ p \neq q \in [X^{1/2}, 2X^{1/2}[,\ p, q \text{ primes}\}|} \geqslant 0,51$$

hence

$$|\{(p, q),\ p \neq q \in [X^{1/2}, 2X^{1/2}[,\ p, q \text{ primes } |\mathrm{Kl}_2(1; pq)| \geqslant \kappa^2\}| \geqslant (0,01 + o(1))\frac{X}{(\frac{1}{2}\log X)^2}. \tag{12.4}$$

12.2.2. *Kloosterman sums change sign.* We now discuss briefly the proof the remaining two statements: to establish the existence of sign changes is suffice to prove that given $V \in \mathcal{C}_c^\infty(]1, 2[)$ some non-negative smooth function, there exist $u > 0$ such that, for X large enough

$$\Big|\sum_{\substack{c \geqslant 1 \\ p|c \Rightarrow p \geqslant X^{1/u}}} \mathrm{Kl}_2(1; c)V(\frac{c}{X})\Big| < \sum_{\substack{c \geqslant 1 \\ p|c \Rightarrow p \geqslant X^{1/u}}} |\mathrm{Kl}_2(1; c)|V(\frac{c}{X}). \tag{12.5}$$

which will shows the existence of sign changes for Kloosterman sums $\mathrm{Kl}_2(1; c)$ whose modulus has at most $1/u$ prime factors. Using sieve methods and the Kuznetzov formula (expressing sums of Kloosterman sums in terms of fourier coefficients of modular forms) one can show that (we refer to [FM07] for a proof)

Proposition 12.4. *For any $\eta > 0$, there exists $u = u(\eta) > 0$ such that*

$$\Big| \sum_{\substack{c\geqslant 1\\ p|c\Rightarrow p\geqslant X^{1/u}}} \mathrm{Kl}_2(1;c)V(\frac{c}{X})\Big| \leqslant \eta\frac{X}{\log X}$$

for X large enough (depending on η and V).

To conclude it is sufficient to show that for some $u = u_0$, one has

$$\sum_{\substack{c\geqslant 1\\ p|c\Rightarrow p\geqslant X^{1/u}}} |\mathrm{Kl}_2(1;c)|V(\frac{c}{X}) \gg_V \frac{X}{\log X} \tag{12.6}$$

(the left-hand side is an increasing function of u so the above inequality remains valid for any $u \geqslant u_0$.) The inequality (12.4) points in the right direction (for $u_0 = 2$), however as stated it is off by a factor $\log X$. One can however recover this factor $\log X$ entierely and prove the lower bound

$$\sum_{\substack{c\geqslant 1\\ p|c\Rightarrow p\geqslant X^{3/8}}} |\mathrm{Kl}_2(1;c)|V(\frac{c}{X}) \gg_V \frac{X}{\log X}$$

the reason is Thm. 8.1 applies also when p is significantly smaller than q (if $q \simeq X^{1/2+\delta}$ on can obtain a non-trivial bound in (8.2) for p of size $X^{1/2-\delta}$ for $\delta \in [0, 1/8[$). The details involve making a partition of unity and we leave it to the interested reader. Another possibility (the one followed originally in [FM07]) is to establish the lower bound (12.6) for a suitable u by restricting to moduli c which are products of exactly three prime factors) using the techniques discussed so far.

13. Multicorrelation of trace functions

So far we have mainly discussed the evaluation of correlation sums associated to two trace functions K_1 and K_2 (especially the case $K_1 = K$ and $K_2 = \gamma^* K$)

$$\mathcal{C}(K_1, K_2) = \frac{1}{q}\sum_x K_1(x)\overline{K_2(x)}.$$

In several applications, multiple correlation sums occur: sums of the shape

$$\mathcal{C}(K_1, K_2, \cdots, K_L) := \frac{1}{q}\sum_x K_1(x)K_2(x)\cdots K_L(x)$$

where the $K_i,\ 1 = 1\cdots L$ are trace functions; of course rewriting the inner term of the sum above as a product of two factors reduce to evaluating a double correlation sum, say associated to the sheaves

$$\mathcal{F} = \mathcal{K}_1 \otimes \cdots \mathcal{K}_l,\ \mathcal{G} = \mathcal{K}_{l+1} \otimes \cdots \mathcal{K}_L$$

but it would remain to determine if $\mathcal{F}$ and $\mathcal{G}$ share a common irreducible component and this may be a hard task. In practice the multicorrelation sums that occur (due to the application of some Hölder inequality and of the Polya-Vinogradov method) are of the shape

$$\mathcal{C}(K, \gamma, h) = \frac{1}{q}\sum_x K(\gamma_1.x)\cdots K(\gamma_l.x)\overline{K(\gamma_1'.x)\cdots K(\gamma_l'.x)}\mathrm{e}_q(xh)$$

for K the trace function of some geometrically irreducible, pure of weight 0, sheaf $\mathcal{F}$,

$$\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_l, \gamma_1', \cdots, \gamma_l') \in \mathrm{PGL}_2(\mathbf{F}_q)^{2l}$$

and $h \in \mathbf{F}_q$.

This sum is the correlation associated to the trace functions of the sheaves

$$\gamma_1^*\mathcal{F}\otimes\cdots\otimes\gamma_l^*\mathcal{F} \text{ and } {\gamma'}_1^*\mathcal{F}\otimes\cdots\otimes{\gamma'}_l^*\mathcal{F}\otimes\mathcal{L}_\psi$$

whose conductors are bounded polynomially in $C(\mathcal{F})$. If $\mathcal{F}$ has rank one, the two sheaves above have rank one and it is not difficult to determine whether these sheaves are geometrically isomorphic or not.

For $\mathcal{F}$ of higher rank we describe a method due to Katz [FI92] which has been axiomatized in [FKM15]: this method rest on the notion of geometric monodromy group which we discussed in the previous section.

13.1. **A theorem on sums of products of trace functions.** In this section we discuss some general result making it possible to evaluate multicorrelations sums of trace function of interest for analytic number theory. The method is basically due to Katz ([FI92]) and was used on several occasions, for instance in [Mic95, FM98] and the general result presented here is a special case of the results of [FKM15]. For this we need to introduce the following variants of the group of automorphism of a sheaf: one is the group of projective automorphisms

$$\mathrm{Aut}^p_{\mathcal{F}}(\mathbf{F}_q)=\{\gamma\in\mathrm{PGL}_2(\mathbf{F}_q),\ \exists \text{ some rank one sheaf } \mathcal{L} \text{ s.t. } \gamma^*\mathcal{F}\simeq_{geom}\mathcal{F}\otimes\mathcal{L}\},$$

the other is the right $\mathrm{Aut}^p_{\mathcal{F}}(\mathbf{F}_q)$-orbit

$$\mathrm{Aut}^d_{\mathcal{F}}(\mathbf{F}_q)=\{\gamma\in\mathrm{PGL}_2(\mathbf{F}_q),\ \exists \text{ some rank one sheaf } \mathcal{L} \text{ s.t. } \gamma^*\mathcal{F}\simeq_{geom}D(\mathcal{F})\otimes\mathcal{L}\}.$$

Let $\mathcal{F}$ be a weight 0, rank k, irreductible sheaf: we assume that

- the geometric monodromy group equals $G_{\mathcal{F},\mathrm{geom}}=\mathrm{SL}_k$ or Sp_k, (we then say that $\mathcal{F}$ is of SL or Sp type),
- the inclusion (12.1) holds,
- $\mathrm{Aut}^p_{\mathcal{F}}(\mathbf{F}_q)=\{\mathrm{Id}\}$; in particular $\mathrm{Aut}^d_{\mathcal{F}}(\mathbf{F}_q)$ is either empty or is reduced to a single element, $\xi_{\mathcal{F}}$ which is a possibly trivial involution ($\xi_{\mathcal{F}}^2=\mathrm{Id}$) and is called the special involution.

Example 13.1. The Kloosterman sheaves $\mathcal{K}\ell_k$ have this property [Kat88]. The special involution is either Id if k is even ($\mathcal{K}\ell_k$ is self-dual) or the matrix $\xi=\begin{pmatrix}-1&\\&1\end{pmatrix}$ for k odd.

Finally we introduce the following ad-hoc definition:

Definition 13.2. Given

$$\boldsymbol{\gamma}=(\gamma_1,\cdots,\gamma_l,\gamma'_1,\cdots,\gamma'_l)\in\mathrm{PGL}_2(\mathbf{F}_q)^{2l},$$

one say that

- γ is normal if there is $\gamma\in\mathrm{PGL}_2(\mathbf{F}_q)$ such that
$$|\{i,\ \gamma_i=\gamma\}|+|\{j,\ \gamma'_j=\gamma\}|\equiv 1\,(\mathrm{mod}\,2).$$
- For $k\geqslant 3$, $\boldsymbol{\gamma}$ is k-normal if there exist $\gamma\in\mathrm{PGL}_2(\mathbf{F}_q)$ such that
$$|\{i,\ \gamma_i=\gamma\}|-|\{\gamma'_j=\gamma\}|\not\equiv 0\,(\mathrm{mod}\,k).$$
- For $k\geqslant 3$, and $\xi\in\mathrm{PGL}_2(\mathbf{F}_q)$ a non-trivial involution, $\boldsymbol{\gamma}$ is k-normal w.r.t. ξ if there exist $\gamma\in\mathrm{PGL}_2(\mathbf{F}_q)$ such that
$$|\{i,\ \gamma_i=\gamma\}|+|\{j,\ \gamma'_j=\xi\gamma\}|-|\{j,\ \gamma'_j=\gamma\}|-|\{i,\ \gamma_i=\xi\gamma\}|\not\equiv 0\,(\mathrm{mod}\,k).$$

Theorem 13.1. *Let K be the trace function of a sheaf $\mathcal{F}$ as above, $l\geqslant 1$ $\boldsymbol{\gamma}\in\mathrm{PGL}_2(\mathbf{F}_q)^{2l}$ and $h\in\mathbf{F}_q$. We assume that either $h\neq 0$ or*

(1) *the sheaf $\mathcal{F}$ is self-dual (so that K is real-valued) and $\boldsymbol{\gamma}$ is normal*

(2) *the $\mathcal{F}$ is of SL_k-type with $k \geqslant 3$, $q > r$, and $\boldsymbol{\gamma}$ is k-normal or k-normal w.r.t. the special involution of $\mathcal{F}$, if it exists. We have*

$$\mathcal{C}(K, \gamma, h) = \frac{1}{q} \sum_x K(\gamma_1.x) \cdots K(\gamma_l.x) \overline{K(\gamma_1'.x) \cdots K(\gamma_l'.x)} \mathrm{e}_q(xh) \ll_{l, C(\mathcal{F})} \frac{1}{q^{1/2}}.$$

Proof. We discuss the proof only in the self-dual case for simplicity. We group together identical γ_i, γ_j' and the sum becomes

$$\frac{1}{q} \sum_x K(\gamma_1''.x)^{m_1} \cdots K(\gamma_r''.x)^{m_r} \mathrm{e}_q(xh)$$

where the γ_i'' are distinct and by hypothesis one of the m_i is odd. The above sum is associated to the trace function of the sheaf

$$\bigotimes \mathrm{Std}({\gamma''}_i^* \mathcal{F})^{\otimes m_i} \otimes \mathcal{L}_\psi$$

where $\psi(\cdot) = \mathrm{e}_q(h\cdot)$ and Std is the tautological representation. We decompose each representation into irreducible

$$\varrho_{m,0} = \mathrm{Std}(G)^{\otimes m} = \sum_r m_r(\varrho_{m,0}) r$$

and are reduced to consider the sheaves

$$\bigotimes_i r_i({\gamma''}_i^* \mathcal{F}) \otimes \mathcal{L}_\psi$$

where the r_i range over irreducible representations of G; by our hypothesis, we know that either $\mathcal{L}_\psi$ is not trivial or at least one of the r_i is non trivial (any necessarily of dimension > 1).

It is then sufficient to show that under these hypotheses, these sheaves are irreducible. For this we consider the direct sum sheaf

$$\bigoplus_i {\gamma''}_i^* \mathcal{F}$$

and let $G_{\oplus,\mathrm{geom}} \subset \prod_i G$ be the Zariski closure of the image of G^{geom} under the sum of representations. The following very criterion is due to Katz

Theorem 13.2 (Goursat-Kolchin-Ribet criterion)**.** *Let $(\mathcal{F}_i)_i$ be a tuple of geometrically irreducible sheaves lisse on $U \subset \mathbf{A}^1_{\mathbf{F}_q}$, pure of weight 0 each with geometric monodromy groups G_i. We assume that*

- *For every i, $G_i = \mathrm{Sp}_{k_i}$ or SL_{k_i},*
- *for any rank 1 sheaf $\mathcal{L}$ and any $i \neq j$ there is no geometric isomorphism between $\mathcal{F}_i \otimes \mathcal{L}$ and $\mathcal{F}_j$,*
- *for any rank 1 sheaf $\mathcal{L}$ and any $i \neq j$ there is no geometric isomorphism between $\mathcal{F}_i \otimes \mathcal{L}$ and $D(\mathcal{F}_j)$,*

then the geometric monodromy group of the sheaf $\bigoplus_i \mathcal{F}_i$ equals $\prod_i G_i$ (obviously it is contained in that product).

Our assumptions (the projective automorphism group of $\mathcal{F}$ is trivial, $\boldsymbol{\gamma}$ is normal and the geometric monodromy group is either SL or Sp) imply that the above criterion hold and this implies that

$$\bigotimes_i r_i({\gamma''}_i^* \mathcal{F}) \otimes \mathcal{L}_\psi$$

is always irreducible. □

13.2. **Application to non-vanishing of Dirichlet L-functions.** We now discuss a beautiful result involving of these techniques due to R. Khan and H. Ngo [KN16] and concern the proportion of non-vanishing of Dirichlet L-functions at the central point $1/2$. The interest in this kind of problems from analytic number theory was renewed with the work of Iwaniec and Sarnak in their celebrated attempt to prove the non-existence of a Landau-Siegel zero [IS00]. Their approach was based on the following general problem: *given*

$$\{L(f,s)=\sum_{n\geqslant 1}\frac{\lambda_f(n)}{n^s},\ f\in\mathcal{F}\}$$

a family of L-function indexed by a "reasonable" family of automorphic forms $\mathcal{F}$[15]*, show that for many $f\in\mathcal{F}$, one has*

$$L(f,1/2)\neq 0.$$

Indeed their work [IS00], Iwaniec and Sarnak showed specifically that when $\mathcal{F}=\mathcal{S}_2(q)$ the set of holomorphic new-forms of weight 2 and prime level q (with trivial nebentypus) if one could show that for q large enough at least $25\%+2017^{-2017}$ of the Hecke L-values $L(f,1/2)$ do not vanish[16] then there is no Landau-Siegel zero; unfortunately they eventually proved

Theorem 13.3 ([IS00])**.** *As $q\to\infty$ along the primes on has*

$$\frac{|\{f\in\mathcal{S}_2(q),\ L(f,1/2)\neq 0\}|}{|\mathcal{S}_2(q)|}\geqslant 1/4-o(1).$$

which is "just" at the limit.

The possibility of producing a positive proportion of non-vanishing is not limited to this specific family and one of the most powerful and general method to achieve this is via the *mollification method*. The principle of mollification method is as follows: given the family $\mathcal{F}$, one consider for some parameter $L\geqslant 1$ and some suitable vector $\boldsymbol{x}_L=(x_\ell)_{\ell\leqslant L}\in\mathbf{C}^\ell$ the linear form

$$\mathcal{L}(\mathcal{F},\boldsymbol{x}_L):=\frac{1}{|\mathcal{F}|}\sum_{f\in\mathcal{F}}L(f,1/2)M(f,\boldsymbol{x}_L) \tag{13.1}$$

and the quadratic form

$$\mathcal{Q}(\mathcal{F},\boldsymbol{x}_L):=\frac{1}{|\mathcal{F}|}\sum_{f\in\mathcal{F}}|L(f,1/2)M(f,\boldsymbol{x}_L)|^2 \tag{13.2}$$

where $M(f,\boldsymbol{x}_L)$ is the linear form (called "mollifier")

$$M(f,\boldsymbol{x}_L)=\sum_{\ell\leqslant L}\frac{\lambda_f(\ell)}{\ell^{1/2}}x_\ell$$

and the x_ℓ are almost bounded coefficients to be chosen in an optimal way:

$$\forall\varepsilon>0,\ x_\ell\ll|\mathcal{F}|^{o(1)}.$$

By Cauchy's inequality one has

$$\frac{|\{f\in\mathcal{F},\ L(f,1/2)\neq 0\}|}{|\mathcal{F}|}\geqslant\frac{|\mathcal{L}(\mathcal{F},\boldsymbol{x}_L)|^2}{\mathcal{Q}(\mathcal{F},\boldsymbol{x}_L)}.$$

[15] a reasonable definition of the notion of "reasonable" can be found in [Kow13, SST16]

[16] in fact something slightly stronger is necessary

For suitable families one can evaluate asymptotically $\mathcal{L}(\mathcal{F}, \boldsymbol{x}_L)$ and $\mathcal{Q}(\mathcal{F}, \boldsymbol{x}_L)$ (the hard case being $\mathcal{Q}$) where $L = |\mathcal{F}|^\lambda$ for $\lambda > 0$ some fixed constant and (upon minimizing $\mathcal{Q}(\mathcal{F}, \boldsymbol{x}_L)$ with respect to $\mathcal{L}(\mathcal{F}, \boldsymbol{x}_L)$) one usually shows that

$$\frac{|\mathcal{L}(\mathcal{F}, \boldsymbol{x}_L)|^2}{\mathcal{Q}(\mathcal{F}, \boldsymbol{x}_L)} = F(\lambda) + o(1)$$

for F some increasing rational fraction with $F(0) = 0$. Before [IS00], Iwaniec and Sarnak had implemented this strategy for the (simpler) family of Dirichlet L-functions of modulus q

$$\{L(\chi, s) = \sum_{n \geqslant 1} \frac{\chi(n)}{n^s},\ \chi \in (\widehat{\mathbf{Z}/q\mathbf{Z}})^\times\}$$

and were able to evaluate (13.1) and (13.2) for any $\lambda < 1/2$ with

$$F(\lambda) = \frac{\lambda}{\lambda + 1}$$

hence

Theorem 13.4 ([IS99]). *As $q \to \infty$ along the primes on has*

$$\frac{|\{\chi \,(\mathrm{mod}\, q),\ L(\chi, 1/2) \neq 0\}|}{|\{\chi \,(\mathrm{mod}\, q)\}|} \geqslant 1/3 - o(1).$$

proving that the proportion of non-vanishing can be taken arbitrarily close to 33%. Shortly after, Michel and Vanderkam [MV00] obtained the same proportion by a slightly different method: taking into account the fact that for a complex character, the L-function $L(\chi, s)$ is not self-dual ($L(\chi, s) \neq L(\overline{\chi}, s)$) and has root number

$$\varepsilon_\chi = i^{\mathfrak{a}} \frac{\tau(\chi)}{q^{1/2}},\ \mathfrak{a} = \frac{\chi(-1) - 1}{2}$$

were $\tau(\chi)$ is the Gauss sum, they introduced a symmetrized mollifier of the shape

$$M^s(\chi, \boldsymbol{x}_L) = M(\chi, \boldsymbol{x}_L) + \overline{\varepsilon_\chi} M(\overline{\chi}, \boldsymbol{x}_L) = \sum_{\ell \leqslant L} \frac{\chi(\ell) + \overline{\varepsilon_\chi}.\overline{\chi}(\ell)}{\ell^{1/2}} x_\ell.$$

Because of the oscillation of the root number ε_χ, they could evaluate (13.2) only in the shorter range $\lambda < 1/4$. However this weaker range is offset by the fact that the symmetrized mollifier is more effective: indeed the rational fraction $F(\lambda)$ is then replaced by

$$F^s(\lambda) = \frac{2\lambda}{2\lambda + 1}$$

which takes value $1/3$ at $\lambda = 1/4$. Recently R. Khan and H. Ngo founds a better method to bound the exponential sums considered in [MV00] building on Theorem 13.1. In that way they increased the allowed range of the mollifier $M^s(\chi, \boldsymbol{x}_L)$ from the range $\lambda < 1/4$ to $\lambda < 3/10$ and obtained the following improvement on the proportion:

Theorem 13.5 ([KN16]). *As $q \to \infty$ along the primes on has*

$$\frac{|\{\chi \,(\mathrm{mod}\, q),\ L(\chi, 1/2) \neq 0\}|}{|\{\chi \,(\mathrm{mod}\, q)\}|} \geqslant 3/8 - o(1).$$

the keystep in the proof is the asymptotic evaluation of the second mollified moment

$$\frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} |L(\chi, 1/2)|^2 |M^s(\chi, \boldsymbol{x}_L)|^2 \tag{13.3}$$

for $L = q^\lambda$, and any fixed $\lambda < 3/10$. By (now) standard methods[17] the L-value $L(\chi, 1/2)$ can be written as a sums of rapidly converging series (cf. [IK04, Thm. 5.3]): for q prime and $\chi \neq 1$

$$|L(\chi,1/2)|^2 = 2\sum_{n_1,n_2\geqslant 1}\frac{\chi(n_1)\overline{\chi}(n_2)}{(n_1n_2)^{1/2}}V(\frac{n_1n_2}{q})$$

where V is a rapidly decreasing function which depends on χ only through its parity $\chi(-1) = \pm 1$. Plugging this expression in the second moment (13.3) and unfolding, one finds that the key point is to obtain a bound of the following shape[18]

$$\sum_{\substack{\ell_1,\ell_2\leqslant L, n_1,n_2\\ (l_1l_2n_1n_2,q)=1}}\sum\frac{x_{l_1}\overline{x_{l_2}}}{(ql_1l_2n_1n_2)^{1/2}}V(\frac{n_1n_2}{q})e(\frac{n_2\overline{l_1l_2n_1}}{q}) \ll q^{-\delta} \tag{13.4}$$

for some $\delta = \delta(\lambda) > 0$ for any fixed $\lambda < 3/10$. This sum can then be decomposed in various subs-ums in which the variables are localized to specific ranges: The becomes essentially that of bounding by $O(q^{-\delta})$ the family of bilinear sums

$$\Sigma(L_1,L_2,N_1,N_2) = \frac{1}{(qL_1L_2N_1N_2)^{1/2}}\sum_{\substack{l_i\sim L_i, i=1,2\\ n_1,n_2}}\sum x_{l_1}\overline{x_{l_2}}W(\frac{n_1}{N_1})W(\frac{n_2}{N_2})e(\frac{n_2\overline{l_1l_2n_1}}{q})$$

where $W \in \mathcal{C}_c(]1/2,2[)$, $L_1, L_2 \leqslant L$ and $N_1N_2 \leqslant q$.

The n_2 sum is essentially a geometric series bounded by

$$\ll \min(N_2, \|\overline{l_1l_2n_1}/q\|^{-1})$$

where $\|\cdot\|$ is the distance to the nearest integer. Hence

$$\Sigma(L_1,L_2,N_1,N_2) \ll \frac{q^\varepsilon}{(qL_1L_2N_1N_2)^{1/2}}\sum_{m\approx L_1L_2N_1}\min(N_2,\|\overline{m}/q\|^{-1})$$

$$\ll \frac{q^{2\varepsilon}}{(qL_1L_2N_1N_2)^{1/2}}\sum_{m\approx L_1L_2N_1}\min(N_2,\|\overline{m}/q\|^{-1})$$

$$\ll \frac{q^{3\varepsilon}}{(qL_1L_2N_1N_2)^{1/2}}\max_{1\leqslant U\leqslant q/2}\min(N_2,\frac{q}{U})|\{(m,u),\ m\approx L_1L_2N_1,\ u\sim U,\ um\equiv\pm 1\,(\mathrm{mod}\,q)\}|$$

$$\ll \frac{q^{3\varepsilon}}{(qL_1L_2N_1N_2)^{1/2}}\max_{1\leqslant U\leqslant q/2}\min(N_2,\frac{q}{U})(\frac{L_1L_2N_1U}{q}+1)$$

$$\ll \frac{q^{3\varepsilon}}{(qL_1L_2N_1N_2)^{1/2}}\max_{1\leqslant U\leqslant q/2}\min(N_2,\frac{q}{U})\frac{L_1L_2N_1U}{q} \ll q^{3\varepsilon}\frac{L}{q^{1/2}}(\frac{N_1}{N_2})^{1/2}. \tag{13.5}$$

(observe that for $\frac{L_1L_2N_1U}{q} \ll 1$ the equation $um \equiv \pm 1\,(\mathrm{mod}\,q)$ has no solution unless $L_1L_2N_1U \ll 1$

Alternatively, applying the Poisson summation formula to the n_1 variable we obtain a sum of the shape

$$\Sigma(L_1,L_2,N_1,N_2) = \frac{1}{(qL_1L_2N_1N_2)^{1/2}}\frac{N_1}{q^{1/2}}\sum_{\substack{l_i\sim L_i, i=1,2\\ n_1,n_2}}\sum x_{l_1}\overline{x_{l_2}}\widetilde{W}(\frac{n_1}{q/N_1})W(\frac{n_2}{N_2})\,\mathrm{Kl}(\overline{l_1l_2}n_1n_2;q)$$

[17] inappropriately called "approximate function equation"

[18] for simplicity we ignore the dependency of V in the parity of the χ's

where $\widetilde{W}$ is bounded and rapidly decreasing. Bounding this sum trivially (using that $|\operatorname{Kl}(m;q)| \leqslant 2$) yields

$$\Sigma(L_1, L_2, N_1, N_2) \ll q^{\varepsilon} L (\frac{N_2}{N_1})^{1/2}. \tag{13.6}$$

The expression $\min(\frac{L}{q^{1/2}}(\frac{N_1}{N_2})^{1/2}, L(\frac{N_2}{N_1})^{1/2})$ is maximal for $\frac{N_1}{N_2} = q^{1/2}$ and equals $L/q^{1/4}$ which is $O(q^{-\delta})$ if $\lambda < 1/4$.

The bound (13.6) did not exploit cancellation from the n_1, n_2, l_1, l_2 averaging and indeed this is not evident because in the limiting case $N_1 = q^{3/4},\ N_2 = q/N_1 = q^{1/4}$, $L_1 = L_2 = L = q^{1/4}$, one has

$$n_1 \approx n_2 \approx l_1 \approx l_2 \approx q^{1/4}$$

which is pretty short. Nevertheless Khan and Ngo where able to detect further cancellation from summing of these short variables. The idea, which we have met already, is to group some of these variables to form longer variables. One possibility could be to group together n_1, n_2 on the one hand and l_1, l_2 on the other hand with the idea of applying the methods of §9. Unfortunately the new variables would have size $q^{1/2}$ which is the Polya-Vinogradov range at which point the standard completion method just fails-. Instead one group n_1, n_2 and l_2 together and leave l_1 alone. The variable $r = n_1 n_2 \overline{l_2} \pmod q$ takes essentially $q^{3/4}$ distinct values but over all of $\mathbf{F}_q^{\times}$ and does not vary along an interval. To counter this defect one uses the Holder inequality instead of Cauchy-Schwarz.

Proceeding as above we write

$$\Sigma(L_1, L_2, N_1, N_2) = \frac{1}{(qL_1L_2N_1N_2)^{1/2}} \frac{N_1}{q^{1/2}} \sum_{r \in \mathbf{F}_q^{\times}} \sum_{l_1} x_{l_1} \nu(r) \operatorname{Kl}(\overline{l_1} r; q)$$

where

$$\nu(r) = \sum_{\substack{l_2 \sim L_2 \\ n_1, n_2}} \overline{x_{l_2}} \widetilde{W}(\frac{n_1}{q/N_1}) W(\frac{n_2}{N_2}).$$

Under the assumption

$$L_2 \frac{q}{N_1} N_2 < q/100 \Longrightarrow L \frac{N_2}{N_1} < 1/100 \tag{13.7}$$

we have

$$\sum_r |\nu(r)| + \sum_r |\nu(r)|^2 \ll q^{\varepsilon} L_2 \frac{q}{N_1} N_2.$$

Indeed under (13.7) one has

$$\overline{l}_2 n_1 n_2 \equiv \overline{l}_2' n_1' n_2' \pmod q \Longleftrightarrow l_2' n_1 n_2 \equiv l_2 n_1' n_2' \pmod q \Longleftrightarrow l_2' n_1 n_2 = l_2 n_1' n_2'$$

and the choice of l_2', n_1, n_2 determine l_2, n_1', n_2' up to $O(q^{\varepsilon})$ possibilities. Hence, applying Cauchy's inequality twice we obtain

$$\Sigma(L_1, L_2, N_1, N_2) = \frac{q^{\varepsilon}}{(qL_1L_2N_1N_2)^{1/2}} \frac{N_1}{q^{1/2}} (L_2 \frac{q}{N_1} N_2)^{3/4} (\sum_{r \in \mathbf{F}_q^{\times}} |\sum_{l \sim L_1} x_l \operatorname{Kl}(\overline{l} r; q)|^4)^{1/4}.$$

Now (using that $\operatorname{Kl}(n;q) \in \mathbf{R}$)

$$\sum_{r \in \mathbf{F}_q^{\times}} |\sum_{l \sim L_1} x_l \operatorname{Kl}(\overline{l} r; q)|^4 \ll q^{\varepsilon} \sum_{\boldsymbol{l}} |\sum_{r \in \mathbf{F}_q^{\times}} \prod_{i=1}^{4} \operatorname{Kl}(\overline{l_i} r; q)|$$

where $\boldsymbol{l} = (l_1, l_2, l_3, l_4) \in [L_1, 2L_1[^4$.

The Theorem 13.1 applied to the Kloosterman sheaf gives

$$\sum_{r\in\mathbf{F}_q^\times}\prod_{i=1}^4 \mathrm{Kl}(\overline{l}_i r;q)\ll q^{1/2}$$

unless there exist a partition $\{1,2,3,4\}=\{i,j\}\sqcup\{k,l\}$ such that

$$l_i=l_j,\ l_k=l_l$$

in which case we use the trivial bound

$$\sum_{r\in\mathbf{F}_q^\times}\prod_{i=1}^4 \mathrm{Kl}(\overline{l}_i r;q)\ll q.$$

Hence

$$\sum_{\boldsymbol{l}}|\sum_{r\in\mathbf{F}_q^\times}\prod_{i=1}^4 \mathrm{Kl}(\overline{l}_i r;q)|\ll L_1^2q+L_1^4q^{1/2}$$

and

$$\begin{aligned}\Sigma(L_1,L_2,N_1,N_2)&\ll\frac{q^\varepsilon}{(qL_1L_2N_1N_2)^{1/2}}\frac{N_1}{q^{1/2}}(L_2\frac{q}{N_1}N_2)^{3/4}(L_1^{1/2}q^{1/4}+L_1q^{1/8})\\&\ll q^\varepsilon L(\frac{N_2}{N_1})^{1/2}(Lq\frac{N_2}{N_1})^{-1/4}(L^{-1/2}q^{1/4}+q^{1/8}).\end{aligned}\tag{13.8}$$

For $L\geqslant q^{1/4}$ (the range one would like to improve) one obtains under (13.7)

$$\Sigma(L_1,L_2,N_1,N_2)\ll q^\varepsilon L(\frac{N_2}{N_1})^{1/2}(Lq^{1/2}\frac{N_2}{N_1})^{-1/4}.\tag{13.9}$$

Suppose now we are in a limiting case for (13.6), $L^2N_2/N_1=1$; then (13.7) holds as long as $L\gg 1$ and (13.9) improves over (13.6) by a factor $(q^{1/2}/L)^{1/4}$ which is <1 as long as $L<q^{1/2}$.

A more detailed analysis combining (13.5), (13.6) and (13.9) shows that (13.4) holds for any fixed $\lambda<3/10$ and hence to Theorem 13.5.

14. Advanced completions methods: the q-van der Corput method

In this section and the next ones, we discuss general methods to evaluate trace function along intervals of length smaller than the Polya-Vinogradov range discuss in §6.

14.1. **The q-van der Corput method.** The next method we are going to describe is an arithmetic analog of the above technique : the q-Van der Corput method due to Heath-Brown. That method concern c-periodic functions for c a *composite number*. Suppose (to simplify the presentation) that $c=pq$ for two primes p and q and let

$$K_c=K_pK_q:\mathbf{Z}/c\mathbf{Z}\to\mathbf{C}$$

somme function modulo c which is the product of two trace functions modulo p and q (of conductor bounded by some constant C). We consider the partial sum

$$S_V(K,N):=\sum_n K_c(n)V(\frac{n}{N})=\sum_n K_p(n\,(\mathrm{mod}\,p))K_q(n\,(\mathrm{mod}\,q))V(\frac{n}{N})$$

where $V\in\mathcal{C}^\infty(]1,2[)$ and $2N<pq$. We will explain the proof of the following result

Theorem 14.1 (q van der corput method)**.** *Assume that $\mathcal{F}$ is geometrically irreducible and not geometrically isomorphic to a linear or quadratic phase (ie. not of the shape $[P]^*\mathcal{L}_\psi$ for P a polynomial of degre $\leqslant 2$) then*

$$S_V(K,N) \ll_C N^{1/2}(p+q^{1/2})^{1/2}$$

Remark. This bound is non trivial as long as

$$N \geqslant \max(p, q^{1/2})$$

which is a weaker condition than $N \geqslant (pq)^{1/2}$ as long as

$$1 < p < q.$$

We have therefore improved over the Polya-Vinogradov range; moreover the range of non triviality is maximal when $p \approx c^{1/3}$ and $q \approx c^{2/3}$ and in that case one has

$$S_V(K,N) \ll_C N^{1/2}c^{1/6} \tag{14.1}$$

which is non-trivial as long as

$$N \geqslant c^{1/3}.$$

Proof. The proof make use of the (semi-)invariance of K under translations:

$$K(n+ph) = K_p(n)K_q(n+qh).$$

For $H \leqslant N/100p$ we have

$$S_V(K,N) = \frac{1}{2H+1}\sum_{|h|\leqslant H}\sum_n K_p(n)K_q(n+ph)V(\frac{n+ph}{N})$$

$$= \frac{1}{2H+1}\sum_{|n|\leqslant 3N} K_p(n)\sum_{\substack{|h|\leqslant H\\ n}} K_q(n+ph)V(\frac{n+ph}{N})$$

$$\ll \frac{1}{2H+1}N^{1/2}\Big(\sum_{|n|\leqslant 3N}\Big|\sum_{|h|\leqslant H} K_q(n+ph)V(\frac{n+ph}{N})\Big|^2\Big)^{1/2}$$

$$= \frac{N^{1/2}}{H}\Big(\sum_{|h|,|h'|\leqslant H}\sum_n K_q(n+ph)\overline{K_q(n+ph')}W_{p,h,h'}(\frac{n}{N})\Big)^{1/2}$$

where

$$W_{p,h,h'}(\frac{n}{N}) = V(\frac{n+ph}{N})\overline{V(\frac{n+ph'}{N})}.$$

We split the h,h'-sum into its diagonal and non-diagonal contribution

$$\sum_{|h|,|h'|\leqslant H}\sum\cdots = \sum_{\substack{|h|,|h'|\leqslant H\\ h=h'}}\sum\cdots + \sum_{\substack{|h|,|h'|\leqslant H\\ h\neq h'}}\sum\cdots.$$

The diagonal sum contributes by $O(NH)$ and it remains to consider the correlation sum

$$\mathcal{C}(K_q,h,h') := \sum_n K_q(n+ph)\overline{K_q(n+ph')}W(\frac{n}{N},\frac{ph}{N},\frac{ph'}{N})$$

for $h \neq h'$.

Observe that this is the sum of a trace function of modulus q of length $\approx N$. By comparison with the initial sum, we had a trace function of modulus pq of length $\approx N$ so the relative length

of n compared to the modulus has increased ! By the Polya-Vinogradov method, it is sufficient to determine whether the sheaf

$$[+ph]^*\mathcal{F} \otimes [+ph']^*D(\mathcal{F})$$

has an Artin-Schreier sheaf in its irreducible components; this is equivalent to whether one has an isomorphism

$$[+p(h-h')]^*\mathcal{F} \simeq \mathcal{F} \otimes \mathcal{L}_\psi$$

for some Artin-Schreier sheaf. We will answer this question in a slighly more general form:

Definition 14.1. A polynomial phase sheaf of degree d is a sheaf of the shape $[P]^*\mathcal{L}_\psi$ for P a polynomial of degree d and ψ a non-trivial additive character. It is lisse on $\mathbf{A}^1_{\mathbf{F}_q}$, ramified at infinity with Swan conductor equal to d and its trace function equals

$$x \mapsto \psi(P(x)).$$

We can now invoque the following

Proposition 14.2 ([Pol14])**.** *Suppose that $\mathcal{F}$ is geometrically irreducible, not isomorphic to a polynomial phase of degree $\leqslant d$ and that $C(\mathcal{F}) \leqslant q^{1/2}$, then for any $h \in \mathbf{F}_q - \{0\}$ and any polynomial P of degree $\leqslant d-1$ there is no geometric isomorphism of the shape*

$$[+h]^*\mathcal{F} \simeq \mathcal{F} \otimes [P]^*\mathcal{L}_\psi.$$

Proof. We will only give the easiest part of it and refer to [Pol14] for the complete argument. Suppose that $\mathcal{F}$ is ramified at some point $x_0 \in \mathbf{A}^1(\overline{\mathbf{F}_q})$, since polynomial phases are ramified only at ∞ the isomorphism

$$[+h]^*\mathcal{F} \simeq \mathcal{F} \otimes [P]^*\mathcal{L}_\psi$$

restricted to the inertia group I_x implies that $\mathcal{F}$ is ramified at $x_0 - h$ and iterating at $x_0 - nh$ for any $n \in \mathbf{Z}$, this would imply that $C(\mathcal{F}) \geqslant q$ which is excluded. It remains to deal with the case where $\mathcal{F}$ is ramified only at ∞. □

Under our assumptions the above proposition implies that for $h \neq h'$

$$\mathcal{C}(K_q, h, h') = O(q^{1/2})$$

and that

$$S_V(K, N) \ll N^{1/2}(\frac{N}{H} + q^{1/2})^{1/2}$$

and we choose $H = N/100p$ to conclude the proof.

□

14.2. **iterating the method.** Suppose more generally that c is a squarefree number and that

$$K_c = \prod_{q|c} K_q$$

is a product of trace functions associated to sheaves not containing any polynomial phases. One can repeat the above argument after factoring c into a product of squrefree coprime moduli $r.s$ and decompose accordingly

$$K_c = K_r.K_s.$$

By the exact same method we reach sums of the shape

$$\sum_n K_s(n+rh)\overline{K_s(n+rh')}W_{r,h,h'}(\frac{n}{N}) \tag{14.2}$$

This time we need to be a bit more careful and decompose the h, h' sum according to the gcd $(h-h', s)$. After applying the Poisson summation formula (cf. (6.2)) we can factor the resulting Fourier transform modulo s into sums over prime moduli $q|s$:

$$\widehat{K_s}(y) = \prod_{q|c} \widehat{K_q}(\overline{s_q} y \,(\mathrm{mod}\, q)),\ y \in \mathbf{Z}/s\mathbf{Z},\ s_q = s/q.$$

If $q|h-h'$ we use the trivial bound $\widehat{K_q}(\overline{s_q} y \,(\mathrm{mod}\, q)) \ll q^{1/2}$ and if $q \not| h-h'$ we use the non-trivial bound $\widehat{K_q}(\overline{s_q} y \,(\mathrm{mod}\, q)) \ll 1$. We eventually obtain (see [Pol14])

Theorem 14.2. *Let $C \geqslant 1$ and let c be squarefree and $K_c : \mathbf{Z}/c\mathbf{Z} \to \mathbf{C}$ be a product of trace functions K_q such that for any prime $q|c$ the underlying sheaf $\mathcal{F}_q$ is of conductor $\leqslant C$, geometrically irreducible, not isomorphic to geometrically isomorphic to any polynomial phase of degree $\leqslant 2$ then*

$$S_V(K, N) \ll_C c^{\varepsilon} N^{1/2}(r + s^{1/2})^{1/2}$$

for any $\varepsilon > 0$.

If s is not a prime we could also iterate, factor s into $s = r_2 s_2$ and instead of applying the Polya-Vinogradov completion method to the sum (14.2) one could also apply the q-van der Corput method with the trace functions

$$n \to K_q(n + rh)\overline{K_q(n + rh')},\ q|s_1.$$

This lead us to the quadruple correlation sum

$$\mathcal{C}(K_q, \gamma, \alpha) = \frac{1}{q}\sum_x K_q(\gamma_1.x)K(\gamma_2.x)\overline{K(\gamma_1'.x)K(\gamma_2'.x)}\mathrm{e}_q(\alpha.x)$$

where the $\gamma_i, \gamma_j',\ i, j = i, 2$ are unipotent matrices

$$\gamma_i = \begin{pmatrix} 1 & h_i \\ 0 & 1 \end{pmatrix},\ \gamma_i' = \begin{pmatrix} 1 & h_j' \\ 0 & 1 \end{pmatrix}$$

In suitable situations we can then apply Theorem 13.1 from the previous section.

An important example is when

$$K_c(n) = \mathrm{Kl}_k\, n; c)$$

is an hyper-Kloosterman sums: for any $q|c$, one has

$$K_q(y) = \mathrm{Kl}_k(\overline{c_q}^k y; q),\ c_q = c/q$$

and the sheaf underlying sheaf is the multiplicatively shifted Kloosterman sheaf $\mathcal{F}_q = [\times \overline{c_q}^k]^* \mathcal{K}\ell_k$: in that case Theorem 13.1 and we eventually obtain the bound

$$S_V(\mathrm{Kl}_k(\cdot; c), N) \ll_k c^{\varepsilon} N^{1/2}(r + (N^{1/2}(s_1 + s_2^{1/2}))^{1/2})^{1/2}.$$

for any factorisation $c = rs_1s_2$. In particular, if there exists a factorisation $c = rs_1s_2$ such that

$$r \approx c^{1/4},\ s_1 \approx c^{1/4},\ s_2 \approx c^{1/2}$$

we obtain

$$S_V(\mathrm{Kl}_k(\cdot; c), N) \ll_k N^{1-\eta}$$

for some $\eta = \eta(\delta) > 0$ as long as

$$N \geqslant c^{1/4+\delta}.$$

Iterating once more we see that for any factorisation $c = rs_1s_2s_3$ one has

$$S_V(\mathrm{Kl}_k(\cdot; c), N) \ll_{k,\varepsilon} c^{\varepsilon} N^{1/2}(r + (N^{1/2}(s_1 + (N^{1/2}(s_2 + s_3^{1/2}))^{1/2}))^{1/2})^{1/2} \tag{14.3}$$

so if there exists a factorisation $c = rs_1s_2s_3$ such that

$$r \approx c^{1/5},\ s_1 \approx c^{1/5},\ s_2 \approx c^{1/5},\ s_3 \approx c^{2/5}$$

then

$$S_V(\mathrm{Kl}_k(\cdot;c), N) \ll_{k,\varepsilon} N^{1-\eta}$$

for some $\eta = \eta(\delta) > 0$ as long as

$$N \geqslant c^{1/5+\delta}$$

and we can cary-on that way as long as enough factorisation for c are available. Such availability is garanteed by the notion of smoothness:

Definition 14.3. An integer $c \neq 0$ is Δ-smooth if

$$q|c\ (q \text{ prime }) \Rightarrow q \leqslant \Delta.$$

Using the reasoning above Irving [Irv15] proved (for $k = 2$) the following result (in a quantitative form):

Theorem 14.3. *For any $L \geqslant 2$ there exists $l = l(L) \geqslant 1$ and $\eta = \eta(L) > 0$ such that for c a squarefree integer which is $c^{1/l}$-smooth and any $k \geqslant 2$, one has,*

$$S_V(\mathrm{Kl}_k(\cdot;c), N) \ll_{k,V} N^{1-\eta}$$

whenever $N \geqslant c^{1/L}$.

Therefore one can obtain non-trivial bounds for extremely short Kloosterman sums as long as their modulus is smooth enough. In particular for $k = 2$ we have seen in Remark 11.1 that improving on Selberg's 2/3-exponent for the distribution of the divisor function in large arithmetic progression (Theorem 11.2) was essentially equivalent to bounding non-trivially sums of the shape

$$\sum_{n_1,n_2} \mathrm{Kl}_2(an_1n_2;c)V(\frac{n_1}{N_1^*})V(\frac{n_2}{N_2^*})$$

for

$$N_1^*N_2^* \approx c^{1/2}.$$

If $N_1^*N_2^* \approx c^{1/2}$ then $\max(N_1^*, N_2^*) \gg c^{1/4}$ and we can use the (14.3) to bound non-trivially the above sum granted that c is smooth enough:

Theorem 14.4. [Irv15] *There exists $L \geqslant 4$ and $\eta > 0$ such that for any $c \geqslant 1$ which is squarefree and $c^{1/L}$-smooth and any a coprime with c, one has for $x \geqslant c^{2/3+\eta}$ and any $A \geqslant 0$*

$$E(d_2; c, a) \ll_A \frac{x}{c}(\log x)^{-A}.$$

See [Irv16] and [XW16] for further applications.

15. Around Zhang's theorem on bounded gaps between primes

the methods of the previous chapter have been applied in a spectacular way by Yitang Zhang on his proof of the existence of bounded gaps between primes:

Theorem 15.1 ([Zha14])**.** *Let $(p_n)_{n\geqslant 1}$ be the sequence of primes in increasing order ($p_1 = 2, p_2 = 3, p_3 = 5, \cdots$) there exists an absolute constant C such that*

$$p_{n+1} - p_n \leqslant C$$

for infinitely many n.

Besides Zhang's original paper, we refer to [Gra15, Kow15] for a detailed description of Zhang's proof and the methods involved and historical background. Let us however mention a few important facts

- The question of the existence of small gaps between primes has occupied analytic number theorists for a while and has been the motivations for the invention of many techniques in analytic number theory in particular the *sieve method* to detect primes with additional constraints. A conceptual breakthrough occurred with the work of Goldston, Yildirim and Pintz [GPY09] who proved the weaker result

$$\liminf_n \frac{p_{n+1} - p_n}{\log p_n} = 0$$

  and who on this occasion provided a technique which is be key to Zhang's approach (see Soundararajan's account of their works [Sou07b].)
- Zhang's theorem can be seen as an approximation to the twin prime conjecture:

  There exists infinitely many primes p such that $p + 2$ is prime.

  Indeed Zhang's theorem with $C = 2$ is equivalent to the twin prime conjecture.
- A value for the constant C can be explicited : Zhang himself gave

$$C = 70.10^6$$

  and mentioning this could certainly be improved. Improving the value of this constant was the objective of the Polymath8 project: following and optimizing Zhang's method on several aspect (some to be explained below) the value was reduced to

$$C = 4680.$$

  However James Maynard [May16] made independently another conceptual breakthrough, simplifying the whole proof, making it possible to obtain stronger result and improving the constant to

$$C = 600.$$

  Eventually the Polymath8 project jointed with Maynard and optimizing Maynard argument reached

$$C = 246.$$

  A side effect of Maynard result is that what we are going to describe plays no role anymore in this specific application. Nevertheless it adresses another important question in analytic number theory.

15.1. **The Bombieri-Vinogradov theorem and beyond.** The breakthrough of Goldston, Yildirim and Pintz itself at the origin of Zhang's works builds on the use of sieve methods to detect the existence of infinitely many pairs of primes at distance $\leqslant C$ from one another. The fuel to be put in this sieve machine are results concerning the distribution of primes in arithmetic progressions of moduli large with respect to the size of the primes which are sought after: given $x \geqslant 2$, $q \geqslant 1$ and integer and $a \in \mathbf{Z}$ coprime to q, let

$$\psi(x; q, a) := \sum_{\substack{n \leqslant x \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n), \ \psi(x; q) := \sum_{\substack{n \leqslant x \\ (n,q)=1}} \Lambda(n) \sim x$$

where $\Lambda(n)$ is the von Mangolt function: one seek the equivalence

$$\psi(x; q, a) \sim \frac{\psi(x; q)}{\varphi(q)}, \ x \to \infty \tag{15.1}$$

for q as large as possible compared to x. The very first result in that direction is of course Dirichlet's theorem (on which occasion the concept of L-function was invented) and the prime number Theorem in arithmetic progressions shows (15.1) as long as $q \leqslant (\log x)^{O(1)}$. The *Generalized Riemann Hypothesis (GRH)* would provide the much stronger uniformity $q \ll_\varepsilon x^{1/2-\varepsilon}$, for any fixed $\varepsilon > 0$ but this is highly conditional. Fortunately the Bombieri-Vinogradov theorem provides an unconditional substitute to GRH, on average over q, which is essentially as strong for most sieve purposes:

Theorem 15.2 (Bombieri-Vinogradov)**.** *For any $A > 0$ there is $B = B(A) > 0$ such that for $x \geqslant 2$*

$$\sum_{q \leqslant x^{1/2}/\log^B x} \max_{(a,q)=1} |\psi(x;q,a) - \frac{\psi(x;q)}{\varphi(q)}| \ll \frac{x}{\log^A x}.$$

The exponent $1/2$ in the constraint $q \leqslant x^{1/2}/\log^B x$ turns out to be crucial in Zhang's approach to the existence of small gaps: Goldston-Yildirim-Pintz had already pointed out that this statement with $1/2$ replaced by any strictly larger exponent would be sufficient to show the existence of infinitely many bounded gaps between primes. This is not unexpected as the Elliott-Halberstam conjecture predicts that any fixed exponent < 1 could replace $1/2$.

That this is not hopeful thinking comes from the work of Fouvry, Iwaniec and Bombieri-Friedlander-iwaniec from the late 80's [FI92, ?BFI] who proved analogs of the Bombieri-Vinogradov theorem with exponents $> 1/2$ but unfortunately for "fixed congruences" classes (for instance with the sum involving the difference $|\phi(x;q,1) - \frac{\psi(x;q)}{\varphi(q)}|$ instead of the max $\max_{(a,q)=1} |\psi(x;q,a) - \frac{\psi(x;q)}{\varphi(q)}|$.) Zhang's groundbreaking insight has been to nailed down a post-Bombieri-Vinogradov type theorem that could be established unconditionally and would be sufficient to establish the existence of bounded gaps between primes. The following theorem is a variant of Zhang's beyond-Bombieri-Vinogradov theorem ([Pol14, Thm 1.1]). Let us recall that an integer $q \geqslant 1$ is Δ-smooth if any prime p dividing it is $\leqslant \Delta$.

Theorem 15.3. *Let $(\boldsymbol{a}) = (a_p)_{p\in\mathcal{P}}$ be a sequence of integers indexed by the primes such that for any $p \in \mathcal{P}$ $(a_p, p) = 1$; for any squarefree integer $q = \prod_{p|q} p$ let $a_q \pmod q$ be the unique congruence class modulo q such that*

$$\forall p|q,\ a_q \equiv a_p \pmod p;$$

in particular $a_q \in (\mathbf{Z}_q\mathbf{Z})^\times$. There exists absolute constants $\theta > 1/2$ and $\delta > 0$ (independent of $\boldsymbol{a}$) such that for any $A > 0$, $x > 2$ one has

$$\sum_{\substack{q\leqslant x^\theta,\ sqfree\\ q\ x^\delta-smooth}} |\psi(x;q,a_q) - \frac{\psi(x;q)}{\varphi(q)}| \ll \frac{x}{\log^A x}.$$

Here the implicit constant depends only on A, k (but not on $(\boldsymbol{a})$).

Remark. Zhang essentially proved this theorem for $\theta = 1/2 + 1/585$ and in an effort to improve Zhang's constant, the Polymath8 project improved $1/585$ to $7/301$.

We will now describe some of the principles of the proof of this theorem and especially at the points algebraic exponential sums occur. We refer to the introduction of [Pol14] and to E. Kowalski's account in the Bourbaki seminar [Kow15].

Let us write $c(q)$ for $\mu^2(q)$ times the sign of the difference $\psi(x;q,a_q) - \frac{\psi(x;q)}{\varphi(q)}$. The above sum equals

$$\sum_{\substack{q\leqslant x^\theta\\ q\ x^\delta-\text{smooth}}} c(q) \sum_{n\leqslant x} \Lambda(n)\Delta_{\boldsymbol{a}}(n;q).$$

where

$$\Delta_{\boldsymbol{a}}(n) := \delta_{n\equiv a_q \,(\mathrm{mod}\, q)} - \frac{\delta_{(n,q)=1}}{\varphi(q)}$$

As is usual when counting primes numbers, the next step is to decompose the vom Mangolt function $\Lambda(n)$ into a sum of convolution of arithmetic functions (for instance by using Heath-Brown's identity Lemma 8.2 as in §8): we essentially arrive at the problem of bounding $(\log x)^{O_J(1)}$ of the following model sums (for $j \leqslant J$ and J is a fixed and large integer)

$$\Sigma(\mathbf{M};\boldsymbol{a},Q) := \sum_{\substack{q\sim Q\\ q\ x^\delta-\text{smooth}}} c(q) \sum_{m_1,\cdots m_{2j}} \mu(m_1)\cdots\mu(m_j) V_1\Big(\frac{m_1}{M_1}\Big)\cdots V_{2j}\Big(\frac{m_{2j}}{M_{2j}}\Big)\Delta_{a_q}(m_1.\cdots.m_{2j})$$

where $V_i,\ i=1,\cdots 2j$ are smooth functions compactly supported in $]1,2[$, $\mathbf{M}=(M_1,\cdots,M_{2j})$ is a tuple satisfying

$$Q\leqslant x^\theta,\ M_i =: x^{\mu_i},\ \forall i\leqslant j,\ \mu_i\leqslant 1/J,\ \sum_{i\leqslant 2j}\mu_i = 1+o(1).$$

Our target is the bound

$$\Sigma(\mathbf{M};\boldsymbol{a},Q) \overset{?}{\ll} \frac{x}{\log^A x}. \tag{15.2}$$

The most important case is when

$$Q = x^\theta = x^{1/2+\varpi}$$

for some fixed sufficiently small $\varpi>0$.

The variables with index $i\in\{j+1,2j\}$ are called *smooth* because they are weigthed by smooth functions and this makes it possible to use Poisson summation formula on them to analyse the congruence condition mod q. This is going to be efficient if the range M_i is sufficiently big relative to $q\sim Q$. The variables with indices $i\in\{1,j\}$ are weighted by the Moebius function but (at least as long as some strong form of Generalized Riemann Hypothesis is not available) we cannot exploit this information and we will consider these like arbitrary bounded functions. The tradeoff to non-smoothness is that the range of these variables is pretty short $M_i\leqslant x^{1/J}$ especially if J is choosen large.

As we did before we will regroup some the variables $m_i,\ i=1,\cdots,2j$ so as to form two new variables whose ranges are located adequately (similarly to what we did in §8) and will use different methods to bound the sums depending on the size and the type of these new variables.

More precisely we define

$$\alpha_i(m) = \begin{cases} \mu(m)V_i\Big(\frac{m}{M_i}\Big) & 1\leqslant i\leqslant j\\ V_i\Big(\frac{m}{M_i}\Big) & j+1\leqslant i\leqslant 2j\end{cases}$$

Given some partition of the set of m-indices

$$\{1,\cdots,2j\} = \boldsymbol{I}\sqcup\boldsymbol{J}$$

let

$$M=\prod_{i\in\boldsymbol{I}} M_i,\ N=\prod_{j\in\boldsymbol{J}} M_j$$

and

$$\mu_{\boldsymbol{I}} := \sum_{i\in\boldsymbol{I}}\mu_i,\ \mu_{\boldsymbol{J}} := \sum_{i\in\boldsymbol{J}}\mu_i;$$

we have

$$\mu_{\boldsymbol{I}}+\mu_{\boldsymbol{J}} = 1+o(1),\ M=x^{\mu_{\boldsymbol{I}}},\ N=x^{\mu_{\boldsymbol{J}}};$$

in the sequel we will always make the convention that $N\leqslant M$ or equivalently $\mu_{\boldsymbol{I}}\geqslant\mu_{\boldsymbol{J}}$.

Finally we define the Dirichlet convolution functions

$$\alpha(m) := \star_{i\in \boldsymbol{I}}\alpha_i(m),\ \beta(n) := \star_{i\in \boldsymbol{J}}\alpha_i(n).$$

We are reduced to bound sums of the shape

$$\sum_{\substack{q\sim Q\\ x^\delta-\text{smooth}}} c(q)\sum_{\substack{m\sim M\\ n\sim N}}\sum \alpha(m)\beta(n)\Delta_{a_q}(mn) \overset{?}{\ll} \frac{x}{\log^A x}. \tag{15.3}$$

Observe that the functions α, β are essentially bounded

$$\forall \varepsilon > 0, \alpha(m), \beta(n) \ll x^\varepsilon$$

so we need only to improve slightly over the trivial bound.

15.2. **Splitting into types.** These sums will subdivided into three different types and their treatment will depend on which type the sum belong.

This subdivision is along the following simple combinatorial Lemma (cf. [Pol14, Lem. 3.1]):

Lemma 15.1. *Let $1/10 < \sigma < 1/2$ and let $\mu_i,\ i = 1, \cdots 2j$ be some non-negative real numbers such that*

$$\sum_{i=1}^{2j}\mu_i = 1.$$

One of the following holds

- *Type 0: there exist i such that $\mu_i \geqslant 1/2 + \sigma$.*
- *Type II: there exists a partition*

$$\{1, \cdots, 2j\} = \boldsymbol{I} \sqcup \boldsymbol{J}$$

such that

$$1/2 - \sigma \leqslant \sum_{i\in \boldsymbol{J}}\mu_i \leqslant \sum_{i\in \boldsymbol{I}}\mu_i < 1/2 + \sigma.$$

- *Type III: there exist distincts i_1, i_2, i_3 such that*

$$2\sigma \leqslant \mu_{i_1} \leqslant \mu_{i_2} \leqslant \mu_{i_3} \leqslant 1/2 - \sigma \text{ and } \mu_{i_1} + \mu_{i_2} \geqslant 1/2 + \sigma.$$

Remark 15.2. If $\sigma > 1/6$ the Type III situation never occurs since $2\sigma > 1/2 - \sigma$.

Given σ such that

$$1/10 < \sigma < 1/2$$

we assume that J is choosen large enough such that

$$1/J \leqslant \min(1/2 - \sigma, \sigma). \tag{15.4}$$

We call a sum as above of

- Type 0, if there exists some i_0 such that $\mu_{i_0} \geqslant 1/2 + \sigma$. We choose

$$\boldsymbol{I} = \{i_0\} \text{ and } \boldsymbol{J} \text{ the complement.}$$

Since for any $i \leqslant j$, one has $\mu_i \leqslant 1/J < 1/2 + \sigma$, necessarily $i_0 \geqslant j + 1$ corresponds to a smooth variable; the corresponding sum therefore equals

$$\sum_{\substack{q\sim Q\\ x^\delta-\text{smooth}}} c(q)\sum_{m\geqslant 1, n\sim N} V(\frac{m}{M_{i_0}})\beta(n)\Delta_{a_q}(mn). \tag{15.5}$$

– Type I/II if one can partition the set of indices

$$\{1, \cdots, 2j\} = \boldsymbol{I} \sqcup \boldsymbol{J}$$

in a way that the corresponding ranges

$$M = \prod_{i \in \boldsymbol{I}} M_i = x^{\mu_{\boldsymbol{I}}} \geqslant N = \prod_{i \in \boldsymbol{J}} M_i = x^{\mu_{\boldsymbol{J}}}$$

satisfy

$$1/2 - \sigma \leqslant \mu_{\boldsymbol{J}} = \sum_{i \in \boldsymbol{J}} \mu_i \leqslant 1/2 \tag{15.6}$$

– Type III if we are neither in the Type 0 or Type I/II situation: there exists distincts indices i_1, i_2, i_3 such that

$$2\sigma \leqslant \mu_{i_1} \leqslant \mu_{i_2} \leqslant \mu_{i_3} \leqslant 1/2 - \sigma \text{ and } \mu_{i_1} + \mu_{i_2} \geqslant 1/2 + \sigma.$$

We choose

$$\boldsymbol{I} = \{i_1, i_2, i_3\} \text{ and } \boldsymbol{J} \text{ to be the complement.}$$

Again, since $1/J < 2\sigma$ by (15.4), the indices i_1, i_2, i_3 are associated to smooth variables and the Type III sums are of the shape

$$\sum_{\substack{q \sim Q \\ x^\delta\text{–smooth}}} c(q) \sum_{\substack{m_1, m_2, m_3 \\ n \sim N}} \sum V(\frac{m_1}{M_{i_1}}) V(\frac{m_2}{M_{i_2}}) V(\frac{m_3}{M_{i_3}}) \beta(n) \Delta_{a_q}(m_1 m_2 m_3 n).$$

Remark. In the paper [Pol14] the "Type II" sums introduced here were split into two further Types called "Type I" and "Type II" there. These are the sums for which the N variable satisfies

$$\text{Type I: } x^{1/2-\sigma} \leqslant N < x^{1/2-\varpi-c}$$

$$\text{Type II: } x^{1/2-\varpi-c} \leqslant N \leqslant x^{1/2}$$

for c some extra parameter satisfying

$$1/2 - \sigma < 1/2 - \varpi - c < 1/2.$$

This distinction was necessary for optimisation purposes and especially to achieve the exponent $1/2 + 7/301$ in Theorem 15.3.

Zhang's theorem now essentially follows from

Theorem 15.4. *There exists $\varpi, \sigma > 0$ with $1/10 < \sigma < 1/2$ such that the bound* (15.3) *holds for the Type 0, II and III sums.*

For the rest of this section we will succinctly describe how each type of sum is handled.

The case of Type 0 sums (15.5) is immediate: one apply the Poisson summation formula to the m variable to decompose the congruence $mn \equiv a_q \pmod q$. The zero frequency contribution is cancelled up to an error terms by the second term of $\Delta_{a_q}(mn)$ while the non-zero frequencies contribution a negligible error term as long as the range of the m variable is larger than the modulus

$$1/2 + \sigma > 1/2 + \varpi$$

which can be assumed.

15.3. **Treatment of type II sums.**

15.3.1. *The art of applying Cauchy-Schwarz.* The Type II sums are more complicated to deal with because we have essentially no control on the shape of the coefficients $\alpha(m), \beta(n)$ (excepted for being essentially bounded). The basic principle is to consider the largest variable $m \sim M$, to make it smooth smooth using the Cauchy-Schwarz inequality and then resolve the congruence

$$m \equiv \overline{n} a_q \,(\mathrm{mod}\, q)$$

using the Poisson summation formula. This is the essence of the *dispersion method* of Linnik.

When implementing this stratgy one has to decide which variables to put "inside" the Cauchy-Schwarz inequality and which to leave "outside": to be more specific, suppose we need to bound a general trilinear sum (K some function)

$$\sum_{m\sim M, n\sim N, q\sim Q} \alpha_m \beta_n \gamma_q K(m,n,q)$$

and wish to smooth the m variable using Cauchy-Schwarz. There are two possibilities, either

$$\sum_{m\sim M, n\sim N, q\sim Q} \alpha_m \beta_n \gamma_q K(m,n,q) \ll \|\alpha\|_2 \|\gamma\|_2 (\sum_{n,q} |\sum_{n\sim N} \beta_n K(m,n,q)|^2)^{1/2}$$

or

$$\sum_{m\sim M, n\sim N, q\sim Q} \alpha_m \beta_n \gamma_q K(m,n,q) \ll \|\alpha\|_2 (\sum_{n} |\sum_{n\sim N, q\sim Q} \beta_n \gamma_q K(m,n,q)|^2)^{1/2}$$

In the first case the inner sum of the second factor equals

$$\sum_{n_1,n_2\sim N} \beta_{n_1} \overline{\beta_{n_2}} \sum_{m\sim M, q\sim Q} K(m,n_1,q) \overline{K(m,n_2,q)}$$

and in the second case

$$\sum_{n_1,n_2\sim N, q_1,q_2\sim Q} \beta_{n_1} \gamma_{q_1} \overline{\beta_{n_2} \gamma_{q_2}} \sum_{m\sim M} K(m,n_1,q_1) \overline{K(m,n_2,q_2)}|^2.$$

In either case, one expect to be able to detect cancellation from the m-sum, at least when the other variables (n_1, n_2) or (n_1, n_2, q_1, q_2) are not located on the diagonal (ie. $n_1 = n_2$ or $n_1 = n_2,\ q_1 = q_2$). If the other variables are the diagonal, no cancellation is possible but the diagonal is small compared to the space of variables.

We are faced with the following tradeoff:

- For the first possibility, the m-sum is simpler (it involves two parameters n_1, n_2) but the ratio "size of the diagonal"/"size of the set of parameters" is $N/N^2 = N^{-1}$.
- For the second possibility, the m-sum is more complicated as it involves more auxiliary parameters n_1, n_2, q_1, q_2 but the ratio "size of the diagonal"/"size of the set of parameters" $NQ/N^2Q^2 = 1/NQ$ is smaller (hence more saving can be obtained from the diagonal part.)

15.3.2. *The Type II sums.* We illustrate this discussion in the base of Type II sums. If we apply Cauchy with the q variable outside the diagonal $n_1 = n_2$ would not provide enough saving. If onthe other hand we apply Cauchy with q inside the diagonale is large but we have to analysize the congruence

$$mn_1 \equiv a \,(\mathrm{mod}\, q)_1,\ mn_2 \equiv a \,(\mathrm{mod}\, q)_2$$

which is a congruence modulo $[q_1, q_2]$. Assuming weare in the generic case of q_1, q_2 coprime the resulting modulus is $q_1 q_2 \sim Q^2 = x^{1+2\varpi}$ while $m \sim M \leqslant x^{1/2}$ which is too small for the Poisson formula to be efficient.

They is fortunately a middleground: we can use the extra flexibility (due to Zhang's wonderful insight) that our problem involves *smooth* moduli: by the greedy algorithm, one can factor $q \sim Q$ into a product $q = rs$ where r and $s \sim Q/r$ vary over ranges that we can essentially choose as we

which (up to a small indeterminacy of x^δ for δ small). In over terms we are reduced to bound sums of the shape

$$\Sigma(M,N;\boldsymbol{a},R,S)=\sum_{\substack{r\sim R,\ s\sim S\\ rs\ x^\delta\text{–smooth}}}c(rs)\sum_{\substack{m\sim M\\ n\sim N}}\sum\alpha(m)\beta(n)\Delta_{a_{rs}}(mn)$$

for any factorisation $RS=Q$ fitting with our needs; but now, when applying Cauchy-Schwarz, we have the extra flexibility of having the r variable "out" and the s variable "in".

We apply the Cauchy-Schwarz getting

$$\sum_s c(rs)\sum_{\substack{m\sim M\\ n\sim N}}\sum\alpha(m)\beta(n)\Delta_{a_{rs}}(mn)=\sum_{r\sim r}\sum_{m\sim M}\alpha(m)\sum_s c(rs)\sum_{n\sim N}\beta(n)\Delta_{a_{rs}}(mn)$$
$$\ll_\varepsilon R^{1/2}M^{1/2+\varepsilon}(\sum_r\sum_{s_1,s_2,n_1,n_2}\sum\sum c(rs_1)\overline{c(rs_2)}\beta(n_1)\overline{\beta(n_2)}\sum_m V(\frac{m}{M})\Delta_{a_{rs_1}}(mn_1)\Delta_{a_{rs_2}}(mn_2))^{1/2}$$

for V a smooth function compactly supported in $[M/4,4M]$. We choose R of the shape

$$R=Nx^{-\varepsilon}\leqslant Mx^{-\varepsilon}$$

for $\varepsilon>0$ but small.

Expanding the square we reach a sum involving four terms the most important one coming from the product

$$\Delta_{a_{rs_1}}(mn_1)\Delta_{a_{rs_2}}(mn_2)=(\delta_{mn_1\equiv a_{rs_1}\,(\mathrm{mod}\,rs_1)}-\frac{\delta_{(n,rs_1)=1}}{\varphi(rs_1)})(\delta_{mn_2\equiv a_{rs_2}\,(\mathrm{mod}\,rs_2)}-\frac{\delta_{(n,rs_2)=1}}{\varphi(rs_2)})\tag{15.7}$$

We will concentrate on the contribution of these terms from now on.

The generic and main case is when $(s_1,s_2)=1$ so that m satisfies a congruence modulo $rs_1s_2\sim RS^2=Mx^{2\varpi+\varepsilon}$ which is not much larger than M if ϖ is small. Observe that

$$mn_i\equiv a_{rs_i}\,(\mathrm{mod}\,rs_i),\ i=1,2\Longrightarrow\ n_1\equiv n_2\,(\mathrm{mod}\,r).$$

We can therefore write $n_1=n,\ n_2=n+rl$ with $|l|\ll N/R=x^\varepsilon$. By the Poisson summation formula we have

$$\sum_m V(\frac{m}{M})\delta_{m\equiv b\,(\mathrm{mod}\,rs_1s_2)}=\frac{M}{rs_1s_2}\widehat{V}(0)+\frac{M}{rs_1s_2}\sum_{h\neq0}\widehat{V}(\frac{h}{rs_1s_2/M})e(\frac{hb}{rs_1s_2})$$

where $b=b(n,l)\,(\mathrm{mod}\,rs_1s_2)$ is such that

$$b\equiv a_{rs_1s_2}\overline{n}\,(\mathrm{mod}\,r),\ b\equiv a_{rs_1s_2}\overline{n}\,(\mathrm{mod}\,s_1),b\equiv a_{rs_1s_2}\overline{n+lr}\,(\mathrm{mod}\,s_2).$$

The $h=0$ contribution provide a main term is cancelled up to an admissible error term by the main contributions coming from the other summands of (15.7). The contribution of the frequencies $h\neq0$ will be prove to be error terms:

$$\sum_r\sum_{s_1,s_2,n,l}\sum\sum c(rs_1)\overline{c(rs_2)}\beta(n)\overline{\beta(n+rl)}\frac{M}{rs_1s_2}\sum_{h\neq0}\widehat{V}(\frac{h}{rs_1s_2/M})e(\frac{hb}{rs_1s_2})\overset{?}{\ll}\frac{MN^2}{R}x^{-\eta}=x^{1-\eta+\varepsilon}$$

for some fixed $\eta>0$. The length of the h sum is essentially

$$H=RS^2/M=Q^2N/(xR)=x^{2\varpi+\varepsilon}$$

which is small (if ϖ and ε are). We essentially need to prove that

$$\frac{1}{H}\sum_{r\sim R}\sum_{l\ll N/R}\sum_{n}\beta(n)\overline{\beta(n+lr)}\sum_{0\neq h\ll H}|\sum_{s_1,s_2}c(rs_1)\overline{c(rs_2)}e(h\frac{a_{rs_1s_2}\overline{n}}{rs_1}+h\frac{a_{rs_1s_2}\overline{n+lr}}{rs_2})| \overset{?}{\ll} x^{1-\eta+\varepsilon}. \tag{15.8}$$

We can now exhibit cancellation in the n-sum by smoothing out the n variable using the Cauchy-Schwarz inequality for any fixed r, l: letting the h variable "in" we obtain exponential sums of the shape

$$\sum_{n\sim N}e(h\frac{a_{rs_1s_2}\overline{n}}{rs_1}-h'\frac{a_{rs_1's_2'}\overline{n}}{rs_1'}+h\frac{a_{rs_1s_2}\overline{n+lr}}{rs_2}-h'\frac{a_{rs_1's_2'}\overline{n+lr}}{rs_2'}).$$

The generic case is when $h-h', s_1, s_2, s_1', s_2'$ are all coprime. In that case the above exponential sum has length

$$N\in[x^{1/2-\sigma}, x^{1/2}]$$

and the involved moduli are of size

$$RS^4=Q^4/R^3=x^{O(\varepsilon)}Q^4/N^3=[x^{1/2+4\varpi+O(\varepsilon)}, x^{1/2++4\varpi+3\sigma+O(\varepsilon)}].$$

Therefore if $\sigma, \varpi, \varepsilon$ are small, the length N is not much smaller than the modulus so we could apply completion methodsto improve over the trivial bound $O(N)$ for the n-sum. If we apply the Polya-Vinogradov method the trivial bound is replaced by $O((RS^4)^{1/2+o(1)})$ and we find that we obtain that the lefthand side of (15.8) is bounded by

$$\frac{1}{H}R.\frac{N}{R}N^{1/2}(H^2S^4(RS^4)^{1/2+o(1)})^{1/2}=x^{O(\varepsilon)+o(1)}N^{3/2}S^3R^{1/4}=x^{\frac{7}{8}+3\varpi+\frac{5}{4}\sigma+O(\varepsilon)+o(1)}$$

which is $\ll x^{1-\eta}$ for some $\eta>0$ whenever $\sigma<1/10$ and ϖ and ε are small enough.

Instead of using the Polya-Vinogradov bound we could take advantage of the fact that the modulus $rs_1s_1's_2s_2'$ is x^δ-smooth (again we can take $\delta>0$ as small as we need) and applythe q-van der Corput method from the previous section. Factoring $rs_1s_1's_2s_2'$ into a product $r's'$ such that $r'\sim(rs_1s_1's_2s_2')^{1/3+O(\delta)}$, $s'\sim(rs_1s_1's_2s_2')^{2/3+O(\delta)}$ a suitable variant of (14.1) bounds the n-sum by $O(N^{1/2}(RS^4)^{1/6+O(\delta)+o(1)})$ and the lefthand side of (15.8) is bounded by

$$\frac{R}{H}\frac{N}{R}N^{\frac{1}{2}}(H^2S^4N^{1/2}(RS^4)^{1/6})^{\frac{1}{2}+o(1)+O(\delta)}=x^{O(\varepsilon+\delta)+o(1)}N^{7/4}S^{7/3}R^{1/12}=x^{\frac{11}{12}+\frac{7}{3}\varpi+\frac{1}{2}\sigma+O(\varepsilon+\delta)+o(1)}$$

which is $\ll x^{1-\eta}$ for some $\eta>0$ whenever $\sigma<1/6$ and ϖ and ε are small enough.

15.4. **Treatment of type III sums.** Our objective for the Type III sums is the following bound: for some $\eta>0$

$$\sum_{\substack{q\sim Q\\ x^\delta-\text{smooth}}}c(q)\sum_{n\sim N}\beta(n)\sum_{m}\tau_{3,\mathbf{M}}(m)\Delta_{a_q}(m_1m_2m_3n)\overset{?}{\ll}x^{1-\eta}; \tag{15.9}$$

here $\mathbf{M}=(M_{i_1}, M_{i_2}, M_{i_3})$ and

$$\tau_{3,\mathbf{M}}(m):=\sum_{m_1m_2m_3=m}V(\frac{m_1}{M_{i_1}})V(\frac{m_2}{M_{i_2}})V(\frac{m_3}{M_{i_3}})$$

and $M_{i_1}, M_{i_2}, M_{i_3}$ satisfy

$$M=M_{i_1}M_{i_2}M_{i_3}\geqslant x^{1/2+3\sigma}.$$

The function

$$m\mapsto\tau_{3,\mathbf{M}}(m)$$

is basically a smoothed version of the ternary divisor function $m \mapsto \tau_3(m)$ we have discussed in §11.

In fact, while describing the proof of Thm. 11.3, we have shown that for $M = x$, and for q a prime satisfying

$$q \sim x^{1/2+\varpi},\ \varpi = 1/47$$

one has

$$\sum_m \tau_{3,\mathbf{M}}(m)\Delta_{a_q}(m_1m_2m_3n) \ll \frac{x^{1-\eta}}{q}$$

for some $\eta > 0$. We have therefore the required bound but for individual moduli instead of having it on average.

As we have observed when discussing Type II sums, the parameter σ can be taken as close to $1/6$ as we wish and in particular $M \in [x^{1+3(\sigma-\frac{1}{6})}, x]$ can be made as close as we wish from x and $N \in [1, x^{3(\frac{1}{6}-\sigma)}]$ as we wish from x (in the logarithmic scale) and in particular this establishes (15.9) for prime moduli $q \sim Q$ for some value of σ (close enough to $1/6$), some value of ϖ (close enough to 0) and some $\eta > 0$.

The case of x^δ-smooth moduli uses similar methods and (besides some elementary technical issues) is maybe simpler than in the prime modulus case because of the extra flexibility provided by the smooth moduli.

Remark. By a more elaborate treatment, involving different uses of the Cauchy-Schwarz inequality and iterations of the q-van der Corput method, it is possible to bounds succesfuly all the Type II sums associated to some explicit parameter $\sigma > 1/6$. As pointed out in Remark 15.2, this makes the section devoted to Type III sums and in particular the theory of hyper-Kloosterman sums $\mathrm{Kl}_3(x; q)$ unnecessary. The interest of this remark comes from the fact that the trace functions occurring in the treatment the sums of Type II are exclusively algebraic exponentials:

$$x \mapsto \mathrm{e}_q(f(x)),\ \text{for } f(X) \in \mathbf{F}_q(X)$$

and for such trace function Corollary 4.2 "only" uses Weil's resolution of the Riemann Hypothesis for curves over finite fields [Wei41] and not the full proof of the Weil conjectures by Deligne [Del80].

16. Advanced completions methods: the $+ab$ shift

In this last section we describe another method allowing to break the Polya-Vinogradov barrier for prime moduli. this method has its origins in the celabrated work of Burgess on short sums of Dirichlet characters [Bur62]

16.1. **Burgess's bound.** Let q be a prime and $\chi : \mathbf{F}_q^\times \to \mathbf{C}^\times$ be a non trivial multiplicative character;

$$S_V(\chi, N) := \sum_n \chi(n)V(\frac{n}{N})$$

where $V \in \mathcal{C}^\infty(]1, 2[)$.

Theorem 16.1 (Burgess)**.** *For any $N \geqslant 1$, $l \geqslant 1$ such that*

$$q^{1/2l} \leqslant N < \frac{1}{2}q^{1/2+1/4l} \tag{16.1}$$

one has

$$S_V(\chi, N) \ll_{V,l} q^{o(1)}N(N/q^{1/4+1/4l})^{-1/l}.$$

Remark. Observe that this bound is non trivial (sharper than $S_V(\chi, N) \ll N$) whenever

$$q^{1/4+1/4l+o(1)} \leqslant N < \frac{1}{2} q^{1/2+1/4l}.$$

Moreover for $N \geqslant \frac{1}{2} q^{1/2+1/4l}$, the Polya-Vinogradov bound $S_V(\chi, N) \ll q^{1/2}$ is non trivial, therefore, we see taking l large enough that (16.1) yield a non-trivial bound for $S_V(\chi, N)$ as long as

$$N \geqslant q^{1/4+\delta}$$

for any fixed $\delta > 0$.

Proof. Burgess's argument exploit two features in a critical way: the first one is that an interval is "essentially" invariant under sufficiently small additive translations; the second is the multiplicativity of the Dirichlet character.

Let $A, B \geqslant 1$ be parameters such that $AB \leqslant N/2$; we will also assume that $2B < q$.

We have $(a \sim A \Longleftrightarrow a \in [A, 2A[)$

$$S_V(\chi, N) = \frac{1}{AB} \sum_{|n| \leqslant 2N} \sum_{a \sim A, b \sim B} \chi(n + ab) V(\frac{n + ab}{N}).$$

The next step is to invoque the Fourier inversion formula to make the variable n and ab independent: one has

$$V(\frac{n + ab}{N}) = \int_{\mathbf{R}} \widehat{V}(t) e(\frac{tn}{N}) e(\frac{tab}{N}) dt.$$

Plugging this formula in our sum, we obtain

$$\begin{aligned} S_V(\chi, N) &= \frac{1}{AB} \int_{\mathbf{R}} \sum_{|n| \leqslant 2N} e(\frac{tn}{N}) \sum_{a \sim A, b \sim B} \chi(n + ab) e(\frac{tab}{N}) \widehat{V}(t) dt \\ &\leqslant \frac{1}{AB} \int_{\mathbf{R}} \sum_{|n| \leqslant 2N} \sum_{a \sim A} \Big| \frac{\chi(a)}{a} \widehat{V}(\frac{t}{a}) \Big| \Big| \sum_{b \sim B} \chi(\overline{a}n + b) e(\frac{tb}{N}) \Big| dt \\ &\leqslant \frac{1}{AB} \int_{\mathbf{R}} \sum_{|n| \leqslant 2N} \sum_{a \sim A} \Big| \sum_{b \sim B} \chi(\overline{a}n + b) e(\frac{tAb}{N}) \Big| |W(t)| dt \end{aligned}$$

for W some bounded rapidly decaying function.

Remark. Observe that the factor $\chi(a)$ coming from the identity

$$\chi(n + ab) = \chi(a(\overline{a}n + b)) = \chi(a)\chi(\overline{a}n + b) \tag{16.2}$$

has been absorbed in the absolute value of the first inequality above.

The innermost sum can be rewritten

$$\sum_{|n| \leqslant 2N} \sum_{a \sim A} \Big| \sum_{b \sim B} \chi(\overline{a}n + b) e(\frac{tAb}{N}) \Big| = \sum_{r \in \mathbf{F}_q^\times} \nu(x) \Big| \sum_{b \sim B} \eta_b \chi(r + b) \Big|$$

where $\eta_b = e(\frac{tAb}{N})$ and

$$\nu(r) := |\{\overline{a}n = r \,(\mathrm{mod}\, q),\ a \in [A, 2A[,\ |n| \leqslant 2N\}|.$$

Consider the map

$$(a, n) \in [A, 2A[\times [-2N, 2N] \to \overline{a}n \,(\mathrm{mod}\, q) = r \in \mathbf{F}_q.$$

the function $\nu(r)$ is the size of the fiber of that map above r. We will show that this map is essentially injective (has small fibers on average): suppose that A is chosen such that $4AN < q$; one has

$$\sum_r \nu(r) \ll AN,\ \sum_r \nu^2(r) \ll (AN)^{1+o(1)}$$

the first bound is obvious while for the second

$$\sum_r \nu^2(r) = |\{(a,a',n,n'), an' \equiv an \,(\mathrm{mod}\, q)\ a,a' \in [A,2A[,\ |n|,|n'| \ll N\}|$$

uses the fact that $AN < q$ and that the integer an' has at most $(an')^{o(1)}$ decomposition of the shape $an' = a'n$.

This map however is not surjective nor even close so in general so that the change of variable $\overline{a}.n \leftrightarrow x$ is not very effective. A way to moderate ineffectiveness is to use Hölder inequality.

Let $l \geqslant 1$ be some integer parameter, applying Hölder inequality with $p = 1 - 1/2l,\ q = 1/2l$ and the above estimate one obtains

$$\begin{aligned}\sum_{x\in\mathbf{F}_q^\times} \nu(x)|\sum_{b\sim B}\eta_b\chi(x+b)\big| &\leqslant (\sum_x \nu(x)^{\frac{2l}{2l-1}})^{1-1/2l}(\sum_x|\sum_{b\sim B}\eta_b\chi(x+b)\big|^{2l})^{1/2l}\\ &\ll (AN)^{1-1/2l+o(1)}(\sum_x|\sum_{b\sim B}\eta_b\chi(x+b)\big|^{2l})^{1/2l}\end{aligned}$$

The x-sum in the rightmost factor equals

$$\sum_{\boldsymbol{b}} \varepsilon_{\boldsymbol{b}} \sum_{r\in\mathbf{F}_q} \chi(\frac{\prod_{i=1}^l (r+b_i)}{\prod_{i=i}^l (r+b_{k+i})})$$

where $\boldsymbol{b} = (b_1,\cdots,b_{2l}) \in [B,2B[^{2l}$. Consider the fraction

$$F_{\boldsymbol{b}}(X) := \frac{\prod_{i=1}^l (X+b_i)}{\prod_{i=i}^l (X+b_{k+i})} \in \mathbf{Q}(X)$$

and the function on $\mathbf{F}_q$

$$r \in \mathbf{F}_q \mapsto \chi(F_{\boldsymbol{b}}(r))$$

(extended by 0 for $r = -b_i \,(\mathrm{mod}\, q),\ i = 1,\cdots,2l$). This function is the trace function of the rank one sheaf $[F_{\boldsymbol{b}}]^*\mathcal{L}_\chi$ whose conductor is bounded in terms of l only and (because it is of rank 1) is geometrically irreducible if not-geometrically constant. If not geometrically constant one has[19]

$$\sum_{r\in\mathbf{F}_q} \chi(F_{\boldsymbol{b}}(r)) \ll_l q^{1/2}.$$

If $q > \max(l,2B)$ this occurs precisely when $F_{\boldsymbol{b}}(X)$ is not constant nor a k-th power where k is the order of χ: this occursfor $\boldsymbol{b}$ outside an explicit set $\mathcal{B}^{bad} \subset [B,2B[^{2l}$ of size bounded by $O(B^l)$. If $\boldsymbol{b} \in \mathcal{B}^{bad}$ one then use the trivial bound

$$|\sum_{r\in\mathbf{F}_q} \chi(F_{\boldsymbol{b}}(r))| \leqslant q.$$

All in all we eventually obtain

$$\sum_{\boldsymbol{b}} \varepsilon_{\boldsymbol{b}} \sum_x \chi(\frac{\prod_{i=1}^l (x+b_i)}{\prod_{i=i}^l (x+b_{k+i})}) \ll |\mathcal{B}^{bad}|q + |\mathcal{B} - \mathcal{B}^{bad}|q^{1/2} \ll B^l q + B^{2l}q^{1/2}.$$

[19] if is not necessary to invoque Deligne's main theorem here: this follows from A. Weil's proof of the Riemann hypothesis for curves [Wei41]

Choosing $B = q^{1/2l}$ (so as to equal the two terms in the bound above) and $A \approx Nq^{-1/2l}$ and such that $4AN < q$ (which is equivalent to (16.1)) we obtain that

$$S_V(\chi, N) \ll_l \frac{q^{o(1)}}{AB}(AN)^{1-1/2l}(q^{3/2})^{1/2l} \ll q^{o(1)}N^{1-1/l}q^{3/4l-(1-1/2l)/2l} = q^{o(1)}N(N/q^{1/4+1/4l})^{-1/l}$$

□

16.2. **The $+ab$-shift for type I sums.** It is natural to try to extend this method to other trace functions; unfortunately the above argument breaksdown because the identity (16.2) is not valid in general. It is however possible mitigate this problem by introducing an extra average in the above sum.

This technique goes back to Karacuba and Vinogradov (for the function $x \mapsto \chi(x+1)$); it was also used by Friedlander-Iwaniec [FI85] (for the function $x \mapsto e(\frac{\overline{x}}{q})$), Fouvry-Michel [FM98] and Kowalski-Michel-Sawin [KMS17, KMS18].

Instead of a single sum $S_V(K, N)$, one considers instead the following average of multiplicative shifts

$$B_V(K, \boldsymbol{\alpha}, N) := \sum_{m\sim M} \alpha_m \sum_n V(\frac{n}{N})K(mn)$$

where $1 \leqslant M < q$ and $(\alpha_m)_{m\sim M}$ is a sequence of complex numbers of modulus $\leqslant 1$ (this includes the averaged sum $\sum_{m\sim M}\big|\sum_n K(mn)V(\frac{n}{N})\big| = \sum_m |S_V([\times m]^*K, N)|$). The objective here is to improve over the trivial bound

$$B_V(K, \boldsymbol{\alpha}, N) \ll MN.$$

Proceeding as above we have

$$\begin{aligned} B_V(K, \boldsymbol{\alpha}, N) &= \frac{1}{AB}\sum_m \alpha_m \sum_n \sum_{a\sim A, b\sim B} K(m(n+ab))V(\frac{n+ab}{N}) \\ &\leqslant \frac{1}{AB}\int_{\mathbf{R}} \sum_{m\sim M} \alpha_m \sum_{|n|\leqslant 2N} \sum_{a\sim A}\Big|\sum_{b\sim B} K(am(\overline{a}n+b))e(\frac{tAb}{N})\Big||W(t)|dt \end{aligned}$$

We have

$$\sum_{m\sim M} \alpha_m \sum_{|n|\leqslant 2N} \sum_{a\sim A}\Big|\sum_{b\sim B} K(am(\overline{a}n+b))e(\frac{tAb}{N})\Big| = \sum_{\substack{r,s\in\mathbf{F}_q\\ \nu}}\sum (r,s)\Big|\sum_{b\sim B}\eta_b K(s(r+b))\Big|$$

with

$$\nu(r,s) = \sum_{m\sim M}\sum_{|n|\leqslant 2N}\sum_{a\sim A}\alpha_m \delta_{\overline{a}n=r, am=s \,(\mathrm{mod}\, q)}.$$

Assuming that $4AN < q$ and evaluating the number of solutions to the equations

$$am = a'm',\ a\overline{n} \equiv a'\overline{n}' \,(\mathrm{mod}\, q),\ (a,m,n) \in [A, 2A[\times[M, 2M[\times[N, 2N[$$

one finds that

$$\sum_{r,s\in\mathbf{F}_q}\sum |\nu(r,s)| \ll AMN,\ \sum_{r,s\in\mathbf{F}_q}\sum |\nu(r,s)|^2 \ll q^{o(1)}AMN$$

which we interpret as saying that the map

$$(a,m,n) \in [A, 2A[\times[M, 2M[\times[N, 2N[\to (r,s) = (\overline{a}.n, am) \in \mathbf{F}_q \times [AM, 4AM[$$

is essentially injective (ie. has small fibers on average). As above this map is far from being surjective but one can dampen this with Hölder inequality:

$$\sum_{\substack{r\in\mathbf{F}_q\\ 1\leqslant s\leqslant 4AM}}\sum \nu(r,s)\Big|\sum_{b\sim B}\eta_b K(s(r+b))\Big| \ll \Big(\sum_{r,s}\sum |\nu(r,s)|^{\frac{2l}{2l-1}}\Big)^{1-1/2l}\Big(\sum_{r,s}\sum\Big|\sum_{b\sim B}\eta_b K(s(r+b))\Big|^{2l}\Big)^{1/2l}$$

$$\ll q^{o(1)}(AMN)^{1-1/2l}\big(\sum_{\boldsymbol{b}}\eta_{\boldsymbol{b}}\sum_{r,s}\prod_{i=1}^{l}K(s(r+b_i))\overline{K(s(r+b_{i+l}))}\big)^{1/2l}.$$

We are now reduced to the problem of bounding the two variable sum

$$\sum_{r,s}\prod_{i=1}^{l}K(s(r+b_i))\overline{K(s(r+b_{i+l}))}=\sum_{r}\sum_{s}\mathbf{K}(sr,s\boldsymbol{b})=\sum_{r}\mathbf{R}(r,\boldsymbol{b}) \tag{16.3}$$

(say) where

$$\mathbf{K}(r,\boldsymbol{b}):=\prod_{i=1}^{l}K(r+b_i)\overline{K(r+b_{i+l})},\ \mathbf{R}(r,\boldsymbol{b})=\sum_{s}\mathbf{K}(sr,s\boldsymbol{b}). \tag{16.4}$$

The bound will depend on the vector $\boldsymbol{b}\in[B,2B[^{2l}$. To get a feeling of what is going on, let us consider one of the very special cases of [FM98]: let

$$K(x)=\mathrm{e}_q(\overline{x}+x).$$

We have

$$\mathbf{R}(sr,s\boldsymbol{b})=\sum_{s\in\mathbf{F}_q^\times}\mathrm{e}_q(\overline{s}(\sum_{i=1}^{l}\overline{r+b_i}-\overline{r+b_{i+l}})+s(\sum_{i=1}^{l}b_i-b_{i+l})).$$

This sum is either

- Equal to $q-1$ if and only if the vector $(b_1,\cdots,b_l)$ equals the vector $(b_{l+1},\cdots,b_{2l})$ up to permutation of the entries
- Equals to -1 if $\boldsymbol{b}$ is not as above but is in the hyperplane with equation $\sum_{i=1}^{l}b_i-b_{i+l}=0$,
- The Kloosterman sum

$$\mathbf{R}(r,\boldsymbol{b})=q^{1/2}\,\mathrm{Kl}_2(\frac{\sum_{i=1}^{l}\overline{r+b_i}-\overline{r+b_{i+l}}}{\sum_{i=1}^{l}b_i-b_{i+l}};q)$$

otherwise.

The last case is the most interesting: given $\boldsymbol{b}$ as in the last situation we have to evaluate

$$q^{1/2}\sum_{r}\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)$$

where $G_{\boldsymbol{b}}(X)$

$$G_{\boldsymbol{b}}(X)=\frac{\sum_{i=1}^{l}\overline{X+b_i}-\overline{X+b_{i+l}}}{\sum_{i=1}^{l}b_i-b_{i+l}}.$$

Lemma 16.1. *For $\boldsymbol{b}=(b_1,\cdots,b_{2l})\in{\mathbf{F}_q}^{2l}$ such that*

$$(b_1,\cdots,b_l)\text{ is not equal to }(b_{l+1},\cdots,b_{2l})\text{ up to permutation and }\sum_{i=1}^{l}b_i-b_{i+l}\neq 0, \tag{16.5}$$

one has

$$\sum_{r}\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)\ll_l q^{1/2}.$$

Proof. The function

$$r\mapsto\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)$$

is the trace function of the rank 2 sheaf $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ obtained by pull-back from the Kloosterman sheaf $\mathcal{K}\ell_2$ of the non constant (because of our assumptions) map on $\mathbf{P}^1$

$$x\to G_{\boldsymbol{b}}(x).$$

One can show that the conductor of $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ is bounded in terms of l only more over the geometric monodromy group of $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ is obtained as the (closure of the) image of the representation $\varrho_{\mathcal{K}\ell_2}$ restricted to a finite index subgroup of $\mathrm{Gal}(K^{\mathrm{sep}}/\overline{\mathbf{F}_q}.K)$. Since the geometric monodromy group of $\mathcal{K}\ell_2$ is SL_2 which has no finite index subgroup geometric monodromy group of $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ is SL_2 as well. It follow that the sheaf $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ is geometrically irreducible (and not geometrically trivial because of rank 2) and the estimate follows by Deligne's theorem. □

It follows from this analysis that

$$\sum_{r,s}\sum\Big|\sum_{b\sim B}\eta_b K(s(r+b))\Big|^{2l} \ll B^l q^2 + B^{2l} q$$

hence choosing $B = q^{1/l}$, $AB \approx N$ and $A \approx Nq^{-1/l}$ we obtain

$$B_V(K,\boldsymbol{\alpha},N) \ll \frac{q^{o(1)}}{AB}(AMN)^{1-1/2l}q^{3/2l} = q^{o(1)}MN(\frac{N^2M}{q^{1+1/l}})^{-1/2l}$$

To resume we have therefore proven the

Theorem 16.2. *Let $K(x) = \mathrm{e}_q(\overline{x}+x)$ and $M,N,l \geqslant 1$ and $(\alpha_m)_{m\sim M}$ be a sequence of complex numbers of modulus bounded by 1, assuming that*

$$q^{1/l} \leqslant N < \frac{1}{2}q^{1/2+1/2l}$$

on a has

$$\sum_{m\sim M}\alpha_m\sum_n V(\frac{n}{N})K(mn) \ll q^{o(1)}MN(\frac{N^2M}{q^{1+1/l}})^{-1/2l}.$$

Observe that this bound is non trivial (sharper than $\ll MN$) as long as[20]

$$N^2M \geqslant q^{1+1/l}.$$

For instance if $M = q^\delta$ for some $\delta > 0$, the above bound is nontrivial for l large enough and $N \geqslant q^{1/2+\delta/3}$; alternatively if $M = N$, this bound is non trivial as long as

$$N = M \geqslant q^{1/3+\delta}$$

(if l is large enough). Therefore this method improves the range of non-triviality in Theorem 9.1.

16.3. **The $+ab$-shift for type II sums.** With the above method it is also possible to deal with the more general (type II) bilinear sums

$$B(K,\boldsymbol{\alpha},\boldsymbol{\beta}) = \sum_{m\sim M,n\sim N}\sum\alpha_m\beta_n K(mn)$$

where $(\alpha_m)_{m\sim M}$, $(\beta_n)_{n\sim N}$ are sequences of complex numbers of modulus bounded by 1.

We leave it to the interested reader to fill the details (or to look at [FM98, KMS17] or the forthcoming [KMS18]. The first step is to apply Cauchy-Schwarz to smooth out the n variable: for a suitable smooth, compactly supported in $[1/2,5/2]$, bounded by 1 function V one has

$$\sum_{m\sim M,n\sim N}\sum\alpha_m\beta_n K(mn) \leqslant \Big(\sum_{m_1,m_2\sim M}\alpha_{m_1}\overline{\alpha_{m_2}}\sum_n V(\frac{n}{N})K(mn_1)\overline{K(mn_2)}\Big)^{1/2};$$

the next step is to perform the $+ab$-shift on the n variable and to make the change of variables

$$(a,m_1,m_2,n) \in [A,2A[\times[M,2M[^2\times[N,2N[\longleftrightarrow (\overline{a}n, am_1, am_2) \,(\mathrm{mod}\, q) = (r,s_1,s_2) \in {\mathbf{F}_q}^3.$$

[20] if $N \geqslant \frac{1}{2}q^{1/2+1/2l}$ the Polya-Vinogradov inequality is non trivial already

Considering the fiber counting function for that map

$$\nu(r,s_1,s_2) := \sum_{\substack{(a,n,m_1,m_2)\\ a\sim A,|n|\leqslant 2N,\ m_i\simeq M}}\sum\alpha_{m_1}\overline{\alpha_{m_2}}\delta_{\overline{a}n=r,\ am_i=s_i\,(\mathrm{mod}\,q)}$$

one show that for $AN<q/2$ one has

$$\sum_{(r,s_1,s_2)\in\mathbf{F}_q{}^3}\sum|\nu(r,s_1,s_2)|\ll AM^2N,\quad \sum_{(r,s_1,s_2)\in\mathbf{F}_q{}^3}\sum|\nu(r,s_1,s_2)|^2\leqslant q^{o(1)}AM^2N.$$

Applying Hölder's inequality lead us to the problem of bounding the following complete sum index by the parameter $\boldsymbol{b}$

$$\sum_{r\in\mathbf{F}_q}|\mathbf{R}(r,\boldsymbol{b})|^2-q\sum_{r\in\mathbf{F}_q}|\mathbf{K}(r,\boldsymbol{b})|^2. \tag{16.6}$$

We will explain what is expect in general in a short moment but let us see what happens for our previous case $K(x)=\mathrm{e}_q(\overline{x}+x)$: for $\boldsymbol{b}=(b_1,\cdots,b_{2l})\in\mathbf{F}_q{}^{2l}$ satisfying (16.5) the sum (16.6) equals

$$q\sum_{\substack{r\in\mathbf{F}_q\\ r\neq -b_i}}|\,\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)|^2-q\sum_{\substack{r\in\mathbf{F}_q\\ r\neq -b_i}}1=q\sum_{\substack{r\in\mathbf{F}_q\\ r\neq -b_i}}(|\,\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)|^2-1)+O_l(q).$$

Lemma 16.2. *For $\boldsymbol{b}=(b_1,\cdots,b_{2l})\in\mathbf{F}_q{}^{2l}$ satisfying (16.5), one has*

$$\sum_r(|\,\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)|^2-1)\ll_l q^{1/2}.$$

Proof. This follows from the fact that $[G_{\boldsymbol{b}}]^*\mathcal{K}\ell_2$ is geometrically irreducible with geometric monodromy group equal to SL_2: since the tensor product of the standard representation of SL_2 with itself equals the trivial representation plus the symmetric square of the standard representation which is non-trivial and irreducible,

$$x\to|\,\mathrm{Kl}_2(G_{\boldsymbol{b}}(r);q)|^2-1$$

is the trace function of a geometrically irreducible sheaf. □

Using this bound and trivial estimates for $\boldsymbol{b}$ not stisfying (16.5) one eventually obtains

Theorem 16.3. *Let $K(x)=\mathrm{e}_q(\overline{x}+x)$, $1\leqslant M,N<q$ and $l\geqslant 1$ some integer; assuming that*

$$N<\frac{1}{2}q^{1/2+1/2l}$$

one has

$$B(K,\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{m\sim M,n\sim N}\sum\alpha_m\beta_nK(mn)\ll q^{o(1)}MN(\frac{1}{M}+(\frac{MN}{q^{3/4+3/4l}})^{-1/4l})^{1/2}.$$

Remark. For l large enough, this bound is non-trivial as long as $M\geqslant q^{\delta}$ and $MN\geqslant q^{3/4+\delta}$ again improving on 9.1 in this specific case.

16.4. **The $+ab$-shift for more general trace functions.** For applications to analytic number theory it is highly desirable to extend the methods of the previous section to trace functions as general as possible. the above method be axiomatized in the following way; first we recall some notations

For q be a prime, $K : \mathbf{F}_q \to \mathbf{C}$ a complex valued function bounded by 1 in absolute value, $1 \leqslant M, N < q$ some parameters and $\boldsymbol{\alpha} = (\alpha_m)_{m\sim M}$, $\boldsymbol{\beta} = (\beta_n)_{n\sim N}$ sequences of complex number bounded by 1, we define the type I sum

$$B(K, \boldsymbol{\alpha}, 1_N) = \sum_{m\sim M}\sum_{n\sim N} \alpha_m K(mn)$$

and the type II sum

$$B(K, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{m\sim M}\sum_{n\sim N} \alpha_m \beta_n K(mn).$$

For $l \geqslant 1$ an integer, let $\mathbf{K}(r, \boldsymbol{b})$ and $\mathbf{R}(r, \boldsymbol{b})$ be the functions in the variable $(r, \boldsymbol{b}) \in \mathbf{F}_q \times {\mathbf{F}_q}^{2l}$ given in (16.4). For $B \geqslant 1$ we set

$$\mathcal{B} = \mathbf{Z}^{2l} \cap [B, 2B[^{2l}.$$

An axiomatic treatment of the type I sums $B(K, \boldsymbol{\alpha}, 1_N)$ is provided by the following

Theorem 16.4. *Notations as above, for $B, C \geqslant 1$ and $\gamma \in [0, 2]$ be some real numbers,*

– *let $\mathcal{B}^\Delta \subset \mathcal{B}$ be the complement of the set of $\boldsymbol{b} \in \mathcal{B}$ satisfying*

$$\forall r \in \mathbf{F}_q,\ |\mathbf{R}(r, \boldsymbol{b})| \leqslant Cq^{1/2}. \tag{16.7}$$

– *Let $\mathcal{B}^\Delta \subset \mathcal{B}_I^{bad} \subset \mathcal{B}$ be the complement in $\mathcal{B}$ of the set of $\boldsymbol{b} \in \mathcal{B} - \mathcal{B}^\Delta$ satisfying*

$$\big|\sum_{r\in\mathbf{F}_q} \mathbf{R}(r, \boldsymbol{b})\big| \leqslant Cq. \tag{16.8}$$

Assume that for any $1 \leqslant B < q/2$ one has

$$|\mathcal{B}^\Delta| \leqslant CB^l,\ |\mathcal{B}_I^{bad}| \leqslant B^{(2-\gamma)l} \tag{16.9}$$

If N satisfies

$$q^{1/l} \leqslant N \leqslant \frac{1}{2}q^{1/2+1/2l},$$

one has for any $\varepsilon > 0$

$$B(K, \boldsymbol{\alpha}, 1_N) \ll_{C,l,\varepsilon} q^\varepsilon MN\Big(\frac{q^{1+1/l}}{MN^2} + \frac{q^{3/2-\gamma+1/l}}{MN^2}\Big)^{1/2l}. \tag{16.10}$$

An axiomatic treatment of the type II sums $B(K, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is provided by the following

Theorem 16.5. *Notations as above, for $B, C \geqslant 1$ and $\gamma \in [0, 2]$ be some real numbers,*

– *Let $\mathcal{B}^\Delta \subset \mathcal{B}$ be the complement of the set of $\boldsymbol{b} \in \mathcal{B}$ satisfying*

$$\forall r \in \mathbf{F}_q,\ |\mathbf{R}(r, \boldsymbol{b})| \leqslant Cq^{1/2}.$$

– *Let $\mathcal{B}^\Delta \subset \mathcal{B}_{II}^{bad} \subset \mathcal{B}$ be the complement (in $\mathcal{B}$) of the set of $\boldsymbol{b} \in \mathcal{B} - \mathcal{B}^\Delta$ satisfying*

$$\big|\sum_{r\in\mathbf{F}_q} |\mathbf{R}(r, \boldsymbol{b})|^2 - q\sum_{r\in\mathbf{F}_q} |\mathbf{K}(r, \boldsymbol{b})|^2\big| \leqslant Cq^{3/2} \tag{16.11}$$

Assume that for any $B \in [1, q/2[$ one has

$$|\mathcal{B}^\Delta| \leqslant CB^l,\ |\mathcal{B}_{II}^{bad}| \leqslant CB^{(2-\gamma)l}. \tag{16.12}$$

If N satisfies

$$q^{3/2l} \leqslant N \leqslant \frac{1}{2}q^{1/2+3/4l},$$

one has for any $\varepsilon > 0$,

$$B(K, \boldsymbol{\alpha}, \boldsymbol{\beta}) \ll_{C,l,\varepsilon} q^\varepsilon MN\Big(\frac{1}{M} + \big(\frac{q^{1-\frac{3}{4}\gamma+\frac{3}{4l}}}{MN} + \frac{q^{\frac{3}{4}+\frac{3}{4l}}}{MN}\big)^{\frac{1}{l}}\Big)^{1/2}. \tag{16.13}$$

We conclude these lectures with a few remark concerning these two theorems

(1) In the case $K(x) = \mathrm{e}_q(\overline{x} + x)$ we have just verified that the conditions (16.9) and (16.12) hold with $\gamma = 1$. In [FM98] this was shown to hold also for the trace functions

$$K(x) = \mathrm{e}_q(x^{-k} + ax),\ a \in \mathbf{F}_q,\ k \geqslant 1.$$

(2) For more general trace functions, the first condition in (16.9) and (16.12) can be verified using some variant of the "sums of products" Theorem 13.1 and does not constitute a main obstacle. One also should notice that Theorem 13.1 also implies that for any $\boldsymbol{b} = (b_1, \cdots, b_{2l})$ on the "first" diagonal (ie. $b_1 = b_{l+1}, \cdots, b_l = b_{2l}$) one has

$$\mathbf{R}(r, \boldsymbol{b}) = \sum_s \prod_{i=1}^{l} |K(s(r + b_i))|^2 = |K(0)|^{2l} + \sum_{s \neq 0} \prod_{i=1}^{l} |K(s(r + b_i))|^2 \gg_l q$$

and therefore

$$|\mathcal{B}^{\Delta}| \geqslant B^l.$$

It follows that the first bound in (16.9) and (16.12) is sharp and for the second condition one cannot expect γ to be greater than 1.

(3) That said, in order to reach the best available bound by the above method, it is not necessary to aim for $\gamma = 1$: it is sufficient to establish (16.9) with $\gamma \geqslant 1/2$ and (16.12) with $\gamma \geqslant 1/3$. In such a case the above bounds are non trivial as long as

$$MN^2 \geqslant q^{1+1/l}\ MN \geqslant q^{3/4+3/4l}.$$

(4) Checking the second bound in (16.9) and (16.12) for general trace functions is much more difficult. In [KMS17], with specific applications in mind, these bounds have been established for $l = 2$ and $\gamma = 1/2$ for the hyper-Kloosterman sums

$$K(x) = \mathrm{Kl}_k(x; q),\ k \geqslant 2.$$

Because $l = 2$ is too small, this alone is not sufficient to improve over the Polya-Vinogradov type bound of Theorem 9.1 (one would have needed $l \geqslant 4$). A more refined treatment is necessary: instead of letting (somewhat wastefully) the variables $s = am \,(\mathrm{mod}\, q)$ or $s_1 = am_1, s_2 = am_2 \,(\mathrm{mod}\, q)$ vary freely over the whole interval $[0, q-1] \simeq \mathbf{F}_q$ one use the fact that s, s_1, s_2 are varying along the shorter interval $[AM, 4AM[$. Applying the Polya-Vinogradov completion method to detect this inclusion through additive characters this lead to bounds for complete sums analogous to (16.8) and (16.11) but for the additively twisted variant of $\mathbf{R}(r, \boldsymbol{b})$,

$$\mathbf{R}(r, \lambda, \boldsymbol{b}) = \sum_s \mathbf{K}(sr, s\boldsymbol{b}) e(\frac{\lambda s}{q}),\quad \text{for } \lambda \in \mathbf{F}_q.$$

Specifically the bounds are: $\forall \boldsymbol{b} \in \mathcal{B} - \mathcal{B}^{\Delta}$,

$$\forall \lambda \in \mathbf{F}_q,\ |\mathbf{R}(r, \lambda, \boldsymbol{b})| \leqslant Cq^{1/2},$$

and $\forall \boldsymbol{b} \in \mathcal{B} - \mathcal{B}_I^{bad}$,

$$\forall \lambda \in \mathbf{F}_q,\ |\sum_r \mathbf{R}(r, \lambda, \boldsymbol{b})| \leqslant Cq,$$

and $\forall \boldsymbol{b} \in \mathcal{B} - \mathcal{B}_{II}^{bad}$

$$\forall \lambda, \lambda' \in \mathbf{F}_q,\ \Big|\sum_r \mathbf{R}(r, \lambda, \boldsymbol{b})\overline{\mathbf{R}(r, \lambda', \boldsymbol{b})} - q\delta_{\lambda=\lambda'} \sum_s \prod_{i=1}^{l} |K(s(r + b_i))|^2\Big| \leqslant Cq^{3/2}$$

and in [KMS17] these bounds were established for $l = 2$ and $\boldsymbol{b}$ outside the sets $\mathcal{B}^{\Delta}$, $\mathcal{B}_I^{bad}$ and $\mathcal{B}_{II}^{bad}$ satisfying

$$|\mathcal{B}^{\Delta}| \leqslant B^2,\ |\mathcal{B}_{I,II}^{bad}| \leqslant CB^3.$$

(5) In the forthcoming paper [KMS18] the bounds (16.9) and (16.12) are established for the hyper-Kloosterman sums and more general hypergeometric type sums for every $l \geqslant 2$, $\gamma = 1/2$.

16.5. **Some applications of the $+ab$-shift bounds.** The problem of estimating bilinear sums of trace functions below the critical Polya-Vinogradov range has had several applications in analytic number theory. We list some of them below with references for the interested remaining reader(s).

– This method was used by Karacuba and Vinogradov for the function

$$K(n) = \chi(n + a)$$

$(a, q) = 1$ and $\chi \pmod q$ a non-trivial Dirichlet character to bound non-trivially its sum along the primes over short intervals (now a special case of Theorem 8.1): in particular, Karacuba [?Kar] proved for any $\varepsilon > 0$, the bound

$$\sum_{n \leqslant x} \chi(n + a) \ll x^{1-\varepsilon^2/1024}$$

whenever $x \geqslant q^{1/2+\varepsilon}$; this bound is therefore non-trivial in a range which is wider than the $x \geqslant q^{3/4+\varepsilon}$ established in Theorem 8.1 for general trace functions.

– The method was used by Friedlander-Iwaniec for the function

$$K(n) = e_q(\overline{n}),\ n.\overline{n} \equiv 1 \pmod q.$$

to show that the ternary divisor function $d_3(n),\ n \leqslant n$ is well distributed in arithmetic progressions of modulus $q \leqslant x^{1/2+1/230}$, passing for the first time the Bombieri-Vinogradov barrier (see Theorem 11.3).

– The bound in case of the Kloosterman sums

$$K(n) = \mathrm{Kl}_2(n; q)$$

established in [KMS17] (see above) together with [BFK+17] gave a sharp asymptotic formula for the second moment of character twists of modular L-functions of prime modulus q: for f a fixed Hecke-eigen cuspform one has

$$\frac{1}{q-1} \sum_{\chi \pmod q} |L(f \otimes \chi, 1/2)|^2 = MT_f(\log q) + O_f(q^{-1/145})$$

where $MT_f(\log q)$ is a polynomial in $\log q$ (of degree $\leqslant q$) depending on f, completing the work of Young for f an Eisenstein series [You11] and of Blomer-Milicevic for f cuspidal and q a suitably composite modulus[21].

– Using that method, Nunes [Nun17] obtained non-trivial bounds, below the Polya-Vinogradov range, for the following (smooth) bilinear sum

$$\sum_{\substack{m \leqslant M \\ n \leqslant N}} \sum K(mn^2)$$

where K is the Kloosterman-like trace function

$$K(n; q) := \frac{1}{q^{1/2}} \sum_{x \in \mathbf{F}_q^{\times}} e_q(a\overline{x}^2 + bx)$$

[21] who used the q-van der Corput method on this occasion.

(a, b some integral parameters sum that $(ab, q) = 1$) and deduced from this bound, that the characteristic function of squarefree integers

$$n \mapsto \mu^2(n),\ n \leqslant x$$

is well distributed in arithmetic progression of prime modulus

$$q \leqslant x^{2/3+1/57}.$$

The previous best result, due to Prachar [Pra58], was $q \leqslant x^{2/3-\varepsilon}$ (similar to Selberg's Theorem 11.2 for the divisor function $d_2(n)$) was from 1958 !

References

[BM15] V. Blomer and D. Milićević, *The second moment of twisted modular L-functions*, Geom. Funct. Anal. **25** (2015), no. 2, 453–516.

[BFK+17] Valentin Blomer, Étienne Fouvry, Emmanuel Kowalski, Philippe Michel, and Djordje Milićević, *On moments of twisted L-functions*, Amer. J. Math. **139** (2017), no. 3, 707–768. arXiv:1411.4467.

[Bur62] D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. (3) **12** (1962), 179–192.

[Del80] P. Deligne, *La conjecture de Weil, II*, Publ. Math. IHÉS **52** (1980), 137–252.

[Fou85] Étienne Fouvry, *Sur le problème des diviseurs de Titchmarsh*, J. Reine Angew. Math. **357** (1985), 51–76 (French). MR783533

[FI92] Étienne Fouvry and Henryk Iwaniec, *The divisor function over arithmetic progressions*, Acta Arith. **61** (1992), no. 3, 271–287. With an appendix by Nicholas Katz. MR1161479

[FKM15] É. Fouvry, E. Kowalski, and Ph. Michel, *Algebraic twists of modular forms and Hecke orbits*, GAFA **25** (2015), no. 2, 580-657. arXiv:1207.0617.

[FKM13] Étienne Fouvry, Emmanuel Kowalski, and Philippe Michel, *Counting sheaves using spherical codes*, Math. Res. Lett. **20** (2013), no. 2, 305–323.

[FKM15] É. Fouvry, E. Kowalski, and Ph. Michel, *A study in sums of products*, Philos. Trans. A **373** (2015), no. 2040, 20140309, 26pp. arXiv:1304.3199.

[FKM14] É. Fouvry, E. Kowalski, and Ph. Michel, *Algebraic trace functions over the primes*, Duke Math. J. **163** (2014), no. 9, 1683–1736. arXiv:1211.6043.

[FKM15] ———, *On the exponent of distribution of the ternary divisor function*, Mathematika **61** (2015), no. 1, 121-144. arXiv:1304.3199.

[FM98] É. Fouvry and Ph. Michel, *Sur certaines sommes d'exponentielles sur les nombres premiers*, Ann. Sci. École Norm. Sup. (4) **31** (1998), no. 1, 93–130.

[FM07] E. Fouvry and Ph. Michel, *Sur le changement de signe des sommes de Kloosterman*, Ann. of Math. (2) **165** (2007), no. 3, 675-715.

[FKM+17] Étienne Fouvry, Emmanuel Kowalski, Ph. Michel, C. S. Raju, J. Rivat, and K. Soundararajan, *On short sums of trace functions*, Ann. Inst. Fourier (Grenoble) **167** (2017), no. 1, 423–449. arXiv:1508.00512.

[FI85] J.B. Friedlander and H. Iwaniec, *Incomplete Kloosterman sums and a divisor problem*, Ann. of Math. (2) **121** (1985), no. 2, 319–350. (with an appendix by B. J. Birch and E. Bombieri).

[GPY09] Daniel A. Goldston, János Pintz, and Cem Y. Yıldırım, *Primes in tuples. I*, Ann. of Math. (2) **170** (2009), no. 2, 819–862. MR2552109

[Gra15] Andrew Granville, *Primes in intervals of bounded length*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 171–222. MR3312631

[HBP79] D. R. Heath-Brown and S. J. Patterson, *The distribution of Kummer sums at prime arguments*, J. Reine Angew. Math. **310** (1979), 111–130. MR546667

[HB86] D.R. Heath–Brown, *The divisor function $d_3(n)$ in arithmetic progressions*, Acta Arith. **47** (1986), 29–56.

[IT13] Atsushi Ichino and Nicolas Templier, *On the Voronoĭ formula for* GL(n), Amer. J. Math. **135** (2013), no. 1, 65–101. MR3022957

[Irv15] Alastair Irving, *The divisor function in arithmetic progressions to smooth moduli*, Int. Math. Res. Not. IMRN **15** (2015), 6675–6698, DOI 10.1093/imrn/rnu149. MR3384495

[Irv16] ———, *Estimates for character sums and Dirichlet L-functions to smooth moduli*, Int. Math. Res. Not. IMRN **15** (2016), 4602–4633, DOI 10.1093/imrn/rnv285. MR3564622

[Iwa97] Henryk Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17, American Mathematical Society, Providence, RI, 1997. MR1474964

[IK04] H. Iwaniec and E. Kowalski, *Analytic number theory*, Vol. 53, American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2004.

[IS00] Henryk Iwaniec and Peter Sarnak, *The non-vanishing of central values of automorphic L-functions and Landau-Siegel zeros*. part A, Israel J. Math. **120** (2000), no. part A, 155–177, DOI 10.1007/s11856-000-1275-9. MR1815374

[IS99] H. Iwaniec and P. Sarnak, *Dirichlet L-functions at the central point*, Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997), de Gruyter, Berlin, 1999, pp. 941–952. MR1689553

[KL78] G. A. Kabatjanskiĭ and V. I. Levenšteĭn, *Bounds for packings on the sphere and in space*, Problemy Peredači Informacii **14** (1978), no. 1, 3–25 (Russian). MR0514023

[Kat80] N. M. Katz, *Sommes exponentielles*, Astérisque, vol. 79, Société Mathématique de France, Paris, 1980.

[Kat88] ———, *Gauss sums, Kloosterman sums, and monodromy groups*, Annals of Mathematics Studies, vol. 116, Princeton University Press, Princeton, NJ, 1988.

[Kat90a] ———, *Exponential sums and differential equations*, Annals of Mathematics Studies, vol. 124, Princeton University Press, Princeton, NJ, 1990.

[Kat90b] Nicholas M. Katz, *Exponential sums over finite fields and differential equations over the complex numbers: some interactions*, Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 2, 269–309.

[Kat96] N. M. Katz, *Rigid local systems*, Annals of Mathematics Studies, vol. 139, Princeton University Press, Princeton, NJ, 1996.

[Kat05a] ———, *Moments, monodromy, and perversity: a Diophantine perspective*, Annals of Mathematics Studies, vol. 159, Princeton University Press, Princeton, NJ, 2005.

[Kat05b] ———, *Twisted L-Functions and Monodromy*, Annals of Mathematics Studies, vol. 150, Princeton University Press, Princeton, NJ, 2005.

[Kat12] Nicholas M. Katz, *Convolution and equidistribution: Sato-Tate theorems for finite-field Mellin transforms*, Annals of Mathematics Studies, vol. 180, Princeton University Press, Princeton, NJ, 2012. MR2850079

[KN16] Rizwanur Khan and Hieu T. Ngo, *Nonvanishing of Dirichlet L-functions*, Algebra Number Theory **10** (2016), no. 10, 2081–2091, DOI 10.2140/ant.2016.10.2081.

[KZ16] Eren Mehmet Kıral and Fan Zhou, *The Voronoi formula and double Dirichlet series*, Algebra Number Theory **10** (2016), no. 10, 2267–2286. MR3582019

[Kow13] E. Kowalski, *Families of cusp forms*, Actes de la Conférence "Théorie des Nombres et Applications", Publ. Math. Besançon Algèbre Théorie Nr., vol. 2013, Presses Univ. Franche-Comté, Besançon, 2013, pp. 5–40. MR3220018

[Kow15] Emmanuel Kowalski, *Gaps between prime numbers and primes in arithmetic progressions [after Y. Zhang and J. Maynard]*, Astérisque **367-368** (2015), Exp. No. 1084, ix, 327–366.

[KMS17] Emmanuel Kowalski, Philippe Michel, and Will Sawin, *Bilinear forms with Kloosterman sums and applications*, Ann. of Math. (2) **186** (2017), no. 2, 413–500. `arXiv:1511.01636`.

[KMS18] ———, *Bilinear forms with Kloosterman sums II* (2018). (in preparation).

[KMV02] E. Kowalski, Ph. Michel, and J. VanderKam, *Rankin–Selberg L-functions in the level aspect*, Duke Math. Journal **114** (2002), 123–191.

[Lau87] G. Laumon, *Transformation de Fourier, constantes d'équations fonctionnelles et conjecture de Weil*, Inst. Hautes Études Sci. Publ. Math. **65** (1987), 131–210 (French).

[Mat11] Kaisa Matomäki, *A note on signs of Kloosterman sums*, Bull. Soc. Math. France **139** (2011), no. 3, 287–295 (English, with English and French summaries). MR2869308

[May16] James Maynard, *Large gaps between primes*, Ann. of Math. (2) **183** (2016), no. 3, 915–933.

[Mic95] Ph. Michel, *Autour de la conjecture de Sato-Tate pour les sommes de Kloosterman. I*, Invent. Math. **121** (1995), no. 1, 61–78.

[Mic98] ———, *Minorations de sommes d'exponentielles*, Duke Math. J. **95** (1998), no. 2, 227–240.

[MV00] Philippe Michel and Jeffrey VanderKam, *Non-vanishing of high derivatives of Dirichlet L-functions at the central point*, J. Number Theory **81** (2000), no. 1, 130–148. MR1743500

[MS06] Stephen D. Miller and Wilfried Schmid, *Automorphic distributions, L-functions, and Voronoi summation for* GL(3), Ann. of Math. (2) **164** (2006), no. 2, 423–488. MR2247965

[Nun17] R. M. Nunes, *On the least squarefree number in an arithmetic progression*, Mathematika **63** (2017), no. 2, 483–498.

[Pol14] D.H.J. Polymath, *New equidistribution estimates of Zhang type*, Algebra & Number Theory **8** (2014), no. 9, 2067–2199. `arXiv:1402.0811`.

[Pra58] Karl Prachar, *Über die kleinste quadratfreie Zahl einer arithmetischen Reihe*, Monatsh. Math. **62** (1958), 173–176 (German). MR0092806

[SST16] Peter Sarnak, Sug Woo Shin, and Nicolas Templier, *Families of L-functions and their symmetry*, Families of automorphic forms and the trace formula, Simons Symp., Springer, [Cham], 2016, pp. 531–578. MR3675175

[SF09] Jimena Sivak-Fischler, *Crible asymptotique et sommes de Kloosterman*, Bull. Soc. Math. France **137** (2009), no. 1, 1–62 (French, with English and French summaries). MR2496700

[Sou07a] K. Soundararajan, *The fourth moment of Dirichlet L-functions*, Analytic number theory, Clay Math. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 2007, pp. 239–246.

[Sou07b] ———, *Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 1, 1–18. MR2265008

[Top] B. Topacogullari, *The shifted convolution of divisor functions*. (preprint, 2015; `arXiv:1506.02608`).

[You11] M.P. Young, *The fourth moment of Dirichlet L-functions*, Ann. of Math. (2) **173** (2011), no. 1, 1–50.

[Wei41] André Weil, *On the Riemann hypothesis in functionfields*, Proc. Nat. Acad. Sci. U. S. A. **27** (1941), 345–347.

[XW16] P. Xi and J. Wu, *Arithmetic exponent pairs for algebraic trace functions and applications* (2016). `https://arxiv.org/abs/1603.07060`.

[Xi15] Ping Xi, *Sign changes of Kloosterman sums with almost prime moduli*, Monatsh. Math. **177** (2015), no. 1, 141–163. MR3336337

[Zha14] Yitang Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174.

[SGA1] A. Grothendieck and M. Raynaud, *Revêtements étales et groupe fondamental*, Lecture Notes in Mathematics, vol. 224, Springer-Verlag, Berlin-New York, 1971. Séminaire de Géométrie Algébrique du Bois-Marie 1960–1961 (SGA 1).

[SGA4] M. Artin, A. Grothendieck, and J.-L. Verdier, *Théorie des topos et cohomologie étale des schémas*, Lecture Notes in Mathematics, vol. 269,270,305, Springer-Verlag, Berlin-New York, 1972. Séminaire de Géométrie Algébrique du Bois-Marie (SGA 4).

[SGA4$\frac{1}{2}$] P. Deligne, *Cohomologie étale*, Lecture Notes in Mathematics, vol. 569, Springer-Verlag, Berlin-New York, 1977. Séminaire de Géométrie Algébrique du Bois-Marie (SGA $4\frac{1}{2}$).

[SGA5] A. Grothendieck and L. Illusie, *Cohomologie ℓ-adique et fonctions L*, Lecture Notes in Mathematics, vol. 589, Springer-Verlag, Berlin-New York, 1977. Séminaire de Géometrie Algébrique du Bois-Marie 1965–1966 (SGA 5).

[SGA7] P. Deligne and N.M. Katz, *Groupes de monodromie en géométrie algébrique, II*, Lecture Notes in Mathematics, vol. 340, Springer-Verlag, Berlin-New York, 1973. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II).

EPFL/SB/TAN, Station 8, CH-1015 Lausanne, Switzerland
E-mail address: `philippe.michel@epfl.ch`