Some applications of representation theory to the sum–product phenomenon

I.D. Shkredov

Annotation.

In our paper, we introduce a new method for estimating incidences via representation theory. We obtain several applications to various sums with multiplicative characters and to Zaremba's conjecture from number theory.

1 Introduction

Given two finite sets A and B of an abelian ring, define the sumset, and the product set of A and B as

$$A+B=\{a+b : a \in A,\, b \in B\}, \qquad A \cdot B=\{ab : a \in A,\, b \in B\}\,. \tag{1}$$

The sum-product phenomena was introduced by Erdős and Szemerédi in paper [9] where they proved that for an arbitrary finite subset A of integers one has

$$\max\{|A+A|, |A\cdot A|\} \gg |A|^{1+c}\,. \tag{2}$$

Here $c>0$ is an absolute constant and Erdős and Szemerédi conjectured that any $c<1$ is admissible, at the cost of the implicit constant. As a general heuristic, the conjecture suggests that either $A+A$ or AA is significantly larger then the original set, unless A is close to a subring. Even more generally speaking, the sum–product phenomenon predicts that the an arbitrary subset of a ring cannot have good additive and multiplicative structures simultaneously. The interested reader may consult [36] for a rather thorough treatment of sumsets and related questions, including some prior work on the sum-product problem. The sum-product phenomenon has been extensively studied in the last few decades, the current records as of writing being [30] for real numbers, and [22] for sufficiently small sets in finite fields.

In our paper we consider the case of the ring $\mathbb{Z}_q := \mathbb{Z}/q\mathbb{Z}$ and we have deal with *large* sets $A \subseteq \mathbb{Z}_q$ (basically, it means that $|A| > q^{1-\kappa}$ for a certain constant $\kappa > 0$). In the case of a prime q the behaviour of the maximum from (2) is fully known thanks to the beautiful result of Garaev [13] who used some classical exponential sums bounds in his proof. Another approach was suggested in [40] and in [27] where some finite geometry considerations were applied. For example, Vinh [40] proved that for an arbitrary prime q and any two sets $\mathcal{A} \subseteq \mathbb{Z}_q\times\mathbb{Z}_q$, $\mathcal{B} \subseteq \mathbb{Z}_q\times\mathbb{Z}_q$ one has

$$\left| |\{(a_1,a_2)\in \mathcal{A},\, (b_1,b_2)\in \mathcal{B} \;:\; a_1b_1-a_2b_2 \equiv 1 \pmod q\}| - \frac{|\mathcal{A}||\mathcal{B}|}{q} \right| \leqslant \sqrt{q|\mathcal{A}||\mathcal{B}|}\,. \tag{3}$$

In the proof he used the fact that equation (3) can be interpreted as a question about points/lines incidences. Clearly, the result above has the sum–product flavour and indeed one can use (3) to derive some lower bounds for the maximum from (2) (in the case of large subsets of $\mathbb{Z}_q$, of course).

In this paper we introduce a new method of estimating sum–product quantities as in (3) which does not use any exponential sums, as well as any considerations from the incidence geometry. It turns out that representation theory makes it possible to obtain (almost automatically) asymptotic formulae for the number of solutions to systems of equations that are preserved by the actions of certain groups. For example, equation (3) can be interpreted as the equation $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$, where $(a_1, a_2) \in \mathcal{A}$ and $(b_1, b_2) \in \mathcal{B}$ and hence the equation respects the usual action of $\mathrm{SL}_2(\mathbb{Z}_q)$. The advantage of our approach is its generality and (relative) simplicity. First of all, having a certain equation, the method makes it possible to obtain an asymptotic formula for the number of solutions to the equation for composite q due to the fact that representation theory for composite q is usually not so complicated and can be reduced to the case of prime powers. We should mention that the question about the sum–product phenomenon for general $\mathbb{Z}_q$ and large sets is considered to be difficult and there are few results in this direction, see [37] and paper [28], where the case of finite valuation rings was considered (also, see [8]). Another statement of the problem concerning the sum–product results in $\mathbb{Z}_q$ is contained in [10], [14], [34]. Let us remark that in [10] Fish also uses the property of equation invariance, but combines it with classical Fourier analysis. Secondly, due to the obvious observation that representation theory deals with some facts concerning the acting group but not with sets, in all our results all the sets involved (as $\mathcal{A}, \mathcal{B}$ in (3)) are absolutely general and do not require to have a special structure, for example, to be Cartesian products of some other sets. The last constraint is sometimes crucial for Fourier analysis manipulations, see, e.g., [1], [39], although it usually allows to obtain better error terms in asymptotic formulae.

To be more specific let us mention just one result here (see Theorem 8 of Section 3 below). Given positive integers q,n, m, $d = n + m$, an element $\lambda \in \mathbb{Z}_q$ and sets $\mathcal{A} \subseteq (\mathbb{Z}_q^d)^n$, $\mathcal{B} \subseteq (\mathbb{Z}_q^d)^m$ define by $\mathcal{D}_\lambda(\mathcal{A}, \mathcal{B})$ the number of solutions to the equation

$$\det(a_1, \dots, a_n, b_1, \dots, b_m) \equiv \lambda \pmod q, \qquad (a_1, \dots, a_n) \in \mathcal{A}, \quad (b_1, \dots, b_m) \in \mathcal{B}. \tag{4}$$

We assume that

Theorem 1 *Let q be an odd prime number and $\lambda \neq 0$. Then*

$$\left| \mathcal{D}_\lambda(\mathcal{A}, \mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q-1} \right| \ll q^{d^2/2 - d/4 - 3/4} \sqrt{|\mathcal{A}||\mathcal{B}|}\,. \tag{5}$$

In Section 4 we obtain further applications of our approach to some problems of number theory. Our main observation is that the representation theory of $\mathrm{SL}_2(\mathbb{Z}_q)$ makes it easy to insert *multiplicative* characters into all formulae with incidences and, therefore, to obtain non–trivial estimates for the corresponding exponential sums. In the author opinion this is a rather interesting phenomenon due to the widely–known fact that results with multiplicative characters are usually very difficult to obtain. As an example, we formulate the following theorem concerning summation over a hyperbolic surface. Denote by $\mathcal{D} \subset \mathbb{C}$ the unit disk.

Theorem 2 *Let q be a prime number, $\delta > 0$ be a real number, $A, B, X, Y \subseteq \mathbb{Z}_q$ be sets, let χ be a non–principal multiplicative character and $|X||Y| \geqslant q^{\delta}$. Also, let $c_A : A \to \mathcal{D}$, $c_B : B \to \mathcal{D}$ be some weights. Then there is $\varepsilon(\delta) > 0$ such that*

$$\sum_{a\in A,\, b\in B,\, x\in X,\, y\in Y \;:\; (a+x)(b+y)=1} c_A(a)c_B(b)\chi(a+x) \leqslant \sqrt{|A||B|}(|X||Y|)^{1-\varepsilon(\delta)} \,.$$

Another application of the approach allows us to generalize [33, Theorem 4] (also, see Theorem 33 from this paper). Let χ be a non–principal multiplicative character over a finite field $\mathbb{F}$. Consider the Kloosterman sum twisted by the character χ, namely,

$$K_\chi(n,m) = \sum_{x\in \mathbb{F}\setminus\{0\}} \chi(x)e(nx+mx^{-1})\,,$$

where $e(\cdot)$ is an additive character on $\mathbb{F}$. We are interested in bilinear forms of Kloosterman sums (motivation can be found, say, in [33]) that is, the sums of the form

$$S_\chi(\alpha,\beta) = \sum_{n,m} \alpha(n)\beta(m)K_\chi(n,m)\,,$$

where $\alpha : \mathbb{F} \to \mathbb{C}$, $\beta : \mathbb{F} \to \mathbb{C}$ are arbitrary functions.

Theorem 3 *Let $c > 0$ and q be a prime number. Let $t_1, t_2 \in \mathbb{Z}_p$, N, M be integers, $N, M \leqslant q^{1-c}$ and let $\alpha, \beta : \mathbb{Z}_q \to \mathbb{C}$ be functions supported on $\{1,\dots,N\}+t_1$ and $\{1,\dots,M\}+t_2$, respectively. Then there exists $\delta(c) > 0$ such that*

$$S_\chi(\alpha,\beta) \lesssim \|\alpha\|_2\|\beta\|_2 q^{1-\delta}\,. \tag{6}$$

Finally, we obtain an application to Zaremba's conjecture [42]. Recall the main result of [24].

Theorem 4 *Let q be a positive sufficiently large integer with sufficiently large prime factors. Then there is a positive integer a, $(a,q) = 1$ and*

$$M = O(\log q/\log\log q) \tag{7}$$

such that

$$\frac{a}{q} = [0; c_1, \dots, c_s] = \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{1}{c_3 + \dots + \cfrac{1}{c_s}}}}\,, \qquad c_j \leqslant M\,, \qquad \forall j \in [s]\,. \tag{8}$$

Also, if q is a sufficiently large square–free number, then (7), (8) take place.
Finally, if $q = p^n$, p is an arbitrary prime, then (7), (8) hold for sufficiently large n.

Using an idea from representation theory, one can generalize Theorem 4.

Theorem 5 *Let q be a sufficiently large prime number and $\Gamma \leqslant \mathbb{Z}_q$ be a multiplicative subgroup,*

$$|\Gamma| \gg \frac{q}{\log^{\kappa} q}\,, \tag{9}$$

where $\kappa > 0$ is an absolute constant. Then there is $a \in \Gamma$ and

$$M = O(\log q / \log\log q) \tag{10}$$

such that

$$\frac{a}{q} = [0; c_1, \dots, c_s]\,, \qquad c_j \leqslant M\,, \qquad \forall j \in [s]\,.$$

Some results of this type concerning restrictions of the numerators of fractions (8) to multiplicative subgroups were obtained in [7], [25] and [26].

We thank Nikolai Vavilov for useful discussions and references.

2 Definitions and preliminaries

Let $\mathbf{G}$ be a group (commutative or not) and A, B be some subsets of $\mathbf{G}$. The sumset (and the product set) of A and B was defined in (1). Let us write $A \dotplus B$ if for finite sets A, B one has $|A+B| = |A||B|$. We use a representation function notation such as $r_{AB}(x)$ or $r_{AB^{-1}}(x)$, which counts the number of ways $x \in \mathbf{G}$ can be expressed as the product ab or ab^{-1} with $a \in A$, $b \in B$, respectively. For example, $|A| = r_{AA^{-1}}(1)$. Let us write $r^{(k)}_A$ for $r_{A\dots A}$, where the set A is taken k times. Having real functions $f_1, \dots, f_{2k} : \mathbf{G} \to \mathbb{C}$ (let k be an even number for concreteness), we put

$$\mathsf{T}_k(f_1, \dots, f_{2k}) = \sum_{a_1 a_2^{-1} \dots a_{k-1} a_k^{-1} = a_{k+1} a_{k+2}^{-1} \dots a_{2k-1} a_{2k}^{-1}} f_1(a_1) \dots f_{2k}(a_{2k})\,.$$

In this paper we use the same letter to denote a set $A \subseteq \mathbf{G}$ and its characteristic function $A : \mathbf{G} \to \{0,1\}$. Finally, if $|\mathbf{G}| < \infty$, then we consider the balanced function f_A of A, namely, $f_A(x) := A(x) - |A|/|\mathbf{G}|$.

In this paper we have deal with the group $\mathrm{SL}_2(\mathbb{Z}_q) \leqslant \mathrm{GL}_2(\mathbb{Z}_q)$ of matrices

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ab|cd) = (a,b|c,d)\,, \qquad a,b,c,d \in \mathbb{Z}_q\,, \qquad \det(g) = ad - bc = 1\,,$$

which acts on the project line (in the case of a prime number q) via the formula $gx = \frac{ax+b}{cx+d}$ and naturally acting on $\mathbb{Z}_q \times \mathbb{Z}_q$ for an arbitrary q.

Now we give a simplified version of the special case of [20, Theorem 6] (also, see [17, Theorem 3]). Let p be a prime number, d be a positive integer, $V(\mathbb{Z}_{p^d})$ be a vector space over $\mathbb{Z}_{p^d}$, $\dim V(\mathbb{Z}_{p^d}) = n$ on which a non–degenerate symmetric bilinear form $\Phi(\cdot,\cdot)$ is given. The group of isometries of V is called the *orthogonal group of* $V(\mathbb{Z}_{p^d})$, $O_n(\mathbb{Z}_{p^d})$ and the subgroup of isometries with determinant one is called the *special orthogonal group of* $V(\mathbb{Z}_{p^d})$, $SO_n(\mathbb{Z}_{p^d})$.

Theorem 6 *Let p be a prime number, $p \geqslant 5$, d be a positive integer, $V(\mathbb{Z}_{p^d})$ be a vector space over $\mathbb{Z}_{p^d}$, and $\Phi(x_1,\dots,x_n;y_1,\dots,y_n) = x_1y_1+\dots+x_ny_n$ defined on $V(\mathbb{Z}_{p^d})\times V(\mathbb{Z}_{p^d})$. Suppose that Γ is a normal subgroup of $SO_n(\mathbb{Z}_{p^d})$, where $n \geqslant 3$, $n \neq 4$. Then Γ is a congruence subgroup with the quotient isomorphic to $SO_n(\mathbb{Z}_{p^r})$, $r < d$.*

Indeed, in [20, Theorem 6] it requires to calculate the center of $SO_n(\mathbb{Z}_{p^d})$, which is trivial as one can easily check (or consult [17, Lemma 1] for general Φ). Further one needs to find an isotropic vector $x = (x_1,\dots,x_n) \neq 0$ such that $\Phi(x,x) = 0$ and this is an obvious task to do as the equation $x_1^2+x_2^2+x_3^2 \equiv 0 \pmod p$ has a nonzero solution (and hence a solution modulo p^d by Hensel's lemma). Finally, notice that in the case $n = 4$ one can in principle use [17, Remark 2] (in [20, Theorem 6] the author considers the case $n = 4$ under some additional assumptions which exclude the case of the sum of two hyperbolic planes).

Basic facts of representation theory can be found in [18]. Recall that a representation ρ of a group $\mathbf{G}$ is called *faithful* if it is injective. We need some number–theoretic functions. Given a positive integer n we write $\tau(n)$ for the number of all divisors of n and by $\omega(n)$ denote the number of all prime divisors. Also, denote by $J_k(n) = n^k\prod_{p|n}(1-p^{-k})$ the Jordan totient function equals the number of k–tuples of positive integers that are less than or equal to n and that together with n form a coprime set of $k+1$ integers. For example, it is easy to see that $|\mathrm{SL}_2(\mathbb{Z}_q)| = qJ_2(q)$.

The signs $\ll$ and $\gg$ are the usual Vinogradov symbols. When the constants in the signs depend on a parameter M, we write $\ll_M$ and $\gg_M$. If $a \ll_M b$ and $b \ll_M a$, then we write $a \sim_M b$. All logarithms are to base 2. We write $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$ and let $\mathbb{Z}_q^*$ be the group of all invertible elements of $\mathbb{Z}_q$. By $\mathbb{F}_p$ denote $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ for a prime p. Finally, let us denote by $[n]$ the set $\{1,2,\dots,n\}$.

3 Applications to incidence problems

We start with the simplest question about points/hyperplanes incidences (see equation (11) below). This problem was considered before in [41], [38], where the authors obtained better asymptotic formulae for the quantity $\mathcal{I}_\lambda(\mathcal{A},\mathcal{B})$ using other approaches. We commence with equation (11) because it allows us to transparently demonstrate our method, and because we will use some of the calculations from the proof below. As we will see the proof of Theorem 7 exploits some facts about representation theory of $SO_n(\mathbb{Z}_q)$, which preserves the distance $x_1^2+\dots+x_n^2$ in $\mathbb{Z}_q^n$. Thus, our approach is applicable in principle to all distance problems, for example, to the well–known Erdős–Falconer distance problem see, e.g., [16].

Given positive integers $q,n \geqslant 2$, an element $\lambda \in \mathbb{Z}_q$ and sets $\mathcal{A} \subseteq \mathbb{Z}_q^n$, $\mathcal{B} \subseteq \mathbb{Z}_q^n$ consisting of tuples all coprime to q, define by $\mathcal{I}_\lambda(\mathcal{A},\mathcal{B})$ the number of solutions to the equation

$$a_1b_1+\dots+a_nb_n \equiv \lambda \pmod q\,. \tag{11}$$

Theorem 7 *Let $q,n \geqslant 2$ be positive integers, $\mathcal{A} \subseteq \mathbb{Z}_q^n$, $\mathcal{B} \subseteq \mathbb{Z}_q^n$ be sets and $\lambda \in \mathbb{Z}_q^*$. Let m be the least prime divisor of q and suppose that $m \geqslant 5$. Then*

$$\left|\mathcal{I}_\lambda(\mathcal{A},\mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q\prod_{p|q}(1-p^{-n})}\right| \leqslant 2q^{n-1}\sqrt{|\mathcal{A}||\mathcal{B}|}\cdot(\Theta(n)m^{-n_*})^{1/4}\,, \tag{12}$$

where $n_* = 1$ *for* $n = 2,3$ *and* $n_* = n-3$ *for* $n \geqslant 4$ *and further,* $\Theta(2) \ll \min\{\tau(q), \log_m q\}$, $\Theta(3) \ll \min\{\log \omega(q), 1 + \omega(q)/m\}$ *and* $\Theta(n) \ll 1$ *for* $n \geqslant 4$.

Proof. Let $q = p_1^{\omega_1} \dots p_t^{\omega_t}$, where $m = p_1 < p_2 < \dots < p_t$ are primes and ω_j are positive integers. Also, let $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ and let $M(a,b) = 1$ iff the pair (a,b) satisfies our equation (11). Considering the unitary decomposition of the hermitian matrix $M(a,b)$, we obtain

$$M(a,b) = \sum_{j=1}^{q^n} \mu_j u_j(a) \overline{u}_j(b)\,, \tag{13}$$

where $\mu_1 \geqslant \mu_2 \geqslant \dots$ are the eigenvalues and u_j are correspondent orthonormal eigenfunctions. Clearly,

$$\mathcal{I}_\lambda(\mathcal{A}, \mathcal{B}) = \sum_{a \in \mathcal{A}, b \in \mathcal{B}} M(a,b) = \sum_{j=1}^{q^n} \mu_j \langle \mathcal{A}, u_j \rangle \overline{\langle \mathcal{B}, u_j \rangle}\,.$$

Let $N = J_n(q)$. By the definition of the Jordan totient function the number of vectors $a = (a_1, \dots, a_n)$ such that $a_1, \dots, a_n, q$ are coprime is exactly N. It is easy to see that $\mu_1 = q^{n-1}$ and $u_1(x) = N^{-1/2}(1, \dots, 1) \in \mathbb{R}^N$. Indeed, we fix b and thanks to the Chinese remainder theorem we need to we solve linear equation (11) modulo $p_j^{\omega_j}$, $j \in [t]$. Since $\lambda \in \mathbb{Z}_q^*$ and hence $\lambda \in \mathbb{Z}^*_{p_j^{\omega_j}}$ for all $j \in [t]$, it follows that not all coefficients of (11) are divided by p_j and hence there are $p_j^{\omega_j(n-1)}$ solutions modulo $p_j^{\omega_j}$. Hence there are q^{n-1} solutions in total. Thus we obtain

$$\mathcal{I}_\lambda(\mathcal{A}, \mathcal{B}) - \frac{q^{n-1}|\mathcal{A}||\mathcal{B}|}{N} = \sum_{j=2}^{q^n} \mu_j \langle \mathcal{A}, u_j \rangle \overline{\langle \mathcal{B}, u_j \rangle} := \mathcal{E}\,. \tag{14}$$

By the orthonormality of u_j and the Hölder inequality, we get

$$|\mathcal{E}| \leqslant |\mu_2| \sqrt{|\mathcal{A}||\mathcal{B}|}\,. \tag{15}$$

Thus it remains to estimate the second eigenvalue μ_2 and to do this we calculate the rectangular norm of the matrix M, that is

$$\sum_{j=1}^{q^n} |\mu_j|^4 = \sum_{a,a'} \left| \sum_b M(a,b) M(a',b) \right|^2 := \sigma\,,$$

and then $|\mu_2|$. Fixing a pair $(a, a') \in \mathbb{Z}_q^n \times \mathbb{Z}_q^n$, we need to solve the system of two linear equations

$$a_1 b_1 + \dots + a_n b_n \equiv \lambda \pmod q\,, \qquad a'_1 b_1 + \dots + a'_n b_n \equiv \lambda \pmod q\,. \tag{16}$$

It implies, in particular, that

$$\sum_{j=2}^{n} b_j (a'_1 a_j - a_1 a'_j) \equiv \lambda (a'_1 - a_1) \pmod q\,, \tag{17}$$

and if $a_1' - a_1 \in \mathbb{Z}_q^*$, say, then we obtain q^{n-2} solutions by the previous argument. If not, then consider all possible determinants of 2×2 matrices consisting of the elements of the matrix $(1, a_1, \dots, a_n | 1, a_1', \dots, a_n')$. Further given a tuple $(r_1, \dots, r_n)$, where $0 \leqslant r_j \leqslant \omega_j$ we consider the set $\mathcal{A}(r_1, \dots, r_n)$ of pairs $(a, a') \in \mathbb{Z}_q^n \times \mathbb{Z}_q^n$ such that $p_j^{r_j}$ is the maximal divisor of all these determinants. If $(a, a') \in \mathcal{A}(r_1, \dots, r_n)$, then $a_j \equiv a_j' \pmod{p_j^{r_j}}$ and hence

$$|\mathcal{A}(r_1, \dots, r_n)| \leqslant \frac{q^{2n}}{\prod_{j=1}^t p_j^{r_j n}} \,. \tag{18}$$

To solve (17) (recall that we consider the case when $(a, a') \in \mathcal{A}(r_1, \dots, r_n)$) one can use the Chinese remainder theorem again and we see that there are

$$\prod_{j=1}^t p_j^{\omega_j(n-2)+r_j} = q^{n-2} \prod_{j=1}^t p_j^{r_j}$$

solutions to equation (17). Combining the last bound with (18), we obtain

$$\sigma \leqslant q^{4n-4} \sum_{r_1 \leqslant \omega_1, \dots, r_t \leqslant \omega_t} \prod_{j=1}^t p_j^{-r_j(n-2)} = q^{4n-4} \Theta(n) \,,$$

where $\Theta(n) = O(1)$ for $n \geqslant 4$, $\Theta(3) = O(\log t)$ and $\Theta(2) \ll \prod_{j=1}^t (1 + \omega_j) = \tau(q)$. Let us remark other bounds for $\Theta(2)$ and for $\Theta(3)$, namely, from $m^{\tau(q)} \leqslant q$ one has $\Theta(2) \ll \tau(q) \ll \log_m q$ and, clearly, $\Theta(3) \ll 1 + t/m$.

It is instructive to consider the case $n = 2$ separately. Redefining the set $\mathcal{A}$, we need to solve the equation

$$a_1 b_1 - a_2 b_2 \equiv \lambda \pmod q \,, \quad (a_1, a_2) \in \mathcal{A}, \quad (b_1, b_2) \in \mathcal{B} \,, \tag{19}$$

where $a = (a_1, a_2)$ and $b = (b_1, b_2)$. It is clear that our equation (19) has the form $\det(a|b) \equiv \lambda \pmod q$ and hence we enjoy the following invariance property

$$M(a, b) = M(ga, gb) \,, \qquad \forall g \in \mathrm{SL}_2(\mathbb{Z}_q) \,. \tag{20}$$

Hence if f is an eigenfunction of M with the eigenvalue μ, then for $f^g(x) := f(gx)$ one has

$$\sum_{a,b} M(a, b) f^g(b) = \sum_{a,b} M(a, b) f(gb) = \sum_{a,b} M(ga, gb) f(gb) = \mu f(ga) = \mu f^g(a) \,,$$

where we have used (20) and the transitivity of the natural action of $\mathrm{SL}_2(\mathbb{Z}_q)$. In other words, $\mathrm{SL}_2(\mathbb{Z}_q)$ preserves the eigenspace L_μ, which corresponds to μ. Now consider an arbitrary eigenfunction u_j, $j > 1$. We know that $\sum_x u_j(x) = 0$ and hence u_j is not a constant function. There are many ways to see that $\dim(L_{\mu_j}) > 1$ or, in other words, that $\langle \{u_j^g\}_{g \in \mathrm{SL}_2(\mathbb{Z}_q)} \rangle \neq \langle u_j \rangle = L_{\mu_j}$. For example, one can use the transitivity again. Another approach is to notice that the group $\mathrm{SL}_2(\mathbb{Z}_q)$ has no non–trivial one–dimensional representations but, on the other hand, any one–dimensional eigenspace would give us a character (the same holds in the general case which will be considered below). Thus anyway we conclude that for any $j > 1$ the multiplicity of each μ_j is at least the minimal dimension of non–trivial representations of $\mathrm{SL}_2(\mathbb{Z}_q)$.

Now we essentially repeat the argument from [34, Theorem 12]. Another way is to use the first part of [3, Theorem 1] which says exactly the same. So, let us repeat what is known about representation theory of the group $\mathrm{SL}_2(\mathbb{Z}_q)$, see [5, Sections 7, 8]. First of all, for any irreducible representation ρ_q of $\mathrm{SL}_2(\mathbb{Z}_q)$ we have $\rho = \rho_q = \rho_{p_1^{\rho_1}} \otimes \dots \otimes \rho_{p_t^{\rho_t}}$ and hence it is sufficient to understand representation theory for $\mathrm{SL}_2(\mathbb{Z}_{p^d})$, where p is a prime number and d is a positive integer. Now by [5, Lemma 7.1] we know that for any odd prime the dimension of any faithful irreducible representation of $\mathrm{SL}_2(\mathbb{Z}_{p^d})$ is at least $2^{-1}p^{d-2}(p-1)(p+1)$ (a similar proof for $\mathrm{SL}_n(\mathbb{Z}_{p^d})$, $n \geqslant 2$ can be found in [3, Theorem 1]). If $d = 1$, then the classical result of Frobenius [11] says that the minimal dimension of any non–trivial representation is at least $(p-1)/2$. For an arbitrary positive integer $r \leqslant d$ we can consider the natural projection $\pi_r : \mathrm{SL}_2(\mathbb{Z}_{p^d}) \to \mathrm{SL}_2(\mathbb{Z}_{p^r})$ and let $H_r = \mathrm{Ker}\, \pi_r$. One can show that the set $\{H_r\}_{r \leqslant d}$ gives all normal subgroups of $\mathrm{SL}_2(\mathbb{Z}_{p^d})$ and hence any nonfaithful irreducible representation arises as a faithful irreducible representation of $\mathrm{SL}_2(\mathbb{Z}_{p^r})$ for a certain $r < d$. Anyway, we see that the multiplicity (dimension) d_ρ of any non–trivial irreducible representation ρ of $\mathrm{SL}_2(\mathbb{Z}_{p^d})$ is at least $(m-1)/2 \geqslant m/3$.

Returning to the quantity σ, we get (below $n = 2$)

$$|\mu_2|^4 m \leqslant 3\sigma \leqslant 3q^{4n-4}\Theta(n)$$

and hence

$$|\mu_2| \leqslant \left(3m^{-1}q^{4n-4}\Theta(n)\right)^{1/4} = q^{n-1}(3\Theta(n)m^{-1})^{1/4} \tag{21}$$

Recalling (14), (15), we obtain the required result for $n = 2$.

Now let $n > 2$. It remains only to find a good lower bound for the multiplicity of μ_j, $j > 1$ (the fact $\dim(L_{\mu_j}) > 1$ is immediate consequence that $SO_n(\mathbb{Z}_q)$ has no non–trivial one–dimensional representations or thus see paper [28]). In the higher–dimensional case $n > 2$ our form $\Phi(a, b) = a_1b_1 + \dots + a_nb_n$ is preserved by the group of orthogonal transformations $O_n(\mathbb{Z}_q)$ (as well as $SO_n(\mathbb{Z}_q)$) and hence our task is to find a good lower bound for the dimension of any non–trivial irreducible representation of $SO_n(\mathbb{Z}_{p^d})$. Using Theorem 6 and the arguments as above, we see that it is enough to have deal with faithful representations and this problem was solved in [2]. The authors prove that the minimal dimension of any faithful representations coincides (up to constants) with the classical lower bound for minimal dimension of an arbitrary non–trivial representation for split Chevalley groups over $\mathbb{F}_{p^d}$, see [21], [32]. These results combining with the existence of isomorphisms between low–dimensional classical groups (see [19, Proposition 2.9.1], for example) give us $d_\rho \geqslant 2^{-2}p^{n-3}$ for $n \geqslant 4$ and $d_\rho \geqslant 2^{-2}p$ for $n = 3$. For $n = 4$ one cannot apply Theorem 6 but it is easy to see that in this case the multiplicity of μ_2 is at least $d_\rho \geqslant 2^{-1}(p-1)$ due to the fact that the group $\mathrm{SL}_2(\mathbb{Z}_{p^d}) \times \mathrm{SL}_2(\mathbb{Z}_{p^d})$ acts on the quadruples $(a_1, \dots, a_4)$ and we can use previous arguments concerning $\mathrm{SL}_2(\mathbb{Z}_{p^d})$ and the case $n = 2$. It follows that for any $n \geqslant 2$ the multiplicity of μ_2 is at least $\Omega(m^{-n_*})$. This completes the proof. $\square$

Thus, as the reader can see, our method almost automatically gives some asymptotic formulae for the number of solutions to systems of equations that are preserved by the actions of certain groups. The only thing we need to calculate is the first eigenfunction of the correspondent operator and its rectangular norm. After that we use quasi–random technique in the spirit of papers [12], [15] and [31].

Now we are ready to obtain Theorem 1 from the introduction and for simplicity we consider the case of a prime number q. We assume that the sets $\mathcal{A}$, $\mathcal{B}$ consisting of linearly independent tuples because otherwise there is no solutions to equation (4).

Theorem 8 *Let q be an odd prime number and $\lambda \neq 0$. Then*

$$2^{-3} \left| \mathcal{D}_\lambda(\mathcal{A},\mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q} \right| \leqslant q^{d^2/2-d/4-3/4} \sqrt{|\mathcal{A}||\mathcal{B}|} + \frac{|\mathcal{A}||\mathcal{B}|}{q^2} \,. \tag{22}$$

Proof. The case $n = m = 1$ was considered in Theorem 7, so we assume that $\min\{n,m\} \geqslant 2$. Let $a = (a_1,\dots,a_n)$, $b = (b_1,\dots,b_m)$ and let $M(a,b) = 1$ iff the pair (a,b) satisfies our equation (4). Considering the singular decomposition of the matrix $M(a,b)$, we obtain

$$M(a,b) = \sum_{j=1}^{q^d} \lambda_j u_j(a) \overline{v}_j(b) \,,$$

where $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant 0$ are the singularvalues and u_j, v_j are correspondent orthonormal singularfunctions. Let

$$\mathcal{N} = (q^d-q^m)(q^d-q^{m+1})\dots(q^d-q^{d-1}) = q^{dn} \prod_{j=1}^{n} (1-q^{-j}) \quad \text{and} \quad \mathcal{M} = q^{dm} \prod_{j=1}^{m} (1-q^{-j}) \,.$$

It is easy to calculate λ_1 and to show that $u_1(a) = \mathcal{N}^{-1/2}(1,\dots,1) \in \mathbb{R}^{\mathcal{N}}$, as well as $v_1(b) = \mathcal{M}^{-1/2}(1,\dots,1) \in \mathbb{R}^{\mathcal{M}}$. Indeed, for any fixed a or b we need to solve the equation $\det(a|b) = \lambda$ in b or a, correspondingly. It is easy to see that the equation $\det(a|b) = \lambda$, a is fixed, has $q^{dm-1} \prod_{j=2}^m \left(1-q^{-j}\right) = \frac{\mathcal{M}}{q-1}$ solutions due to the number of independent vectors over $\mathbb{Z}_q$. Similarly, the second equation has $q^{dn-1} \prod_{j=2}^n \left(1-q^{-j}\right) = \frac{\mathcal{N}}{q-1}$ solutions in a. Thus, these numbers do not depend on a and b and hence, indeed we have $u_1(a) = \mathcal{N}^{-1/2}(1,\dots,1) \in \mathbb{R}^{\mathcal{N}}$, $v_1(b) = \mathcal{M}^{-1/2}(1,\dots,1) \in \mathbb{R}^{\mathcal{M}}$ and

$$\lambda_1 = \langle M u_1, v_1 \rangle = \frac{\mathcal{M}}{q-1} \cdot \mathcal{N} \cdot (\mathcal{M}\mathcal{N})^{-1/2} = \frac{\sqrt{\mathcal{M}\mathcal{N}}}{q-1} \,.$$

Thus we get

$$\left| \mathcal{D}_\lambda(\mathcal{A},\mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q-1} \right| = \left| \sum_{j=2}^{q^d} \lambda_j \langle \mathcal{A}, u_j \rangle \overline{\langle \mathcal{B}, v_j \rangle} \right| \leqslant \lambda_2 \sqrt{|\mathcal{A}||\mathcal{B}|} \,. \tag{23}$$

As above we need to estimate the rectangular norm of the matrix M that is

$$\sum_{j=1}^{q^d} \lambda_j^4 = \sum_{a,a'} \left| \sum_b M(a,b) M(a',b) \right|^2 \,,$$

and thus we arrive to the system of equations $\det(a'|b) = \det(a|b) = \lambda$ with fixed a and a'. Fixing vectors $b_1, \dots, b_{m-1}$ we have exactly equation (16) which has at most q^{md-2} solutions. Thus

$$\sum_{j=1}^{q^d} \lambda_j^4 \leqslant q^{2nd} q^{2md-4} = q^{2d^2-4} \,. \tag{24}$$

Now it is easy to see that

$$M(ga, gb) = M(ga_1, \dots, ga_n, gb_1, \dots, gb_m) = M(a,b)$$

for an arbitrary $g \in \mathrm{SL}_d(\mathbb{Z}_q)$ and thus any λ_j, $j > 1$ has multiplicity equals the minimal dimension of any non–trivial irreducible representation of $\mathrm{SL}_d(\mathbb{Z}_q)$. Thus the multiplicity of λ_2 is at most $2^{-2} q^{d-1}$ and hence

$$\lambda_2 \leqslant 2 q^{d^2/2-1} q^{-(d-1)/4} = 2 q^{d^2/2-d/4-3/4} \,.$$

Using the last estimate, and returning to formula (23), we obtain

$$2^{-3} \left| \mathcal{D}_\lambda(\mathcal{A}, \mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q} \right| \leqslant q^{d^2/2-d/4-3/4} \sqrt{|\mathcal{A}||\mathcal{B}|} + \frac{|\mathcal{A}||\mathcal{B}|}{q^2}$$

as required. $\Box$

Finally, we consider an example with the cross–ratio $[a,b,c,d] := \frac{(a-c)(b-d)}{(a-d)(b-c)}$. As one can see, representation theory almost immediately gives asymptotic formula (26) with an acceptable error term. Let q be a prime number, $\lambda \in \mathbb{Z}_q$ and $\mathcal{A} \subseteq \mathbb{Z}_q \times \mathbb{Z}_q$, $\mathcal{B} \subseteq \mathbb{Z}_q \times \mathbb{Z}_q$ be sets. Define

$$\mathcal{C}_\lambda(\mathcal{A}, \mathcal{B}) := |\{(a_1,a_2) \in \mathcal{A},\, (b_1,b_2) \in \mathcal{B} \;:\; [a_1,a_2,b_1,b_2] \equiv \lambda \pmod q\}| \,. \tag{25}$$

Theorem 9 *Let q be a prime number, $\lambda \in \mathbb{Z}_q$, $\lambda \neq 0,1$ and $\mathcal{A} \subseteq \mathbb{Z}_q \times \mathbb{Z}_q$, $\mathcal{B} \subseteq \mathbb{Z}_q \times \mathbb{Z}_q$ be sets. Then*

$$\left| \mathcal{C}_\lambda(\mathcal{A}, \mathcal{B}) - \frac{|\mathcal{A}||\mathcal{B}|}{q} \right| \leqslant 4 q^{3/4} \sqrt{|\mathcal{A}||\mathcal{B}|} \,. \tag{26}$$

P r o o f. As usual let $a = (a_1, a_2)$, $b = (b_1, b_2)$ and let $M(a,b) = 1$ iff the pair (a,b) satisfies our equation (25). It is well–known that $\mathrm{SL}_2(\mathbb{Z}_q)$ preserves the cross–ratio in the sense

$$M(ga, gb) = M(ga_1, ga_2, gb_1, gb_2) = M(a,b) \,. \tag{27}$$

Considering the unitary decomposition of the hermitian matrix $M(a,b)$ as in (13) we see that the property $u_1(a) = q^{-1}(1,\dots,1) \in \mathbb{R}^{q^2}$ automatically follows from (27) and 2–transitivity of $\mathrm{SL}_2(\mathbb{Z}_q)$ on the projective line. It remains to calculate the rectangular norm of the matrix M, that is to solve the system $[x,y,c,d] = [x,y,c',d'] = \lambda$. It follows that

$$xy(1-\lambda) + (\lambda c - d)x + (\lambda d - c)y + dc(1-\lambda) = 0 \,, \tag{28}$$

and

$$xy(1-\lambda)+(\lambda c'-d')x+(\lambda d'-c')y+d'c'(1-\lambda)=0\,. \tag{29}$$

Subtracting (29) from (28) we arrive to the equation

$$x(\lambda(c-c')+d'-d)+y(\lambda(d-d')+c'-c)+(1-\lambda)(dc-d'c')=0 \tag{30}$$

and this is a non–trivial equation excluding two cases: $c=c'$, $d=d'$ and $\lambda=-1$, $c=d'$, $d=c'$. If equation (30) is non–trivial, then we substitute, say, x into (28) and obtain at most 4 solutions in x,y (one can check that we obtain a non–trivial equation thanks to our condition $\lambda\neq 0,1$). In the exceptional cases we have just one equation, say, (28), and it is easy to see that our equation has at most $2q$ solutions. Thus

$$\sum_{j=1}^{q^2}\mu_j^4=\sum_{a,a'}\left|\sum_b M(a,b)M(a',b)\right|^2\leqslant 16q^4+2q^2(2q)^2=24q^4\,.$$

It remains to use the Frobenius Theorem [11] about minimal representations of $\mathrm{SL}_2(\mathbb{Z}_q)$. This result gives us the bound $\mu_2\leqslant 4q^{3/4}$ and we can apply the arguments as in the proofs of Theorems 7, 8. This completes the proof. $\Box$

4 On sums with multiplicative characters over some manifolds and other applications

In this section we want to extend representation theory methods to some sums with multiplicative characters. Below p is a prime number and $\mathbb{F}$ is a finite field of characteristic p. Let us consider a basic example. We know that $\mathrm{SL}_2(\mathbb{F})$ acts on the projective line and it gives us an irreducible representation of this group but from [18], say, it is well–known that there are other irreducible representations of $\mathrm{SL}_2(\mathbb{F})$ and a half of them are connected with "projective lines" equipped with multiplicative characters χ. More precisely, it means that we consider the family of functions $f:\mathbb{F}\times\mathbb{F}\to\mathbb{C}$ such that

$$f(\lambda x,\lambda y)=\chi(\lambda)f(x,y)\,,\qquad \forall\lambda\in\mathbb{F}^*\quad\text{ and }\quad\forall(x,y)\in(\mathbb{F}\times\mathbb{F})\setminus\{0\}\,, \tag{31}$$

and now $\mathrm{SL}_2(\mathbb{F})$ acts on this family, as well as on $\mathbb{F}\times\mathbb{F}$ in a natural way. In our results below we do not need to use the knowledge of concrete irreducible representations of $\mathrm{SL}_2(\mathbb{F})$ (and other groups) but we will use only definition (31) somehow.

Let us start with the following auxiliary proposition concerning summation over a hyperbolic surface (twisted by a multiplicative character) in the spirit of paper [35], say.

Proposition 10 *Let $A,B\subseteq\mathbb{F}_p$ and $G\subseteq\mathrm{GL}_2(\mathbb{F}_p)$ be sets and χ be a non–trivial multiplicative character. Also, let $c_A:A\to\mathcal{D}$, $c_B:B\to\mathcal{D}$ be some weights. Then for any integer $k\geqslant 2$ the following holds*

$$2^{-2}\left|\sum_{a,b}c_A(a)c_B(b)\sum_{g\in G\,:\,ga=b}\chi(\gamma a+\delta)\right|$$

$$\leqslant \sqrt{|A||B||G|} \cdot \mathsf{T}_{2k}^{1/8k}(f_G) + \sqrt{|A||B|}|G| \cdot (\max\{|A|,|B|\})^{-1/2k}\,. \tag{32}$$

P r o o f. Consider the functions $\mathcal{A}(\lambda a,\lambda) = c_A(a)\overline{\chi(\lambda)} = \mathcal{A}(\overline{x})$, $\mathcal{B}(\mu b,\mu) = c_B(b)\chi(\mu) = \mathcal{B}(\overline{y})$, where $a \in A$, $b \in B$, $\overline{x} = (x_1,x_2)$, $\overline{y} = (y_1,y_2)$ and μ,λ run over $\mathbb{F}_p^*$. It is easy to see that we always have $\sum_{a,\lambda} \mathcal{A}(\lambda a,\lambda) = 0$, as well as $\sum_{b,\mu} \mathcal{B}(\mu b,\mu) = 0$ since χ is a non–trivial character. Notice that

$$\sigma := \sum_{\overline{x},\overline{y}} \mathcal{A}(\overline{x})\mathcal{B}(\overline{y}) \sum_{g\in G\,:\,g\overline{x}=\overline{y}} 1 = (p-1)\sum_{a,b} c_A(a)c_B(b) \sum_{g\in G\,:\,ga=b} \chi(\gamma a+\delta) \tag{33}$$

for any trivial/non–trivial multiplicative character χ. We can interpret the left–hand side of (33) as the number of some points on a hyperbolic surface counting with weights $\mathcal{A}(\overline{x})$, $\mathcal{B}(\overline{y})$. The Hölder inequality (see [33, Lemma 13]) gives us

$$\sigma^{2k} \leqslant \|\mathcal{A}\|_2^{2k} \|\mathcal{B}\|_2^{2k-2} \sum_h f^{(k)}_{GG^{-1}}(h) \sum_x \mathcal{B}(x)\mathcal{B}(hx)\,.$$

Applying identity (33), it is easy to see that the contribution of the terms with $\sum_x \mathcal{B}(x)\mathcal{B}(hx) \leqslant 32p$, say, corresponds to the second term from (32). Now using [33, Lemma 12] (we notice that, say, 4 different points uniquely determine the transformation from $\mathrm{GL}_2(\mathbb{F}_p)$), combining with the Hölder inequality again, we derive

$$\sigma^{2k} \leqslant (|A|(p-1))^k(|B|(p-1))^{k-1} \left(\sum_h |f^{(k)}_{GG^{-1}}(h)|^{4/3}\right)^{3/4} \left(\sum_h \left(\sum_x \mathcal{B}(x)\mathcal{B}(hx)\right)^4\right)^{1/4}$$

$$\leqslant (|A|(p-1))^k(|B|(p-1))^{k-1} \left(\sum_h (f^{(k)}_{GG^{-1}}(h))^2\right)^{1/4} \left(\sum_h |f^{(k)}_{GG^{-1}}(h)|\right)^{1/2} \left(\sum_h \left(\sum_x \mathcal{B}(x)\mathcal{B}(hx)\right)^4\right)^{1/4}$$

$$\leqslant 4^k(|A|(p-1))^k(|B|(p-1))^{k-1}\mathsf{T}_{2k}^{1/4}(f_G)|G|^k \cdot |B|(p-1)\,. \tag{34}$$

Recalling (33), we see that estimate (34) is equivalent to the required bound (32). This completes the proof. □

Now we obtain some concrete applications of Proposition 10, which correspond to Theorems 2, 3 of the introduction. Let $A,B,X,Y \subseteq \mathbb{F}_p$ be sets. Consider the equation

$$(a+x)(b+y) \equiv 1 \pmod{p} \tag{35}$$

or, in other words, $y = -b + 1/(a+x) = g_{a,b}x$, where $\det(g_{a,b}) = -1$. The energy $\mathsf{T}_{2k}(f_G)$ of the correspondent family of transformations $G = \{g_{a,b}\}_{a\in A, b\in B}$ was estimated many times see, e.g., paper [33]. Applying Proposition 10 to this particular case of equation (35), we obtain

Corollary 11 *Let $\delta > 0$ be a real number, $A, B, X, Y \subseteq \mathbb{F}_p$ be sets, let χ be a non–principal multiplicative character and $|X||Y| \geqslant p^\delta$. Also, let $c_A : A \to \mathcal{D}$, $c_B : B \to \mathcal{D}$ be some weights. Then there is $\varepsilon(\delta) > 0$ such that*

$$\sum_{a,b,x,y \,:\, (a+x)(b+y)=1} c_A(a) c_B(b) X(x) Y(y) \chi(a+x) \leqslant \sqrt{|A||B|}(|X||Y|)^{1-\varepsilon(\delta)}\,.$$

The above corollary immediately implies Theorem 3 from the introduction (compare with [33, Theorems 4, 33]) which we recall here for the reader's convenience. Other results of paper [33] can be obtained in a similar way for bilinear sums $S_\chi(\alpha,\beta)$ with non–trivial characters χ.

Corollary 12 *Let $c > 0$ and p be a prime number. Let $t_1, t_2 \in \mathbb{F}_p$, N, M be integers, $N, M \leqslant p^{1-c}$ and let $\alpha, \beta : \mathbb{F}_p \to \mathbb{C}$ be functions supported on $\{1, \dots, N\} + t_1$ and $\{1, \dots, M\} + t_2$, respectively. Then there exists $\delta(c) > 0$ such that*

$$S_\chi(\alpha,\beta) \lesssim \|\alpha\|_2 \|\beta\|_2 p^{1-\delta}\,. \tag{36}$$

Now consider the case when our set A is a collection of disjoint intervals. It is an important family of sets, including discrete fractal sets see, e.g., papers [4], [7], [23]—[26] and [42].

Theorem 13 *Let $\Lambda \subset \mathbb{F}_p$, $I = [N]$, $A = I \dotplus \Lambda$, $|A| > p^{1-\epsilon}$, and χ be a non-principal multiplicative character. Then there is an absolute constant $c_* > 0$ such that*

$$|\sum_{x \in A \cap A^{-1}} \chi(x)| \leqslant |A \cap A^{-1}| \cdot N^{-c_*} \ll \frac{|A|^2}{p} \cdot N^{-c_*}\,, \tag{37}$$

provided $N \geqslant p^{\epsilon/c_*}$.

P r o o f. We combine an appropriate version of Corollary 11 and the well–known Bourgain–Gamburd machine [6] applied to equation (35) see, e.g., [23]. Indeed, for any $x \in A \cap A^{-1}$, we have $x = i + \lambda$ such that $1 = (i+\lambda)(i'+\lambda')$, where $i, i' \in I$ and $\lambda, \lambda' \in \Lambda$. Thus we in very deed arrive to equation (35). Now $I(i) \leqslant N^{-1}(I * \overline{I})(i)$, where $\overline{I} = [-N, N]$ and hence the number of solutions to the equation $1 = (i+\lambda)(i'+\lambda')$ can be bounded above as $1 = (j+a)(j'+a')$ with $a, a' \in A$ and $j, j' \in \overline{I}$ (times N^{-2}, of course). In particular (see [23] or just Proposition 10 and Corollary 11 above), we get for an absolute constant $c \in (0, 1]$ that

$$|A \cap A^{-1}| \leqslant \frac{|A|^2 |\overline{I}|^2}{N^2 p} + O(N^{-2}|A| \cdot N^{2-c}) \ll \frac{|A|^2}{p} \tag{38}$$

and hence the second estimate of (37) follows from the first one. Here we have used the conditions that $|A| > p^{1-\epsilon}$ and $N \geqslant p^{\epsilon/c}$, which is satisfied if we put $c_* = c/4$, say.

Similarly, let $h \in [N]$ be an integer parameter and write $I(i) = h^{-1}(H * I)(i) + \varepsilon(i)$, where $H = [h]$ and $\|\varepsilon\|_\infty = 1$, $|\mathsf{supp}(\varepsilon)| \leqslant 2h$. In particular, we have $\|\varepsilon\|_2^2 \leqslant 2h$ and one can threat ε as a sum of two functions $\varepsilon_1, \varepsilon_2$ with supports on some shifts of the interval H. Put $\tilde{\varepsilon} = \varepsilon_1 + \varepsilon_2 : H \to [-1, 1]$. As always let us write

$$\sum_{x \in A \cap A^{-1}} \chi(x) = \sum_{1=(i+\lambda)(i'+\lambda')} \chi(\lambda + i)\Lambda(\lambda)\Lambda(\lambda')I(i)I(i')$$

$$= \sum_{1=(i+\lambda)(i'+\lambda')} \chi(\lambda + i)\Lambda(\lambda)\Lambda(\lambda')(h^{-1}(H * I)(i) + \varepsilon(i))(h^{-1}(H * I)(i') + \varepsilon(i'))$$

$$= h^{-2} \sum_{a,a',h,h' \,:\, (a+h)(a'+h')=1} A(a)A(a')H(h)H(h')\chi(a+h) + \mathcal{E} = \sigma + \mathcal{E}\,,$$

where the error term $\mathcal{E}$ can be estimated as (there are better bounds as the set A is I–invariant and not just H–invariant)

$$|\mathcal{E}| \leqslant 2h^{-1} \sum_{a,a',h,h' \,:\, (a+h)(a'+h')=1} \Lambda(a)A(a')|\tilde{\varepsilon}(h)|H(h') + \sum_{a,a',h,h' \,:\, (a+h)(a'+h')=1} \Lambda(a)\Lambda(a')|\tilde{\varepsilon}(h)\tilde{\varepsilon}(h')|$$

$$\ll \frac{|A|^2}{p}\left(\frac{h}{N} + \frac{h^2}{N^2}\right) + |A| \cdot \left(\frac{h^{1-c}}{\sqrt{N}} + \frac{h^{2-c}}{N}\right) \ll \frac{|A|^2 h}{pN} + \frac{|A|h^{1-c}}{\sqrt{N}}\,. \tag{39}$$

Here we have assumed that $h \leqslant \sqrt{N}$ and applied the well–known Bourgain–Gamburd machine [6], [23]. Recall that this result replaces Corollary 11 in the case when X, Y are intervals and $\chi \equiv 1$ (that is why we need two additional main terms in (39)). It remains to estimate the sum σ and to do this we can use the Bourgain–Gamburd machine one more time, namely, we apply our Corollary 11 and get $\sigma \ll |A|h^{-c}$. Finally, combining the estimate for σ and bound (39) for the error term $\mathcal{E}$, choosing the parameter $h = [\sqrt{N}]$, we obtain

$$\sum_{x \in A \cap A^{-1}} \chi(x) \ll |A|h^{-c} + \frac{|A|^2 h}{pN} + \frac{|A|h^{1-c}}{\sqrt{N}} \ll |A|h^{-c} \ll |A|N^{-c/2} \ll \frac{|A|^2}{p} \cdot N^{-c/4}$$

thanks to our assumptions $|A| > p^{1-\epsilon}$ and $N \geqslant p^{4\epsilon/c}$. The same calculations show that there is an asymptotic formula for $|A \cap A^{-1}|$ and, in particular, the inverse inequality to (38) takes place. It gives us the first inequality in (37). This completes the proof. $\Box$

It is well–known and it is easy to see that the multiplicative equation (35) is almost coincides (up to some transformation) with the additive equation

$$\frac{1}{x+a} - \frac{1}{y+b} \equiv 1 \pmod p\,,$$

where $a \in A$, $b \in B$, $x \in X$, $y \in Y$. Thus we obtain an analogue of Theorem 13.

Theorem 14 *Let $\Lambda \subset \mathbb{F}_p$, $I = [N]$, $A = I \dotplus \Lambda$, $|A| > p^{1-\epsilon}$, and χ be a non–principal multiplicative character. Then there is an absolute constant $c_* > 0$ such that*

$$|\sum_{x \in A^{-1} \cap (A^{-1}+1)} \chi(x)| \leqslant |A^{-1} \cap (A^{-1}+1)| \cdot N^{-c_*} \ll \frac{|A|^2}{p} \cdot N^{-c_*}\,, \tag{40}$$

provided $N \geqslant p^{\epsilon/c_*}$.

Now we are ready to prove Theorem 5 from the introduction.

P r o o f. We follow the scheme and the notation of the proof from paper [24, Pages 3–7]. It was shown that the set of $a \in A$ satisfying (8) contains a set of the form $Z_M \cap Z_M^{-1}$, $|A| \sim |Z_M|^2/p$, $|Z_M| \sim p^{w_M + 2\varepsilon(1-w_M)}$ and the set Z_M is a disjoint union of some shifts of an interval of length $N \sim p^{2\varepsilon}$, where $\varepsilon \gg 1/M$ is a parameter and Hausdorff dimension w_M enjoys the asymptotic formula $w_M = 1 - O(1/M)$, $M \to \infty$. Thus we can apply Theorem 13 and write

$$|A \cap \Gamma| = (p-1)^{-1} \sum_{\chi} \left(\sum_{x\in A} \chi(x) \right) \left(\sum_{x \in \Gamma} \overline{\chi}(x) \right) \geqslant \frac{|A||\Gamma|}{p-1} - C|A|N^{-c_*} > 0 \,,$$

where $C, c_* > 0$ are some absolute constants. Here we have used conditions (9), (10), the fact that $M \sim \frac{\log p}{\log\log p}$ and $\varepsilon \gg 1/M$. It remains to check that $N \geqslant p^{\epsilon/c_*}$ or, in other words, that $\varepsilon \gg \epsilon$. Since $|Z_M| \sim p^{w_M+2\varepsilon(1-w_M)} = p^{1-\epsilon}$, it follows that $\epsilon = (1-w_M)(1-2\varepsilon) \ll 1/M$ and thus the required condition takes place. This completes the proof. $\Box$

Let us make a final remark. Loosely, Theorem 5 gives us a non–trivial bound for the *multiplicative energy* of the set Z_M, see formula (42) below. Nevertheless, the last fact follows from the circumstance that Z_M is an Ahlfors–David set, [4], that is for an arbitrary $z \in Z_M$ one has

$$|Z_M \cap (\mathcal{D} + z)| \sim_M |\mathcal{D}|^{w_M} N^{1-w_M} \tag{41}$$

for any interval $\mathcal{D}$, $|\mathcal{D}| \geqslant N$ with the center at the origin. Recall that in [4] a non–trivial upper bound was obtained for the *additive energy* of any Ahlfors–David set. Let us briefly prove an upper estimate for the multiplicative energy of an arbitrary Ahlfors–David set Z_M, having large Hausdorff dimension w_M. The advantage of bound (43) below that our power saving can be expressed in terms of $|Z_M|$ but not just N.

Namely, write $Z = Z_M$, $w = w_M$ and then $|Z| \sim_M p^w N^{1-w}$. Also, put $\delta \sim_M \Delta^w N^{1-w}$, where Δ is a parameter. By the points/planes incidences in $\mathbb{F}_p$ (see [29]) and property (41) one has

$$\begin{aligned} \mathsf{E}^\times(Z) &:= |\{(z_1,z_2,z_3,z_4) \in Z^4 ~:~ z_1 z_2 = z_3 z_4\}| \\ &\ll \delta^{-2} |\{(z_1,z_2,z'_1,z'_2,d,d') \in Z^4 \times [\Delta]^2 ~:~ z_1(z_2+d) \equiv z'_1(z'_2+d')\}| \\ &\ll \delta^{-2} \left(\frac{|Z|^4\Delta^2}{p} + |Z|^3 \Delta^{3/2} \right) \ll \delta^{-2} p^3 \,, \end{aligned} \tag{42}$$

where the optimal choice for Δ is $\Delta = (p/|Z|)^2$. Thus

$$\mathsf{E}^\times(Z) \ll_M |Z|^3 (p/|Z|)^{3-4w} N^{-2(1-w)} \sim_M |Z|^3 \cdot |Z|^{-\frac{(4w-3)(1-w)}{w}} N^{\frac{-(1-w)(3-2w)}{w}} < |Z|^3 \tag{43}$$

for $w > 3/4$. Thus, we have a power saving in terms of $|Z|$ for the multiplicative energy of any Ahlfors–David set.

5 Data availability

No data was used for the research described in the article.

References

[1] O. Ahmadi and I. E. Shparlinski. Distribution of matrices with restricted entries over finite fields. *Indagationes Mathematicae*, 18(3):327–337, 2007.

[2] M. Bardestani, C. Karimianpour, K. Mallahi-Karai, and H. Salmasian. Faithful representations of Chevalley groups over quotient rings of non-Archimedean local fields. *Groups, Geometry, and Dynamics*, 11(1):57–74, 2017.

[3] M. Bardestani and K. Mallahi-Karai. Quasi-random profinite groups. *Glasgow Mathematical Journal*, 57(1):181–200, 2015.

[4] J. Bourgain and S. Dyatlov. Spectral gaps without the pressure condition. *Annals of Mathematics*, 187(3):825–867, 2018.

[5] J. Bourgain and A. Gamburd. Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$: I. *Journal of the European Mathematical Society*, 10(4):987–1011, 2008.

[6] J. Bourgain and A. Gamburd. Uniform expansion bounds for cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Annals of Mathematics*, pages 625–642, 2008.

[7] M.-C. Chang. Partial quotients and equidistribution. *Comptes Rendus Mathematique*, 349(13-14):713–718, 2011.

[8] D. Covert, A. Iosevich, and J. Pakianathan. Geometric configurations in the ring of integers modulo p^l. *Indiana university mathematics journal*, pages 1949–1969, 2012.

[9] P. Erdős and E. Szemerédi. On sums and products of integers. *Studies in pure mathematics: to the memory of Paul Turán*, pages 213–218, 1983.

[10] A. Fish. On product of difference sets for sets of positive density. *Proceedings of the American Mathematical Society*, 146(8):3449–3453, 2018.

[11] G. Frobenius. Über Gruppencharaktere, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin. 1896.

[12] A. Gamburd. On the spectral gap for infinite index "congruence" subgroups of $\mathrm{SL}_2(\mathbb{Z})$. *Israel Journal of Mathematics*, 127(1):157–200, 2002.

[13] M. Garaev. The sum-product estimate for large subsets of prime fields. *Proceedings of the American Mathematical Society*, 136(8):2735–2739, 2008.

[14] K. Gyarmati and A. Sárközy. Equations in finite fields with restricted solution sets. II (Algebraic equations). *Acta Mathematica Hungarica*, 119(3):259–280, 2008.

[15] M. N. Huxley. Exceptional eigenvalues and congruence subgroups. *The Selberg trace formula and related topics (Brunswick, Maine, 1984)*, 53:341–349, 1986.

[16] A. Iosevich and M. Rudnev. Erdös distance problem in vector spaces over finite fields. *Transactions of the American Mathematical Society*, 359(12):6127–6142, 2007.

[17] D. James. On the structure of orthogonal groups over local rings. *American Journal of Mathematics*, 95(2):255–265, 1973.

[18] A. A. Kirillov. *Elements of the Theory of Representations*, volume 220. Springer Science & Business Media, 2012.

[19] P. B. Kleidman and M. W. Liebeck. *The subgroup structure of the finite classical groups*, volume 129. Cambridge University Press, 1990.

[20] W. Klingenberg. Orthogonal groups over local rings. 1961.

[21] V. Landazuri and G. M. Seitz. On the minimal degrees of projective representations of the finite chevalley groups. 1974.

[22] A. Mohammadi and S. Stevens. Attaining the exponent 5/4 for the sum-product problem in finite fields. *International Mathematics Research Notices*, 2023(4):3516–3532, 2023.

[23] N. Moshchevitin, B. Murphy, and I. Shkredov. Popular products and continued fractions. *Israel J. Math.*, 238(2):807–835, 2020.

[24] N. Moshchevitin, B. Murphy, and I. Shkredov. On Korobov bound concerning Zaremba's conjecture. *arXiv:2212.14646*, 2022.

[25] N. G. Moshchevitin. Sets of the form $A + B$ and finite continued fractions. *Sbornik: Mathematics*, 198(4):537–557, 2007.

[26] N. G. Moshchevitin and D. M. Ushanov. On Larcher's theorem concerning good lattice points and multiplicative subgroups modulo p. *Uniform Distribution Theory*, 5(1):45–52, 2010.

[27] B. Murphy and G. Petridis. A point-line incidence identity in finite fields, and applications. *arXiv preprint arXiv:1601.03981*, 2016.

[28] B. Nica. Unimodular graphs and Eisenstein sums. *Journal of Algebraic Combinatorics*, 45:423–454, 2017.

[29] M. Rudnev. On the number of incidences between points and planes in three dimensions. *Combinatorica*, 38:219–254, 2018.

[30] M. Rudnev and S. Stevens. An update on the sum-product problem. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 173, pages 411–430. Cambridge University Press, 2022.

[31] P. Sarnak and X. Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 1991.

[32] G. M. Seitz and A. E. Zalesskii. On the minimal degrees of projective representations of the finite Chevalley groups, II. *Journal of Algebra*, 158(1):233–243, 1993.

[33] I. D. Shkredov. Modular hyperbolas and bilinear forms of Kloosterman sums. *Journal of Number Theory*, 220:182–211, 2021.

[34] I. D. Shkredov. On some multiplicative properties of large difference sets. *arXiv:2301.09206v1*, 2023.

[35] A. N. Skorobogatov. Exponential sums, the geometry of hyperplane sections, and some Diophantine problems. *Israel Journal of Mathematics*, 80:359–379, 1992.

[36] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[37] P. V. Thang and L. A. Vinh. Erdős–Rényi graph, Szemerédi–Trotter type theorem, and sum-product estimates over finite rings. *Forum Mathematicum*, 27(1):331–342, 2015.

[38] N. V. The and L. A. Vinh. Dot-Product Sets and Simplices Over Finite Rings. *Journal of Fourier Analysis and Applications*, 28(2):38, 2022.

[39] L. A. Vinh. Distribution of determinant of matrices with restricted entries over finite fields. *arXiv preprint arXiv:0903.2508*, 2009.

[40] L. A. Vinh. The Szemerédi–Trotter type theorem and the sum-product estimate in finite fields. *European Journal of Combinatorics*, 32(8):1177–1181, 2011.

[41] L. A. Vinh. Product graphs, sum-product graphs and sum-product estimates over finite rings. In *Forum Mathematicum*, volume 27, pages 1639–1655. De Gruyter, 2015.

[42] S. K. Zaremba. La méthode des "bons treillis" pour le calcul des intégrales multiples. In *Applications of number theory to numerical analysis*, pages 39–119. Elsevier, 1972.