The large sieve and random walks on left cosets of arithmetic groups

Comment. Math. Helv. 85 (2010), 647–704
DOI 10.4171/CMH/207

Commentarii Mathematici Helvetici

The large sieve and random walks on left cosets of arithmetic groups

Florent Jouve

Abstract. Building on Kowalski's work on random walks on the groups $\mathrm{SL}(n,\mathbb{Z})$ and $\mathrm{Sp}(2g,\mathbb{Z})$, we consider similar problems (we try to estimate the probability with which, after k steps, the matrix obtained has a characteristic polynomial with maximal Galois group or has no nonzero squares among its entries) for more general classes of sets: in $\mathrm{GL}(n,A)$, where A is a subring of $\mathbb{Q}$ containing $\mathbb{Z}$ that we specify, we perform a random walk on the set of matrices with fixed determinant $D \in A^{\times}$. We also investigate the case where the set involved is any of the two left cosets of the special orthogonal group $\mathrm{SO}(n,m)(\mathbb{Z})$ with respect to the spinorial kernel $\Omega(n,m)(\mathbb{Z})$.

Mathematics Subject Classification (2010). Primary 15A36, 15A52, 11N36; Secondary 15A33, 12E05, 11E08.

Keywords. Random walks on arithmetic groups, Property (τ), large sieve, polynomials and orthogonal matrices over finite fields.

Introduction and statement of the results

For G a fixed subgroup of $\mathrm{GL}(n,\mathbb{Q})$, it is natural to wonder what the typical behavior of an element $g \in G$ chosen at random should be. That kind of question is investigated by Kowalski in [KoSieve, Chapter 7]. Having in mind such intuitive facts as: a random element should have, with high probability, an irreducible characteristic polynomial (or indeed, one with large splitting field) and no square among its entries, Kowalski shows that the k-th step of a random walk lies in the set of the exceptional elements of G (i.e., the elements which do not satisfy the desired property) with probability tending to zero exponentially fast as k grows to infinity.

In loc. cit., these results are obtained, in the case where $G = \mathrm{SL}(n,\mathbb{Z})$ or $\mathrm{Sp}(2g,\mathbb{Z})$ (for $n \geqslant 2$ and $g \geqslant 2$), as an application of the very general large sieve framework exposed in the first chapters of [KoSieve].

In this paper, we answer the same type of questions (i.e., we try to detect similar properties) for sets Y being either left cosets $\alpha\mathrm{SL}(n,A)$ of $\mathrm{GL}(n,A)$ (where $n \geqslant 3$ and A is a subring of $\mathbb{Q}$ containing $\mathbb{Z}$ which we will specify) or left cosets of $\mathrm{SO}(n,m)(\mathbb{Z})$

for $n+m \geq 6$ and $nm \neq 0$ (i.e., we will fix an indefinite quadratic form with signature (n,m) when seen as defined over a $(n+m)$-dimensional space over $\mathbb{R}$) with respect to the normal subgroup $\Omega(n,m)(\mathbb{Z})$ (which is to be defined later). The method used is that of the "coset sieve" described by Kowalski in [KoSieve, Chapter 3.3] (see also [KoZeta] where that idea already appears to study properties of the numerator of zeta functions of curves over finite fields).

Let us now define what is needed to give the precise statements for the main results of this paper. The first kind of subgroups G of $\mathrm{GL}(n,\mathbb{Q})$ we consider are of the type $G = \mathrm{GL}(n,A)$ where, if $\mathcal{P}$ denotes a (possibly infinite) set of primes with complement having positive Dirichlet density, then A is taken to be equal to the ring $\mathbb{Z}[1/\mathcal{P}]$ which is the smallest subring of $\mathbb{Q}$ containing $\mathbb{Z}$ in which every $p \in \mathcal{P}$ is invertible. The left coset to which we apply large sieve techniques in that first case is a fixed element of $\mathrm{GL}(n,A)/\mathrm{SL}(n,A)$.

We also consider the case where G is the subgroup of integral points of a special orthogonal group. For $n+m \geq 6$, let (M,Q) be a quadratic module over $\mathbb{Z}$ such that, seen over $\mathbb{R}$, the quadratic form Q is indefinite with signature (n,m). The group of automorphisms of M preserving Q can be seen as the subgroup of integral points (denoted $\mathrm{O}(n,m)(\mathbb{Z})$) of the algebraic group $\mathbf{O}(n,m)/\mathbb{Q}$. We will restrict ourselves to the case where $G = \mathrm{SO}(n,m)(\mathbb{Z})$, the subgroup of integral points of the algebraic group $\mathbf{SO}(n,m)/\mathbb{Q}$. In $\mathrm{SO}(n,m)(\mathbb{Z})$ lies the normal subgroup $\Omega(n,m)(\mathbb{Z})$ (see [HM, Section 7.2C, pp. 422–424] where that subgroup is denoted $\mathrm{O}'(M)$). A precise definition for that group will be given in Section 2 in the case where M is a vector space over a finite field and in Section 3.1 in the general case. However, to state our results, the important thing is that the fixed coset we consider in that case corresponds to an element in the quotient $\mathrm{SO}(n,m)(\mathbb{Z})/\,\Omega(n,m)(\mathbb{Z})$ (an abelian quotient; see [HM, 7.2.21]).

In the sequel we emphasize the case where (M,Q) is a *free hyperbolic module* over $\mathbb{Z}$ (see [HM, p. 197]), i.e., M is a $\mathbb{Z}$-module of rank $2n$ equipped with a quadratic form Q (with attached bilinear form denoted h) such that there exists a basis of isotropic vectors $\mathcal{X} = (x_1,\ldots,x_{2n})$ with respect to which the matrix of h is

$$\mathrm{Mat}_{\mathcal{X}} h = \begin{pmatrix} 0 & \mathrm{Id} \\ \mathrm{Id} & 0 \end{pmatrix},$$

where the inner blocks are $n \times n$ matrices.

Seen over $\mathbb{R}$, such a quadratic form has signature (n,n) and we will restrict ourselves to such quadratic forms to state Theorem 1 (which is a sample of Theorem 18 in which the case of more general quadratic modules is handled).

The question of the irreducibility of the characteristic polynomial of an element chosen "at random", in one of the two types of groups we have just described is only relevant if no trivial factorisation pattern is imposed by the definition of the groups. If we only suppose that $g \in \mathrm{GL}(n,A)$, there is no a priori imposed factor for the

characteristic polynomial $P_g(T) = \det(T - g)$, but things are different if g is an orthogonal matrix. Indeed P_g verifies in that case the functional equation

$$P_g(T) = \det(-g)T^N P_g\left(\frac{1}{T}\right), \tag{1}$$

where g is assumed to be a $N \times N$ matrix.

Therefore it seems natural to wonder about the factorisation properties of the *reduced* characteristic polynomial which is defined by

$$\det(T-g)_{\mathrm{red}} = \begin{cases} \det(T-g)/(T-\det(g)) & \text{if } N \text{is odd,} \\ \det(T-g)/(T^2-1) & \text{if } N \text{ is even and } \det(g) = -1, \\ \det(T-g) & \text{otherwise.} \end{cases}$$

Here, the matrix g will always lie in the special orthogonal group attached to Q, so that in the case where N is even, we will always have $\det(T-g)_{\mathrm{red}} = \det(T-g)$. Notice moreover that the degree N_{red} of $\det(T-g)_{\mathrm{red}}$ is always *even*.

Now, with the above notation, let G be the group $\mathrm{GL}(n, A)$, for $n \geqslant 3$ (resp. $\mathrm{SO}(n,n)(\mathbb{Z})$, for $n \geqslant 3$), and G^g the normal subgroup $\mathrm{SL}(n, A)$ (resp. $\Omega(n,n)(\mathbb{Z})$) of G. Let S be a symmetric generating system for G^g (i.e., for any $s \in S$, we have $s^{-1} \in S$). Notice here that we *do not* assume G^g to be finitely generated. In other words S could well be infinite (but still countable). Let $(p_s)_{s\in S}$ be a sequence of *strictly positive* real numbers indexed by S satisfying $\sum_{s\in S} p_s = 1$ and $p_s = p_{s^{-1}}$ for any $s \in S$. Finally let α be a fixed element of G.

Suppose a probability space $(\Psi, \Sigma, \mathbf{P})$ is given and let $(X_k)_k$ be the (left invariant) random walk on the left coset αG^g defined by

$$X_0 = \alpha, \quad X_{k+1} = X_k \xi_{k+1},$$

where $(\xi_k)_{k\geqslant 1}$ is a sequence of independent uniformly distributed random variables with values in S and law

$$\mathbf{P}(\xi_k = s) = \mathbf{P}(\xi_k = s^{-1}) = p_s \quad \text{for any } s \in S.$$

Our main result quantifies the speed of rarefaction of "non-typical" elements reached by the k-th step of the random walk as k grows. In order to state it in a unified way, the reduced polynomial $\det(T-g)_{\mathrm{red}}$ denotes, in the first case, nothing but the usual characteristic polynomial $\det(T-g)$; while in the second case, the ring A denotes nothing but the ring $\mathbb{Z}$. A "weak" version of our result can be stated as follows.

Theorem 1. *With notation as above, there exists a $\beta_1 > 0$ such that for all $k \geqslant 1$ we have*

$$\mathbf{P}(\det(T - X_k)_{\mathrm{red}} \in A[T] \textit{ is reducible}) \ll \exp(-\beta_1 k),$$

with β_1 depending only on the underlying algebraic group $\mathbf{G}/\mathbb{Q}$, on the generating set S and on the sequence $(p_s)_s$ (i.e., on the distribution of the ξ_k). Moreover the implied constant depends only on $\mathbf{G}$ and the density of $\mathcal{P}$ in the set of all rational prime numbers (in the case where $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$).

There exists a $\beta_2 > 0$ such that for all $k \geqslant 1$, we have

$$\mathbf{P}(\textit{an entry of the matrix } X_k \textit{ is a square in } A) \ll \exp(-\beta_2 k),$$

with the same dependency for β_2 as for β_1 and the same dependency for the implied constant as in the previous case.

In the above statement, the *underlying algebraic group* is $\mathbf{SL}(n)$ in the case $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$ (with $n \geqslant 3$) and $\mathbf{SO}(n, n)$ if $G = \mathrm{SO}(n, n)(\mathbb{Z})$ (with $n \geqslant 3$).

Of course the second statement of Theorem 1 is only a very special case of the kind of properties that can be investigated using large sieve techniques. In Section 3, we give a more general statement of the above theorem in which the common points of the properties that are likely to be successfully studied via our method appear clearly.

Acknowledgement. This paper contains the main part of the author's PhD Thesis. The author is grateful to his PhD adivsor E. Kowalski for generously sharing his deep ideas on sieve methods.

1. Estimates for the large sieve constants

In this section we will often need to refer to results coming from the large sieve techniques exposed in the appendix. So, before getting into the proof of Theorem 1, the reader might either want to check the appendix or assume Propositions 26 and 27 (which are self-contained statements) to be true and postpone the reading of the whole appendix.

With notation as in Propositions 26 and 27, let Λ be a set consisting of odd primes (with the additional condition $\Lambda \cap \mathcal{P} = \emptyset$ if $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$) and let $\mathcal{L}^*$ be the finite subset of Λ consisting of the elements smaller than a given integer $L \geqslant 1$.

In both Propositions 26 and 27 (note that it is natural to emphasize the case where the sets Θ_ℓ are conjugacy invariant as the estimates are improved when using this special property), the heart of the large sieve method lies in the following inequality

$$\mathbf{P}\left(\rho_\ell(X_k) \notin \Theta_\ell \text{ for all } \ell \leqslant L\right) \leqslant \Delta(X_k, L) H^{-1}, \tag{2}$$

where

$$H = \sum_{\substack{\ell \leqslant L \\ \ell \in \Lambda}} |\Theta_\ell| (|G_\ell^g| - |\Theta_\ell|)^{-1}$$

is the saving factor which depends only on the sieving sets Θ_ℓ and where we denote $\Delta(X_k, L)$ for the constant Δ of Propositions 26 and 27 with the above choice of $\mathcal{L}^*$. In this section though, we focus on the large sieve constant $\Delta(X_k, L)$ for which we give an upper bound in the case where G is one of the two groups that Theorem 1 deals with.

The possibility to obtain the sort of quantitative information stated in Theorem 1 for the random walk (X_k) defined in the introduction depends crucially on the sharpness of the upper bound we can find for the large sieve constants involved. It is not realistic to hope for any useful explicit bound without any assumption on the group G we are working with. As the sums (17) and (19) involve representations of the group G (that factor through finite groups), the fact that Lubotzky's Property (τ) comes into play is not surprising. Let us first review some definitions and facts concerning that property.

1.1. Lubotzky's Property (τ). We first recall the definition of the property, as stated in [LZ, 1.4] or [Lu, p. 49].

Definition 2. Let G be a topological group and $\mathcal{N} = \{N_i\}_i$ a family of normal finite index subgroups of G, indexed by a set I. The group G has *Property* (τ) *with respect to* $\mathcal{N}$ if there exists a finite set S and $\varepsilon > 0$ such that for any unitary irreducible continuous representation ρ of G on a Hilbert space $\mathcal{H}$ satisfying $\ker(\rho) \supset N_i$ for some i and that leaves no nonzero vector invariant, we have

$$\max_{s\in S} \|\rho(s)v - v\| > \varepsilon\|v\|,$$

for all nonzero $v \in \mathcal{H}$. The pair (ε, S) is called a (τ)-constant for G and S is called a (τ)-set for G.

Note that Kazhdan's Property (T) is defined in the exact same way except for the fact that ρ need not factor through some prescribed subgroup. Thus, as is obvious from the definition, Lubotzky's Property is a "weak" version of Kazhdan's. That means of course that any group with Property (T) also has Property (τ) with respect to any family of its finite index subgroups. However, that property is indeed strictly weaker. For instance $\mathrm{SL}(2, \mathbb{Z})$ does not have (T) (see [HV, Proposition 6, p. 34]) but has (τ) with respect to the family of its congruence subgroups

$$\big(\Gamma(d) = \ker(\rho_d \colon \mathrm{SL}(2, \mathbb{Z}) \to \mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z}))\big)_{d\geq 1}.$$

That last fact though does not come for free, as it requires Selberg's result on the eigenvalues of the hyperbolic laplacian acting on $L^2(\Gamma(d)\backslash\mathbf{H})$ (see, e.g. [Lu, 4.4]).

Remark. A useful combinatorial interpretation for Property (τ) is the following: if G is a group and S is a subset of G, recall that the *Cayley graph* $\mathcal{C}(G, S)$ is the oriented

graph with vertex set equal to the set of elements in G and where x is connected to y if there exists an $s \in S$ such that $xs = y$. If we suppose that S is symmetric, then $\mathcal{C}(G, S)$ can be considered as a non oriented graph and if S spans G, that graph is connected. Now with these two additional assumptions, the fact that G has Property (τ) with respect to a family $\mathcal{N} = \{N_i\}_i$ of subgroups is equivalent to the property of *expansion* of the family of Cayley graphs $(\mathcal{C}(G/N_i, S_i))_i$, where for each i, S_i is the projection of S to the corresponding quotient. More generally a graph $X = (V, E)$ is said to be a δ-expander graph (where $\delta > 0$), if for any subset A of V containing less than one half of the elements of V, the number of vertices in $V \setminus A$ which are neighbors of elements of A is at least $\delta|A|$. Moreover, the expansion ration δ is explicitly related to the (τ)-constant for G with respect to $\mathcal{N}$. The notion of expander, born in the 1970's in order to solve problems linked to networks, has motivated lots of mathematical researches. The beautiful constructions of such families that can be found in [Chung], [Mo] or [LPS], rely heavily on deep mathematical tools.

If G is finitely generated (which is the case in the applications developed by Kowalski in [KoSieve, Chapter 7]) and has Property (τ) with respect to $\mathcal{N}$, it can be shown that any generating set S can be chosen to be the (τ)-set for G (see [LZ, Proposition 1.2] or [Lu, Theorem 4.3.2]). For the applications we have in mind, however, we need to work with groups which are not necessarily finitely generated (this is obviously the case for $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ if $\mathcal{P}$ is infinite). The following result, explained by M. Burger, shows that in this case, we can also choose a (τ)-set among the elements of a generating system for G.

Proposition 3. *Let G be a group with Property (τ) with respect to a family of finite index subgroups $\mathcal{N} = \{N_i\}_i$. Let S be a generating system for G, then there exists a finite subset S_0 of S which is a (τ)-set for G.*

Proof. Let F be a (finite) (τ)-set for G and $\delta > 0$ such that (δ, F) is a (τ)-constant for G. As F is finite there exists a subset S_0 of S and an integer $n \geqslant 1$ such that $F \subset S_0^n$ (i.e., each element in F can be written as the product of at most n elements of S_0). Now let $\pi : G \to U(\mathcal{H})$ be a continuous unitary representation of G which factors through N_i for some i (i.e., $\ker \pi \supset N_i$) and without any nonzero invariant vectors in $\mathcal{H}$. If $v \in \mathcal{H}$ has norm 1, then there exists $f = s_0^1 \dots s_0^n \in F$ such that

$$\delta \leqslant \|\pi(f)(v) - v\| \leqslant \|\pi(s_0^1 \dots s_0^n)(v) - v\|.$$

Using the fact that the representation π is unitary, the right-hand side of the above inequality can be written

$$\left\| \sum_{j=0}^{n-1} \left(\pi(s_0^1 \dots s_0^{j+1})(v) - \pi(s_0^1 \dots s_0^j)v\right) \right\| \leqslant \sum_{j=0}^{n-1} \left\| \pi(s_0^1 \dots s_0^j)(\pi(s_0^{j+1})(v) - v) \right\|$$

$$\leqslant \sum_{j=0}^{n-1} \left\| \pi(s_0^{j+1})(v) - v \right\|.$$

Combining these last two series of inequalities we deduce there exists a $t_0 \in S_0$ such that

$$\frac{\delta}{n} \leqslant \left\| \pi(t_0)(v) - v \right\|. \qquad \square$$

Before explaining how Property (τ) yields the kind of upper bound we need for the large sieve constants, let us first give (an infinite family of) examples of groups having Property (τ) (with respect to a certain family of subgroups for each of these examples). Note moreover that in the case where $n = 2$ the following lemma provides us with infinitely many examples of groups having Property (τ) (with respect to a suitably chosen family of subgroups) without being Kazhdan groups. These groups of course are directly involved in the proof of Theorem 1.

Lemma 4. *Let $\mathcal{P}$ be a proper subset of the rational primes. For any $n \geqslant 2$ the group $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ has Property (τ) with respect to the family of its congruence subgroups*

$$\left(\ker \pi_d \colon \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]) \to \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]/d\mathbb{Z}[1/\mathcal{P}])\right)_{\{d \geqslant 1 \mid p \nmid d \text{ if } p \in \mathcal{P}\}}.$$

Proof. Let S_1 be a finite generating system for $\mathrm{SL}(n, \mathbb{Z})$ (the elementary transformations for instance). As already mentioned, S_1 is a (τ)-set for $\mathrm{SL}(n, \mathbb{Z})$. The natural inclusion

$$\mathrm{SL}(n, \mathbb{Z}) \hookrightarrow \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$$

enables us to consider a generating set $S \supset S_1$ for $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$. Let m be an integer without prime factors in $\mathcal{P}$. We consider the projection

$$\pi_m \colon \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]) \to \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]/m\mathbb{Z}[1/\mathcal{P}]) \simeq \mathrm{SL}(n, \mathbb{Z}/m\mathbb{Z}).$$

The restriction of the morphism π_m to $\mathrm{SL}(n, \mathbb{Z})$ is surjective thus, since the family of Cayley graphs $\left(\mathcal{C}(\mathrm{SL}(n, \mathbb{Z})/(\ker \pi_m \cap \mathrm{SL}(n, \mathbb{Z})), \pi_m(S_1))\right)_m$ (indexed by the integers m coprime to any element in $\mathcal{P}$) is an expander family (see the discussion preceding the above remark and recall that, for $n \geqslant 3$, the group $\mathrm{SL}(n, \mathbb{Z})$ is a Kazhdan group), then so is the family

$$\left(\mathcal{C}(\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])/\ker \pi_m, \pi_m(S_1))\right)_m.$$

A fortiori the family $\left(\mathcal{C}(\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])/\ker \pi_m, \pi_m(S))\right)_m$ forms of family of expanders. In other words the group $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ has Property (τ) with respect to the family of its congruence subgroups. $\square$

1.2. Upper bounds for $\Delta(X_k, L)$. Coming back to our sieve framework (see the appendix), we now state the key proposition making precise all the assumptions that will need to be verified for the arithmetic groups we consider, in order to obtain a sufficiently sharp bound for the large sieve constants Δ and H of Propositions 26 and 27 (this is the analogue of [KoSieve, Proposition 7.2]).

Proposition 5. *We recall G is a discrete group, G^g a normal subgroup of G with abelian quotient G/G^g and let T be a finite subset of a set of representatives of G/G^g in which we let α vary. For a fixed symmetric generating system S of G^g we consider the random walk (X_k), defined in the introduction, on the coset αG^g (with $\alpha \in T$) and we suppose*

- *there exists a relation of **odd** length c among the elements of S:*

$$s_1 \dots s_c = 1,$$

- *the steps ξ_k, for $k \geqslant 1$, are independent and independent of X_0,*
- *G^g has Property (τ) with respect to a family $(N_i)_{i \in I}$ of finite index subgroups,*

then there exists $\eta > 0$ such that, for any finite dimensional representation

$$\pi : G \to \mathrm{GL}(V),$$

satisfying $\ker \pi_{|G^g} \supset N_i$, *for some i and without any nonzero G^g-invariant vector, the inequality*

$$|\mathbf{E}(\langle \pi(X_k)e; f \rangle)| \leqslant \|e\| \, \|f\| \exp(-\eta k),$$

holds for all vectors e, f in V and all $k \geqslant 0$; $\langle\, ;\, \rangle$ denoting a G-invariant inner product on V.

The constant η only depends on the (τ)-constant associated to $(G^g, S, (N_i))$, on the distribution of the ξ_k and on the length c of a fixed relation in S.

Proof. The proof is quite similar to that of [KoSieve, Proposition 7.2]. Still, as many technical points need to be modified we give the full detail of the arguments here.

We fix an index $i \in I$ and a representation

$$\pi : G \to \mathrm{GL}(V),$$

such that the restriction $\pi_{|G^g}$ factors through N_i and has no nonzero G^g-invariant vector. Consider

$$M = \mathbf{E}(\pi(\xi_k)) = \sum_{s \in S} p(s)\pi(s).$$

The linear operator M is a well-defined element of $\mathrm{End}(V)$ since the series defining M converges absolutely (because π is a unitary representation and $\sum_s p(s) = 1$). From M we can then define two other elements of $\mathrm{End}(V)$:

$$M^+ = \mathrm{Id} - M, \quad M^- = \mathrm{Id} + M.$$

Note that these formulæ define two operators which are both independent of k and self adjoint. Indeed the set S as well as the distribution of the ξ_k are symmetric. Moreover the mapping associating its adjoint to an operator is linear and continuous. We also need to define

$$N_0 = \mathbf{E}(\pi(X_0)) = \sum_{t \in T} \mathbf{P}(X_0 = t)\pi(t) \in \mathrm{End}(V).$$

For $k \geqslant 1$, the random variables X_0 and ξ_k are independent therefore we have

$$\mathbf{E}(\pi(X_k)) = N_0 M^k,$$

thus by linearity,

$$\mathbf{E}(\langle \pi(X_k)e; f\rangle) = \langle M^k e; N_0^* f\rangle,$$

where N_0^* denotes the adjoint of N_0.

As $\sum_{s\in S} p(s) = 1$ and since $\pi(s)$ is a unitary operator for every $s \in S$, the eigenvalues of M are in the interval $[-1; 1]$. Now, let ρ be the spectral radius of M:

$$\rho = \max\{|\gamma| \mid \gamma \text{ is an eigenvalue of } M\}.$$

We have the inequality

$$|\langle M^k e; N_0^* f\rangle| \leqslant \|e\| \, \|f\| \, \rho^k,$$

since the norm of N_0 is smaller than 1.

We need to exhibit a $\delta > 0$ independent of i and π such that $0 \leqslant \rho \leqslant 1 - \delta$. We will then be able to choose $\eta = -\log(1 - \delta) > 0$ for the constant we are looking for. We use the fact that $\rho = \max(\rho^+, \rho^-)$ where ρ^+ (resp. ρ^-) is the real number equal to the greatest positive eigenvalue of M (resp. equal to the opposite of the smallest negative eigenvalue of M). It is enough to prove that $\rho^\pm < 1 - \delta_\pm$ for some constants $\delta_\pm$ which are independent of i and π. To that purpose we use the variational interpretation for the eigenvalues of a self adjoint operator on a finite dimensional Hilbert space. Indeed $1 - \rho^+$ (resp. $1 + \rho^-$), which is the smallest eigenvalue of M^+ (resp. of M^-), is equal to

$$\lambda = \min_{v \neq 0} \frac{\langle Tv; v\rangle}{\|v\|^2},$$

where $T = M^\pm$. Now applying Proposition 3, we know that there exists for the

group G^g a finite (τ)-set S_0 included in S. This yields

$$\begin{aligned}\frac{\langle M^+v;v\rangle}{\|v\|^2} &= \frac{1}{2}\sum_{s\in S} p_s \frac{\|\pi(s)v-v\|^2}{\|v\|^2}\\ &\geqslant \frac{1}{2}\sum_{s\in S_0} p_s \frac{\|\pi(s)v-v\|^2}{\|v\|^2}\\ &\geqslant \frac{{p_0}^+}{2}\inf_{\varpi}\inf_{v\neq 0}\max_{s\in S_0}\frac{\|\varpi(s)v-v\|^2}{\|v\|^2},\end{aligned}$$

where $p_0^+ = \min_{s\in S_0} p(s) > 0$ and ϖ runs over the representations of G^g without any nonzero G^g-invariant vector and which factorize through N_j for some index j. Let $\kappa > 0$ be such that (κ, S_0) is a (τ)-constant for G^g with respect to $(N_i)_{i\in I}$, we can choose

$$\delta^+ = \frac{\kappa p_0^+}{2}.$$

The argument used to determine δ^- is very close to that of [KoSieve, Proposition 7.2]. Since there exists by assumption a relation of odd length c among the elements of S, we can write for $v \in V$:

$$v = \frac{1}{2}\big((v+\pi(s_1)v) - (\pi(s_1)v + \pi(s_1s_2)v) + \cdots + (\pi(s_1\ldots s_{c-1})v + \pi(1)v)\big).$$

Then, invoking the Cauchy–Schwarz inequality and using the G-invariance of the inner product,

$$|v\|^2 \leqslant \frac{c}{4}\sum_{i=0}^{c-1}\|\pi(r_i)v + \pi(r_is_{i+1})v\|^2 \leqslant \frac{c}{4}\sum_{i=0}^{c-1}\|v+\pi(s_{i+1})v\|^2,$$

where $r_0 = 1$ and $r_i = s_1\ldots s_i$ for $i \geqslant 1$. In particular we deduce

$$\|v\|^2 \leqslant \frac{c}{4}\max_{1\leqslant i\leqslant c}\frac{1}{p(s_i)}\sum_{i=0}^{c-1} p(s_{i+1})\,\|v+\pi(s_{i+1})v\|^2,$$

then, taking into account possible repetitions of generators in the sequence $(s_1,\ldots,s_c)$,

$$\begin{aligned}\|v\|^2 &\leqslant \frac{c^2}{4}\frac{1}{\min\{p(s_i)\mid 1\leqslant i\leqslant c\}}\sum_{s\in S} p(s)\,\|\pi(s)v+v\|^2\\ &\leqslant \frac{c^2}{2}\big(\min\{p(s_i)\mid 1\leqslant i\leqslant c\}\big)^{-1}\langle M^-v;v\rangle.\end{aligned}$$

Therefore we can choose $\delta^- = \frac{2}{c^2}\min\{p(s_i)\mid 1\leqslant i\leqslant c\} > 0$. □

Under the assumptions of Proposition 5 together with an additional hypothesis of *linear disjointness* (which really is a property of "independence of ℓ of the setting"), the next two propositions give the upper bound we need for the two large sieve constants Δ we are working with.

To begin with, we consider the case of the conjugacy coset sieve which is somewhat simpler to handle.

Proposition 6. *Let $(Y, \Lambda, (\rho_\ell : Y \to Y_\ell))$ be the conjugacy coset sieve of the appendix (see Proposition 26). We suppose that*

- *the assumptions of Proposition 5 are verified,*
- *the system $(\rho_\ell)_{\ell\in\Lambda}$ is **linearly disjoint**, i.e., the restricted product map defined for $\ell, \ell' \in \Lambda$, $\ell \neq \ell'$, by*

$$\rho_{\ell,\ell'} = \rho_\ell \times \rho_{\ell'} : G^g \to G^g_{\ell,\ell'} = G^g_\ell \times G^g_{\ell'}$$

is surjective.

Then, with notation as in Proposition 26, there exists $\eta > 0$ such that

$$\Delta(X_k, L) \leqslant 1 + L^A \exp(-\eta k),$$

where $\eta > 0$ depends only on G, S and the distribution of the ξ_k and $A = (3d+2)/2$, with $d = n^2-1$ in the case where $G = \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ and $d = (n+m)(n+m-1)/2$ if $G = \mathrm{SO}(n,m)(\mathbb{Z})$.

Proof. From Proposition 26 we have

$$\Delta(X_k, L) \leqslant \max_{\ell\in\mathfrak{L}^*} \max_{\pi\in\Pi^*_\ell} \sum_{\ell'\in\mathfrak{L}^*} \sum_{\tau\in\Pi^*_{\ell'}} |W(\pi,\tau)|.$$

For $\ell, \ell' \in \Lambda$, we need to give an upper bound for the sums

$$W(\varphi_\pi, \varphi_\tau) = \frac{1}{\sqrt{|\hat{\Gamma}^\pi_\ell||\hat{\Gamma}^\tau_{\ell'}|}} \mathbf{E}(\mathrm{Tr}[\pi, \bar{\tau}]\rho_{\ell,\ell'}(X_k)),$$

once rewritten using the trick explained in the appendix after Lemma 25.

With notation as in the appendix (see the discussions preceding and following Lemma 25), if $[\pi, \bar{\tau}]\rho_{\ell,\ell'}$ has no G^g-invariant vector then Proposition 5 yields an upper bound for

$$|\mathbf{E}(\mathrm{Tr}[\pi, \bar{\tau}]\rho_{\ell,\ell'}(X_k))|.$$

Indeed it is enough to choose $e = f$ running over an orthonormal basis of the representation space V of $[\pi, \bar{\tau}]\rho_{\ell,\ell'}$ and then to sum up the terms obtained over e.

We are reduced to computing the multiplicity of the trivial representation of G^g in the restriction of $[\pi, \bar{\tau}]\rho_{\ell,\ell'}$ to G^g. As the sieve setting we work with is assumed to be linearly disjoint, that quantity is the same as the multiplicity of the trivial representation of $G^g_{\ell,\ell'}$ in $[\pi, \bar{\tau}]_{|G^g_{\ell,\ell'}}$. From Lemma 25 of the appendix we know that multiplicity is zero unless $(\ell, \pi) = (\ell', \tau)$ in which case its value is $|\hat{\Gamma}^{\pi}_{\ell}|$. Thus, denoting $[\pi, \bar{\tau}]_0$ the part of $[\pi, \bar{\tau}]$ without any nonzero G^g-invariant vector, we deduce

$$\mathrm{Tr}[\pi, \bar{\tau}]\rho_{\ell,\ell'}(X_k) = |\hat{\Gamma}^{\pi}_{\ell}|\delta((\ell, \pi), (\ell', \tau)) + \mathrm{Tr}[\pi, \bar{\tau}]_0\rho_{\ell,\ell'}(X_k).$$

Applying Proposition 5 to the representation $[\pi, \bar{\tau}]_0\rho_{\ell,\ell'}$ of G^g yields

$$|W(\pi, \tau) - \delta((\ell, \pi), (\ell', \tau))| \leqslant (\dim \pi)(\dim \tau)\exp(-\eta k),$$

where we use the trivial upper bound $\left(\sqrt{|\hat{\Gamma}^{\pi}_{\ell}||\hat{\Gamma}^{\tau}_{\ell'}|}\right)^{-1} \leqslant 1$.

The result follows from exploiting such trivial bounds as

$$\dim \pi \leqslant \sqrt{|G|}, \quad \sum_{\pi \in \mathrm{irr}(G)} \dim \pi \leqslant |G|,$$

for any irreducible complex representation π of a finite group G (see [KoSieve, Chapter 5] for better bounds for such quantities). □

In the next proposition we give an upper bound for $\Delta(X_k, \mathcal{L}^*)$ in the case of the non-conjugacy coset sieve. It is very close to the one provided by Proposition 6. However, needing to use another equivalence relation to define the orthonormal basis for the L^2 space involved (compare the statements of Proposition 26 and 27), the above proof cannot be directly adapted to the case of the non-conjugacy coset sieve. Indeed, in order to prove the following result, we use the remark following Proposition 27 about the generalisation of the sieve statements of the appendix to a framework in which we do not only use primes but more generally squarefree integers to perform the sieve.

Proposition 7. *Let $(Y = \alpha G^g, \Lambda, (\rho_\ell : Y \to Y_\ell))$ be the non-conjugacy coset sieve of the appendix (see, e.g., Proposition 27). For any fixed integer $L \geqslant 1$ and under the same assumptions as in Proposition 6, there exists $\eta' > 0$ such that*

$$\Delta(X_k, L) \leqslant 1 + L^{A'}\exp(-\eta' k),$$

where $\eta' > 0$ depends only on G, the (τ)-constant for G^g, the distribution of the ξ_k and $A' = (17d + 4)/4$ with $d = n^2 - 1$ if $G = \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ and $d = (n + m)(n + m - 1)/2$ if $G = \mathrm{SO}(n, m)(\mathbb{Z})$.

Proof. As in the proof of the previous proposition we need to evaluate a sum involving group characters. The point is that the maximal contribution in those sums comes from the function corresponding to the trivial representation. Following that idea, we apply an "equidistribution approach" in order to obtain the estimate we are aiming at:

$$\mathbf{E}(\langle [\pi,\bar{\tau}](\rho_{\ell,\ell'}(X_k))\tilde{e}\,;\tilde{f}\rangle_{[\pi,\bar{\tau}]}) = \sum_{y\in Y_{\ell,\ell'}} \langle [\pi,\bar{\tau}](y)\tilde{e}\,;\tilde{f}\rangle_{[\pi,\bar{\tau}]} I_y,$$

where the notation are those of (15) and where I_y, on the right-hand side of the equality is defined by

$$I_y = \mathbf{P}(\rho_{\ell,\ell'}(X_k) = y).$$

To evaluate I_y we decompose the characteristic function χ_y of $\{y\}$ in Fourier series. To that purpose, we need to extend by multiplicativity the result of Proposition 27 to the case of squarefree integers (not only primes, see the remark following Proposition 27 in the appendix). In $L^2(Y_{\ell,\ell'},\nu_{\ell,\ell'})$, the following equality holds:

$$\chi_y = \sum_{\varphi\in\mathcal{B}_{\ell,\ell'}} \langle\varphi;\chi_y\rangle\varphi = \sum_{\varphi\in\mathcal{B}_{\ell,\ell'}} \varphi(y)|G^g_{\ell,\ell'}|^{-1}\varphi.$$

Thus, we obtain

$$\begin{aligned} I_y = \mathbf{E}(\chi_y(\rho_{\ell,\ell'}X_k)) &= \frac{1}{|G^g_{\ell,\ell'}|}\sum_{\varphi\in\mathcal{B}_{\ell,\ell'}} \varphi(y)\mathbf{E}(\varphi\rho_{\ell,\ell'}(X_k)) \\ &= \frac{1}{|G^g_{\ell,\ell'}|} + \frac{1}{|G^g_{\ell,\ell'}|}\sum_{\varphi\in\mathcal{B}_{\ell,\ell'}\setminus\{1\}} \varphi(y)\mathbf{E}(\varphi\rho_{\ell,\ell'}(X_k)). \end{aligned}$$

Now if $\varphi = \varphi_{\pi_{\ell,\ell'},e',f'}$ is an element of $\mathcal{B}_{\ell,\ell'}\setminus\{1\}$ (up to a suitable normalisation, see Lemma 24), we know in particular that $\pi_{\ell,\ell'}$ is an irreducible representation of $G_{\ell,\ell'}$ and the quantity for which we need to find an upper bound is

$$|\mathbf{E}(\langle\pi_{\ell,\ell'}\rho_{\ell,\ell'}(X_k)e';f'\rangle_{\pi_{\ell,\ell'}})|.$$

In order to apply Proposition 5, we need to determine the multiplicity of the trivial representation of $G^g_{\ell,\ell'}$ in the restriction of $\pi_{\ell,\ell'}$ to $G^g_{\ell,\ell'}$. As we assume the sieve setting to be linearly disjoint, this multiplicity is the same as that of 1_{G^g} in $\pi_{\ell,\ell'}\rho_{\ell,\ell'}$ for the group G^g. Applying Lemma 25 (once extended by multiplicativity) with $\pi = \pi_{\ell,\ell'}$, $\tau = 1_{G_\ell}$ (note that the assertion of Lemma 25 remains valid for the trivial representation, see [KoSieve, Proof of Lemma 3.2]), and using the fact that ℓ is a prime factor of $\mathrm{ppcm}(\ell,\ell')$, we see that multiplicity is zero unless $\mathrm{ppcm}(\ell,\ell') = \ell$ (i.e., $\ell' = \ell$) and $\pi_{\ell,\ell'|G^g_{\ell,\ell'}} = 1_{|G^g_{\ell,\ell'}}$ (or more precisely, applying [KoSieve, Lemma 3.2], $\pi_{\ell,\ell'}\otimes\psi \simeq 1_{G_{\ell,\ell'}}$ for a certain character ψ of $G_{\ell,\ell'}/G^g_{\ell,\ell'}$). In particular, $\pi_{\ell,\ell'}$ has

dimension 1 and for every vector e' with norm 1 spanning the representation space of $\pi_{\ell,\ell'}$ and every $g \in G^g_{\ell,\ell'}$,

$$\langle \pi_{\ell,\ell'}(g)e'; e' \rangle = \langle e'; e' \rangle = 1,$$

where the index $\pi_{\ell,\ell'}$ is purposely omitted to avoid the use of too much notation. Thus, with notation as in Lemma 24, we deduce that $\varphi_{\pi_{\ell,\ell'},e',e'} \sim 1$, which is a contradiction.

Invoking Proposition 5 now yields an $\eta' > 0$ such that for all $\varphi = \varphi_{\pi_{\ell,\ell'},e',f'} \in \mathcal{B}_{\ell,\ell'} \setminus \{1\}$,

$$|\mathbf{E}(\langle \pi_{\ell,\ell'}\rho_{\ell,\ell'}(X_k)e'; f' \rangle_{\pi_{\ell,\ell'}})| \leqslant \exp(-\eta' k).$$

Finally, for the quantity

$$(\dim \pi)^{(-1/2)}(\dim \tau)^{(-1/2)} \big| W(\varphi_{\pi,e,f}, \varphi_{\tau,\varepsilon,\phi}) - \delta(\varphi_{\pi,e,f}, \varphi_{\tau,\varepsilon,\phi}) \big|,$$

we obtain the following upper bound (note that the inverse of the denominator of the normalisation factor is trivially smaller than 1):

$$\left| \frac{1}{|G^g_{\ell,\ell'}|} \sum_{y \in Y_{\ell,\ell'}} \left(\langle [\pi, \bar{\tau}](y)\tilde{e}; \tilde{f} \rangle_{[\pi,\bar{\tau}]} \sum_{\varphi \in \mathcal{B}_{\ell,\ell'} \setminus \{1\}} \varphi(y)\mathbf{E}(\varphi\rho_{\ell,\ell'}(X_k)) \right) \right|.$$

Applying the Cauchy–Schwarz inequality we obtain

$$\begin{aligned} |\langle [\pi, \bar{\tau}](y)\tilde{e}; \tilde{f} \rangle_{[\pi,\bar{\tau}]}| &\leqslant \|[\pi, \bar{\tau}](y)\tilde{e}\|_{[\pi,\bar{\tau}]} \, \|\tilde{f}\|_{[\pi,\bar{\tau}]} \\ &\leqslant 1 \end{aligned}$$

and more generally $|\varphi(y)| \leqslant 1$, for all $y \in Y_{\ell,\ell'}$ and all $\varphi \in \mathcal{B}_{\ell,\ell'} \setminus \{1\}$. Using the triangle inequality we deduce an upper bound for

$$(\dim \pi)^{(-1/2)}(\dim \tau)^{(-1/2)} \big| W(\varphi_{\pi,e,f}, \varphi_{\tau,\varepsilon,\phi}) - \delta(\varphi_{\pi,e,f}, \varphi_{\tau,\varepsilon,\phi}) \big|,$$

namely

$$(|\mathcal{B}_{\ell,\ell'}| - 1)\exp(-\eta' k).$$

Now, by classical group representation theory,

$$|\mathcal{B}_{\ell,\ell'}| \leqslant \sum_{\pi \in \mathrm{irr}(G_{\ell,\ell'})} (\dim \pi)^2 = |G_{\ell,\ell'}|.$$

We finally deduce an upper bound for the large sieve constant:

$$\Delta(X_k, L) \leqslant 1 + L^{A'} \exp(-\eta k),$$

with $A' = (17d + 4)/4$ and either $d = n^2 - 1$ if $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$, or $d = (n + m)(n + m - 1)/2$ if $G = \mathrm{SO}(n, m)(\mathbb{Z})$. Indeed, from the argument above, we just need to use the same kind of trivial bounds as the ones at the end of the proof of Proposition 6 as well as the obvious inequality $|G_{\ell,\ell'}| \leqslant |G_\ell||G_{\ell'}|$. □

2. Local densities for polynomials and orthogonal matrices

In this section, which is independent of the others, we compute different densities in subsets of the ring $\mathbb{F}_\ell[T]$ or of the orthogonal group $\mathrm{O}(N, \mathbb{F}_\ell)$ (notice that we do not assume anything here on the integer N and, in particular, we do not distinguish between the split and non split model for the orthogonal group, in the case N is even).

The goal of this section is to give enough quantitative information in order to find a useful lower bound for the constant H appearing in (2). However this section has an interest of its own and we do not restrict ourselves to the computations that are strictly needed for the purpose of the paper. The style in which we expose the different estimates we are interested in is very much inspired by [Chav, Section 3]. Doing so, it is easy to point out similarities as well as the differences between the symplectic case (treated by Chavdarov in loc. cit.) and the orthogonal case. We will notably highlight the lack of good topological properties for the orthogonal group. That feature imposes that we be very careful in the statement of our results (such precautions need not be taken in the case of the symplectic group).

2.1. Review of orthogonal groups over finite fields. We briefly recall some basic facts and notation about orthogonal groups, as exposed in Section 6 of [KaL]. The proofs and details can be found e.g. in [ABS].

Let V be a vector space with dimension greater or equal to 2 over a fixed finite field $\mathbb{F}_q$ with characteristic different from 2. We assume we are given a non degenerate quadratic form Q on V (we will denote by Φ the bilinear form attached to Q). If $T(V)$ denotes the tensor algebra associated to V, we can consider the ideal $\Im(Q)$ generated by the elements $x \otimes x - Q(x).1$, where x runs over the vectors of V. The quotient algebra $\mathrm{Cl}(V, Q) = T(V)/\Im(Q)$ is the *Clifford algebra* of V with respect to Q. That construction yields a natural injection

$$i_Q \colon V \to T(V) \to \mathrm{Cl}(V, Q),$$

which enables us to see $\mathrm{Cl}(V, Q)$ as the solution of the following universal problem: for every morphism of $\mathbb{F}_q$-vector spaces $f \colon V \to A$, where A is an $\mathbb{F}_q$-algebra satisfying $f(x)^2 = Q(x).1_A$, there exists a unique $\mathbb{F}_q$-algebra homomorphism $\tilde{f} \colon \mathrm{Cl}(V, Q) \to A$ such that $\tilde{f} \circ i_Q = f$.

The involution $v \mapsto -v$ of V can be extended to an involution denoted I of $\mathrm{Cl}(V, Q)$. Another morphism plays a crucial role in the theory of Clifford algebras: it is the antiautomorphism $t \colon \mathrm{Cl}(V, Q) \to \mathrm{Cl}(V, Q)$ coming from the natural antiautomorphism defined on the k-th tensor power of V by

$$v_1 \otimes v_2 \otimes \cdots \otimes v_k \mapsto v_k \otimes \cdots \otimes v_2 \otimes v_1.$$

Let $\mathrm{Cl}^\times$ be the group of invertible elements of $\mathrm{Cl}(V, Q)$. It acts on $\mathrm{Cl}(V, Q)$ via

the morphism ρ defined by

$$\rho(u)x = I(u)xu^{-1},$$

for every $u \in \mathrm{Cl}^\times$ and $x \in \mathrm{Cl}(V, Q)$. The elements of $\mathrm{Cl}^\times$ that leave V globally invariant form a subgroup of $\mathrm{Cl}^\times$. We denote it $C^\times$. It is the *the Clifford group* of $\mathrm{Cl}^\times$. Typical elements of $C^\times$ are the images in $\mathrm{Cl}(V, Q)$ of non isotropic vectors $v \in V$, since the transformation $x \mapsto I(u)xu^{-1}$ is then the reflection with respect to the hyperplane which is orthogonal to v. In fact any element of $C^\times$ is a scalar multiple of a product of such vectors (the transformation associated to that element being an automorphism of the quadratic module (V, Q)).

Finally we define the map Norm: $u \in C^\times \mapsto t(u)u$ which takes its values in $\mathbb{F}_q^\times$ (see [ABS, Propositions 3.3 and 3.8] where the proof, given in the case where the base field is $\mathbb{R}$, can be easily adapted to the finite field case). The *spinor group* $\mathrm{Spin}(V, Q)$ is then defined as the subgroup

$$\mathrm{Spin}(V, Q) = (\ker \mathrm{Norm})^I$$

of the elements of $C^\times$ that are fixed by I. When V is N-dimensional and there is no ambiguity on the chosen quadratic form Q, we will denote that group $\mathrm{Spin}(N, \mathbb{F}_q)$ instead of $\mathrm{Spin}(V, Q)$. With such notation the group $\mathrm{Spin}(N, \mathbb{F}_q)$ can in fact be seen as the group of $\mathbb{F}_q$-rational points of an algebraic group defined over $\mathbb{F}_q$. Unless the context is ambiguous that algebraic group will also be called the spinor group. It will be denoted $\mathbf{Spin}(N)$. It is well known that the spinor group is a connected simply-connected algebraic group and that it is in fact the universal cover of the special orthogonal group $\mathbf{SO}(N)/\mathbb{F}_q$. In other words ([Hu, p. 189]) there exists an isogeny φ such that we have an exact sequence of algebraic groups

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathbf{Spin}(N) \xrightarrow{\varphi} \mathbf{SO}(N) \longrightarrow 1. \tag{3}$$

Thus the spinor group shares the same dimension as the special orthogonal group $N(N-1)/2$ and the same rank $\lfloor N/2 \rfloor$.

Remark. In all of the above, we do not need to assume that the base field is a finite field. Every construction and definition we have recalled can in fact be stated for quadratic modules over any perfect field.

Next let $\overline{\mathbb{F}_q}$ denote a fixed separable closure of $\mathbb{F}_q$. The short exact sequence (3) gives rise to the following exact sequence of $\mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$-invariant groups:

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Spin}(N, \mathbb{F}_q) \xrightarrow{\rho} \mathrm{SO}(N, \mathbb{F}_q) \xrightarrow{\mathrm{N_{Spin}}} \{\pm 1\} \longrightarrow 1, \tag{4}$$

where the group homomorphism $\mathrm{N_{Spin}}$ is called the *spinor norm* and can be defined as follows. For a non isotropic vector $v \in V$, the image of the reflection with respect

to v by $\mathrm{N}_{\mathrm{Spin}}$ is the class of $Q(v)$ in $\mathbb{F}_q^\times/(\mathbb{F}_q^\times)^2$ (the group of classes modulo nonzero squares of $\mathbb{F}_q^\times$). The morphism $\mathrm{N}_{\mathrm{Spin}}$ is completely determined by the images of those elements and it can be shown (see [KaL, Section 6]) that $\mathrm{N}_{\mathrm{Spin}}$ can be extended to a surjective morphism from $\mathrm{O}(N,\mathbb{F}_q)$ onto $\{\pm 1\}$ since we have supposed $N \geqslant 2$.

Finally we will denote $\Omega(N,\mathbb{F}_q)$ the image of ρ in (4) (i.e., the kernel of $\mathrm{N}_{\mathrm{Spin}}$). That group which is of great importance in our sieving context, is easily seen to be the derived group of both $\mathrm{SO}(N,\mathbb{F}_q)$ and $\mathrm{O}(N,\mathbb{F}_q)$ (see [Art, 5.17]). We note that it can be directly defined from $\mathrm{O}(N,\mathbb{F}_q)$ by saying that it is the simultaneous kernel of the determinant and the spinor norm.

To perform the density computations we need, we will have to estimate the cardinality of sets of orthogonal matrices with fixed determinant and/or spinor norm. In practice it will be convenient to relate those quantities to the number of polynomials with coefficients in the base field that can be realized as characteristic polynomials of such matrices. To exhibit the link between these cardinalities the crucial point lies in the possibility to "see" the value of the spinor norm and of the determinant of a matrix g in the coefficients of its characteristic polynomial. This is, of course, very easy in the case of the determinant. But as far as the spinor norm is concerned, we do not see any obvious reason for such an explicit link to exist. However the following beautiful result of Zassenhaus (which we recall here, in view of its importance) gives us under certain conditions the kind of link we need:

Theorem 8 (Zassenhaus,1962). *If g is an element of the orthogonal group associated to the quadratic space (V, Q) (satisfying the same assumptions as above) then, provided -1 is not an eigenvalue of g, we have*

$$\det\left(\frac{g+1}{2}\right) = \mathrm{N}_{\mathrm{Spin}}(g),$$

modulo squares.

In [Za], Zassenhaus first defines the spinor norm via the formula of Theorem 8 (see (2.1) in loc. cit.) and then proves that definition coincides we the one we gave earlier in this section (see the corollary of [Za, Theorem, p. 446]).

Remark. Over finite fields of odd characteristic we know that there are two isomorphism classes of quadratic forms (classified by the value modulo nonzero squares of the discriminant of the quadratic form). In the case where the dimension is odd these two classes give rise to the same orthogonal group, but this does not hold if the dimension is even. Indeed if $N = 2n$, two distinct models for the orthogonal group $\mathrm{O}(2n,\mathbb{F}_q)$ need to be distinguished: they are respectively called the *split* and *nonsplit* model, referring to the algebraic group $\mathbf{O}(2n)$ being split or not over $\mathbb{F}_q$ (see [KaL, Section 6] for examples of quadratic forms corresponding to each of these

two models). Note that the computations in the sequel are performed independently of the chosen model of orthogonal group. However we will see later how a result of Baeza makes the choice of the split or of the nonsplit model come back into play.

2.2. Characteristic polynomials of orthogonal matrices over finite fields. Let ℓ be an odd prime number and $N \geqslant 2$ an integer. As in the previous subsection (V, Q) denotes a quadratic space over $\mathbb{F}_\ell$ and Q is still assumed to be non degenerate. For g an element of $\mathrm{O}(N, \mathbb{F}_\ell)$ we denote by P_g the *reversed* characteristic polynomial of g:

$$P_g(T) = \det(1 - Tg).$$

This polynomial also satisfies the functional equation (1). A short proof of that fact goes as follows: if σ is the automorphism of the ambiant quadratic space (V, Q) attached to the matrix g, we denote by $\mathcal{Q}$ the matrix of the quadratic form Q written in the basis in which the matrix of σ equals g. Then we have $g\mathcal{Q}({}^t g) = \mathcal{Q}$. We deduce

$$\begin{aligned} P_g(T) &= \det(1 - Tg) = \det(1 - T({}^t g)) \\ &= \det(1 - T(\mathcal{Q}^{-1} g^{-1} \mathcal{Q})) \\ &= T^N \det(g^{-1}) \det(g/T - 1) = (-T)^N \det(g) P_g(1/T). \end{aligned}$$

Here one should be careful that P_g no longer designates the "usual" characteristic polynomial $\det(T - g)$, as in the introduction. What motivates the change in the notation is that Chavdarov works with reversed characteristic polynomials in [Chav], so that we can easily understand to what extent his results can be transposed to the orthogonal case.

The functional equation (1) imposes P_g to be "almost self-reciprocal", i.e., its roots $\alpha_1, \ldots, \alpha_N$ can be reordered in such a way that

$$\begin{cases} \alpha_i \alpha_{N+1-i} = 1, \ 1 \leqslant i \leqslant n, & \text{if } N = 2n, \\ \alpha_i \alpha_{N+1-i} = 1, \ 1 \leqslant i \leqslant n, \text{ and } \alpha_{n+1} = \det(g) & \text{if } N = 2n+1, \end{cases}$$

That leads us to consider the set of polynomials

$$M_{N,\ell} = \{1 + b_1 T + \cdots + b_N T^N \mid b_i \in \mathbb{F}_\ell, b_N^2 = 1 \text{ and } b_{N-i} = b_N b_i \text{ if } 0 \leqslant i \leqslant \lfloor N/2 \rfloor\}.$$

Before studying certain subsets of $M_{N,\ell}$, we recall a result due to Edwards which gives in the case where N is even a (quite surprising at first) link between the discriminant of a polynomial $f \in M_{N,\ell}$ and the values f takes at ± 1.

Lemma 9 (Edwards). *Let N be an even integer and let $f \in M_{N,\ell}$ be a monic separable polynomial, then we have*

$$\operatorname{disc}(f) \equiv f(1)f(-1) \pmod{(\mathbb{F}_\ell^\times)^2}.$$

That result is obtained by combining Theorem 1 and Theorem 2 of [E] (note that the definition of the discriminant of a polynomial used in loc. cit. is not the standard one hence the statement of Lemma 9 only coincides with the one of [E] up to sign).

We are interested in the cardinality of certain subsets of $M_{N,\ell}$. Our first result in this direction is very close to [Chav, Lemma 3.2]. It deals with the subset of irreducible polynomials in $M_{N,\ell}$ or rather with the subset of those of them which are irreducible *once reduced*. Indeed, using the notation of the introduction and remembering we consider reversed characteristic polynomials here, we will use the following notation:

- if $N = 2n$,

$$K_{N,\ell}^{\varepsilon} = \left\{ f(T) \in M_{N,\ell} \mid f(T)_{\mathrm{red}} = \frac{f(T)}{1-\varepsilon T^2} \text{ is irreducible} \right\},$$

  with $\varepsilon = 1$ if $b_N = -1$ and $\varepsilon = 0$ otherwise,

- if $N = 2n+1$,

$$K_{N,\ell}^{\varepsilon} = \left\{ f(T) \in M_{N,\ell} \mid f(T)_{\mathrm{red}} = \frac{f(T)}{1-\varepsilon T} \text{ is irreducible} \right\},$$

  with $\varepsilon = \pm 1$ if $b_N = \mp 1$.

In both cases we will denote by N_{red} the degree of the reduced polynomial f_{red}.

For those sets we have the following estimates:

Proposition 10.

- *If N is even,*

$$\frac{1+\ell^{N/2-\varepsilon}}{N-2\varepsilon} \geqslant |K_{N,\ell}^{\varepsilon}| \geqslant \frac{\ell^{N/2-\varepsilon}}{N-2\varepsilon} - \sqrt{\ell^{N/2-\varepsilon}}.$$

- *If N is odd,*

$$\frac{1+\ell^{(N-1)/2}}{N-1} \geqslant |K_{N,\ell}^{\varepsilon}| \geqslant \frac{\ell^{(N-1)/2}}{N-1} - \sqrt{\ell^{(N-1)/2}}.$$

Proof. It is enough to consider the case where N is even and $\varepsilon = 0$. Indeed, if N is even and $\varepsilon = 1$ (which means that the leading coefficient of f is -1), we are interested in the irreducibility of $g(T) = f(T)/(1-T^2)$. The leading coefficient

of g is obviously 1 and it is clear that $g \in M_{N-2,\ell}$. Thus $g \in K^0_{N-2,\ell}$ as soon as $f \in K^1_{N,\ell}$.

In the case where N is odd the polynomial $g(T) = f(T)/(1-\varepsilon T)$ has leading coefficient 1 and $g \in M_{N-1,\ell}$ so that $g \in K^0_{N-1,\ell}$ as soon as $f \in K^\varepsilon_{N,\ell}$. In other words,

- if N is even, $|K^1_{N,\ell}| = |K^0_{N-2,\ell}|$, and
- if N is odd, $|K^\varepsilon_{N,\ell}| = |K^0_{N-1,\ell}|$.

We are now reduced to computing the number of elements of

$$K^0_{N,\ell} = \{f(T) \in M_{N,\ell} \mid N \text{ even}, \ b_{N-i} = b_i, \ 0 \leqslant i \leqslant N/2, \\ f \text{ is monic and irreducible}\}.$$

This happens to be exactly the set K^1_g studied by Chavdarov (see Lemma 3.2 in [Chav]) for $g = N/2$. In loc. cit., Chavdarov proves that for N even,

$$\frac{1+\ell^{N/2}}{N} \geqslant |K^0_{N,\ell}| \geqslant \frac{\ell^{N/2}}{N} - \ell^{N/4},$$

and this completes the proof. □

Restricting ourselves to polynomials f satisfying $f = f_{\text{red}}$ (i.e., $N = \deg f$ is even and f is monic), we are interested in the same kind of computation as above with the extra condition that f must be such that $f(1)$ and $f(-1)$ are in a given class (not necessarily the same for $f(1)$ and $f(-1)$) of $\mathbb{F}_\ell^\times$ modulo its subgroup of nonzero squares, i.e., by Lemma 9 and Theorem 8, we work with polynomials having fixed discriminant and that can only be characteristic polynomials for elements with fixed spinor norm. At first it seems likely that we will get a "good" proportion of such polynomials among the set of irreducible self-reciprocal polynomials of even degree (that proportion should roughly be $1/4$). The following result of Meyn (see [Me], Theorem 8) tells us that this intuition is wrong:

Proposition 11 (Meyn). *Let f be a self-reciprocal monic polynomial of even degree N over $\mathbb{F}_\ell$. Let us write*

$$f = x^{N/2} h(x + x^{-1})$$

*with h monic of degree $N/2$. If h is an irreducible polynomial then f is irreducible iff $h(2)h(-2)$ is a **nonsquare** of $\mathbb{F}_\ell$.*

Moreover with notation as in the above statement it is easy to see that if we start with an irreducible f, the attached polynomial h will also be irreducible. So any self-reciprocal monic polynomial f with $f(1)$ and $f(-1)$ chosen in such a way that $(-1)^{N/2} f(1) f(-1)$ is a square in $\mathbb{F}_\ell$ is *not* irreducible. This means that out of the

four classes determined by the imposed value at 1 and -1 modulo squares, only two are non empty.

The following result asserts that the expected equidistribution property holds for the two non empty classes.

Lemma 12. *Let $N \geq 4$ be an even integer. If $\varepsilon_\ell^{(1)}, \varepsilon_\ell^{(2)}$ are (non necessarily distinct) fixed elements of a fixed set $\{1, \varepsilon_0\}$ of representatives of $\mathbb{F}_\ell^\times/(\mathbb{F}_\ell^\times)^2$ and if $(-1)^{N/2}\varepsilon_\ell^{(1)}\varepsilon_\ell^{(2)}$ is a nonsquare of $\mathbb{F}_\ell$, then we have*

$$\left|\{f \in M_{N,\ell} \mid f \text{ is irreducible and } f(a) \equiv \varepsilon_\ell^1,\ f(b) \equiv \varepsilon_\ell^2\}\right| \geq \frac{\ell^{N/2}}{2N}\left(1 - \frac{2(1+N)}{\ell}\right).$$

where, in the set of the left-hand side, congruences are taken modulo the group of nonzero squares of $\mathbb{F}_\ell$.

Proof. By Meyn's Theorem and the discussion preceding Lemma 12, the set we are interested in is in one-to-one correspondence with

$$\left\{h \in \mathbb{F}_\ell[T] \text{ monic of degree } N/2 \mid h \text{ is irreducible and } h(2) \equiv \varepsilon_\ell^{(1)},\ h(-2) \equiv (-1)^{N/2}\varepsilon_\ell^{(2)}\right\}.$$

If we vary $\varepsilon_\ell^{(1)}$ and $\varepsilon_\ell^{(2)}$ in $\{1, \varepsilon_0\}$, we see that computing the cardinality of the above set amounts to evaluating the four following character sums

$$\frac{1}{4}\sum_h (1 \pm \chi_\ell(h(2)))(1 \pm \chi_\ell(h(-2))),$$

where the sum is taken over all monic irreducible polynomials of $\mathbb{F}_\ell[T]$ with degree $N/2$ and where χ_ℓ denotes the Legendre character of $\mathbb{F}_\ell$.

It is enough to focus on the study of the sum

$$\begin{aligned} S &= \sum_h (1 + \chi_\ell(h(2)))(1 + \chi_\ell(h(-2))) \\ &= \sum_h 1 + \sum_h \chi_\ell(h(2)) + \sum_h \chi_\ell(h(-2)) + \sum_h \chi_\ell(h(2)h(-2)). \end{aligned}$$

The first sum of the right-hand side is nothing but the number of irreducible monic polynomials of degree $N/2$ in $\mathbb{F}_\ell[T]$. There are well-known lower bounds for that quantity. For our purpose, it is enough to use the inequality (see, e.g., Lemma 3.1 in [Chav])

$$\sum_h 1 \geq \frac{2\ell^{N/2}}{N} - \ell^{N/4}.$$

Next we consider both the sums $\sum_h \chi_\ell(h(2))$ and $\sum_h \chi_\ell(h(-2))$. As $N/2 \geqslant 2$ (and thus an irreducible h of degree $N/2$ cannot be $X \pm 2$) we have from the definition of χ_ℓ that

$$\sum_h \chi_\ell(h(2)) = 2 \sum_{a \text{ nonzero square}} |\{h \mid h(2) = a\}| - \sum_h 1.$$

Notice that the summand of the right-hand side does not depend on the point of $\mathbb{F}_\ell$ at which h is evaluated (i.e., imposing the value of h at any point of $\mathbb{F}_\ell$ yields a set with the same cardinality as the analogue set where the value imposed is $h(0)$), so using on the one hand the lower bound (see [KoSieve, Appendix B, formula B.8])

$$|\{h \in \mathbb{F}_\ell[T] \mid h \text{ monic irreducible }, \deg h = N/2, h(2) = a\}| \geqslant \frac{2\ell^{N/2-1}}{N} - \ell^{N/4},$$

and on the other hand the upper bound (see [Chav], Lemma 3.1, or Lemma B.1 in Appendix B of [KoSieve])

$$|\{h \in \mathbb{F}_\ell[T] \mid h \text{ monic irreducible }, \deg h = N/2\}| \leqslant \frac{2\ell^{N/2}}{N},$$

we get

$$\begin{aligned} -\sum_h \chi_\ell(h(2)) &\leqslant \frac{2\ell^{N/2}}{N} - (\ell-1)\left(\frac{2\ell^{N/2-1}}{N} - \ell^{N/4}\right) \\ &\leqslant \frac{2\ell^{N/2-1}}{N} + \ell^{N/4+1}. \end{aligned}$$

For the remaining sum, we need to be more careful (as the value of the polynomials considered at two different elements of $\mathbb{F}_\ell$ are involved). To begin with, the sum can be expressed as follows:

$$\sum_h \chi_\ell(h(2)h(-2)) = \frac{2}{N} \sum_{\alpha, \deg \alpha = N/2} \chi_\ell(\mathrm{Norm}((-1)^{N/2}(2-\alpha)(2+\alpha))),$$

where Norm denotes the norm map with respect to the extension $\mathbb{F}_{\ell^{N/2}}/\mathbb{F}_\ell$. Now, using the inclusion-exclusion principle, we get

$$\sum_{\alpha, \deg \alpha = N/2} \chi_\ell(\mathrm{Norm}(4-\alpha^2)) = \sum_{d \mid N/2} (-1)^{N/2-d} \sum_{\alpha \in \mathbb{F}_{\ell^d}} \chi_\ell(\mathrm{Norm}(4-\alpha^2)).$$

For each divisor d of $N/2$, if we set $\chi_{\ell,N} = \chi_\ell \circ \mathrm{Norm}$ (which is a multiplicative character of $\mathbb{F}_{\ell^d}$), we need to evaluate $\sum_{\alpha \in \mathbb{F}_{\ell^d}} \chi_{\ell,N}(4-\alpha^2)$. From the Riemann Hypothesis for curves over finite fields, we derive

$$\left| \sum_{\alpha \in \mathbb{F}_{\ell^d}} \chi_{\ell,N}(4-\alpha^2) \right| \leqslant \ell^{d/2},$$

since the polynomial $X^2 - 4$ has distinct roots in $\mathbb{F}_\ell$ (see [KaMul, Introduction and Theorem 1] for the statement and the proof of a more general result handling the case of higher dimensional varieties). Using the trivial fact that the number of divisors of $N/2$ is less than $N/2$, we get the upper bound:

$$-\sum_{\alpha,\deg\alpha=N/2} \chi_\ell(\mathrm{Norm}(4-\alpha^2)) \leqslant \frac{N}{2}\ell^{N/4}.$$

Thus

$$\sum_h \chi_\ell(h(2)h(-2)) \geqslant -\ell^{N/4}.$$

Finally, putting the above estimates together, we get

$$\mathcal{S} \geqslant \frac{2\ell^{N/2}}{N} - \frac{4\ell^{N/2-1}}{N} - 2\ell^{N/4+1} - 2\ell^{N/4} \geqslant \frac{2\ell^{N/2}}{N}\Big(1 - \frac{2(1+N)}{\ell}\Big). \quad \square$$

In the last lemma of this subsection, we are interested in counting monic polynomials h of degree $N/2$ with certain imposed factorization patterns and imposed value modulo squares at 2 and -2. Indeed it will be convenient, in the sections to come, to use such information in order to prove the existence of certain elements in the Galois group of an integral polynomial f whose reduction $f \pmod{\ell}$ can be written $f = x^{N/2}h(x + x^{-1})$.

Lemma 13. *Let $N \geqslant 4$ be an even integer and $\ell \geqslant 5$ be prime. With notation as above (e.g. all the congruences are taken modulo the subgroup of nonzero squares of $\mathbb{F}_\ell$), if we denote*

(i) $\tilde{\Theta}_{\ell,3}$ *for the set of monic polynomials $f \in M_{N,\ell}$ such that the corresponding h is separable, has an irreducible factor of prime degree $> N/4$ and such that $h(2) \equiv \varepsilon_\ell^{(1)}$, $h(-2) \equiv \varepsilon_\ell^{(2)}$, then we have*

$$|\tilde{\Theta}_{\ell,3}| \geqslant \frac{\ell^{N/2}}{N}\Big(\frac{7}{2} - \frac{1}{\ell}\Big(7 + \frac{15N}{2}\Big)\Big),$$

(ii) $\tilde{\Theta}_{\ell,4}$ *for the set of monic polynomials $f \in M_{N,\ell}$ such that the corresponding h is separable, has a unique irreducible quadratic factor, no other irreducible factor of even degree and such that $h(2) \equiv \varepsilon_\ell^{(1)}$, $h(-2) \equiv \varepsilon_\ell^{(2)}$, then we have*

$$|\tilde{\Theta}_{\ell,4}| \gg \frac{\ell^{N/2}}{N},$$

with an absolute implied constant.

Proof. (i) Let ℓ' be a prime such that $N/4 < \ell' \leqslant N/2$. The cardinality we are computing is greater than that of the set of monic polynomials h of degree $N/2$ which factor as the product of a monic irreducible polynomial of degree ℓ' with imposed values modulo squares at 2 and -2 with any monic irreducible polynomial of degree $N/2-\ell'$ (note that $N/2-\ell' < N/4$ so that no double root may occur in this way). Applying Lemma 12 (more precisely, using the arguments of the proof) and using once more the lower bound of [Chav, Lemma 3.1] we get

$$|\tilde{\Theta}_{\ell,3}| \geqslant \frac{\ell^{\ell'}}{4\ell'}\Big(1-\frac{2(1+2\ell')}{\ell}\Big)\Big(\frac{\ell^{N/2-\ell'}}{N/2-\ell'}-\ell^{N/4-\ell'/2}\Big).$$

As $N/4 < \ell' \leqslant N/2$, we have on the one hand

$$\frac{\ell^{\ell'}}{4\ell'}\Big(1-\frac{2(1+2\ell')}{\ell}\Big)\frac{\ell^{N/2-\ell'}}{N/2-\ell'} \geqslant \frac{4\ell^{N/2}}{N}\Big(1-\frac{2(1+N)}{\ell}\Big),$$

and on the other hand

$$\frac{\ell^{\ell'}}{4\ell'}\Big(1-\frac{2(1+2\ell')}{\ell}\Big)\ell^{N/4-\ell'/2} \leqslant \frac{\ell^{N/2}}{2N}\Big(1-\frac{2+N}{\ell}\Big).$$

Gathering those two inequalities we get the estimate we wanted to establish.

(ii) We consider separately the case where $N/2$ is odd and the case where $N/2$ is even. In the former case the cardinality of $\tilde{\Theta}_{\ell,4}$ is greater than that of the set of polynomials factoring as the product of an irreducible quadratic polynomial having imposed values modulo squares at 2 and -2 with any monic irreducible polynomial of degree $N/2-2$. Thus

$$|\tilde{\Theta}_{\ell,4}| \geqslant \frac{\ell^2}{8}\Big(1-\frac{10}{\ell}\Big)\Big(\frac{\ell^{N/2-2}}{N/2-2}-\ell^{N/4-1}\Big).$$

Now if $N/2$ is even the set we consider contains all the polynomials which, once divided by their quadratic factor (still with imposed values modulo squares at ± 2) are products of a polynomial of degree 1 (different from $X \pm 2$) with any irreducible polynomial of degree $N/2-3$ (note that if $N=4$ such a polynomial does not exist and if $N=8$ the other factor of degree 1 as well as the polynomials $X \pm 2$ must be removed from the set from which that polynomial of degree $N/2-3$ is picked). Thus, using the same inequalities as above,

$$\begin{cases} \text{if } N=4, & |\tilde{\Theta}_{\ell,4}| \geqslant \frac{\ell^2}{8}\big(1-\frac{10}{\ell}\big), \\ \text{if } N=8, & |\tilde{\Theta}_{\ell,4}| \geqslant \frac{\ell^2}{8}\big(1-\frac{10}{\ell}\big)(\ell-2)(\ell-3), \\ \text{if } N \geqslant 12, & |\tilde{\Theta}_{\ell,4}| \geqslant \frac{\ell^2}{8}\big(1-\frac{10}{\ell}\big)(\ell-2)\big(\frac{\ell^{N/2-3}}{N/2-3}-\ell^{(N-6)/4}\big). \end{cases}$$

In particular we get the estimate stated. □

2.3. Number of matrices with prescribed characteristic polynomial. In our sieving context we need a result which would be the analogue, in the orthogonal case, of [Chav, Theorem 3.5]. The point is that we need to know, for a fixed $f \in M_{N,\ell}$, how many matrices in $\mathrm{O}(N, \mathbb{F}_\ell)$ have a reversed characteristic polynomial equal to f. Towards such a computation our first task is to show that *there exists at least one* matrix $g \in \mathrm{O}(N, \mathbb{F}_\ell)$ such that $P_g = f$. This is a crucial step of the proof if we try to follow Chadarov's method. Unfortunately Chavdarov's proof relies heavily on the fact that the symplectic group $\mathbf{Sp}(2g)$ (as an algebraic group over $\mathbb{F}_\ell$) is simply connected and thus (by a Theorem of Steinberg), that the centralizer under $\mathbf{Sp}(2g)$ of any semisimple element in $\mathrm{Sp}(2g, \mathbb{F}_\ell)$ is a connected algebraic group to which Lang's Rationality Theorem may be applied. As we already mentioned neither $\mathbf{O}(N)$ nor $\mathbf{SO}(N)$ is a simply connected algebraic group. As a matter of fact we can easily construct examples of polynomials in $M_{N,\ell}$ which *are not* the reversed characteristic polynomials of any matrix in $\mathrm{O}(N, \mathbb{F}_\ell)$. Take, e.g., the polynomial $f(T) = T^2 + T + 1 \in M_{2,\ell}$ and suppose that the quadratic structure on the ambiant space $\mathbb{F}_\ell \times \mathbb{F}_\ell$ is given by the standard scalar product: $\Phi((x_1, y_1), (x_2, y_2)) = x_1x_2 + y_1y_2$. A straightforward computation shows that any matrix in $\mathrm{O}(2, \mathbb{F}_\ell)$ having f as its reversed characteristic polynomial must have its non diagonal coefficients equal to half a square root of 3, and this is obviously not always possible for matrices with coefficients in $\mathbb{F}_\ell$ (problems already occur for $\ell = 5$...).

While the direct adaptation of Chavadarov's method to the orthogonal case seems to be hopeless, we can however use a result of Baeza (see [Ba]) that gives a very useful criterion in the case where the dimension is even, to decide whether an $f \in M_{N,\ell}$ is or is not the reversed characteristic polynomial of a $g \in \mathrm{O}(N, \mathbb{F}_\ell)$. From Baeza's result we derive the following proposition:

Proposition 14. *Let N be even and $f \in M_{N,\ell}$ such that*

(1) *f is monic,*

(2) *f is separable,*

(3) $\mathrm{disc}(f) = \mathrm{disc}(Q)$.

Then there exists a semisimple element $A \in \mathrm{SO}(N, \mathbb{F}_\ell)$ such that

$$\det(1 - TA) = f.$$

The equality of discriminants (condition (3)) is seen as an equality of residue classes in $\mathbb{F}_\ell^\times/(\mathbb{F}_\ell^\times)^2$ (an equality of discriminants will always be understood in this way from now on). This condition is crucial and we easily see that, in our counterexample, $\mathrm{disc}(P) = -1$ while $\mathrm{disc}(Q) = 1$ in the case where $\ell = 5$.

Proof. From [Ba, Theorem 3.7], we know that the quadratic forms Q' on V such that there exists a $\sigma \in \mathrm{SO}(V, Q')$ satisfying

$$\det(1 - T\sigma) = \det(\sigma - T) = f(T),$$

are exactly those that can be written $s_*(K,n)$, in the notation of loc. cit. We briefly recall how these quadratic spaces are constructed: the separable $\mathbb{F}_\ell$-algebra $K = \mathbb{F}_\ell[T]/(f(T)) = \mathbb{F}_\ell[x]$ is equipped with the involution

$$x \mapsto x^{-1}.$$

If we consider the subalgebra $L = \mathbb{F}_\ell(x + x^{-1})$, then we have a norm map

$$n\colon K \to L,$$

that defines a non degenerate quadratic form (K,n) with coefficients in L. For any trace map $s\colon L \to \mathbb{F}_\ell$, we can then consider the composition $s \circ n\colon K \to \mathbb{F}_\ell$; it defines a non degenerate quadratic form on K. We denote $s_*(K,n)$ the quadratic space obtained (see [Ba, discussion following Proposition 3.6] and the references therein for details, notably concerning trace maps).

For such a fixed quadratic space $s_*(K,n)$, let us consider a $\sigma \in \mathrm{SO}(s_*(K,n))$ such that

$$\det(1 - T\sigma) = \det(\sigma - T) = f(T).$$

From [Ba, Theorem 1.2], we then have $\mathrm{disc}(f) = \mathrm{disc}(s_*(K,n))$ thus, by assumption, $\mathrm{disc}(s_*(K,n)) = \mathrm{disc}(Q)$. But we know that quadratic forms over $\mathbb{F}_\ell$ are classified by their discriminant: so $s_*(K,n) \simeq (V,Q)$. Finally, to the element σ corresponds a matrix $A \in \mathrm{SO}(V,Q) \simeq \mathrm{SO}(N,\mathbb{F}_\ell)$ such that

$$\det(1 - TA) = f(T).$$

The semisimplicity of A is obvious from the separability assumption on f. □

In that proof we see how the distinction between the two models of orthogonal groups (in the case N is even) naturally appears. In particular we notice that, with the notation of Baeza, the quadratic space $s_*(K,n)$ is precisely the one chosen by Katz in [KaL, Section 6] to describe a model for the nonsplit orthogonal group.

We are now ready to prove the main result of this section. Provided the assumptions of Proposition 14 are verified, the statement is the analogue in the orthogonal case of [Chav, Theorem 3.5] and it can be interpreted as a property of equidistribution of characteristic polynomials of orthogonal matrices among the polynomials of $M_{N,\ell}$. Apart from the use of Proposition 14 the arguments developed in the proof are quite close to those of loc. cit.

Theorem 15. *Let N be even and let $f \in M_{N,\ell}$ be such that*

(1) *f is monic,*

(2) *f is separable,*

(3) $\mathrm{disc}(f) = \mathrm{disc}(Q)$.

Then,

$$\left|\left\{B \in \mathrm{O}(N, \mathbb{F}_\ell) \mid f(T) = \det(1 - TB)\right\}\right| \gg \ell^{N^2/2-N},$$

with an implied constant independent of N.

Proof. Let A be a semisimple element of $\mathrm{SO}(N, \mathbb{F}_\ell)$ with reversed characteristic polynomial equal to f (the existence of such an A is justified by Proposition 14). Let

$$\Delta = \{B \in \mathrm{O}(N, \mathbb{F}_\ell) \mid \det(1 - TB) = f\}.$$

If $B \in \Delta$, its Jordan decomposition can be written

$$B = B_s B_u, \quad B_s B_u = B_u B_s, \quad B_s, B_u \in \mathrm{O}(N, \mathbb{F}_\ell),$$

with B_s semisimple and B_u unipotent.

In particular $\det(1 - TB_s) = \det(1 - TA)$, therefore the set of matrices B_s satisfying $\det(1 - TB_s) = f(T)$ contains the set of all semisimple matrices which are $\mathrm{SO}(N, \mathbb{F}_\ell)$-conjugate to A. Hence the lower bound

$$|\Delta| \geq \left|\left\{(B_s, B_u) \mid B_s \ \mathrm{SO}(N, \mathbb{F}_\ell)\text{-conjugate to } A, B_u \in (C_{\mathrm{SO}(N)}(B_s))_u(\mathbb{F}_\ell)\right\}\right|,$$

where $C_{\mathrm{SO}(N)}(B_s)$ (resp. $(C_{\mathrm{SO}(N)}(B_s))_u$) denotes the centralizer of B_s, seen as an algebraic group, under the action of $\mathbf{SO}(N)$ (resp. the unipotent part of that centralizer).

As already mentioned we cannot guarantee that the algebraic group $C_{\mathrm{SO}(N)}(A)$ is connected and we will denote $C_{\mathrm{SO}(N)}(A)^0$ its (connected) identity component. Then we argue as in [Chav, proof of Theorem 3.5] to obtain

$$|\{\text{Unipotent elements of } (C_{\mathrm{SO}(N)}(A))^0(\mathbb{F}_\ell)\}| = \ell^{\dim(C_{\mathrm{SO}(N)}(A))^0 - \mathrm{rk}(C_{\mathrm{SO}(N)}(A))^0}.$$

From [Bo, II.12.2, prop.], $C_{\mathrm{SO}(N)}(A)^0$ is a maximal torus in $\mathbf{SO}(N)$ (indeed $\mathbf{SO}(N)$ is a reductive group and from the separability assumption on f we know A is regular semisimple), thus

$$\mathrm{rk}(C_{\mathrm{SO}(N)}(A))^0 = \mathrm{rkSO}(N) = \frac{N}{2}.$$

This finally yields the lower bound:

$$|\Delta| \geq \ell^{\dim(C_{\mathrm{SO}(N)}(A))^0 - N/2} \frac{|\mathrm{SO}(N, \mathbb{F}_\ell)|}{|C_{\mathrm{SO}(N)}(A)(\mathbb{F}_\ell)|}.$$

Moreover a theorem of Steinberg asserts that the group of connected components of an algebraic group is always a subgroup of its fundamental group (see [SpSt, II

Corollary 4.4]). We deduce that $C_{\mathrm{SO}(N)}(A)$ has at most 2 connected components. Adapting the result of Nori (see [No]) used by Chavdarov ([Chav, p. 160]), we obtain

$$|C_{\mathrm{SO}(N)}(A)(\mathbb{F}_\ell)| \leqslant 2(\ell+1)^{\dim(C_{\mathrm{SO}(N)}(A))^0}.$$

Thanks to the formula (see, e.g., [Art, p. 147]) on the cardinality of the special orthogonal group over $\mathbb{F}_\ell$ in the even dimensional case, we deduce

$$(\ell-1)^{N(N-1)/2} \leqslant |\mathrm{SO}(N,\mathbb{F}_\ell)| \leqslant \ell^{N(N-1)/2}.$$

Combining those last inequalities, we get

$$|\Delta| \gg \ell^{-N/2}(\ell-1)^{N(N-1)/2},$$

thus $|\Delta| \gg \ell^{N^2/2-N}$, where in these last two inequalities the implied constant does not depend on N. □

The purpose of the last result we give in this section is to relate the cardinalities of a given conjugacy invariant subset of $\mathrm{O}(N,\mathbb{F}_\ell)$ and of the set of corresponding reversed characteristic polynomials in $M_{N,\ell}$. To that extent its main interest lies in how we can apply it to our sieving problem. The arguments being very close to those used in the proofs of Theorem 15 and Proposition 14, it seems fair to include it in this "independent section".

Lemma 16. *Let $N \geqslant 2$ be an integer and ℓ be an odd prime number. Consider a subset $\tilde{\Theta}_\ell$ with cardinality $\tilde{\theta}_\ell$ of $M_{N_{\mathrm{red}},\ell}$ such that the elements $f \in \tilde{\Theta}_\ell$ satisfy*

(1) *f is monic,*

(2) *f is separable,*

(3) $\mathrm{disc}(f) = \mathrm{disc}(Q)$,

(4) *either $f(-1)$ is a nonzero square of $\mathbb{F}_\ell$ for every $f \in \tilde{\Theta}_\ell$, or $f(-1)$ is a non square for every $f \in \tilde{\Theta}_\ell$.*

Moreover let $\varepsilon_\ell^{(1)}, \varepsilon_\ell^{(2)}$ be two elements of $\mathbb{F}_\ell$ each equal to ± 1 and such that $\varepsilon_\ell^{(2)}$ "is" the residue class in $\mathbb{F}_\ell^\times/\mathbb{F}_\ell^{\times 2}$ defined by condition (3) above and Lemma 9. Let

$$\Theta_\ell = \{g \in \mathrm{O}(N,\mathbb{F}_\ell) \mid (\det, \mathrm{N}_{\mathrm{Spin}})(g) = (\varepsilon_\ell^{(1)}, \varepsilon_\ell^{(2)}),\ \det(1-Tg)_{\mathrm{red}} \in \tilde{\Theta}_\ell\}.$$

If $\theta_\ell = |\Theta_\ell|$, we then have

$$\theta_\ell |\Omega(N,\mathbb{F}_\ell)|^{-1} \geqslant \tilde{\theta}_\ell \ell^{-N_{\mathrm{red}}/2}\left(1-\frac{1}{\ell+1}\right)^{N_{\mathrm{red}}(N_{\mathrm{red}}-1)/2}.$$

At first the above statement can look ambiguous as the integer N_{red} is not entirely defined by N but also depends on the matrix we consider. The point is that once the determinant is fixed (which is the case in the lemma since we restrict ourselves to matrices with determinant $\varepsilon_\ell^{(1)}$), there is only one integer N_{red} that can correspond to N, so that, a posteriori, the assertion of the lemma makes sense.

Proof. Let us first consider the auxiliary set

$$\{h \in \mathrm{O}(N_{\mathrm{red}}, \mathbb{F}_\ell) \mid (\det, \mathrm{N_{Spin}})(h) = (1, \tilde{\varepsilon}_\ell^{(2)}), \det(1 - Th) = \det(1 - Th)_{\mathrm{red}} \in \tilde{\Theta}_\ell\}, \tag{5}$$

where $\tilde{\varepsilon}_\ell^{(2)}$ is a fixed element of $\{1, \varepsilon_0\}$ (which still denotes a set of representatives for $\mathbb{F}_\ell^\times/\mathbb{F}_\ell^{\times 2}$). We can trivially inject the set (5) in Θ_ℓ via the map $h \mapsto h \oplus u$ where u is any representative of a fixed class of $\mathrm{O}(N - N_{\mathrm{red}}, \mathbb{F}_\ell)$ modulo $\Omega(N - N_{\mathrm{red}}, \mathbb{F}_\ell)$ (that class corresponding to the couple $(x, y) \in \{\pm 1\}$ such that $(1, \tilde{\varepsilon}_\ell^{(2)}) \times (x, y) = (\varepsilon_\ell^{(1)}, \varepsilon_\ell^{(2)})$) and the quadratic structure on the corresponding $(N - N_{\mathrm{red}})$-dimensional vector space being chosen with discriminant 1. Imbedding (5) in Θ_ℓ that way we end up with a N_{red}-dimensional quadratic space having the same discriminant as the ambiant N-dimensional quadratic space (V, Q). Moreover, from a fixed h in the set (5) we construct $(\Omega(N, \mathbb{F}_\ell) : \Omega(N_{\mathrm{red}}, \mathbb{F}_\ell))$ distinct elements of Θ_ℓ. Therefore, following the same idea as in [KoZeta, Lemma 7.2], we can now compute a lower bound for θ_ℓ involving $\tilde{\theta}_\ell$. First, we have

$$\theta_\ell \geqslant \frac{|\Omega(N, \mathbb{F}_\ell)|}{|\Omega(N_{\mathrm{red}}, \mathbb{F}_\ell)|} \sum_{f \in \tilde{\Theta}_\ell} |\{g \in \mathrm{O}(N_{\mathrm{red}}, \mathbb{F}_\ell) \mid (\det, \mathrm{N_{Spin}})(g) = (1, \varepsilon_\ell^{(2)}), \det(1 - Tg) = f\}|,$$

Thanks to the assumption (3) we know that Proposition 14 can be applied and hence that each summand of the right-hand side of the above inequality is nonzero. More precisely each of these quantities is equal to the cardinality of the set Δ (depending on the polynomial f) of the proof of Theorem 15. Following that proof and using the above inequality, we get

$$\theta_\ell \geqslant \ell^{-N_{\mathrm{red}}/2} \frac{|\Omega(N, \mathbb{F}_\ell)|}{|\Omega(N_{\mathrm{red}}, \mathbb{F}_\ell)|} \sum_{f \in \tilde{\Theta}_\ell} \frac{\ell^{d_f} |\mathrm{SO}(N_{\mathrm{red}}, \mathbb{F}_\ell)|}{|C_{\mathrm{SO}(N_{\mathrm{red}})}(A_f)(\mathbb{F}_\ell)|},$$

where, for each $f \in \tilde{\Theta}_\ell$, the matrix A_f is the semisimple element whose existence is guaranteed by Proposition 14 , $C_{\mathrm{SO}(N_{\mathrm{red}})}(A_f)$ denotes the centralizer of A_f under the action of $\mathbf{SO}(N_{\mathrm{red}})$ and d_f is the dimension of the identity component of that algebraic group. The proof of Theorem 15 yields

$$|C_{\mathrm{SO}(N_{\mathrm{red}})}(A_f)| \leqslant 2(\ell + 1)^{d_f}.$$

Now the derived group $\Omega(N_{\mathrm{red}}, \mathbb{F}_\ell)$ has index 2 in $\mathrm{SO}(N_{\mathrm{red}}, \mathbb{F}_\ell)$ so we have

$$\frac{\theta_\ell}{|\Omega(N, \mathbb{F}_\ell)|} \geqslant 2\ell^{-N_{\mathrm{red}}/2} \sum_{f \in \tilde{\Theta}_\ell} \frac{\ell^{d_f}}{2(\ell+1)^{d_f}}.$$

Thus

$$\begin{aligned}\theta_\ell |\Omega(N, \mathbb{F}_\ell)|^{-1} &\geqslant \ell^{-N_{\mathrm{red}}/2} \sum_{f \in \tilde{\Theta}_\ell} \left(1 - \frac{1}{\ell+1}\right)^{d_f} \\ &\geqslant \ell^{-N_{\mathrm{red}}/2} \left(1 - \frac{1}{\ell+1}\right)^{N_{\mathrm{red}}(N_{\mathrm{red}}-1)/2} \tilde{\theta}_\ell,\end{aligned}$$

since we have $d_f \leqslant \dim \mathrm{SO}(N_{\mathrm{red}}) = N_{\mathrm{red}}(N_{\mathrm{red}} - 1)/2$. □

Remark. One can wonder why, in all the computations performed in this section, we always give uniform bounds (with respect to the parameter N) for the quantities studied rather than asymptotic estimates which, most likely, would have been easier to establish and would suffice for the argument needed in the proof of Theorems 1 and 18. This is because we have in mind another possible application of the same kind of sieving method to the study of L-functions of families of elliptic curves over function fields (it is the main subject of [J]). In that work we obtain a lower bound for the proportion of elliptic curves with irreducible (up to trivial factors) L-function (seen as a $\mathbb{Q}$-polynomial) when the curve varies in a suitable algebraic family. That bound is uniform with respect to the common conductor of the family (provided the related estimates for the local densities involved are uniform as well).

3. Statement and proof of the main result

In this last section we state the main result of this paper which generalizes Theorem 1 in two different ways. To that purpose we show that for the two different kinds of group considered, Proposition 6 and Proposition 7 hold. Then, to derive our results from the large sieve inequality (2), we need to find a suitable lower bound for the constant H. That issue, in the case where the groups involved are orthogonal groups, can be handled thanks to the results of Section 2 and to a lemma proved in [J]. In the other case we consider ($G = \mathrm{GL}(n, A)$, $G^g = \mathrm{SL}(n, A)$), the question of finding a lower bound for H turns out to be easier, as we do not have as many constraints on the matrices considered as in the former case.

Let us first explain the additional terminology needed to state our result in full generality. Indeed, we need not (as Theorem 1 would suggest) restrict ourselves to the study of the irreducibility of characteristic polynomials of "random matrices", but

the sieve setting we are working with enables us to study the Galois group of those $\mathbb{Q}$-polynomials (see [KoSieve, Chapter 7] for the analogous study for the groups $\mathrm{SL}(n,\mathbb{Z})$ and $\mathrm{Sp}(2g,\mathbb{Z})$). The Galois group of the (splitting field over $\mathbb{Q}$ of) the characteristic polynomial of a matrix of $\mathrm{GL}(n,\mathbb{Z}[1/\mathcal{P}])$ can a priori be as big as the whole symmetric group $\mathfrak{S}_n$, but in the case of orthogonal matrices, (1) imposes conditions on the roots. If we denote $N = n + m$ and $N_{\mathrm{red}} = 2\lfloor (n+m)/2 \rfloor$, then the Galois group of (the splitting field over $\mathbb{Q}$ of) the reduced characteristic polynomial of a matrix $g \in \mathrm{SO}(n,m)(\mathbb{Z})$ is contained in the group denoted $W_{N_{\mathrm{red}}}$ which can be seen as the subgroup of $\mathfrak{S}_{N_{\mathrm{red}}}$ acting on $N_{\mathrm{red}}/2$ pairs of elements of $\{1,2,\ldots,N_{\mathrm{red}}\}$. However, as noticed in [J, Section 4.1], the biggest group the Galois group investigated can be equal to might be smaller than the full group $W_{N_{\mathrm{red}}}$. Indeed, depending on the sign of the functional equation (1) and on the fact that the underlying quadratic form splits over $\mathbb{Q}$ or not, the Galois group we study might be included a priori in a subgroup of index 2 of $W_{N_{\mathrm{red}}}$. More precisely, denoting by p the natural projection obtained by identifying each of the $N_{\mathrm{red}}/2$ pairs with a single letter, the group $W_{N_{\mathrm{red}}}$ fits the exact sequence

$$1 \longrightarrow \{\pm 1\}^{N_{\mathrm{red}}/2} \longrightarrow W_{N_{\mathrm{red}}} \xrightarrow{p} \mathfrak{S}_{N_{\mathrm{red}}/2} \longrightarrow 1,$$

whereas the "right" maximal Galois group might only be the subgroup of index two $W^+_{N_{\mathrm{red}}}$ of $W_{N_{\mathrm{red}}}$ fitting the exact sequence

$$1 \longrightarrow \{\pm 1\}^{N_{\mathrm{red}}/2-1} \longrightarrow W^+_{N_{\mathrm{red}}} \xrightarrow{p} \mathfrak{S}_{N_{\mathrm{red}}/2} \longrightarrow 1.$$

Remark. The group $W_{N_{\mathrm{red}}}$ is the group of all permutations of $N_{\mathrm{red}}/2$ pairs of letters combined with all possible sign changes of the corresponding pairs of letters. Its subgroup $W^+_{N_{\mathrm{red}}}$ is the group of permutations of $N_{\mathrm{red}}/2$ pairs of letters combined with evenly many sign changes among the pairs. Those two groups correspond respectively to the Weyl group of the root system of type $B_{N_{\mathrm{red}}}$ and $D_{N_{\mathrm{red}}}$.

Because of that ambiguity on what the maximal Galois group for the splitting field of $\det(T - X_k)_{\mathrm{red}}$ over $\mathbb{Q}$ should be if $X_k \in \mathrm{SO}(n,m)(\mathbb{Z})$, we define a notion of "small" Galois group (see [J, Definition 4.2]) as follows:

Definition 17. With notation as above let M be a matrix in $\mathrm{GL}(n,\mathbb{Z}[1/\mathcal{P}])$ (resp. $\mathrm{SO}(n,m)(\mathbb{Z})$). Let K_M be the splitting field over $\mathbb{Q}$ of the reduced characteristic polynomial $\det(T - M)_{\mathrm{red}}$. We say that $\det(T - M)_{\mathrm{red}}$ has *small Galois group* if $\mathrm{Gal}(K_M/\mathbb{Q})$ is a proper subgroup of $\mathfrak{S}_n$ (resp. $W^+_{N_{\mathrm{red}}}$).

Note that, as pointed in [J, discussion following Theorem 4.3], the group $W^+_{N_{\mathrm{red}}}$ is a transitive subgroup of $\mathfrak{S}_{N_{\mathrm{red}}}$. In particular the fact that $\det(T - M)_{\mathrm{red}}$ does *not* have small Galois group implies that $\det(T - M)_{\mathrm{red}}$ is $\mathbb{Q}$-irreducible.

We can also state a generalized version of the second part of Theorem 1. To do so, it is convenient to use (some of the basics of) the language of logic (as done in [KoDef]). Recall (see Section 2 of loc. cit.) that a *term* in the language of rings is simply a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$ and that an *atomic formula* φ is a formula of the form $f = g$ where f and g are polynomials (non necessarily in the same variables). Now if φ is an atomic formula, A is a ring and if we assign the family of elements $a = (a_i)$ to the set of variables involved in the definition of φ, we say that $\varphi(a)$ *is satisfied in* A and we denote

$$A \models \varphi(a),$$

if the equality which "is" φ is satisfied in A when the variables are given the values a_i. From atomic formulæ we can build the so-called *first order formulæ* by induction, using the symbols $\neg$, $\vee$, $\wedge$ and the quantifiers $\exists$, $\forall$ (we refer the reader to loc. cit. for examples of quite complicated formulæ that can be obtained in this way). Next, if $\varphi(x)$ is a first order formula with respect to the variables $x = (x_1, \ldots, x_n)$ and if A is a ring, then we denote

$$\varphi(A) = \{x \in A^n \mid A \models \varphi(x)\}.$$

With such a terminology, we can state the main result of this paper:

Theorem 18. *With the above notation (as well as those used in the introduction) and assuming that the first condition of Proposition* 5 *holds and that* $n + m \geq 6$, $nm \neq 0$ *(resp.* $n \geq 3$*) in the case* $G = \mathrm{SO}(n, m)(\mathbb{Z})$ *(resp.* $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$*), we have the following.*

(i) *There exists a* $\beta_3 > 0$ *such that for all* $k \geq 1$*, we have*

$$\mathbf{P}(\det(T - X_k)_{\mathrm{red}} \in A[T] \text{ is reducible or has small Galois group}) \ll \exp(-\beta_3 k),$$

with β_3 *depending only on the underlying algebraic group* $\mathbf{G}/\mathbb{Q}$*, on the generating set* S *and on the sequence* $(p_s)_s$ *(i.e., on the distribution of the* ξ_k*). Moreover the implied constant depends only on* $\mathbf{G}$ *and the density of* $\mathcal{P}$ *in the set of all rational prime numbers (in the case where* $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$*).*

(ii) *Let* φ *be a first order formula in the language of rings with respect to the variables* $x = (x_{i,j})_{1 \leq i,j \leq N}$ *(where* $N = n$ *if* $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$ *and* $N = n + m$ *if* $G = \mathrm{SO}(n, m)(\mathbb{Z})$*). Set*

$$\Lambda_\delta = \left\{\ell \text{ prime} \mid \left|\left(\neg\varphi(\mathbb{F}_\ell)\right) \cap Y_\ell\right| \cdot |G_\ell^g|^{-1} \geq \delta\right\},$$

and assume

$$\text{there exists } \delta > 0 \text{ such that } \Lambda_\delta \text{ has strictly positive Dirichlet density,} \qquad (6)$$

then there exists a $\beta_4 > 0$ *such that for all* $k \geqslant 1$, *we have*

$$\mathbf{P}(A \models \varphi(X_k)) \ll \exp(-\beta_4 k),$$

with the same dependency for β_4 *as for* β_3 *and the same dependency for the implied constant as in the previous case.*

Remarks. (i) If G is the subgroup of integral points of a special orthogonal group, the fact that we emphasize (e.g. in Theorem 1) the case where the quadratic structure is hyperbolic (i.e., the signature of the corresponding form is (n, n)) comes from the first condition of Proposition 5. Indeed, in the hyperbolic case, that condition is always fulfilled (as we will see in Lemma 19). Another (somewhat artificial) way to ensure we can find in the general case a relation of odd length among the elements of a generating set S could be to add the identity element to S. Doing so we would end up with a "lazy" random walk for which we would have at each step a probability $p_{\mathrm{Id}} > 0$ to stay at the same point.

The same problem occurs in the case where $G = \mathrm{SL}(2, \mathbb{Z}[1/\mathcal{P}])$ (and this explains why this case is omitted in the statement of Theorem 1); indeed, while the periodicity condition is always fulfilled for any generating system S of $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ if $n \geqslant 3$ (as will be proved in Lemma 20 (ii)), there are examples of such sets S for which that property does not hold in the case $n = 2$ (see [KoSieve, Section 7.4] for further details).

(ii) The fact that the Galois groups we investigate can be embedded in the Weyl group of the underlying algebraic group (at least in the split case) seems to be a quite general fact (it is proven by Kowalski for $\mathbf{SL}(n)$ and $\mathbf{Sp}(2g)$ in [KoSieve, Chapter 7] and the case of (the split form of) the exceptional group $\mathbf{E}_8/\mathbb{Q}$ is treated in [JKZ]).

(iii) In the case of orthogonal groups, the ambiguity as for what the maximal Galois group should be might appear as a bit unsatisfactory. That issue is addressed in [J], Remark (iv) following Lemma 4.6. The point here is that figuring what the "right" maximal Galois group ($W_{N_{\mathrm{red}}}$ or $W^+_{N_{\mathrm{red}}}$) is depends on the way the quadratic form considered splits modulo ℓ (where many values of ℓ are to be considered). That feature is particularly tricky because a nonsplit quadratic form over $\mathbb{Q}$ might give rise after reduction to a split quadratic form over $\mathbb{F}_\ell$ for every odd prime ℓ (see Remark (iv) following Lemma 4.6 in [J] for an example).

Before getting into the details of the proof, let us give a few more remarks on part (ii) of the statement. First, for a fixed first order formula φ, the set

$$\big(\neg\varphi(\mathbb{F}_\ell)\big) \cap Y_\ell$$

(recall that $Y_\ell = \rho_\ell(\alpha) G_\ell^g$) in the above statement is in fact the sieving set Θ_ℓ with index ℓ, in the notation of (2) and of the appendix. Next, to deduce the second part

of Theorem 1 from (ii) of Theorem 18, we choose for φ the formula

$$\varphi(x) : \bigvee_{1\leqslant i,j\leqslant N} \exists y, y^2 = x_{i,j}.$$

Thus, assuming Theorem 18, proving the second part of Theorem 1 is equivalent (once shown that the hypotheses of Proposition 7 are satisfied for the groups we study) to the fact that (6) holds for that choice of φ.

Finally, let us give examples of situations (i.e., choices of φ) for which (6) holds/does not hold: consider for instance the case where $\varphi(\mathbb{F}_\ell)$ is the set of $\mathbb{F}_\ell$-rational points of a subvariety $V/\mathbb{F}_\ell$ with codimension $\geqslant 1$ in $\mathbf{G}$. Then (6) clearly holds since $|V(\mathbb{F}_\ell)|$ is trivially bounded by $C\ell^{\dim V}$ where C is an absolute constant. In the opposite direction, if we investigate the probability with which the trace of a matrix in αG^g is sum of two squares, we quickly see that our method does not yield any quantitative information at all: indeed, if ℓ is an odd prime number, any element in $\mathbb{F}_\ell$ is the sum of two squares (a quite classical application of the pigeonhole principle) so that (6) is false for the choice

$$\varphi((x_{i,j})) : \exists a, \exists b, \sum_{i=1}^{N} x_{i,i} = a^2 + b^2.$$

The remaining sections are devoted to the proof of Theorem 1 and its generalization Theorem 18. In order to handle the case of orthogonal groups, we first review some useful facts concerning certain quadratic modules over $\mathbb{Z}$.

3.1. Quadratic modules over $\mathbb{Z}$. Let n, m be integers such that $n + m \geqslant 4$ and let (M, Q) be a quadratic module over $\mathbb{Z}$ with signature (n, m). The notion of spinor group we reviewed at the beginning of Section 2 can be extended to the case of quadratic modules (see [HM, Section 7.2A]): in that more general case, we still have a morphism

$$\mathrm{Spin}(M) \to \mathrm{O}(M),$$

the image of which is denoted $\Omega(M)$ and the kernel of which is ± 1 (see [HM, Theorem 7.2.21]). In other words, these groups fit the exact sequence

$$1 \to \{\pm 1\} \to \mathrm{Spin}(M) \to \Omega(M) \to 1, \tag{7}$$

where $\Omega(M)$ can once more be seen as the simultaneous kernel in $\mathrm{O}(M)$ of the determinant and the spinor norm (as defined in [HM, p 419]). In the arithmetic context we are interested in we will respectively denote by $\mathrm{O}(n, m)(\mathbb{Z})$, $\mathrm{SO}(n, m)(\mathbb{Z})$ and $\Omega(n, m)(\mathbb{Z})$ for $\mathrm{O}(M)$, $\mathrm{SO}(M)$ and $\Omega(M)$. Moreover, it will be convenient sometimes to see the first two of these groups as the groups of integral points of the algebraic groups $\mathbf{O}(n, m)/\mathbb{Q}$ and $\mathbf{SO}(n, m)/\mathbb{Q}$ respectively. The properties we need

$\Omega(n,m)(\mathbb{Z})$ to verify, in order to apply Propositions 6 and 7 are contained in the following lemma

Lemma 19. (i) *For integers n, m such that $n+m \geqslant 4$, we have*

(1) *If $d \geqslant 1$ is a squarefree integer whose only prime factors are outside a fixed finite set $\mathcal{S}$, then the reduction modulo d: $\Omega(n,m)(\mathbb{Z}) \to \Omega(n,m)(\mathbb{Z}/d\mathbb{Z})$ is onto.*

(2) *Property (τ) holds for $\Omega(n,m)(\mathbb{Z})$ with respect to the family of its congruence subgroups $\left(\ker(\rho_d : \Omega(n,m)(\mathbb{Z}) \to \Omega(n,m)(\mathbb{Z}/d\mathbb{Z}))\right)_{d\geqslant 1}$.*

(ii) *If in addition the quadratic module considered is hyperbolic (in the sense of the introduction, in particular $n = m$), and if $n \geqslant 3$, we have*

(1) *$\Omega(n,n)(\mathbb{Z})$ is finitely generated.*

(2) *For every symmetric generating system S of $\Omega(n,n)(\mathbb{Z})$ there exits a relation of odd length c inside S:*

$$s_1 \dots s_c = 1, \quad s_i \in S.$$

Proof. For (ii) we use the useful elements of the automorphism group of $\mathrm{O}(n,n)(\mathbb{Z})$ called *Eichler transformations* (see [HM, 5.2.9]). As the ambiant quadratic module (M, Q) we consider here is hyperbolic, we know from Theorem 9.2.14 of loc. cit. that these transformations span the subgroup $\Omega(n,n)(\mathbb{Z})$. This proves (ii)(1). Moreover the same result asserts that the *unitary elementary transformations* $E_{i,j}(1)$, where $i \neq j$ run over a finite set of indices, span the group $\Omega(n,n)(\mathbb{Z})$ and satisfy the commutator relation

$$[E_{i,j}(1) : E_{k,l}(1)] = E_{i,l}(1),$$

if i, j, k, l are distinct and run over the same (suitably chosen) set of indices (see [HM, Theorem 9.2.14]). Hence $\Omega(n,n)(\mathbb{Z})$ equals its own derived group. Now to prove (ii)(2) let us assume, by contradiction, that there is no relation of odd length among a fixed symmetric generating system S of $\Omega(n,n)(\mathbb{Z})$ and let $F(S)$ denote the free group generated by S. The morphism

$$\begin{aligned} F(S) &\to \{\pm 1\} \\ s &\mapsto -1, \end{aligned}$$

induces a surjective morphism $\Omega(n,n)(\mathbb{Z}) \to \{\pm 1\}$. Thus a quotient of $\Omega(n,n)(\mathbb{Z})$ is isomorphic to $\{\pm 1\}$; this contradicts the fact that $\Omega(n,n)(\mathbb{Z})$ is its own derived group.

For (i)(1) we first use strong approximation to justify the surjectivity of the reduction

$$\pi_p : \mathrm{Spin}(n,m)(\mathbb{Z}) \to \mathrm{Spin}(n,m)(\mathbb{F}_p),$$

where p runs over the set of all prime numbers which do not lie in a fixed finite set $\mathcal{S}$. Indeed, using the fact that the group $\mathrm{Spin}(n,m)(\mathbb{R})$ is not compact (the algebraic group $\mathbf{Spin}(n,n)$ is said to be *of non compact type*), we can apply Borel's Density Theorem (see [PR, Theorem 4.10]). Now $\mathrm{Spin}(n,m)(\mathbb{Z})$ is a Zariski dense subgroup of the simply connected algebraic group $\mathbf{Spin}(n,m)/\mathbb{Q}$, so strong approximation can be applied; more precisely, thanks to [PR, Theorem 7.15], we deduce the surjectivity of π_p provided p remains outside of a finite set $\mathcal{S}$ of prime numbers.

Then, using (7), we can consider the diagram with exact rows

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \{\pm 1\} & \longrightarrow & \mathrm{Spin}(n,m)(\mathbb{Z}) & \longrightarrow & \Omega(n,m)(\mathbb{Z}) & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \pi_p} & & \downarrow{\scriptstyle \pi_p} & & \downarrow & & \\ 1 & \longrightarrow & \{\pm 1\} & \longrightarrow & \mathrm{Spin}(n,m)(\mathbb{F}_p) & \longrightarrow & \mathrm{Spin}(n,m)(\mathbb{F}_p)/\{\pm 1\} & \longrightarrow & 1. \end{array}$$

If p is a prime not lying in $\mathcal{S}$, the two left vertical arrows of the above diagram are onto (as we have assumed to be working only with odd prime numbers). Then, if we define the last vertical arrow in such a way that the diagram commutes, this map must be onto as well. Moreover, it is easily seen that this last arrow also corresponds to the usual reduction modulo p

$$\Omega(n,m)(\mathbb{Z}) \to \Omega(n,m)(\mathbb{F}_p).$$

Now, for a squarefree integer d without any prime factor in $\mathcal{S}$, we invoke Goursat–Ribet's lemma (as stated in [Chav, Proposition 5.1]). Indeed, for $p \notin \mathcal{S}$, the group $\Omega(n,m)(\mathbb{F}_p)$ has no non central proper normal subgroup and the group $\Omega(n,m)(\mathbb{F}_p)$ modulo its center is a simple group. So, for such a d, we have a surjective morphism

$$\Omega(n,m)(\mathbb{Z}) \to \Omega(n,m)(\mathbb{Z}/d\mathbb{Z}).$$

Finally, for (3), we can apply [HV, Theorem 8, p. 23], thanks to which we know that, as $n+m \geqslant 4$, the group $\mathrm{SO}(n,m)(\mathbb{R})$ has Kazhdan's Property (T). Combining this with [HV, Corollary 5 to Theorem 4, p. 33] we deduce first that $\mathrm{SO}(n,m)(\mathbb{Z})$ has Property (T) and then that $\Omega(n,m)(\mathbb{Z})$ also has Property (T). The weaker Property (τ) for $\Omega(n,m)(\mathbb{Z})$ with respect to the family of its congruence subgroups follows immediately. □

Remark. In the above proof, we use the notation $\mathrm{Spin}(n,m)(\mathbb{F}_p)$ or $\Omega(n,m)(\mathbb{F}_p)$ just to keep track of the indefinite quadratic form over $\mathbb{Q}$ giving rise to the matrix groups for which we then take the reduction modulo p.

3.2. Proof of the main theorem. Recall that we are working out the two following cases.

• Either, for $n+m \geqslant 6$ and with notation as above, $G = \mathrm{SO}(n,m)(\mathbb{Z})$, $G^g = \Omega(n,m)(\mathbb{Z})$, and $\Gamma = \mathrm{SO}(n,m)(\mathbb{Z})/\Omega(n,m)(\mathbb{Z})$, in which case, these groups fit the following commutative diagram with exact rows and surjective *left* vertical map

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \Omega(n,m)(\mathbb{Z}) & \longrightarrow & \mathrm{SO}(n,m)(\mathbb{Z}) & \longrightarrow & \Gamma & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \pi_p} & & \downarrow{\scriptstyle \pi_p} & & \downarrow & & \\ 1 & \longrightarrow & \Omega(n,m)(\mathbb{F}_p) & \longrightarrow & \mathrm{SO}(n,m)(\mathbb{F}_p) & \longrightarrow & \Gamma_p & \longrightarrow & 1, \end{array}$$

provided $p \notin \mathcal{S}$ (see Lemma 19), and where Γ_p denotes the abelianization of $\mathrm{SO}(n,m)(\mathbb{F}_p)$.

• Or, for $n \geqslant 2$ and with notation as in the introduction, we set $G = \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}])$, $G^g = \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$, $\Gamma = \mathbb{Z}[1/\mathcal{P}]^\times$, in which case these groups fit the following commutative diagram with exact rows

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]) & \longrightarrow & \mathrm{GL}(n, \mathbb{Z}[1/\mathcal{P}]) & \longrightarrow & \Gamma & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \pi_p} & & \downarrow{\scriptstyle \pi_p} & & \downarrow & & \\ 1 & \longrightarrow & \mathrm{SL}(n, \mathbb{F}_p) & \longrightarrow & \mathrm{GL}(n, \mathbb{F}_p) & \longrightarrow & \mathbb{F}_p^\times & \longrightarrow & 1, \end{array}$$

provided $p \notin \mathcal{P}$. The question of the surjectivity of the downward arrows π_p (which still denote reduction modulo p) is easily answered here. Indeed, it is straightforward to check that $\mathrm{GL}(n,\mathbb{Z}) \to \mathrm{GL}(n,\mathbb{F}_p)$ is surjective and the fact that its restriction $\mathrm{SL}(n,\mathbb{Z}) \to \mathrm{SL}(n,\mathbb{F}_p)$ is also surjective is well known (see, for instance [Sh, Lemma 1.38] where the surjectivity is proved in the more general case of the reduction $\mathrm{SL}(n,\mathbb{Z}) \to \mathrm{SL}(n,\mathbb{Z}/d\mathbb{Z})$ modulo any positive integer d). As we want to apply Propositions 5, 6 and 7 to this case, we need the following analogue for the points of Lemma 19

Lemma 20. (i) *For $n \geqslant 2$, we have the following properties:*

(1) $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ *surjects onto* $\mathrm{SL}(n, \mathbb{Z}/d\mathbb{Z})$ *for each squarefree d without prime factors in $\mathcal{P}$.*

(2) $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ *has Property (τ) with respect to the family of its congruence subgroups*

$$\big(\ker(\rho_d \colon \mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}]) \to \mathrm{SL}(n, \mathbb{Z}/d\mathbb{Z}))\big)_d,$$

where d runs over the set of squarefree integers without prime factors in $\mathcal{P}$.

(ii) *If we suppose $n \geqslant 3$ then, for every symmetric generating system S of $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$, there exists a relation of odd length c:*

$$s_1 \dots s_c = 1, \quad s_i \in S.$$

Proof. We have just discussed point (i) (1) and (i) (2) is proven in Lemma 4.

For (ii), we note that, as the ring $\mathbb{Z}[1/\mathcal{P}]$ is euclidian, the group $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ is generated by the (infinite) set of transvection matrices $T_{i,j}(a)$ (the sum of the identity matrix and the matrix with all entries equal to zero except for the one in position (i, j) which equals a), where $a \in \mathbb{Z}[1/\mathcal{P}]^\times$ and $1 \leqslant i \neq j \leqslant n$. Such matrices satisfy the commutator relation

$$[T_{i,j}(a_1), T_{j,k}(a_2)] = T_{i,k}(a_1 a_2),$$

as soon as i, j, k are pairwise distinct (such a choice is indeed possible since $n \geqslant 3$). From that equality, we deduce that $\mathrm{SL}(n, \mathbb{Z}[1/\mathcal{P}])$ equals its own commutator subgroup. The end of the proof is then exactly the same as for the last point of Lemma 19. □

We are now ready to prove Theorem 18. Thanks to Lemmas 19 and 20, Propositions 6 and 7 hold with data corresponding to the two cases described above. To prove the exponential decrease of the probabilities investigated as k grows, we need to give a suitable lower bound for the constant

$$H = \sum_{\ell \in \mathcal{L}^*} \frac{\nu_\ell(\Theta_\ell)}{1 - \nu_\ell(\Theta_\ell)},$$

where $\mathcal{L}^*$ is the set of primes in Λ up to a fixed $L \geqslant 1$ and Λ is a set of primes with strictly positive Dirichlet density that we will make precise in due course.

Both for the conjugacy and the non-conjugacy coset sieve, we have

$$H \geqslant \sum_{\ell \in \mathcal{L}^*} \frac{|\Theta_\ell|}{|G_\ell^g|}.$$

We shall prove that

$$\frac{|\Theta_\ell|}{|G_\ell^g|} \gg 1, \tag{8}$$

with an implied constant depending only on n and on the underlying algebraic group $\mathbf{G}$, for our different choices of sieving sets Θ_ℓ and groups G_ℓ^g. Such an estimate will turn out to be sufficient to prove Theorem 18 (and deduce Theorem 1).

First we note this inequality is quite obvious in the setting of (ii) of Theorem 18 with the choice $\Lambda = \Lambda_\delta$ (the hypotheses are chosen purposely for that estimate to be true). Indeed, in that case, we have $\Theta_\ell = \left(\neg\varphi(\mathbb{F}_\ell)\right) \cap Y_\ell$, so that, by assumption, there exists a $\delta > 0$ such that

$$\frac{|\Theta_\ell|}{|G_\ell^g|} \geqslant \delta,$$

for all $\ell \in \Lambda_\delta$. Then, looking back at (2) and using Proposition 7 as well as the Prime Number Theorem, we can easily get (ii). Indeed, we have

$$\mathbf{P}\big(\mathbb{F}_\ell \models \varphi(\rho_\ell(X_k)) \text{ for all } \ell \leqslant L\big) \ll (1 + L^{(17d+4)/4}\exp(-\eta k))L^{-1}\log L,$$

with an implied constant depending only on n and the density of Λ_δ as a subset of the rational primes. Setting $L = \exp(\frac{4\eta k}{17d+4})$, we obtain (ii) of Theorem 18.

Now as far as (i) of Theorem 18 is concerned, the set Λ we choose is, depending on the type of group considered, either the set of primes which are not in $\mathcal{S}$ (see Lemma 19) or the complement of $\mathcal{P}$ in a subset (with finite complement as we will see later) of all prime numbers (see Lemma 20). The conjugacy coset sieve enables us to study the reduced characteristic polynomial of X_k as k grows. For a fixed $\alpha \in G$ we choose

$$\Theta_\ell = \{g \in \rho_\ell(\alpha)G_\ell^g \mid \det(T-g)_{\mathrm{red}} \in \tilde{\Theta}_\ell\}, \tag{9}$$

where $\tilde{\Theta}_\ell$ is, for each ℓ, a set of polynomials having imposed factorisation patterns. Each Θ_ℓ will be a conjugacy invariant subset of Y_ℓ. Moreover, in the case where $G = \mathrm{SO}(n,m)(\mathbb{Z})$ and $G^g = \Omega(n,m)(\mathbb{Z})$, the set Θ_ℓ can be seen as a subset of $\mathrm{O}(n+m,\mathbb{F}_\ell)$ consisting of matrices with fixed determinant (equal to 1) and fixed spinor norm.

Whereas the estimate (8) will be derived directly, for suitable families (Θ_ℓ) in the case where $G = \mathrm{GL}(n,\mathbb{Z}[1/\mathcal{P}])$, from [KoSieve, Appendix B], the analogue study for $G = \mathrm{SO}(n,m)(\mathbb{Z})$ requires a few additional computations based on the ideas of [KoZeta, Lemmas 7.1 and 7.2]. The strategy of loc. cit. relies on the fact that as soon as a few particular conjugacy classes of $W_{N_{\mathrm{red}}}$, or $W^+_{N_{\mathrm{red}}}$ (where we keep the notation $N_{\mathrm{red}} = 2\lfloor (n+m)/2\rfloor$) are detected in the Galois group of a polynomial $P \in \mathbb{Z}[T]$, the Galois group of P is necessarily isomorphic to the whole group (see [J, Lemma 4.4 and Section 4.2]). It is well known that such conjugacy classes can be detected through the study of the factorisation patterns of $P \pmod{\ell}$ with ℓ taking suitable prime values. As explained in [J, Lemma 4.4 and Section 4.2], it is enough to consider four distinct families (Θ_ℓ), three of which are related (via (9)) to a family of sets of polynomials $(\tilde{\Theta}_\ell)$ (for simplicity we will denote in the sequel $N = n+m$ and $N_{\mathrm{red}} = 2\lfloor (n+m)/2\rfloor$):

(1) Let $\tilde{\Theta}_\ell^{(1)}$ be the set of polynomials f in $M_{N_{\mathrm{red}},\ell}$

- which are irreducible *if N is odd* or which are irreducible with a fixed value modulo nonzero squares of $\mathbb{F}_\ell$ in -1 and satisfy $\mathrm{disc}(f) = \mathrm{disc}(Q)$ *if N is even and* $\mathrm{O}(N_{\mathrm{red}},\mathbb{F}_\ell) = \mathrm{O}(N,\mathbb{F}_\ell)$ *is nonsplit*,
- which factor as a product of two distinct monic irreducible polynomials of degree $N_{\mathrm{red}}/2$ *if* $N = N_{\mathrm{red}}$ *is even,* $\mathrm{O}(N_{\mathrm{red}},\mathbb{F}_\ell)$ *is split and* $\ell \equiv 1 \pmod 4$,
- which factor as a product of an irreducible monic quadratic polynomial and an irreducible polynomial of degree $N_{\mathrm{red}}-2$ *if* $N = N_{\mathrm{red}}$ *is even,* $\mathrm{O}(N_{\mathrm{red}},\mathbb{F}_\ell)$ *is split and* $\ell \equiv 3 \pmod 4$.

(2) Unlike the way the other families of sieving sets are described, we define $\Theta^{(2)}$ directly (i.e., we do not define first the set of characteristic polynomials that the elements of each $\Theta_\ell^{(2)}$ must have).

- If $N = N_{\mathrm{red}}$ is even and $\mathrm{SO}(N, \mathbb{F}_\ell)$ is a model for the split special orthogonal group over $\mathbb{F}_\ell$, we define $\Theta_\ell^{(2)}$ as the set of matrices $g \in \alpha_\ell \Omega(N, \mathbb{F}_\ell)$ such that $\det(1 - Tg) = \det(1 - Tg)_{\mathrm{red}}$ factors as a product of a quadratic irreducible polynomial with an irreducible polynomial of degree 4 as well as pairwise distinct irreducibles with odd degrees.
- Otherwise (i.e., N is odd or $\mathrm{SO}(N_{\mathrm{red}}, \mathbb{F}_\ell)$ is a model for the nonsplit orthogonal group over $\mathbb{F}_\ell$), we define $\Theta_\ell^{(2)}$ as the set of matrices $g \in \alpha_\ell \Omega(N, \mathbb{F}_\ell)$ such that $\det(1-Tg)_{\mathrm{red}}$ factors as a product of an irreducible quadratic polynomial with pairwise distinct irreducible polynomials of odd degrees.

(3) Let $\tilde{\Theta}_\ell^{(3)}$ be the set of polynomials f in $M_{N_{\mathrm{red}},\ell}$ with a fixed value modulo nonzero squares of $\mathbb{F}_\ell$ in -1, which satisfy $\mathrm{disc}(f) = \mathrm{disc}(Q)$ and with associated polynomial h (such that $f = x^n h(x + x^{-1})$) being separable with at least one factor of prime degree $> N_{\mathrm{red}}/4$.

(4) Let $\tilde{\Theta}_\ell^{(4)}$ be the set of polynomials f in $M_{N_{\mathrm{red}},\ell}$ with a fixed value modulo nonzero squares of $\mathbb{F}_\ell$ in -1, which satisfy $\mathrm{disc}(f) = \mathrm{disc}(Q)$ and with associated polynomial h being separable with one irreducible quadratic factor and no other irreducible factor of even degree.

Lemma 21. *For $1 \leqslant i \leqslant 4$ we have, with the above notation,*

$$\frac{|\Theta_\ell^{(i)}|}{|\Omega(N, \mathbb{F}_\ell)|} \gg 1,$$

with an implied constant depending only on N.

Proof. For $(\Theta_\ell^{(1)})_\ell$, let us first consider the case where $N = N_{\mathrm{red}}$ is even. It is enough in the nonsplit case to combine Lemma 12 and Lemma 16. This yields

$$\frac{|\Theta_\ell^{(1)}|}{|\Omega(N, \mathbb{F}_\ell)|} \geqslant \frac{1}{2N}\left(1 - \frac{2(1+N)}{\ell}\right)\left(1 - \frac{1}{\ell+1}\right)^{N(N-1)/2},$$

from which we derive the estimate we want.

In the split case, the estimate we need is exactly the statement of the lemmas 6.5 and 6.6 of [KaL]. Indeed, we deduce directly from loc. cit. that

$$\frac{|\Theta_\ell^{(1)}|}{|\Omega(N, \mathbb{F}_\ell)|} \geqslant \frac{1}{4N^2}.$$

Now if N is odd, we invoke another result of [KaL], namely Lemma 6.4, from which we get

$$\frac{|\Theta_\ell^{(1)}|}{|\Omega(N,\mathbb{F}_\ell)|} \geqslant \frac{1}{2N-2},$$

for $\ell \geqslant \max(7,(N-1)/2)$.

The estimate for $\Theta_\ell^{(2)}$ follows directly from [J, Lemma 4.6] which is itself deduced by combining several lemmas from [KaL, Section 6].

Finally, for $(\Theta_\ell^{(3)})_\ell$ and $(\Theta_\ell^{(4)})_\ell$, it is straightforward to verify that the combination of Lemma 13 and Lemma 16 yields the estimate of Lemma 21. □

Next we turn to the case of the conjugacy coset sieve for $\alpha\mathrm{SL}(n,\mathbb{Z}[1/\mathcal{P}])$. In that case, we have, for each $\ell \notin \mathcal{P}$, $G_\ell^g = \mathrm{SL}(n,\mathbb{F}_\ell)$ and the sieving sets we choose are still given by (9) with this time, for any conjugacy class $c \in \mathfrak{S}_n$ whose elements have a decomposition in disjoint cycles involving n_i cycles of length i for $1 \leqslant i \leqslant r$,

$$\tilde{\Theta}_{\ell,c} = \{f \in \mathbb{F}_\ell[T] \mid f \text{ has factorisation type } c \text{ and } f(0) = \det(\rho_\ell(\alpha))\},$$

where we say that a monic separable polynomial $f \in \mathbb{F}_\ell[T]$ of degree $r \geqslant 1$ has factorisation type $c \in \mathfrak{S}_r$ if f factors as

$$f = f_1 \dots f_r,$$

where f_i is a product of n_i distinct irreducible monic polynomials of degree i and $\sum i n_i = r$.

In the case of the trivial left coset $\rho_\ell(\alpha) = 1$, the estimate we need is given by Kowalski in [KoSieve, Appendix B, Lemmas B.2 and B.5]. For the case of the general left coset, it is straightforward (by performing the obvious change of variable sending $\rho_\ell(\alpha)$ to 1) to verify that the same estimate holds, so that we get the following result:

Lemma 22. *With the above notation, we have*

$$\frac{|\Theta_{\ell,c}|}{|\mathrm{SL}(n,\mathbb{F}_\ell)|} \gg 1,$$

as soon as $\ell > 16n^2$, with an implied constant depending only on n.

Note that, to perform our sieve, the above statement suggests we should remove the primes smaller than $16n^2$ from Λ but this does not affect the final result as the Dirichlet density of Λ remains unchanged.

Now we use the inequality

$$\mathbf{P}(\det(T-X_k)_{\mathrm{red}} \text{ has small Galois group}) \leqslant \sum \mathbf{P}\big(\mathrm{Gal}(\det(T-X_k)) \cap \Theta^\sharp = \emptyset\big),$$

where $\Theta^\sharp$ is the conjugacy class of $W_{N_{\mathrm{red}}}$ (resp. $\mathfrak{S}_n$) determined by the family $\Theta = (\Theta_\ell)_\ell$ and where the sum runs over the family $(\Theta^{(i)})_{1\leqslant i\leqslant 4}$ (resp. $(\Theta_c)_{c\in\mathfrak{S}_n^\sharp}$) if $G = \mathrm{SO}(n,m)(\mathbb{Z})$ and $N_{\mathrm{red}} = 2\lfloor (n+m)/2\rfloor$ (resp. $G = \mathrm{GL}(n,\mathbb{Z}[1/\mathcal{P}])$).

Looking back once more at (2) and applying both Proposition 6 and the Prime Number Theorem, we obtain, for the two types of groups investigated

$$\mathbf{P}(\det(T-X_k)_{\mathrm{red}} \text{ has small Galois group}) \ll (1+L^{(3d+2)/2}\exp(-\eta k))L^{-1}\log L,$$

with an implied constant depending on n and the (strictly positive) density of Λ in the set of all rational primes. If we set $L = \exp(\frac{2k\eta}{3d+2})$, then choosing for β_3 any positive real number smaller than $\frac{2\eta}{3d+2}$ yields (i) of Theorem 18.

3.3. Proof of Theorem 1. As a conclusion, we explain how to deduce Theorem 1 from Theorem 18. Notice first that the first part of Theorem 1 is a trivial consequence of (i) of Theorem 18. As explained in the previous subsection, the only thing we need to prove to get the second inequality of Theorem 1 is that the first order formula

$$\varphi(x): \bigvee_{1\leqslant i,j\leqslant N} \exists y, y^2 = x_{i,j},$$

yields sets $\Theta_\ell = \big(\neg\varphi(\mathbb{F}_\ell)\big)\cap Y_\ell$ (indexed by a set of primes Λ_δ to be determined) such that $|\Theta_\ell|\cdot|G_\ell^g|^{-1} \gg 1$.

For both the case where $G = \mathrm{SO}(n,m)(\mathbb{Z})$ and $G = \mathrm{SL}(n,\mathbb{Z}[1/\mathcal{P}])$, the above sets Θ_ℓ can be expressed in the following way

$$\Theta_\ell = \{g = (g_{i,j}) \in \alpha_\ell G_\ell^g \mid g_{i,j} \text{ is not a quare in } A/(\ell) = \mathbb{F}_\ell\},$$

with notation as in Theorem 1 and where G_ℓ^g denotes $\Omega(n,m)(\mathbb{F}_\ell)$ or $\mathrm{SL}(n,\mathbb{F}_\ell)$ depending on what is G. The element $\alpha_\ell = \rho_\ell(\alpha)$ determines, in both cases, the left coset of G_ℓ (with respect to G_ℓ^g) which contains Θ_ℓ.

We first consider the case where $G = \mathrm{SL}(n,\mathbb{Z}[1/\mathcal{P}])$ (so $G_\ell^g = \mathrm{SL}(n,\mathbb{F}_\ell)$ for $\ell\notin\mathcal{P}$). As a representative of the left coset of $\mathrm{GL}(n,\mathbb{F}_\ell)$ consisting of the matrices with fixed determinant say d_ℓ, we can choose the diagonal matrix α_ℓ given by $\alpha_{\ell,(1,1)} = d_\ell$ and $\alpha_{\ell,(i,i)} = 1$ if $i\geqslant 2$. Using the Legendre character $(\frac{\cdot}{\ell})$ to detect squares, we want to evaluate

$$\frac{1}{2|\mathrm{SL}(n,\mathbb{F}_\ell)|}\sum_{\substack{g\in\alpha_\ell\mathrm{SL}(n,\mathbb{F}_\ell)\\ g_{i,j}\neq 0}}\left(1+\left(\frac{g_{i,j}}{\ell}\right)\right).$$

Thus to obtain the inequality $|\Theta_\ell||\mathrm{SL}(n,\mathbb{F}_\ell)|^{-1}\gg 1$, it is enough to prove

$$\sum_{g\in(\alpha_\ell\mathrm{SL}(n,\mathbb{F}_\ell))_{i,j}}\left(\frac{g_{i,j}}{\ell}\right)\ll \ell^{d-1/2},$$

where we denote $(\alpha_\ell \mathrm{SL}(n,\mathbb{F}_\ell))_{i,j}$ the matrices of $\alpha_\ell \mathrm{SL}(n,\mathbb{F}_\ell)$ with a nonzero entry in position (i,j), and where d equals n^2-1, the dimension of the algebraic group $\mathbf{SL}(n)$.

Now, for each $g \in \alpha_\ell \mathrm{SL}(n,\mathbb{F}_\ell)$, there exists $h \in \mathrm{SL}(n,\mathbb{F}_\ell)$ such that $g = \alpha_\ell h$. The (i,j)-th entry of g is given by

$$g_{i,j} = \sum_k \alpha_{\ell,(i,k)} h_{k,j} = \alpha_{\ell,(i,i)} h_{i,j},$$

since the matrix α_ℓ is diagonal. So it is enough to prove

$$\left(\frac{\alpha_{\ell,(i,i)}}{\ell}\right) \sum_{h \in \mathrm{SL}(n,\mathbb{F}_\ell)_{i,j}} \left(\frac{h_{i,j}}{\ell}\right) \ll \ell^{d-1/2},$$

which can also be written in the following way:

$$\sum_{h \in \mathrm{SL}(n,\mathbb{F}_\ell)_{i,j}} \left(\frac{h_{i,j}}{\ell}\right) \ll \ell^{d-1/2}.$$

This inequality is proved in [KoSieve, Appendix B, Proposition B.4]. Thus (6) of Theorem 18 (ii) holds if we choose for Λ_δ the complementary of $\mathcal{P}$ in the rational primes, and we deduce the second part of Theorem 1 in the case where $G = \mathrm{GL}(n,\mathbb{Z}[1/\mathcal{P}])$.

Finally, in the case where $G = \mathrm{SO}(n,m)(\mathbb{Z})$ (i.e., $G_\ell^g = \Omega(n,m)(\mathbb{F}_\ell)$), things are slightly different as this time, the group $\Omega(n,m)(\mathbb{F}_\ell)$ cannot be seen as the group of $\mathbb{F}_\ell$-points of an algebraic group. In order to end up applying the same techniques as above, we need to relate for fixed indices $1 \leqslant i,j \leqslant n+m$ the cardinality of the sieving set

$$\Theta_\ell = \{g \in \alpha_\ell \Omega(n,m)(\mathbb{F}_\ell) \mid g_{i,j} \text{ is a square in } \mathbb{F}_\ell\},$$

to the cardinality of

$$\Theta_\ell^{\mathrm{SO}} = \{g \in \mathrm{SO}(n,m)(\mathbb{F}_\ell) \mid g_{i,j} \text{ is a square in } \mathbb{F}_\ell\}.$$

To that purpose, we first make a special choice for the basis thanks to which we identify the elements of $\mathrm{SO}(n,m)(\mathbb{F}_\ell)$ with their matrix representation. Indeed the surjectivity of the spinor norm (see Section 2) onto $\{\pm 1\}$ enables us to choose a vector, say e_1, such that $\mathrm{N}_{\mathrm{Spin}}(r_{e_1}) = -1$, where r_{e_1} denotes the reflection with respect to the hyperplane $\mathcal{H}_{e_1}$ which is orthogonal to e_1. We can consider the restriction of the quadratic form Q to $\mathcal{H}_{e_1}$. This restriction is still non degenerate ([OM, p. 139]) and we can choose the second vector e_2 of the basis we are constructing in such a way that the corresponding reflection $r_{e_2}^{\mathcal{H}_{e_1}}$ of $\mathcal{H}_{e_1}$ has spinor norm 1 (see [KaL, Proof of Lemma 6.3]). Then we can complete the basis of $\mathcal{H}_{e_1}$ with vectors $\{e_3,\ldots,e_{n+m}\}$ in

such a way that the matrix of $r_{e_2}^{\mathcal{H}_{e_1}}$ written in the basis $(e_2, \dots, e_{n+m})$ has diagonal coefficients $(-1, 1, \dots, 1)$ and zeros outside the diagonal. Finally we can extend $r_{e_2}^{\mathcal{H}_{e_1}}$ to the whole ambiant space by making it act trivially on e_1. We get a reflection r_{e_2} of the whole space. Now the product $r_{e_1} r_{e_2}$ is an element of $\mathrm{SO}(n,m)(\mathbb{F}_\ell)$ with spinor norm -1 (note that $\mathrm{N}_{\mathrm{Spin}}(r_{e_2}) = \mathrm{N}_{\mathrm{Spin}}(r_{e_2}^{\mathcal{H}_{e_1}})$) and the matrix M_{e_1,e_2} of $r_{e_1} r_{e_2}$ in the basis $(e_i)_i$ is the diagonal matrix with $M_{e_1,e_2}(i,i) = -1$ if $i = 1, 2$ and $M_{e_1,e_2}(i,i) = 1$ otherwise. Now consider the involution

$$\begin{aligned} \Theta_\ell &\to \Theta_\ell^\varepsilon \\ g &\mapsto M_{e_1,e_2} g \end{aligned}$$

where ε is any representative of the left coset of $\mathrm{SO}(n,m)(\mathbb{F}_\ell)$ consisting of elements with spinor norm -1 and

$$\Theta_\ell^\varepsilon = \{g \in (\varepsilon \alpha_\ell)\Omega(n,m)(\mathbb{F}_\ell) \mid g_{i,j} \text{ is a square in } \mathbb{F}_\ell\}.$$

We have $(M_{e_1e_2} g)_{i,j} = -g_{i,j}$ if $j = 1, 2$ and $(M_{e_1e_2} g)_{i,j} = g_{i,j}$ otherwise. So, for primes $\ell \equiv 1 \pmod 4$, $g_{i,j}$ is a square in $\mathbb{F}_\ell$ if and only if $(M_{e_1e_2} g)_{i,j}$ is a square as well. For such primes we deduce that

$$|\Theta_\ell^{\mathrm{SO}}| = |\Theta_\ell| + |\Theta_\ell^\varepsilon| = 2|\Theta_\ell|.$$

So

$$\begin{aligned} &\frac{1}{2|\Omega(n,m)(\mathbb{F}_\ell)|} \sum_{\substack{g \in \alpha_\ell \Omega(n,m)(\mathbb{F}_\ell) \\ g_{i,j} \neq 0}} \left(1 + \left(\frac{g_{i,j}}{\ell}\right)\right) \\ &= \frac{1}{4|\Omega(n,m)(\mathbb{F}_\ell)|} \sum_{\substack{g \in \mathrm{SO}(n,m)(\mathbb{F}_\ell) \\ g_{i,j} \neq 0}} \left(1 + \left(\frac{g_{i,j}}{\ell}\right)\right) \\ &= \frac{1}{2|\mathrm{SO}(n,m)(\mathbb{F}_\ell)|} \sum_{\substack{g \in \mathrm{SO}(n,m)(\mathbb{F}_\ell) \\ g_{i,j} \neq 0}} \left(1 + \left(\frac{g_{i,j}}{\ell}\right)\right), \end{aligned}$$

as soon as we restrict ourselves to primes such that $\ell \equiv 1 \pmod 4$.

To deduce from the last inequality that $|\Theta_\ell||\Omega(n,m)(\mathbb{F}_\ell)|^{-1} \gg 1$ (with an implied constant depending only on n and m), we use the exact same argument as in the previous case (where the finite group involved was $\mathrm{SL}(n, \mathbb{F}_\ell)$). Indeed, we can now see the sum investigated as a character sum over the $\mathbb{F}_\ell$-points of the geometrically irreducible variety $\mathbf{SO}(n,m)/\mathbb{F}_\ell$.

So we can choose $\Lambda_\delta = \{\text{primes congruent to 1 modulo 4}\}$ (which has Dirichlet density $1/2$) and then apply once more (ii) of Theorem 18 to get the second part of Theorem 1 in the case $G = \mathrm{SO}(n,m)(\mathbb{Z})$.

4. Appendix. Coset sieves

The purpose of this appendix is to explain the role that the large sieve plays in the proof of Theorem 18. We give here the full statements with proofs of the different kinds of a priori estimates we need to get the kind of explicit upper bounds of Theorem 18. The results we expose here are very much in the spirit of [KoSieve] (especially Section 3.3 of loc. cit.) and sometimes, we only recall some results of [KoSieve]. Moreover in the last subsection, we give self-contained versions of these statements in order to make it possible for the reader to follow the proof of Theorem 18 without having to get too much involved in the details of the sieving machinery.

4.1. The general framework. Our general sieving context is that of the *coset sieve*. A general description of that sieve goes as follows: we suppose we are given a discrete group G with a normal subgroup G^g such that the quotient $\Gamma = G/G^g$ is abelian. Moreover, we suppose that there exists a subset Λ of the rational primes such that, for any $\ell \in \Lambda$, we have a *surjective* group homomorphism

$$\rho_\ell : G \to G_\ell,$$

where G_ℓ is a *finite* group.

That is, of course, a very natural generalisation of the reduction modulo ℓ morphism from $\mathbb{Z}$ to $\mathbb{Z}/\ell\mathbb{Z}$. We emphasize here the fact that the above data is really all we need to set the sieve for cosets that we apply (and this is really the strong idea underlying [KoSieve, Chapter 3.3]). All the framework we build from there, comes from "natural" deductions. First, we denote

$$G_\ell^g = \rho_\ell(G^g)$$

for $\ell \in \Lambda$. That subgroup is normal in G_ℓ because G^g is a normal subgroup of G and, since for every $\ell \in \Lambda$, the morphism ρ_ℓ is onto. Now, all these groups fit the following commutative diagram with exact rows (such a diagram can already be found, in a geometric context, in [Chav, Theorem 4.1], and is extensively used in [KoZeta]):

$$\begin{array}{ccccccccc} 1 & \longrightarrow & G^g & \longrightarrow & G & \longrightarrow & \Gamma = G/G^g & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \rho_\ell} & & \downarrow{\scriptstyle \rho_\ell} & & \downarrow{\scriptstyle \mathrm{pr}_\ell} & & \\ 1 & \longrightarrow & G_\ell^g & \longrightarrow & G_\ell & \xrightarrow{d} & \Gamma_\ell = G_\ell/G_\ell^g & \longrightarrow & 1 \end{array} \tag{10}$$

where the surjective morphism pr_ℓ is defined in such a way that the diagram commutes. In both rows, the quotient map is denoted d, in order to avoid the introduction of an additional notation.

An important feature of that sieve setting, in view of the proof of (i) of Theorem 18, is that each left coset αG^g of G (α being any fixed element of G) is conjugacy invariant. This comes from the fact that the quotient $\Gamma = G/G^g$ is an abelian group.

We now fix an element $\alpha \in G$, and, use, from now on, the upper index $^\sharp$ to denote the set of conjugacy classes of the (conjugacy invariant) set considered. The two following sieve settings $(Y, \Lambda, (\rho_\ell : Y \to Y_\ell))$ will be useful for our purpose: either

- $Y = (\alpha G^g)^\sharp$, $Y_\ell = (\rho_\ell(\alpha) G_\ell^g)^\sharp$ and ρ_ℓ also denotes the restriction (which remains surjective) $Y \to Y_\ell$, for any $\ell \in \Lambda$ (this will be referred to as the *conjugacy coset sieve*); or
- $Y = \alpha G^g$, $Y_\ell = \rho_\ell(\alpha) G_\ell^g$ and ρ_ℓ also denotes the (surjective) restriction $Y \to Y_\ell$, for any $\ell \in \Lambda$ (this will be referred to as the *non-conjugacy coset sieve*).

As we do in the introduction, we assume we are given a probability space $(\Psi, \Sigma, \mathbf{P})$. The random walk (X_k) we are interested in (see the introduction again), can be seen as an application on Ψ with values in Y (whatever choice we make for the set Y among the two possibilities above). For each k, we end up with a triple $(\Psi, X_k, \mathbf{P})$. Following Kowalski's book [KoSieve], let us denote by $\mathcal{L}^*$ (the *prime sieve support*) the set of elements $\ell \in \Lambda$ that are smaller than a fixed integer $L \geqslant 1$. The large sieve method we use here consists in giving, for any family $\Theta = (\Theta_\ell)$ of subsets of Y_ℓ (the *sieving sets*) indexed by Λ, an upper bound for the probability

$$\mathbf{P}(\{x \in \Psi \mid \rho_\ell(X_k(x)) \notin \Theta_\ell \text{ for all } \ell \in \mathcal{L}^*\}).$$

For a matter of convenience, we will rewrite this probability in the standard way:

$$\mathbf{P}(\rho_\ell(X_k) \notin \Theta_\ell \text{ for all } \ell \in \mathcal{L}^*).$$

We need to make precise the meaning and the definition of the constants Δ and H appearing in the fundamental inequality 2 (the constant Δ is sometimes denoted $\Delta(X_k, \mathcal{L}^*)$ when we want to emphasize the dependency on the parameters). From now on, we will assume we are given a probability density ν_ℓ on Θ_ℓ for any $\ell \in \Lambda$; then the constant H can be taken to be equal to

$$H = \sum_{\ell \in \mathcal{L}^*} \frac{\nu_\ell(\Theta_\ell)}{1 - \nu_\ell(\Theta_\ell)}.$$

To define the large sieve constant $\Delta(X_k, \mathcal{L}^*)$, we should first emphasize that the central issue, in order to get a useful upper bound from (2), is to find a suitable basis for the space $L^2(Y_\ell, \nu_\ell)$ which is the complex Hilbert space with associated inner product (defined for $\mathbb{C}$-valued functions f and g on Y_ℓ):

$$\langle f ; g \rangle = \sum_{y \in Y_\ell} \nu_\ell(y) f(y) \overline{g(y)}.$$

If $\mathcal{B}_\ell$ denotes an orthonormal basis of that space containing the constant function 1 and if $\mathcal{B}_\ell^* = \mathcal{B}_\ell \setminus \{1\}$, then, for any square integrable function $\beta\colon \Psi \to \mathbb{C}$, the large sieve constant $\Delta(X_k, \mathcal{L}^*)$ is defined as the smallest constant Δ satisfying

$$\sum_{\ell\in\mathcal{L}^*}\sum_{\varphi\in\mathcal{B}_\ell^*}\left|\int_\Psi \beta(\omega)\varphi(\rho_\ell(X_k(\omega)))d\mathbf{P}(\omega)\right|^2 \leqslant \Delta\int_\Psi |\beta(\omega)|^2 d\mathbf{P}(\omega),$$

which can also be written, denoting by $\mathbf{E}(X)$ the expectation of a random variable X,

$$\sum_{\ell\in\mathcal{L}^*}\sum_{\varphi\in\mathcal{B}_\ell^*}\left|\mathbf{E}(\beta\cdot\varphi(\rho_\ell(X_k)))\right|^2 \leqslant \Delta\mathbf{E}(|\beta|^2).$$

The proof of the inequality (2) as we state it can be found in [KoSieve, Proposition 2.3]. To find an upper bound for $\Delta(X_k, \mathcal{L}^*)$, we use (see [KoZeta, Section 5] for an analogue given in a geometric context)

$$\Delta \leqslant \max_{\ell\in\mathcal{L}^*}\max_{\varphi'\in\mathcal{B}_\ell^*}\sum_{\ell'\in\mathcal{L}^*}\sum_{\varphi\in\mathcal{B}_\ell^*}|W(\varphi,\varphi')|, \tag{11}$$

where the $W(\varphi,\varphi')$ are the "exponential sums" given by

$$W(\varphi,\varphi') = \mathbf{E}(\varphi(\rho_{\ell'}(X_k))\overline{\varphi'(\rho_\ell(X_k))}). \tag{12}$$

Obviously the usefulness of (11) lies in the fact that it now suffices to give estimates for the individual sums $W(\varphi,\varphi')$ to deduce an upper bound for the large sieve constant. In our context where the sets Y_ℓ are left cosets in finite groups, it is natural to use the irreducible characters of the finite groups G_ℓ in order to construct a suitable basis $\mathcal{B}_\ell$. Moreover, that explains why it seems fair to call the $W(\varphi,\varphi')$ exponential sums.

4.2. Exhibiting orthonormal bases. In what follows, we give the description, for each $\ell \in \Lambda$, of the basis $\mathcal{B}_\ell$ we need (note that we need to handle both the conjugacy coset sieve and the non-conjugacy coset sieve).

First, we recall the following lemma due to Kowalski (see Lemma 3.2 in [KoSieve]) in which a basis for $L^2(Y_\ell, \nu_\ell)$ is described in the case of a conjugacy coset sieve (we recall that in that case, $Y_\ell = (\rho_\ell(\alpha)G_\ell^g)^\sharp$) where the density ν_ℓ is the uniform density defined by $\nu_\ell(y^\sharp) = |y^\sharp||G_\ell^g|^{-1}$.

Lemma 23. *With the same notation as above, let*

$$\varphi_\pi\colon y^\sharp \in G_\ell^\sharp \mapsto \mathrm{Tr}(\pi(y^\sharp)),$$

where $\ell \in \Lambda$ and π is an irreducible representation of G_ℓ. Let $L^2(Y_\ell, \nu_\ell)$ denote the Hilbert space of complex valued square integrable functions with respect to the

scalar product

$$\langle f\,;g\rangle = \frac{1}{|G_\ell^g|}\sum_{y^\sharp\in Y_\ell} \nu_\ell(y^\sharp) f(y^\sharp)\overline{g(y^\sharp)}.$$

Then we have the following:

(1) *If π and τ are irreducible representations of G_ℓ,*

$$\langle \varphi_\pi\,;\varphi_\tau\rangle = \begin{cases} 0 & \text{if } \pi_{|G_\ell^g} \not\simeq \tau_{|G_\ell^g} \text{ or } \varphi_{\pi|Y_\ell} = 0, \\ \overline{\psi(d\circ\rho_\ell(\alpha))}|\hat\Gamma_\ell^\pi| & \text{otherwise,} \end{cases} \tag{13}$$

where $\hat\Gamma_\ell$ is the character group of Γ_ℓ and ψ is an element of the group $\hat\Gamma_\ell^\pi = \{\psi\in\hat\Gamma_\ell \mid \pi\simeq\pi\otimes\psi\}$.

(2) *Let $\mathcal{B}_\ell$ be the family of functions*

$$\begin{aligned} Y_\ell &\to \mathbb{C} \\ y^\sharp &\mapsto |\hat\Gamma_\ell^\pi|^{-1/2}\varphi_\pi(y^\sharp), \end{aligned}$$

where π runs over the subset Π_ℓ^* of a set Π_ℓ of representatives for the irreducible representations of G_ℓ with respect to the equivalence relation

$$\pi\sim\tau \text{ if and only if } \pi_{|G_\ell^g}\simeq\tau_{|G_\ell^g},$$

and where $\pi\in\Pi_\ell^*$ if and only if $\varphi_{\pi|Y_\ell}\neq 0$. Then the family $\mathcal{B}_\ell$ is an orthonormal basis for $L^2(Y_\ell,\nu_\ell)$.

In the case of the non-conjugacy coset sieve, the irreducible representations of G_ℓ cannot be used in such a direct way to construct $\mathcal{B}_\ell$, but in spite of that, they turn out to be very useful once more.

Fix an $\ell\in\Lambda$ and a finite dimensional irreducible representation π_ℓ of G_ℓ. Let

$$B_{\pi_\ell} = (e_{\pi_\ell}^1,\dots,e_{\pi_\ell}^{\dim\pi_\ell})$$

denote an orthonormal basis of the representation space V_{π_ℓ} of π_ℓ, with respect to a G_ℓ-invariant inner product on V_{π_ℓ} denoted $\langle\,;\,\rangle_{\pi_\ell}$ (note that, ρ_ℓ having finite image, it is always possible to assume the existence of such a G_ℓ-invariant inner product). Then for any two elements e and f of $\{e_{\pi_\ell}^1,\dots,e_{\pi_\ell}^{\dim\pi_\ell}\}$, consider the function

$$\varphi_{\pi_\ell,e,f}\colon x\in G_\ell\mapsto\sqrt{\dim\pi_\ell}\langle\pi_\ell(x)e;f\rangle_{\pi_\ell},$$

called a *matrix coefficient.*

Then the family $(\varphi_{\pi_\ell,e,f})$ obtained by varying π_ℓ in Π_ℓ (where Π_ℓ denotes a set of representatives for the isomorphism classes of irreducible representations of G_ℓ)

and e, f in a fixed basis B_{π_ℓ}, forms an orthonormal basis for $L^2(G_\ell, \nu_\ell)$ of square integrable complex valued functions on G_ℓ with respect to the inner product

$$\langle f; g\rangle = \frac{1}{|G_\ell|}\sum_{x\in G_\ell} f(x)\overline{g(x)},$$

corresponding to the uniform density ν_ℓ defined for $y \in G_\ell$ by $\nu_\ell(y) = |G_\ell|^{-1}$ (see [Kn, Section I.5] for a proof).

From that result we can derive, in the non-conjugacy sieve setting, a useful orthonormal basis for $L^2(Y_\ell, \nu_\ell)$, where we recall that $Y_\ell = \rho_\ell(\alpha)G_\ell^g$ and where, for $y \in Y_\ell$, $\nu_\ell(y) = |G_\ell^g|^{-1}$:

Lemma 24. *With notation as above, consider the inner product on $L^2(Y_\ell, \nu_\ell)$ defined by*

$$\langle f; g\rangle = \frac{1}{|G_\ell^g|}\sum_{y\in Y_\ell} f(y)\overline{g(y)},$$

*Let π_ℓ and τ_ℓ be irreducible representations of G_ℓ and (e, f) (resp. (ε, ϕ)) a couple of elements of an orthonormal basis B_{π_ℓ} (resp. B_{τ_ℓ}) of V_{π_ℓ} (resp. V_{τ_ℓ}). The functions $\varphi_{\pi_\ell,e,f}$ and $\varphi_{\tau_\ell,\varepsilon,\phi}$ are said to be **equivalent** (in which case we will note $\varphi_{\pi_\ell,e,f} \sim \varphi_{\tau_\ell,\varepsilon,\phi}$) if the entry (e, f) of $\mathrm{Mat}_{B_{\pi_\ell}}\pi_\ell(g)$ and the entry (ε, ϕ) of $\mathrm{Mat}_{B_{\tau_\ell}}\tau_\ell(g)$ coincide for all $g \in G_\ell^g$.*

Then the following holds.

(1) *If π_ℓ and τ_ℓ are irreducible representations of G_ℓ, and if we denote*

$$\hat\Gamma_\ell^{\varphi_{\pi_\ell,e,f}} = \{\chi \in \hat\Gamma_\ell \mid \varphi_{\pi_\ell,e,f}\otimes\chi = \varphi_{\pi_\ell,e,f} \text{ in } L^2(G_\ell)\},$$

we have

$$\langle\varphi_{\pi_\ell,e,f}; \varphi_{\tau_\ell,\varepsilon,\phi}\rangle = \begin{cases} 0 & \text{if } \varphi_{\pi_\ell,e,f} \not\sim \varphi_{\tau_\ell,\varepsilon,\phi} \text{ or if the entry} \\ & (e, f)\,(\text{resp.}(\varepsilon, \phi)) \text{ of } \mathrm{Mat}_{B_{\pi_\ell}}\pi_\ell(g) \\ & (\text{resp. } \mathrm{Mat}_{B_{\tau_\ell}}\tau_\ell(g)) \text{ is zero for all} \\ & g \in Y_\ell, \\ \overline{\psi(d(\alpha_\ell))}|\hat\Gamma_\ell^{\varphi_{\pi_\ell,e,f}}| & \text{else, where } \alpha_\ell = \rho_\ell(\alpha),\ \psi \in \hat\Gamma_\ell, \\ & \text{and } \varphi_{\pi_\ell,e,f}\otimes\psi \simeq \varphi_{\tau_\ell,\varepsilon,\phi}. \end{cases}$$

(2) *Let $\mathcal{B}_\ell$ be the family of functions*

$$\begin{aligned} Y_\ell &\to \mathbb{C} \\ x &\mapsto |\hat\Gamma_\ell^{\varphi_{\pi_\ell,e,f}}|^{(-1/2)}\varphi_{\pi_\ell,e,f}(x), \end{aligned}$$

where (π_ℓ, e, f) runs over the triples corresponding to a system of representatives for the equivalence relation $\sim$ and where we assume that for every triple (π_ℓ, e, f), there exists an element $g \in Y_\ell$ such that the entry (e, f) of $\mathrm{Mat}_{B_{\pi_\ell}} \pi_\ell(g)$ *is nonzero. Then $\mathcal{B}_\ell$ is an orthonormal basis for $L^2(Y_\ell, \nu_\ell)$.*

Proof. (1) We evaluate the scalar product

$$\begin{aligned}\langle \varphi_{\pi_\ell,e,f}; \varphi_{\tau_\ell,\varepsilon,\phi}\rangle &= \frac{1}{|G_\ell^g|}\sum_{y\in Y_\ell} \varphi_{\pi_\ell,e,f}(y)\overline{\varphi_{\tau_\ell,\varepsilon,\phi}(y)}\\ &= \frac{1}{|G_\ell^g|}\sum_{\substack{y\in G_\ell\\ d(y)=d(\alpha_\ell)}} \varphi_{\pi_\ell,e,f}(y)\overline{\varphi_{\tau_\ell,\varepsilon,\phi}(y)}\\ &= \frac{1}{|G_\ell^g|}\frac{1}{|\hat{\Gamma}_\ell|}\sum_{y\in G_\ell}\Big(\sum_{\psi\in\hat{\Gamma}_\ell}\overline{\psi(\alpha_\ell)}\psi(y)\Big)\varphi_{\pi_\ell,e,f}(y)\overline{\varphi_{\tau_\ell,\varepsilon,\phi}(y)},\end{aligned}$$

where the last inequality is obtained using Frobenius reciprocity.

Obviously the right-hand side of the above equality vanishes as soon as $\varphi_{\pi_\ell,e,f\,|Y_\ell}$ or $\varphi_{\tau_\ell,\varepsilon,\phi|Y_\ell}$ is identically zero (which corresponds respectively to the vanishing of the entry (e, f) of $\mathrm{Mat}_{B_{\pi_\ell}} \pi_\ell(g)$ or (ε, ϕ) of $\mathrm{Mat}_{B_{\tau_\ell}} \tau_\ell(g)$, for all $g \in Y_\ell$). However, if that quantity does not vanish, we have, on the right-hand side,

$$\sum_{\psi\in\hat{\Gamma}_\ell}\overline{\psi(d(\alpha_\ell))}\langle \varphi_{\pi_\ell,e,f}\otimes\psi; \varphi_{\tau_\ell,\varepsilon,\phi}\rangle_{G_\ell}.$$

Now, in $L^2(Y_\ell, \nu_\ell)$, we have the equality of functions $\varphi_{\pi_\ell,e,f}\otimes\psi = \varphi_{\pi_\ell\otimes\psi,e,f}$. Indeed any G_ℓ-invariant scalar product $\langle\,;\,\rangle_{\pi_\ell}$ on $V_{\pi_\ell} \simeq V_{\pi_\ell\otimes\psi}$ (as vector spaces) remains G_ℓ-invariant if G_ℓ acts via $\pi_\ell\otimes\psi$ (which is still an irreducible representation of G_ℓ). We deduce

$$\langle \varphi_{\pi_\ell,e,f}; \varphi_{\tau_\ell,\varepsilon,\phi}\rangle = \sum_{\psi\in\hat{\Gamma}_\ell}\overline{\psi(d(\alpha_\ell))}\delta(\varphi_{\pi_\ell\otimes\psi,e,f}, \varphi_{\tau_\ell,\varepsilon,\phi}),$$

where δ denotes Kronecker's symbol.

The quantity $\delta(\varphi_{\pi_\ell\otimes\psi,e,f}, \varphi_{\tau_\ell,\varepsilon,\phi})$ equals 1 if and only if $\varphi_{\pi_\ell\otimes\psi,e,f} = \varphi_{\tau_\ell,\varepsilon,\phi}$ in $L^2(G_\ell)$. This condition is equivalent to the coincidence of the restrictions $\varphi_{\pi_\ell,e,f\,|G_\ell^g}$ and $\varphi_{\tau_\ell,\varepsilon,\phi|G_\ell^g}$, i.e., the equality between the entry (e, f) of $\mathrm{Mat}_{B_{\pi_\ell}} \pi_\ell(g)$ and the entry (ε, ϕ) of $\mathrm{Mat}_{B_{\tau_\ell}} \tau_\ell(g)$, for all $g \in G_\ell^g$. In that case we deduce, using [KoSieve, Lemma 3.2],

$$\langle \varphi_{\pi_\ell,e,f}; \varphi_{\tau_\ell,\varepsilon,\phi}\rangle = \overline{\psi(d(\alpha_\ell))}|\hat{\Gamma}_\ell^{\varphi_{\pi_\ell,e,f}}|,$$

where ψ is any of the characters of $\hat{\Gamma}_\ell$ such that

$$\varphi_{\pi_\ell,e,f}\otimes\psi \simeq \varphi_{\tau_\ell,\varepsilon,\phi}.$$

The assertion (2) is straightforward using (1) and the above arguments. □

Remarks. (i) In the course of the proof of Lemma 24, we have seen that the relation $\varphi_{\pi_\ell,e,f} \sim \varphi_{\tau_\ell,\varepsilon,\phi}$ is equivalent to the existence of a character $\psi \in \hat{\Gamma}_\ell$ such that

$$\varphi_{\pi_\ell,e,f} \otimes \psi = \varphi_{\pi_\ell\otimes\psi,e,f} = \varphi_{\tau_\ell,\varepsilon,\phi},$$

in $L^2(G_\ell)$. Then we proved that the scalar product

$$\langle \varphi_{\pi_\ell,e,f}; \varphi_{\tau_\ell,\varepsilon,\phi} \rangle,$$

is equal to $\overline{\psi(d(\alpha_\ell))}|\{\chi \in \hat{\Gamma}_\ell \mid \varphi_{\pi_\ell,e,f} \otimes \chi = \varphi_{\pi_\ell,e,f} \text{ in } L^2(G_\ell)\}|$. Thus, if π_ℓ is an irreducible representation of G_ℓ, the group $\hat{\Gamma}_\ell^{\pi_\ell}$, in the sense of Lemma 23, is a subgroup of $\hat{\Gamma}_\ell^{\varphi_{\pi_\ell,e,f}}$ for any choice of vectors e, f in an orthonormal basis of the representation space V_{π_ℓ}. Indeed, if $\psi \in \hat{\Gamma}_\ell^{\pi}$, we have an isomorphism of G_ℓ-representations: $\pi_\ell \otimes \psi \simeq \pi_\ell$, hence,

$$\begin{aligned}(\varphi_{\pi_\ell,e,f} \otimes \psi)(g) &= \varphi_{\pi_\ell\otimes\psi,e,f}(g) = \langle \pi_\ell \otimes \psi(g)e; f\rangle_{\pi_\ell} \\ &= \langle \pi_\ell(g)e; f\rangle_{\pi_\ell} = \varphi_{\pi_\ell,e,f}(g),\end{aligned}$$

for every $g \in G_\ell$. That means that the equivalence relation of Lemma 23 is "stronger" than the one described in Lemma 24 (in the sense that the classes for the former relation, which are contained in those for the latter, may form strict subsets in those classes).

(ii) Using the example of the dihedral group D_n, n even $\geqslant 2$, we see that the equivalence relation $\sim$ defined in Lemma 24 can actually be non trivial, i.e., we can exhibit two non isomorphic irreducible representations (π, V_π) and (τ, V_τ) of D_n and two couples of vectors (e, f) and (ε, ϕ) (respectively in V_π and V_τ) such that $\varphi_{\pi,e,f}(g) = \varphi_{\tau,\varepsilon,\phi}(g)$, for all $g \in G^g$, with a suitable choice of group G^g.

With notation as in [Ser, 5.3], let $G^g = C_n$, be the cyclic group of order n. It is of index 2 in D_n and we have an exact sequence of finite groups:

$$1 \to C_n \to D_n \to \mathbb{Z}/2\mathbb{Z} \to 1.$$

We fix the trivial left coset representative $\alpha = 1$ (with respect to the quotient D_n/C_n). If h_1 and h_2 are two distinct integers such that $h_1 \equiv -h_2 \pmod n$, then the representations ρ^{h_1} and ρ^{h_2} given in the canonical basis of $\mathbb{C}^2$ by

$$\rho^{h_i}(r^k) = \begin{pmatrix} \omega^{h_i k} & 0 \\ 0 & \omega^{-h_i k} \end{pmatrix},$$

(where, as in loc. cit., r is a generator for C_n, $0 \leqslant k \leqslant n-1$ and $\omega = \exp(2i\pi/n)$), are irreducible of degree 2 and are *not isomorphic*. A straightforward computation

shows that the canonical basis (e, f) of $\mathbb{C}^2$ is in fact an orthonormal basis of the representation space $V_{\rho^{h_i}}$, $i = 1, 2$, with respect to the D_n-invariant inner product $\langle\, ;\, \rangle_{\rho^{h_i}}$ constructed from the canonical scalar product on $\mathbb{C}^2$ by averaging over D_n. Then we have, for all $g \in C_n$,

$$\varphi_{\rho^{h_1},e,e}(g) = \varphi_{\rho^{h_2},f,f}(g),$$

since, for all $0 \leqslant k \leqslant n-1$, the equality $\omega^{h_1 k} = \omega^{-h_2 k}$ holds.

Via that example, we also see that there do exist functions $\varphi_{\pi,e,f}$ that vanish identically on a whole left coset of $G = D_n$ with respect to $G^g = C_n$. Indeed, we have, for $0 \leqslant k \leqslant n-1$,

$$\varphi_{\rho^{h_1},e,f}(r^k) = 0,$$

i.e., $\varphi_{\rho^{h_1},e,f}(g) = 0$ for all $g \in C_n$.

A common feature in the two lemmas above is that each individual sum $W(\varphi, \varphi')$ involves two (a priori) distinct representations (of two a priori distinct groups). In order to estimate those sums, it will be convenient to rewrite them in such a way that a single group representation appears for each of the $W(\varphi, \varphi')$. For that purpose, we need to introduce additional notation: for ℓ and ℓ' two elements of Λ, let $G_{\ell,\ell'}$ be the group

$$G_{\ell,\ell'} = \begin{cases} G_\ell \times G_{\ell'} & \text{if } \ell \neq \ell', \\ G_\ell & \text{otherwise.} \end{cases}$$

Now, for π (resp. τ) an irreducible representation of G_ℓ (resp. $G_{\ell'}$), we define the representation of $G_{\ell,\ell'}$:

$$[\pi, \tau] = \begin{cases} \pi \boxtimes \tau & \text{if } \ell \neq \ell' \\ \pi \otimes \tau & \text{otherwise,} \end{cases}$$

where "$\boxtimes$" (resp. "$\otimes$") denotes the external (resp. inner) tensor product of representations. With such notation we can give the statement of [KoSieve, Lemma 3.4] which is useful in the proofs of Propositions 6 and 7:

Lemma 25. *Let ℓ, ℓ' in Λ, π (resp. τ) a non trivial irreducible representation of G_ℓ (resp. of $G_{\ell'}$). The multiplicity of the trivial representation in the restriction of $[\pi, \bar{\tau}]$ to $G^g_{\ell,\ell'}$ is equal to zero if $(\ell, \pi) \neq (\ell', \tau)$, and is equal to $|\hat{\Gamma}^\pi_\ell|$ if $(\ell, \pi) = (\ell', \tau)$.*

If $\ell \neq \ell'$, it is well known that the family $([\pi, \tau])$ of representations of $G_{\ell,\ell'}$ with π (resp. τ) running over a system of representatives of irreducible representations of G_ℓ (resp. $G_{\ell'}$) forms itself a system of representatives for the irreducible representations of $G_{\ell,\ell'}$.

The above notation and the sieving context we would like to work with suggest us to combine the maps ρ_ℓ and $\rho_{\ell'}$ (assuming ℓ and ℓ' are distinct elements of Λ) in a single map

$$\begin{aligned}\rho_{\ell,\ell'}\colon G &\to G_{\ell,\ell'},\\ g &\mapsto (\rho_\ell(g),\rho_{\ell'}(g)),\end{aligned}$$

which is nothing but the product map from G to $G_\ell \times G_{\ell'}$.

Now we claim that the exponential sums (12) can be rewritten, according to the sieve setting considered, in one of the following forms:

- in the case of the conjugacy coset sieve, we have

$$W(\varphi_\pi,\varphi_\tau) = \frac{1}{\sqrt{|\hat{\Gamma}_\ell^\pi||\hat{\Gamma}_{\ell'}^\tau|}}\mathbf{E}(\mathrm{Tr}([\pi,\bar{\tau}]\rho_{\ell,\ell'}(X_k)), \tag{14}$$

  with notation as in Lemma 23,

- in the case of the non-conjugacy coset sieve, we have

$$W(\varphi_{\pi,e,f},\varphi_{\tau,\varepsilon,\phi}) = \sqrt{\frac{(\dim\pi)(\dim\tau)}{|\hat{\Gamma}_\ell^{\varphi_{\pi,e,f}}||\hat{\Gamma}_{\ell'}^{\varphi_{\tau,\varepsilon,\phi}}|}}\mathbf{E}(\langle[\pi,\bar{\tau}](\rho_{\ell,\ell'}(X_k))\tilde{e};\tilde{f}\rangle_{[\pi,\bar{\tau}]}), \tag{15}$$

  with notation as in Lemma 24 and where $\tilde{e} = e\otimes\varepsilon$, $\tilde{f} = f\otimes\phi$.

Both facts are a direct application of [KoSieve, Lemma 2.11].

4.3. Self-contained statements. We finish this appendix by giving self-contained statements (i.e., using no new terminology) for Lemmas 23 and 24 in order to make it possible for the reader to follow the whole proof of Theorem 1 without having to get too much involved (at least for a first reading of the paper) in the details of the sieve. To begin with, we give the following self-contained version of Lemma 23 (this is [KoSieve, Proposition 3.7]):

Proposition 26. *Let G be a group, G^g a normal subgroup of G with abelian quotient Γ; denote $d\colon G\to\Gamma$ the quotient map. Let Λ be a subset of the rational primes and let $\rho_\ell\colon G\to G_\ell$, for $\ell\in\Lambda$, be a family of surjective homomorphisms onto finite groups. Denote $G_\ell^g = \rho_\ell(G^g)$. Let $\alpha\in\Gamma$ be fixed, $Y = d^{-1}(\alpha)\subset G$ and $Y_\ell = \rho_\ell(Y)$. Let $(\Psi,\Sigma,\mathbf{P})$ be a probability space and X a random variable with values in Y. For any $\ell\in\Lambda$ let Π_ℓ be a set of representatives of the set of irreducible representations of G_ℓ modulo equality restricted to G_ℓ^g, containing the constant function 1. Moreover, let $\Pi_\ell^* = \Pi_\ell\setminus\{1\}$ and $\hat{\Gamma}_\ell^\pi$ be the set of characters ψ of $\Gamma_\ell = G_\ell/G_\ell^g$ such that $\pi\otimes\psi\simeq\pi$ for a representation π of G_ℓ.*

Let $\mathcal{L}^*$ be a finite subset of Λ. Then, for any conjugacy invariant subsets $\Theta_\ell \subset Y_\ell$ for $\ell \in \mathcal{L}^*$, we have

$$\mathbf{P}(\rho_\ell(X) \notin \Theta_\ell \text{ for all } \ell \in \mathcal{L}^*) \leqslant \Delta H^{-1},$$

where Δ is the smallest non-negative real number such that

$$\sum_{\ell \in \mathcal{L}^*} \sum_{\pi \in \Pi_\ell^*} \left|\mathbf{E}(\beta \cdot \operatorname{Tr} \pi(\rho_\ell(X)))\right|^2 \leqslant \Delta \mathbf{E}(|\beta|^2)$$

for all square-integrable functions $\beta \in L^2(\Psi, \mathbf{P})$, and

$$H = \sum_{\ell \in \mathcal{L}^*} \frac{|\Theta_\ell|}{|G_\ell^g| - |\Theta_\ell|}.$$

In addition, we have

$$\Delta \leqslant \max_{\ell \in \mathcal{L}^*} \max_{\pi \in \Pi_\ell^*} \sum_{\ell' \in \mathcal{L}^*} \sum_{\tau \in \Pi_{\ell'}^*} |W(\pi, \tau)|, \tag{16}$$

with

$$\begin{aligned} W(\pi, \tau) &= \frac{1}{\sqrt{|\hat{\Gamma}_\ell^\pi||\hat{\Gamma}_{\ell'}^\tau|}} \mathbf{E}(\operatorname{Tr} \pi(\rho_\ell(X)) \overline{\operatorname{Tr} \tau(\rho_{\ell'}(X))}) \\ &= \frac{1}{\sqrt{|\hat{\Gamma}_\ell^\pi||\hat{\Gamma}_{\ell'}^\tau|}} \mathbf{E}(\operatorname{Tr}[\pi, \bar{\tau}](\rho_{\ell,\ell'}(X))), \end{aligned} \tag{17}$$

using the notation $\rho_{\ell,\ell'}$ for the product map $\rho_\ell \times \rho_{\ell'} \colon G \to G_\ell \times G_{\ell'}$ if $\ell \neq \ell'$ and $\rho_{\ell,\ell'} = \rho_\ell$ otherwise, and $[\pi, \bar{\tau}] = \pi \otimes \bar{\tau}$ for the (internal or external, depending on whether $\ell = \ell'$ or not) tensor product of the representations π and $\bar{\tau}$.

The analogue self-contained statement in the non-conjugacy coset sieve setting is the following reformulation of Lemma 24:

Proposition 27. *Let $(G, G^g, \Lambda, (\rho_\ell), (G_\ell), (G_\ell^g))$, $(\Psi, \Sigma, \mathbf{P})$ and $\alpha, Y, (Y_\ell), X$ be as in Proposition 26. Moreover, for each $\ell \in \Lambda$ and each finite dimensional irreducible representation $\pi \in \operatorname{Irr}(G_\ell)$ (the set of isomorphism classes of such irreducible representations), let*

$$B_\pi = (e_\pi^1, \ldots, e_\pi^{\dim \pi})$$

be an orthonormal basis of the space of π with respect to a G_ℓ-invariant inner product $\langle\, ;\, \rangle_\pi$. For the set of triples $\{(\pi, e, f) \mid \pi \in \operatorname{Irr}(G_\ell), e, f \in B_\pi\}$, we denote by Π_ℓ a set of representatives for the following equivalence relation,

$$(\pi, e, f) \sim (\tau, \varepsilon, \phi) \quad \text{if} \quad \langle \pi(g)e; f\rangle_\pi = \langle \tau(g)\varepsilon; \phi\rangle_\tau \text{ for all } g \in G_\ell^g,$$

such that $(1, e, e) \in \Pi_\ell$ *(where* 1 *denotes the trivial representation and* e *is a basis for the* 1*-dimensional space attached to it) and such that there is no* $(\pi, e, f) \in \Pi_\ell$ *satisfying* $\langle \pi(g)e; f\rangle_\pi = 0$ *for all* $g \in G^g$*. Let* $\Pi_\ell^* = \Pi_\ell \setminus \{(1, e, e)\}$ *and let* $\mathcal{L}^*$ *be a finite subset of* Λ*. Then, for any subsets* $\Theta_\ell \subset Y_\ell$ *for* $\ell \in \mathcal{L}^*$*, we have*

$$\mathbf{P}(\rho_\ell(X) \notin \Theta_\ell, \text{ for } \ell \in \mathcal{L}^*) \leqslant \Delta H^{-1},$$

where Δ *is the smallest non-negative real number such that*

$$\sum_{\ell \in \mathcal{L}^*} \sum_{(\pi,e,f)\in\Pi_\ell^*} \sqrt{\dim(\pi)}\, \left|\mathbf{E}(\beta \cdot \langle \pi(\rho_\ell(X))e; f\rangle)\right|^2 \leqslant \Delta \mathbf{E}(|\beta|^2)$$

for all square-integrable functions $\beta \in L^2(\Psi, \mathbf{P})$*, where*

$$H = \sum_{\ell \in \mathcal{L}^*} \frac{|\Theta_\ell|}{|G_\ell| - |\Theta_\ell|}.$$

Moreover we have

$$\Delta \leqslant \max_{\ell \in \mathcal{L}^*} \max_{(\pi,e,f)\in\Pi_\ell^*} \sum_{\ell' \in \mathcal{L}^*} \sum_{(\tau,\varepsilon,\phi)\in\Pi_{\ell'}^*} |W((\pi, e, f), (\tau, \varepsilon, \phi))|, \tag{18}$$

where, with the same notations as in (17)*,*

$$\begin{aligned} &W((\pi, e, f), (\tau, \varepsilon, \phi)) \\ &\quad = \sqrt{\frac{(\dim \pi)(\dim \tau)}{|\hat{\Gamma}_\ell^{(\pi,e,f)}||\hat{\Gamma}_{\ell'}^{(\tau,\varepsilon,\phi)}|}}\, \mathbf{E}(\langle \pi(\rho_\ell(X))e; f\rangle_\pi \overline{\langle \tau(\rho'_\ell(X))\varepsilon; \phi\rangle_\tau}) \\ &\quad = \sqrt{\frac{(\dim \pi)(\dim \tau)}{|\hat{\Gamma}_\ell^{(\pi,e,f)}||\hat{\Gamma}_{\ell'}^{(\tau,\varepsilon,\phi)}|}}\, \mathbf{E}(\langle [\pi, \bar{\tau}](\rho_{\ell,\ell'}(X))(e \otimes \varepsilon); (f \otimes \phi)\rangle_{[\pi,\bar{\tau}]}) \end{aligned} \tag{19}$$

with $\hat{\Gamma}_\ell^{(\pi,e,f)}$ *denoting the set of characters* χ *of* Γ_ℓ *such that*

$$\langle \pi(g)e; f\rangle_\pi \cdot \chi(g) = \langle \pi(g)e; f\rangle_\pi.$$

Remark. In our description of coset sieves, we have restricted ourselves to sieve supports containing only *prime* numbers. Nevertheless as suggested by the discussions preceding and following Lemma 25, we could quite easily extend our sieve method to a framework in which we would use squarefree integers (and not only primes) as a sieve support. As described in [KoSieve, Chapter 3], going from a prime sieve support to a "squarefree" sieve support can be done naturally by extending a few of the definitions we have given in this appendix by multiplicativity. Although using that extended sieve support would surely yield better estimates in Theorem 1, we

prefer working only with a prime sieve support, so that we avoid the use of additional notation. However, for the proof of Proposition 7, it is convenient to use objects defined by multiplicativity from two (not more) primes in Λ. So, for $\ell \neq \ell'$ two such primes, let

$$Y_{\ell,\ell'} = Y_\ell \times Y_{\ell'},$$

on which we have the product density $\nu_{\ell,\ell'}(y, y') = \nu_\ell(y)\nu_{\ell'}(y')$, so that it makes sense to speak about the space $L^2(Y_{\ell,\ell'}, \nu_{\ell,\ell'})$. It is straightforward to check that if $\mathcal{B}_\ell$ (resp. $\mathcal{B}_{\ell'}$) is an orthonormal basis of $L^2(Y_\ell, \nu_\ell)$ (resp. of $L^2(Y_{\ell'}, \nu_{\ell'})$), the family of functions defined by $(y, y') \in Y_{\ell,\ell'} \mapsto \varphi(y)\varphi'(y')$, where $\varphi \in \mathcal{B}_\ell$ and $\varphi' \in \mathcal{B}_{\ell'}$, forms an orthonormal basis of $L^2(Y_{\ell,\ell'}, \nu_{\ell,\ell'})$. Note finally, that, to unify all the possible cases, we can extend the above definitions to the case $\ell = \ell'$ by defining $Y_{\ell,\ell'} = Y_\ell$, $\nu_{\ell,\ell'} = \nu_\ell$ and $\mathcal{B}_{\ell,\ell'} = \mathcal{B}_\ell$.

References

[Art] E. Artin, *Geometric algebra*. Interscience Tracts Pure Appl. Math. 3, Interscience Publishers, New York 1957. Zbl 0077.02101 MR 0082463

[ABS] M. F. Atiyah, R. Bott, and A. Shapiro, Clifford modules. *Topology* **3** (1964), suppl. 1, 3–38. Zbl 0146.19001 MR 0167985

[Ba] R. Baeza, Discriminants of polynomials and of quadratic forms. *J. Algebra* **72** (1981), 17–28. Zbl 0471.10017 MR 0634615

[Bo] A. Borel, *Linear algebraic groups*. Second edition. Grad. Texts in Math. 126, Springer-Verlag, New York 1991. Zbl 0726.20030 MR 1102012

[Ca] L. Carlitz, Some theorems on irreducible reciprocal polynomials over a finite field. *J. Reine Angew. Math.* **227** (1967), 212–220. Zbl 0155.09801 MR 0215815

[Chav] N. Chavdarov, The generic irreducibility of the numerator of the zeta function in a family of curves with large monodromy. *Duke Math. J.* **87** (1997), 151–180. Zbl 0941.14006 MR 1440067

[Chung] F. K. R. Chung, Diameters and eigenvalues. *J. Amer. Math. Soc.* **2** (1989), 187–196. Zbl 0678.05037 MR 0965008

[DeW] P. Deligne, La conjecture de Weil II. *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 137–252. Zbl 0456.14014 MR 0601520

[De] P. Deligne, *Cohomologie étale*. Séminaire de géométrie algébrique du Bois-Marie SGA $4\frac{1}{2}$, Lecture Notes in Math. 569, Springer-Verlag, Berlin 1977. Zbl 0349.14008 MR 0463174

[D] J. Dieudonné, *Sur les groupes classiques*. Hermann, Paris 1958. Zbl 0037.01304 MR 0024439

[E] B. H. Edwards, Rotations and discriminants of quadratic spaces. *Linear and Multilinear Algebra* **8** (1980), 241–246. Zbl 0438.15024 MR 0560566

[HM] A. J. Hahn and O. T. O'Meara, *The classical groups and K-theory*. Grundlehren Math. Wiss. 291, Springer-Verlag, Berlin 1989. Zbl 0683.20033 MR 1007302

[HV] P. de la Harpe and A. Valette, La propriété (T) de Kazhdan pour les groupes localement compacts. *Astérisque* **175** (1989). Zbl 0759.22001 MR 1023471

[Hu] J. E. Humphreys, *Linear algebraic groups*. Grad. Texts in Math. 21, Springer-Verlag, New York 1975. Zbl 0325.20039 MR 0396773

[IK] H. Iwaniec and E. Kowalski, *Analytic number theory*. Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, R.I., 2004. Zbl 1059.11001 MR 2061214

[J] F. Jouve, Maximal Galois group of L-functions of elliptic curves. *Internat. Math. Res. Notices* **2009** (2009), No. 19, 3557–3594. Zbl 05610927 MR 2539184

[JKZ] F. Jouve, E. Kowalski and D. Zywina, An explicit integral polynomial whose splitting field has Galois group $W(E_8)$. *J. Théor. Nombres Bordeaux* **20** (2008), no. 3, 761–782. Zbl 05572700 MR 2523316

[KaMul] N. M. Katz, Estimates for nonsingular multiplicative character sums. *Internat. Math. Res. Notices* **2002** (2002), no. 7, 333–349. Zbl 1028.11075 MR 1883179

[KaL] N. M. Katz, Report on the irreducibility of L-functions. To appear (in the Lang memorial volume).

[Kn] A. Knapp, *Representation theory of semisimple groups*. Princeton Math. Ser. 36, Princeton University Press, Princeton, N.J., 1986. Zbl 0604.22001 MR 0855239

[KoZeta] E. Kowalski, The large sieve, monodromy and zeta functions of curves. *J. Reine Angew. Math.* **601** (2006), 29–69. Zbl 05151112 MR 2289204

[KoDef] E. Kowalski, Exponential sums over definable subsets of finite fields. *Israel J. Math.* **160** (2007), 219–251. Zbl 1146.03020 MR 2342497

[KoSieve] E. Kowalski, *The large sieve and its applications*. Cambridge Tracts in Math. 175, Cambridge University Press, Cambridge 2008; The principle of the large sieve. arXiv:math.NT/0610021. Zbl 1177.11080 MR 2426239

[La] S. Lang, Algebraic groups over finite fields. *Amer. J. Math.* **78** (1956), 555–563. Zbl 0073.37901 MR 0086367

[LPS] A. Lubotzky, R. Philipps and P. Sarnak, Ramanujan graphs. *Combinatorica* **8** (1988), no 3, 261–277. Zbl 0661.05035 MR 0963118

[Lu] A. Lubotzky, *Discrete groups, expanding graphs and invariant measures*. Progr. Math. 125, Birkhäuser, Basel 1994. Zbl 0826.22012 MR 1308046

[LZ] A. Lubotzky and A. Żuk, On property (τ). Draft, online.

[Me] H. Meyn, On the construction of irreducible self-reciprocal polynomials over finite fields. *Appl. Algebra Engrg. Comm. Comput.* **1** (1990), no. 1, 43–53. Zbl 0724.11062 MR 1325510

[Mo] M. Morgenstern, Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q. *J. Combin. Theory Ser. B* **62** (1994), no. 1, 44–62. Zbl 0814.68098 MR 1290630

[No] M. Nori, On subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$. *Invent. Math.* **88** (1987), no. 2, 257–275. Zbl 0632.20030 MR 0880952

[OM] O. T. O'Meara, *Introduction to quadratic forms*. Grundlehren Math. Wiss. 117, Springer-Verlag, Berlin 1963. Zbl 0107.03301 MR 0152507

[PR] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*. Pure Appl. Math. 139, Academic Press, Boston, Mass., 1994. Zbl 0841.20046 MR 1278263

[Ser] J. P. Serre, *Représentations linéaires des groupes finis*. Coll. Méthodes, Hermann, Paris 1967. Zbl 0189.02603 MR 0232867

[Sh] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*. Kanô Memorial Lectures 1, Publ. Math. Soc. Japan 11, Iwanami Shoten Publishers, Tokyo; Princeton University Press, Princeton, N.J., 1971. Zbl 0221.10029 MR 0314766

[Ste1] R. Steinberg, *Endomorphisms of linear algebraic groups*. Mem. Amer. Math. Soc. 80, Amer. Math. Soc., Providence, R.I., 1968. Zbl 0164.02902 MR 0230728

[Ste2] R. Steinberg, *Conjugacy classes in algebraic groups*. Lecture Notes in Math. 366, Springer-Verlag, Berlin 1974. Zbl 0281.20037 MR 0352279

[SpSt] T. A. Springer and R. Steinberg, Conjugacy classes. In *1970 Seminar on algebraic groups and related finite groups*, Lecture Notes in Math. 131, Springer-Verlag, Berlin 1970, 167–266, Zbl 0249.20024 MR 0268192

[Sp] T. A. Springer, Conjugacy classes in algebraic groups. In *Group theory* (Beijing, 1984), Lecture Notes in Math. 1185, Springer-Verlag, Berlin, 175–209. Zbl 0624.20029 MR 0842444

[Za] H. Zassenhaus, On the spinor norm. *Arch. Math.* **13** (1962), 434–451. Zbl 0118.01804 MR 148760

Received July 22, 2008

Florent Jouve, Département de Mathématiques, Bâtiment 425, Faculté des Sciences d'Orsay, Université Paris-Sud 11, 91405 Orsay Cedex, France

E-mail: florent.jouve@math.u-psud.fr