

Elements Powering to Strong Involutions in Finite Classical Groups of Odd Characteristic: Order $1/\log n$ and a Log-Periodic Asymptotic

Alper Ferudun*

October 19, 2026

Abstract

Let G be a finite n -dimensional classical group over a field of odd order q , and let Q be the set of elements g of even order whose involution $g^{|g|/2}$ has a fixed space of dimension in $[n/3, 2n/3]$. Lübeck, Niemeyer and Praeger proved $|Q|/|G| \geq c/\log n$, and Praeger, with Lübeck and Niemeyer, asked at Oberwolfach in 2011 for the asymptotic value of $|Q|/|G|$, in particular whether $|Q|/|G| = O(1/\log n)$ (printed as $O(\log n)$, evidently a misprint). For the groups X of the table of Lübeck, Niemeyer and Praeger and for $\mathrm{Sp}_{2m}(q)$, $\mathrm{SO}_{2m+1}(q)$ and $\mathrm{SO}_{2m}^{\pm}(q)$ we prove $|Q|/|X| \leq C/\log n$ with absolute constants ($C = 100$ for $\mathrm{GL}_n(q)$ and $\mathrm{GU}_n(q)$; 45 for $\mathrm{Sp}_{2m}(q)$, $\mathrm{GSp}_{2m}(q)$, $\mathrm{SO}_{2m+1}(q)$, $\mathrm{GO}_{2m+1}(q)$; 80 for $\mathrm{SO}_{2m}^{\pm}(q)$, $\mathrm{GO}_{2m}^{\pm}(q)^0$), uniformly in q ; with their lower bound, $|Q|/|H| = \Theta(1/\log n)$ for every H between the groups $S \leq X$ of a line of the table, the upper constant carrying a factor at most $q + 1$ for the other H , among them $\mathrm{SL}_n(q)$ and $\mathrm{SU}_n(q)$. For the eight groups listed we also prove $(\log n)|Q|/|X| = \Phi(\log_2 \log n) + o(1)$ uniformly in q , where $\Phi(u) = \frac{\log 2}{3} \sum_{k \in \mathbb{Z}} 2^{u-k} e^{-2^{u-k}/3}$ is 1-periodic with mean 1, not constant, and within 10^{-5} of 1. Thus $|Q|/|X|$ is $1/\log n$ up to a factor that tends to within 10^{-5} of 1, while $(\log n)|Q|/|X|$ has no limit. The set Q depends only on semisimple parts, and the torus formula of Niemeyer and Praeger reduces the problem to random permutations, signed for the symplectic and orthogonal groups, whose cycles carry independent 2-adic levels; what is added here is a level-by-level upper bound, the torus model for the groups of similitudes, and the asymptotic analysis of the levels. The full orthogonal groups $\mathrm{O}_{2m}^{\pm}(q)$ are not covered; for $\mathrm{SL}_n(q)$, $\mathrm{SU}_n(q)$ and the groups strictly between S and X only the order of magnitude is proved, and for Ω only the upper bound; the convergence is slow, and the numerical values are evidence only. This is an unrefereed note.

1 Introduction

1.1 The question

Let q be a power of an odd prime, let G be a finite classical group with natural module V of dimension n (over $\mathbb{F}_{q^2}$ for the unitary groups, over $\mathbb{F}_q$ otherwise), and for $g \in G$ let $|g|$ be its order. Following [EHHO11, LNP09], let

$$Q(G) = \{g \in G : |g| \text{ is even and } n/3 \leq \dim \mathrm{Fix}_V(g^{|g|/2}) < 2n/3\}, \quad P(G) := \frac{|Q(G)|}{|G|}.$$

The involution $t = g^{|g|/2}$ is diagonalisable with eigenvalues ± 1 , so the condition says $n/3 < \dim V_{-1}(t) \leq 2n/3$; involutions with this property are the *strong involutions* of [LNP09], which are

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.
Unrefereed preprint. CC BY 4.0.

used in algorithms for the constructive recognition of classical groups (see [LNP09] for the motivation). Lübeck, Niemeyer and Praeger [LNP09, Theorem 1.1] proved that $P(H) \geq 1/(5000 \log_2 \ell)$ for every group H with $S \leq H \leq X$ in a line of their Table 1, where $\ell \geq 2$ is the Lie rank; the lines (S, X) are $(\mathrm{SL}_n(q), \mathrm{GL}_n(q))$, $(\mathrm{SU}_n(q), \mathrm{GU}_n(q))$, $(\mathrm{Sp}_{2m}(q), \mathrm{GSp}_{2m}(q))$, $(\mathrm{SO}_{2m+1}(q), \mathrm{GO}_{2m+1}(q))$ and $(\mathrm{SO}_{2m}^\pm(q), \mathrm{GO}_{2m}^\pm(q)^0)$. In the problem session of the Oberwolfach workshop on computational group theory of 2011, Praeger, jointly with Lübeck and Niemeyer, recalled this bound and asked [EHHO11, pp. 2154–2155]: “What is the asymptotic value of $|Q|/|G|$?” In particular, is $|Q|/|G| = O(\log(n))$? As printed, the second question is vacuous ($|Q|/|G| \leq 1$); it is evidently a misprint for $O(1/\log n)$, since the preceding sentence gives the lower bound $c/\log n$ and [LNP09] report statistical tests indicating the order $1/\log \ell$. The problem record OWR-11129-006 of the UnsolvedMath corpus reads the question in the same way, and so do we. In [LNP09] the authors state that their method gives no upper bound, and report sampled proportions of about 15% in dimension 1000.

Throughout, $\log$ is the natural logarithm. $\mathrm{GSp}_{2m}(q)$ and $\mathrm{GO}_{2m+1}(q)$ are the groups of similitudes of the forms, and $\mathrm{GO}_{2m}^\pm(q)^0 = \{g \in \mathrm{GO}_{2m}^\pm(q) : \det g = \tau(g)^m\}$, where $\tau(g)$ is the multiplier; this is the group of $\mathbb{F}_q$ -points of the connected component GO_{2m}^0 of the algebraic group of similitudes, and it is the index-2 subgroup of $\mathrm{GO}_{2m}^\pm(q)$ of [LNP09, Table 1].

1.2 Results

Theorem 1.1. *Let $G = \mathrm{GL}_n(q)$ or $G = \mathrm{GU}_n(q)$ with $n \geq 2$ and q odd. Then $P(G) \leq 100/\log n$, and $P(G) \geq 1/(300 \log n)$ if $n \geq 100$.*

Theorem 1.2. *Let X be one of $\mathrm{Sp}_{2m}(q)$, $\mathrm{GSp}_{2m}(q)$, $\mathrm{SO}_{2m}^\pm(q)$, $\mathrm{GO}_{2m}^\pm(q)^0$ (so $n = 2m$), or $\mathrm{SO}_{2m+1}(q)$, $\mathrm{GO}_{2m+1}(q)$ (so $n = 2m + 1$), with $n \geq 2$ and q odd. Then $P(X) \leq C_X/\log n$, where $C_X = 45$ for Sp_{2m} , GSp_{2m} , SO_{2m+1} , GO_{2m+1} and $C_X = 80$ for $\mathrm{SO}_{2m}^\pm$, $\mathrm{GO}_{2m}^\pm(q)^0$.*

Corollary 1.3. *Let $S \leq H \leq X$ be groups in a line of [LNP09, Table 1]. Then*

$$\frac{1}{5000 \log_2 \ell} \leq P(H) \leq \frac{100(q+1)}{\log n} \quad (\ell \geq 2),$$

so $P(H) = \Theta(1/\log n)$ for each such family and each q ; the upper constant is uniform in q when H is one of the groups of Theorems 1.1 and 1.2. For $\Omega_{2m+1}(q) \leq H \leq \mathrm{SO}_{2m+1}(q)$ and $\Omega_{2m}^\pm(q) \leq H \leq \mathrm{SO}_{2m}^\pm(q)$ the upper bound $P(H) \leq 2C_{\mathrm{SO}}/\log n$ holds, where C_{SO} is the constant of Theorem 1.2 for SO_{2m+1} , resp. $\mathrm{SO}_{2m}^\pm$.

The lower bound in Corollary 1.3 is [LNP09, Theorem 1.1]; the lower bound of Theorem 1.1 re-proves it for GL and GU with a better constant. So the answer to the question “is $|Q|/|G| = O(1/\log n)$?” is yes, for all groups of [LNP09, Table 1]; for $\mathrm{SL}_n(q)$, $\mathrm{SU}_n(q)$ and the groups strictly between S and X the upper constant of Corollary 1.3 depends on q . The asymptotic value, for the groups of Theorems 1.1 and 1.2, is given by the next theorem. Let

$$\Phi(u) = \frac{\log 2}{3} \sum_{k \in \mathbb{Z}} 2^{u-k} \exp(-2^{u-k}/3) \quad (u \in \mathbb{R}).$$

Theorem 1.4. (a) Φ is smooth and 1-periodic, with mean 1; its Fourier coefficients are $c_j = 3^{-i\tau_j} \Gamma(1 - i\tau_j)$, $\tau_j = 2\pi j/\log 2$, so Φ is not constant, and $|\Phi(u) - 1| \leq 2 \sum_{j \geq 1} |c_j| < 9.9 \cdot 10^{-6}$ for all u .

(b) Let X be one of the groups of Theorems 1.1 and 1.2. Then, uniformly in the odd prime power q ,

$$(\log n) P(X) = \Phi(\log_2 \log n) + o(1) \quad (n \rightarrow \infty).$$

Consequently $P(X) = (1 + \delta_n)/\log n$ with $\limsup |\delta_n| < 10^{-5}$, and the set of limit points of $(\log n)P(X)$ is the interval $[\min \Phi, \max \Phi]$; in particular $(\log n)P(X)$ has no limit.

The last statement holds because $\log_2 \log n$ increases to infinity with increments tending to 0, so it is dense modulo 1 along every tail. Numerically, $\max \Phi - 1$ and $1 - \min \Phi$ are both $9.884 \cdot 10^{-6}$ to four digits. The constants 100, 45 and 80 are far from optimal (the bounds are non-trivial only for very large n); their point is that they are absolute. The convergence in Theorem 1.4 is slow: for n between 10^6 and 10^7 the values of $(\log n)P(X)$ are near 1.10 (Table 1).

1.3 Method, and what is new

The set Q depends only on the semisimple part of an element and is closed under conjugation; the torus formula of Niemeyer and Praeger [NP10, Theorem 1.3], in the form [LNP09, Lemma 2.3], then expresses $P(X)$ as an average over maximal tori. For the classical groups this average is a statement about a random permutation of $N = n$ or $N = m$ points (with independent signs on its cycles for the symplectic and orthogonal groups), each cycle carrying an independent 2-adic *level*, the 2-adic valuation of the order of a uniform element of a cyclic group (Section 2). A torus element lies in Q if and only if the cycles of maximal level have total length in the middle third. At a level $v = a_1 - 1 + k$, where $2^{a_1} \parallel q^2 - 1$, only the cycles of length divisible by 2^k are active, and these weights do not depend on q . The upper bounds come from a lemma bounding each level separately (Lemma 4.1); the main term comes from the levels with 2^k comparable to $\log n$, and each of them contributes $\frac{\log 2}{3} 2^{-k} e^{-y_k/3}$ with $y_k = \log n / 2^k$, which sums to $\Phi(\log_2 \log n) / \log n$ (Section 6). These are the levels at which the 2-part of the order is of order $t \log n$, where t is the larger of the 2-parts of $q \pm 1$; that the 2-part of the order of a uniform element lies in such a range with probability bounded below was proved by Guest and Praeger [GP12, Theorem 1.2].

The reduction to tori and the description of the tori and of their involutions are from [LNP09, NP10, NPP14], and the lower bound is [LNP09, Theorem 1.1]. New here, to our knowledge, are the upper bounds (Lemma 4.1, Theorems 1.1 and 1.2), the torus model for the groups of similitudes (Lemma 2.5) and Theorem 1.4; we found no printed upper bound or asymptotic in the literature we read (see “Scope and priority” in Section 7), but a search that finds nothing is not a proof of novelty.

2 The torus model

2.1 Reduction to maximal tori

Lemma 2.1. *Let G be one of the groups above and $g = su$ the Jordan decomposition of $g \in G$ (s semisimple, u unipotent, $su = us$). Then $g \in Q(G)$ if and only if $s \in Q(G)$.*

Proof. $|u|$ is a power of the odd characteristic and $|s|$ is prime to it, so $|g| = |s||u|$ is even iff $|s|$ is. If $|s|$ is even, then $g^{[g]/2} = (s^{[s]/2})^{|u|} u^{|u||s|/2} = s^{[s]/2}$, since $s^{[s]/2}$ is an involution and $|u|$ is odd. $\square$

So $Q(G)$ is a union of conjugacy classes which contains g iff it contains the semisimple part of g , a *quokka set* in the sense of [NP10]. Let $\hat{G}$ be a connected reductive group over $\overline{\mathbb{F}}_q$ with a Frobenius morphism F , let T_0 be an F -stable maximal torus and $W = N_{\hat{G}}(T_0)/T_0$. The $\hat{G}^F$ -classes of F -stable maximal tori correspond to the F -classes of W (the classes of the relation $w \sim x^{-1}wF(x)$); let T_C be a torus in the class corresponding to C ([Car85, 3.3.3], as cited in [LNP09]).

Proposition 2.2 ([NP10, Theorem 1.3]; [NPP14, Theorem 2.2]; [LNP09, Lemma 2.3]). *If q is odd, $G = \hat{G}^F$ and $Q \subseteq G$ is a quokka set, then*

$$\frac{|Q|}{|G|} = \sum_C \frac{|C|}{|W|} \cdot \frac{|T_C^F \cap Q|}{|T_C^F|}. \quad (1)$$

For the set Q of strong involutions this is [LNP09, Lemma 2.3]. It applies to all groups of Theorems 1.1 and 1.2: $\mathrm{GL}_n(q)$ and $\mathrm{GU}_n(q)$ are the fixed points of GL_n under the standard and the twisted Frobenius morphism, $\mathrm{Sp}_{2m}(q)$, $\mathrm{SO}_{2m+1}(q)$ and $\mathrm{GSp}_{2m}(q)$ those of the connected groups Sp_{2m} , SO_{2m+1} and GSp_{2m} , $\mathrm{SO}_{2m}^\pm(q)$ and $\mathrm{GO}_{2m}^\pm(q)^0$ those of SO_{2m} and GO_{2m}^0 under the standard and the graph-twisted Frobenius morphism, and $\mathrm{GO}_{2m+1}(q)$ is the group of points of the connected group $\mathbb{G}_m \times \mathrm{SO}_{2m+1}$: a similitude g of multiplier c in odd dimension $2m+1$ satisfies $\det(g)^2 = c^{2m+1}$, so $z = \det(g)/c^m$ has $z^2 = c$, and $z^{-1}g$ is an isometry of determinant 1. Over $\mathbb{F}_q$ this gives $\mathrm{GO}_{2m+1}(q) = \mathbb{F}_q^\times \times \mathrm{SO}_{2m+1}(q)$.

2.2 Levels

Put $a_0 = v_2(q-1)$, $b_0 = v_2(q+1)$ and $a_1 = a_0 + b_0 = v_2(q^2-1) \geq 3$; one of a_0, b_0 equals 1. For $j \geq 1$ let $a(j) = v_2(q^j-1)$, $a^U(j) = v_2(q^j - (-1)^j)$ and $b(j) = v_2(q^j+1)$. Then

$$a(j) = \begin{cases} a_0 & j \text{ odd} \\ a_1 + v_2(j) - 1 & j \text{ even} \end{cases}, \quad a^U(j) = \begin{cases} b_0 & j \text{ odd} \\ a_1 + v_2(j) - 1 & j \text{ even} \end{cases}, \quad b(j) = \begin{cases} b_0 & j \text{ odd} \\ 1 & j \text{ even} \end{cases}. \quad (2)$$

Indeed, for odd j the quotients $(q^j \mp 1)/(q \mp 1)$ are odd; for $j = 2^e j'$ with $e \geq 1$ and j' odd, $v_2(x^{j'}-1) = v_2(x-1)$ and $v_2(x^2-1) = v_2(x-1)+1$ for $x \equiv 1 \pmod{4}$ give $v_2(q^j-1) = v_2(q^2-1) + e - 1$, and $q^j + 1 \equiv 2 \pmod{4}$. In particular $a(j), a^U(j), b(j) \leq a_1 - 1$ for odd j , and $b(j) \leq b_0 \leq a_1 - 1$ always.

Let X be a random variable with $\mathbb{P}(X \geq x) = 2^{-x}$ for $x = 0, 1, 2, \dots$, and write $(y)^+ = \max(y, 0)$.

Lemma 2.3. *If x is uniform in a cyclic group of order N with $v_2(N) = a$, then $v_2(|x|)$ has the law of $(a - X)^+$.*

Proof. The elements whose order has 2-adic valuation t number $\varphi(2^t)N/2^a$, so the probability is 2^{t-1-a} for $1 \leq t \leq a$ and 2^{-a} for $t = 0$; and $\mathbb{P}(a - X = t) = 2^{-(a-t+1)}$, $\mathbb{P}(a - X \leq 0) = 2^{-a}$. $\square$

For a semisimple element t of one of our groups, diagonalisable over $\overline{\mathbb{F}}_q$, put $M(t) = v_2(|t|)$, and let $D(t)$ be the dimension of the sum of the eigenspaces of the eigenvalues ζ with $v_2(|\zeta|) = M(t)$. Since $|t|$ is the least common multiple of the $|\zeta|$, the involution $t^{|t|/2}$ acts as -1 on the eigenspace of ζ if $v_2(|\zeta|) = M(t)$ and as 1 otherwise. Hence

$$t \in Q \iff M(t) \geq 1 \text{ and } n/3 < D(t) \leq 2n/3. \quad (3)$$

2.3 The model

Let $N \geq 1$, let σ be a uniform random permutation of $\{1, \dots, N\}$ with cycles $c_1, \dots, c_r$ of lengths $\lambda_1, \dots, \lambda_r$, let $\varepsilon_1, \varepsilon_2, \dots$ be independent fair signs and $X_1, X_2, \dots$ independent copies of X , all independent. Given levels V_i of the cycles, put $M = \max_i V_i$, $S = \{i : V_i = M\}$, $D = \sum_{i \in S} \lambda_i$, let ν be the number of cycles with $\varepsilon_i = -$, and for a set I of integers let $\mathcal{Q}_I = \{M \geq 1, D \in I\}$. Let

$$I(n) = \{d \in \mathbb{Z} : \frac{n}{3} < d \leq \frac{2n}{3}\}, \quad I_1 = \{d \in \mathbb{Z} : \frac{n}{3} < d \leq \frac{2n}{3}\}, \quad I_2 = \{d \in \mathbb{Z} : \frac{2m+1}{6} < d \leq \frac{2m+1}{3}\}.$$

Proposition 2.4. (a) *If $G = \mathrm{GL}_n(q)$ (resp. $\mathrm{GU}_n(q)$), then $P(G) = \mathbb{P}(\mathcal{Q}_{I(n)})$ for $N = n$ and $V_i = (a(\lambda_i) - X_i)^+$ (resp. $(a^U(\lambda_i) - X_i)^+$).*

(b) *If $G = \mathrm{Sp}_{2m}(q)$ or $\mathrm{SO}_{2m+1}(q)$, then $P(G) = \mathbb{P}(\mathcal{Q}_I)$ for $N = m$, $I = I_1$ or $I = I_2$ respectively, and $V_i = (a(\lambda_i) - X_i)^+$ if $\varepsilon_i = +$, $V_i = (b(\lambda_i) - X_i)^+$ if $\varepsilon_i = -$.*

(c) *If $G = \mathrm{SO}_{2m}^\varepsilon(q)$, then $P(G) = 2\mathbb{E}[\mathbf{1}_{\mathcal{Q}_{I_1}} \mathbf{1}\{\nu \equiv \varepsilon\}]$ in the model of (b), where $\nu \equiv \varepsilon$ means that ν is even for $\varepsilon = +$ and odd for $\varepsilon = -$.*

Proof. We use Lemma 2.1, Proposition 2.2 and the descriptions of the Weyl groups, Frobenius actions and tori in [LNP09, Sections 3 and 5].

(a) $W = S_n$; with the Frobenius morphisms of [LNP09, Section 3.1], which raise the diagonal entries of T_0 to the power q for GL and $-q$ for GU, F acts trivially on W , so in both cases the F -classes are indexed by the cycle types $\lambda \vdash n$, with $|C|/|W| = 1/z_\lambda$, the probability that σ has cycle type λ . The torus is $T^F \cong \prod_i C_{N_i}$ with $N_i = q^{\lambda_i} - 1$ (resp. $q^{\lambda_i} - (-1)^{\lambda_i}$), and the i -th factor acts on the λ_i coordinates of the i -th cycle, with eigenvalues that are images of one another under $x \mapsto x^q$ (resp. $x \mapsto x^{-q}$), hence of the same order [LNP09, Section 5(i)]. A uniform $t \in T^F$ has independent uniform components, so by Lemma 2.3 the levels $v_2(|t_i|)$ have the laws V_i , and $M(t) = M$, $D(t) = D$. By (3) and (1), $P(G) = \sum_\lambda z_\lambda^{-1} \mathbb{P}(\mathcal{Q}_{I(n)} \mid \text{type } \lambda) = \mathbb{P}(\mathcal{Q}_{I(n)})$.

(b) $W = C_2 \wr S_m$ with trivial F ; its classes are the signed cycle types $(\lambda; \mu)$, the lengths of the positive and of the negative cycles, and the centraliser of an element of type $(\lambda; \mu)$ has order $z_\lambda z_\mu 2^{\ell(\lambda) + \ell(\mu)}$ [Car72]. This is also the law of the signed type of (σ, ε) : the probability that σ has type $\lambda \cup \mu$ is $1/z_{\lambda \cup \mu}$, the probability of the given signs is $2^{-\ell(\lambda) - \ell(\mu)} \prod_i \binom{m_i(\lambda) + m_i(\mu)}{m_i(\lambda)}$, where m_i is the multiplicity of i , and the product is $1/(z_\lambda z_\mu 2^{\ell(\lambda) + \ell(\mu)})$. The torus is $\prod_i C_{q^{\lambda_i} - 1} \times \prod_j C_{q^{\mu_j} + 1}$, the factor of a cycle of length l acting on $2l$ coordinates with eigenvalues $x^{\pm q^h}$, all of the order of x , and for SO_{2m+1} one more coordinate with eigenvalue 1 [LNP09, Sections 3 and 5(vi)]. So $D(t) = 2D$, and $n/3 < 2D \leq 2n/3$ means $D \in I_1$ for $n = 2m$ and $D \in I_2$ for $n = 2m + 1$.

(c) $W = D_m$ has index 2 in $C_2 \wr S_m$ and consists of the elements with ν even; F is trivial for $\varepsilon = +$, and for $\varepsilon = -$ it is conjugation by a reflection $r \notin D_m$, and $w \mapsto wr$ maps the F -classes onto the classes of $C_2 \wr S_m$ with ν odd [LNP09, Section 3], [Car72]. For $\varepsilon = +$ a class of $C_2 \wr S_m$ in D_m is a D_m -class or the union of two D_m -classes of equal size; for $\varepsilon = -$ a class of $C_2 \wr S_m$ outside D_m is a single D_m -class, since each of its elements lies outside D_m and centralises itself. The torus attached to w is as in (b) for the signed type of w (resp. wr), and its proportion of elements in Q depends only on that type. As $|W| = |C_2 \wr S_m|/2$, (1) becomes $P(G) = \sum_{\nu \equiv \varepsilon} 2(z_\lambda z_\mu 2^{\ell(\lambda) + \ell(\mu)})^{-1} \mathbb{P}(\mathcal{Q}_{I_1} \mid (\lambda; \mu))$. $\square$

2.4 The groups of similitudes

For $X = \text{GSp}_{2m}(q)$ or $\text{GO}_{2m}^\varepsilon(q)^0$ let $v_* = a_1 + 1$ and $k_0 = 2$, and for $X = \text{GO}_{2m+1}(q)$ let $v_* = a_1$ and $k_0 = 1$. Let N_k be the number of cycles of σ whose length is divisible by 2^k . Write $\mathbb{P}_X(\cdot)$ for the probability in the following experiment: choose (σ, ε) as above, and then a uniform element t of the torus T^F of the signed type of (σ, ε) described in the proof below; for $\text{GO}_{2m}^\varepsilon(q)^0$ replace $\mathbb{P}_X(A)$ by $2\mathbb{P}_X(A \cap \{\nu \equiv \varepsilon\})$.

Lemma 2.5. $P(X) = \mathbb{P}_X(t \in Q)$. For every $v \geq v_*$, $\mathbb{P}_X(t \in Q, M(t) = v)$ equals $\mathbb{P}(\mathcal{Q}_I \cap \{M = v\})$ in the model of Proposition 2.4(b) (with $I = I_1$ for $n = 2m$, I_2 for $n = 2m + 1$, and with the factor 2 and the restriction $\nu \equiv \varepsilon$ for $\text{GO}_{2m}^\varepsilon(q)^0$); and $\mathbb{P}_X(M(t) < v_*) \leq \mathbb{E}[(3/4)^{N_{k_0}}]$, with the factor 2 for $\text{GO}_{2m}^\varepsilon(q)^0$.

Proof. In a basis $e_1, \dots, e_m, f_1, \dots, f_m$ with $\langle e_i, f_j \rangle = \delta_{ij}$, the diagonal torus of GSp_{2m} and of GO_{2m}^0 is $T_0 = \{\text{diag}(t_1, \dots, t_m, ct_1^{-1}, \dots, ct_m^{-1})\}$, where c is the multiplier; the Weyl group $(C_2 \wr S_m, \text{resp. } D_m)$ fixes c , permutes the indices and exchanges t_i and ct_i^{-1} . Solving $wF(t) = t$ cycle by cycle, as in [LNP09, Section 5(i)]: $c \in \mathbb{F}_q^\times$; a positive cycle of length l contributes $x \in \mathbb{F}_{q^l}^\times$ and two blocks of l eigenvalues, x^{q^h} (block A) and cx^{-q^h} (block B), $0 \leq h < l$; a negative cycle of length l contributes $y \in \mathbb{F}_{q^{2l}}^\times$ with $y^{q^l+1} = c$ and one block of $2l$ eigenvalues y^{q^h} , $0 \leq h < 2l$ (note $cy^{-1} = y^{q^l}$). So $T^F \cong \prod_{\text{pos.}} \mathbb{F}_{q^{\lambda_i}}^\times \times \{(c, (y_j)) : c \in \mathbb{F}_q^\times, y_j^{q^{\mu_j}+1} = c\}$, of order $(q-1) \prod_i (q^{\lambda_i} - 1) \prod_j (q^{\mu_j} + 1)$, since the norm $\mathbb{F}_{q^{2l}}^\times \rightarrow \mathbb{F}_{q^l}^\times$ is onto with fibres of size $q^l + 1$. In a uniform element of T^F the x_i are uniform and independent of each other and of $(c, (y_j))$, and c is uniform in $\mathbb{F}_q^\times$. For $\text{GO}_{2m+1} = \mathbb{G}_m \times \text{SO}_{2m+1}$

the torus is $\mathbb{F}_q^\times \times T_{\text{SO}}^F$: with $z \in \mathbb{F}_q^\times$ and the data $x' \in \mathbb{F}_{q^l}^\times$, y' with $y'^{q^l+1} = 1$ of SO_{2m+1} , the blocks are $zx'^{\pm q^h}$ (two blocks), zy'^{q^h} and the single eigenvalue z . The weights $|C|/|W|$ are those of Proposition 2.4(b), (c), which gives $P(X) = \mathbb{P}_X(t \in Q)$.

All eigenvalues of a block have the same order; call v_2 of it the level of the block. Elements of $\mathbb{F}_q^\times$ have level at most a_0 . (i) Negative cycles: $y^{(q^l+1)(q-1)} = c^{q-1} = 1$, so the level is at most $b(l) + a_0 \leq a_1$; for GO_{2m+1} the block zy' has level at most $\max(a_0, b_0) < a_1$, and the block z level at most a_0 . (ii) Positive cycles: write $\pi_2(x)$ for the 2-part of x . If $v_2(|x|) > a_0$, then $\pi_2(cx^{-1}) = \pi_2(c)\pi_2(x)^{-1}$ has the order of $\pi_2(x)$, so blocks A and B have the same level; if $v_2(|x|) < v$ for some $v > a_0$, then both levels are below v . The same holds for the blocks $zx'^{\pm 1}$ of GO_{2m+1} , with x' in place of x .

Let $v \geq v_*$. By (i) and (ii), $M(t) = v$ holds iff the largest of the numbers $v_2(|x_i|)$ (resp. $v_2(|x'_i|)$) over the positive cycles is v , and then the blocks of level v are the two blocks of each positive cycle with $v_2(|x_i|) = v$, so $D(t) = 2 \sum \lambda_i$ over these cycles. Since $v_2(|x_i|)$ has the law $(a(\lambda_i) - \mathbf{X})^+$ by Lemma 2.3, independently over i and of the rest, the event $\{t \in Q, M(t) = v\}$ has the same joint law with (σ, ε) as $\mathcal{Q}_I \cap \{M = v\}$ in model (b), where the negative cycles have levels at most $b_0 < v$. Finally, if $M(t) < v_*$, every positive cycle of length j divisible by 2^{k_0} has $v_2(|x_i|) < a_1 - 1 + k_0 \leq a(j)$, an event of probability $2^{-(a(j)-a_1+2-k_0)} \leq \frac{1}{2}$; as the sign is positive with probability $\frac{1}{2}$, each such cycle contributes a factor at most $\frac{3}{4}$, independently given σ . $\square$

For $m = 1$, $\text{GSp}_2(q) = \text{GL}_2(q)$, and the model above is that of Proposition 2.4(a) for $n = 2$: the positive cycle gives the split torus $\{(x, c/x)\}$ and the negative cycle the non-split torus.

3 Two tools from the cycle index

We use two standard facts on random permutations (see [ABT03]). For weights $w_1, w_2, \dots \geq 0$,

$$\sum_{m \geq 0} x^m \mathbb{E}_{S_m} \left[\prod_{\text{cycles } c} w_{|c|} \right] = \exp \left(\sum_{j \geq 1} \frac{w_j x^j}{j} \right), \quad (4)$$

equivalently $F_0 = 1$ and $mF_m = \sum_{j=1}^m w_j F_{m-j}$ for $F_m = \mathbb{E}_{S_m} [\prod_c w_{|c|}]$. And if $\ell_1, \dots, \ell_t \geq 1$, $m' = m - \sum_i \ell_i \geq 0$ and h is a function of the cycle type of a permutation of the remaining m' points, then

$$\mathbb{E}_{S_m} \left[\sum_{(c_1, \dots, c_t)} \mathbf{1}[|c_i| = \ell_i \ \forall i] h(\sigma|_{\text{rest}}) \right] = \frac{\mathbb{E}_{S_{m'}}[h]}{\ell_1 \dots \ell_t}, \quad (5)$$

the sum running over ordered t -tuples of distinct cycles: the pairs (permutation, tuple) with complement in a set $E \subseteq S_{m'}$ number $\frac{m!}{\ell_1! \dots \ell_t! m'!} \prod_i (\ell_i - 1)! |E|$.

3.1 The witness formula

Consider a model as in Section 2: given σ , the levels V_i are independent, and $\mathbb{P}(V_i < v \mid \sigma) = \alpha_{\lambda_i}(v)$, $\mathbb{P}(V_i = v \mid \sigma) = \beta_{\lambda_i}(v)$ depend only on λ_i . For $v \geq 1$ put

$$A_m(v) = [x^m] \exp \left(\sum_j \frac{\alpha_j(v) x^j}{j} \right) = \mathbb{E}_{S_m} \left[\prod_c \alpha_{|c|}(v) \right], \quad E_d(v) = [x^d] \exp \left(\sum_j \frac{\beta_j(v) x^j}{j} \right).$$

Proposition 3.1. *For every $v \geq 1$ and every set $I \subseteq \{1, \dots, N\}$, $\mathbb{P}(M = v, D \in I) = \sum_{d \in I} E_d(v) A_{N-d}(v)$.*

Proof. $\mathbf{1}\{M = v, D \in I\} = \sum_T \mathbf{1}[V_i = v \ (i \in T), V_i < v \ (i \notin T)] \mathbf{1}[\sum_{i \in T} \lambda_i \in I]$, the sum over the non-empty sets T of cycles, with disjoint events. Given σ the expectation of a term is $\prod_{i \in T} \beta_{\lambda_i}(v) \prod_{i \notin T} \alpha_{\lambda_i}(v)$. Writing each t -set as $t!$ ordered tuples and applying (5) with $h = \prod_c \alpha_{|c|}(v)$ gives $\sum_{t \geq 1} \frac{1}{t!} \sum_{\ell_1 + \dots + \ell_t \in I} \prod_i \frac{\beta_{\ell_i}(v)}{\ell_i} A_{N - \sum \ell_i}(v)$, which is the claim by the exponential formula for $E_d(v)$. $\square$

Lemma 3.2. *Let $k \geq 1$ and $v = a_1 - 1 + k$. Call j active if $2^k \mid j$, and write $j = 2^k \mu$. (a) In the model of Proposition 2.4(a): $\alpha_j(v) = \beta_j(v) = 2^{-(v_2(\mu)+1)}$ for active j , and $\alpha_j(v) = 1, \beta_j(v) = 0$ otherwise. (b) In the model of Proposition 2.4(b): $\alpha_j(v) = \frac{1}{2}(1 + 2^{-(v_2(\mu)+1)}) \leq \frac{3}{4}$ and $\beta_j(v) = 2^{-(v_2(\mu)+2)}$ for active j , and $\alpha_j(v) = 1, \beta_j(v) = 0$ otherwise. (c) $\mathbb{P}(M < a_1) \leq \mathbb{E}_{S_N}[2^{-N_1}]$ in model (a) and $\mathbb{P}(M < a_1) \leq \mathbb{E}_{S_N}[(3/4)^{N_1}]$ in model (b).*

Proof. For a level $(a - X)^+$ and $v \geq 1$: $V < v$ iff $X \geq a - v + 1$, and $V = v$ iff $X = a - v$; so $\alpha = \beta = 2^{-(a-v+1)}$ if $a \geq v$, while $\alpha = 1, \beta = 0$ if $a < v$. At $v = a_1 - 1 + k$ the odd lengths and the negative cycles have $a < v$ by (2), and for even j , $a(j) - v + 1 = a^U(j) - v + 1 = v_2(j) - k + 1$, which is $v_2(\mu) + 1$ if $2^k \mid j$ and at most 0 otherwise. In (b) the two signs are averaged. For (c), at level a_1 every even length j has $a(j) \geq a_1$, so $\mathbb{P}(V_i < a_1) \leq \frac{1}{2}$ for an even cycle (with positive sign in (b)), independently given σ . $\square$

3.2 Two closed forms

Let N_k be the number of cycles of length divisible by 2^k and $\log^+ = \max(\log, 0)$.

Lemma 3.3. (a) For $u \in [0, 1]$, $k \geq 1$, $N = \lfloor m/2^k \rfloor$ and $s = (1 - u)/2^k$:

$$\mathbb{E}_{S_m}[u^{N_k}] = \prod_{i=1}^N \left(1 - \frac{s}{i}\right) \leq (N + 1)^{-s}. \quad (6)$$

(b) In the model of Proposition 2.4(a), for $v = a_1 - 1 + k$ with $k \geq 1$,

$$A_m(v) \geq \exp\left(-\frac{2 \log 2 (1 + \log^+(m/2^k))}{2^k}\right). \quad (7)$$

Proof. (a) By (4) with $w_j = u$ for $2^k \mid j$ and $w_j = 1$ otherwise, and $\sum_{2^k \mid j} x^j/j = -2^{-k} \log(1 - y)$ with $y = x^{2^k}$, the left side is $[x^m](1 - x)^{-1}(1 - y)^s$. As $(1 - x)^{-1} = (1 + x + \dots + x^{2^k-1})(1 - y)^{-1}$, only one term of the first factor contributes, and the coefficient is $[y^N](1 - y)^{s-1} = \prod_{i \leq N} (i - s)/i$. Finally $1 - s/i \leq e^{-s/i}$ and $\sum_{i \leq N} 1/i \geq \log(N + 1)$.

(b) By Lemma 3.2(a), $\prod_c \alpha_{|c|}(v) = 2^{-\sum_{i \geq 0} N_{k+i}}$, because $v_2(\mu) + 1 = \#\{i \geq 0 : 2^i \mid \mu\}$. By Jensen's inequality, $A_m(v) \geq 2^{-\sum_i \mathbb{E} N_{k+i}}$, and $\mathbb{E} N_{k+i} = \sum_{j \leq m, 2^{k+i} \mid j} 1/j = 2^{-(k+i)} H_{\lfloor m/2^{k+i} \rfloor} \leq 2^{-(k+i)} (1 + \log^+(m/2^k))$, where H_N is the harmonic number ($H_0 = 0, H_N \leq 1 + \log N$). Sum over i . $\square$

4 Upper and lower bounds

4.1 A level-by-level upper bound

Lemma 4.1. *Let σ be uniform in S_N , and let $\mathcal{Q}$ be an event and M an integer-valued random variable on a probability space carrying σ . Let $c \in (0, 1]$, $k_0 \in \{1, 2\}$, $w_0 \in \mathbb{Z}$, $0 < \theta_0 < \theta_1 < 1$ and $0 < \theta_2 \leq \theta_1$, and let I be a set of integers such that $\theta_0 N < d \leq \theta_1 N$ and $N - d \geq \theta_2 N$ for all $d \in I$. Assume*

(U1) $\mathbb{P}(M < w_0 + k_0) \leq \mathbb{E}_{S_N}[(1 - c/2)^{N_{k_0}}]$;

(U2) for every $k \geq k_0$ and $v = w_0 + k$: $\mathbb{P}(\mathcal{Q}, M = v) \leq \sum_{d \in I} E_d(v) A_{N-d}(v)$ for weights $\alpha_j(v), \beta_j(v) \in [0, 1]$ with $\alpha_j = 1, \beta_j = 0$ if $2^k \nmid j$, and $\alpha_j \leq 1 - c/2, \beta_j \leq c 2^{-(v_2(\mu)+1)}$ if $j = 2^k \mu$.

Put $r_1 = \theta_1/\theta_0, \kappa = (2 + \log(\theta_1/\theta_2))/3, b = 1 + c\kappa/2$,

$$g(z) = (1 + \log(r_1(b + cz/3)))e^{-cz/6}, \quad H(y) = y g(y/2), \quad \Sigma_H = \sum_{j \geq 0} \sup_{y \in (2^j, 2^{j+1}]} H(y).$$

If $g'(0) < 0$ and $\theta_2 N > 1$, then

$$\mathbb{P}(\mathcal{Q}) \leq (\lfloor N/2^{k_0} \rfloor + 1)^{-c/2^{k_0+1}} + \frac{K}{\log(\theta_2 N)}, \quad K = \frac{c}{2} e^{c\kappa/2} (\Sigma_H + 2g(0)). \quad (8)$$

Proof. $\mathbb{P}(\mathcal{Q}) \leq \mathbb{P}(M < w_0 + k_0) + \sum_{k \geq k_0} \mathbb{P}(\mathcal{Q}, M = w_0 + k)$.

(a) By (U1) and (6) with $u = 1 - c/2$, the first term is at most $(\lfloor N/2^{k_0} \rfloor + 1)^{-c/2^{k_0+1}}$.

(b) Fix $k \geq k_0, v = w_0 + k, d \in I$. As $\prod_c \alpha_{|c|}(v) \leq (1 - c/2)^{N_k}$, (6) with $s = c/2^{k+1}$ and $N - d \geq \theta_2 N$ give $A_{N-d}(v) \leq \max(1, \theta_2 N/2^k)^{-c/2^{k+1}}$, using $\lfloor x \rfloor + 1 \geq \max(1, x)$.

(c) If $\ell_1 + \dots + \ell_t = d \in I$, then all $\ell_i \leq \theta_1 N$ and the largest exceeds $\theta_0 N/t$. Summing over the position of a largest entry,

$$\sum_{d \in I} E_d(v) \leq \sum_{t \geq 1} \frac{t \rho(t) \gamma^{t-1}}{t!} = \sum_{u \geq 0} \frac{\rho(u+1) \gamma^u}{u!}, \quad \rho(t) = \sum_{\theta_0 N/t < \ell \leq \theta_1 N} \frac{\beta_\ell}{\ell}, \quad \gamma = \sum_{\ell \leq \theta_1 N} \frac{\beta_\ell}{\ell}.$$

Write $\ell = 2^k \mu$. With $\beta_\ell \leq c/2$ and $L' = \theta_0 N/(t2^k)$, $\rho(t) \leq \frac{c}{2^{k+1}} \sum_{\mu \in (L', r_1 t L']} \frac{1}{\mu} \leq \frac{c}{2^{k+1}} (1 + \log(r_1 t))$: if $L' \geq 1$, the first term of the sum is at most 1 and the others are at most $\int_{L'}^{r_1 t L'} dx/x \leq \log(r_1 t)$; if $L' < 1$, the sum is empty or at most $1 + \log(r_1 t L')$. Next, grouping $\mu = 2^i o$ with o odd and using $\sum_{o \text{ odd} \leq Y} 1/o \leq 1 + \frac{1}{2} \log Y$ for $Y \geq 1$ (since $1/o \leq \frac{1}{2} \int_{o-2}^o dx/x$ for $o \geq 3$),

$$\gamma \leq \frac{c}{2^k} \sum_{\mu \leq N_0} \frac{2^{-(v_2(\mu)+1)}}{\mu} \leq \frac{c}{2^k} \sum_{i \geq 0} 2^{-(2i+1)} \left(1 + \frac{\log N_0}{2}\right) \leq \bar{\gamma}_k := \frac{c(2 + \log^+(\theta_1 N/2^k))}{3 \cdot 2^k},$$

with $N_0 = \lfloor \theta_1 N/2^k \rfloor$ ($\gamma = 0$ if $N_0 = 0$). Hence, with Y a Poisson variable of mean γ and Jensen's inequality for log,

$$\sum_{d \in I} E_d(v) \leq \frac{c}{2^{k+1}} e^\gamma \mathbb{E}[1 + \log(r_1(Y+1))] \leq \frac{c}{2^{k+1}} e^{\bar{\gamma}_k} (1 + \log(r_1(1 + \bar{\gamma}_k))).$$

(d) Let $L = \log(\theta_2 N) > 0$ and $z_k = \max(0, L - k \log 2)/2^k$. Then $\max(1, \theta_2 N/2^k)^{-c/2^{k+1}} = e^{-cz_k/2}$, and since $\log^+(\theta_1 N/2^k) \leq \log(\theta_1/\theta_2) + 2^k z_k$, we have $\bar{\gamma}_k \leq c\kappa/2^k + cz_k/3 \leq c\kappa/2 + cz_k/3$. By (U2), (b) and (c),

$$\mathbb{P}(\mathcal{Q}, M = w_0 + k) \leq T_k := \frac{c}{2} e^{c\kappa/2} 2^{-k} g(z_k) \quad (k \geq k_0).$$

The function $g'(z)e^{cz/6} = \frac{c/3}{b+cz/3} - \frac{c}{6}(1 + \log(r_1(b + cz/3)))$ is decreasing, so $g'(0) < 0$ implies that g is decreasing on $[0, \infty)$. For the k with $2^k \geq L$ we use $g(z_k) \leq g(0)$ and $\sum_{2^k \geq L} 2^{-k} \leq 2/L$. For the $k \geq 1$ with $2^k < L$ we have $L - k \log 2 > L - \log L \geq L/2$, so $z_k \geq y_k/2$ with $y_k = L/2^k > 1$, and $2^{-k} g(z_k) \leq 2^{-k} g(y_k/2) = H(y_k)/L$; these y_k form a geometric progression of ratio 2 whose smallest term lies in $(1, 2]$, so each interval $(2^j, 2^{j+1}]$ contains at most one of them. Hence $\sum_{k \geq k_0} T_k \leq K/L$. $\square$

The function $\log H$ is concave, since $\frac{d}{dy} \log H(y) = \frac{1}{y} + \frac{c/6}{(b+cy/6)(1+\log(r_1(b+cy/6)))} - \frac{c}{12}$ is decreasing; so each supremum in Σ_H is attained at an endpoint or at the mode of H , and Σ_H is a finite computation (the terms with $j \geq 60$ are below 10^{-300}). The values used below are (Section 7):

$c, (\theta_0, \theta_1, \theta_2)$	used for	Σ_H	$g(0), g'(0)$	K
$1, (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$	$\mathrm{GL}_n, \mathrm{GU}_n$	57.63	2.064, -0.114	48.37
$\frac{1}{2}, (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$	$\mathrm{Sp}_{2m}, \mathrm{SO}_{2m}^\pm, \mathrm{GSp}_{2m}, \mathrm{GO}_{2m}^\pm(q)^0$	116.04	1.896, -0.022	37.50
$\frac{1}{2}, (\frac{1}{3}, \frac{3}{4}, \frac{1}{4})$	$\mathrm{SO}_{2m+1}, \mathrm{GO}_{2m+1}$	121.41	2.041, -0.038	40.62

4.2 Proof of Theorem 1.1

Upper bound. By Proposition 2.4(a), Proposition 3.1 and Lemma 3.2(a), (c), the hypotheses of Lemma 4.1 hold with $N = n$, $c = 1$, $k_0 = 1$, $w_0 = a_1 - 1$, $I = I(n)$ and $(\theta_0, \theta_1, \theta_2) = (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$. So for $n \geq 4$

$$P(G) \leq (\lfloor n/2 \rfloor + 1)^{-1/4} + \frac{48.38}{\log(n/3)}.$$

For $n \geq 9$ we have $\log(n/3) \geq \frac{1}{2} \log n$, and $(\lfloor n/2 \rfloor + 1)^{-1/4} \log n \leq (n/2)^{-1/4} \log n \leq 4 \cdot 2^{1/4}/e < 1.75$, so $P(G) < (1.75 + 96.76)/\log n < 100/\log n$. For $2 \leq n \leq 8$, $100/\log n > 1 \geq P(G)$.

Lower bound. Let $n \geq 100$, $k = \lceil \log_2 \log n \rceil$ (so $\log n \leq 2^k < 2 \log n$) and $v = a_1 - 1 + k$. By Proposition 3.1, keeping only the terms $t = 1$ of $E_d(v)$, $P(G) \geq \sum_{d \in I(n)} \frac{\beta_d(v)}{d} A_{n-d}(v)$. We keep the $d = 2^k \mu$ with μ odd, for which $\beta_d(v) = \frac{1}{2}$ by Lemma 3.2(a); then $d \in I(n)$ iff $\mu \in (R, 2R]$ with $R = n/(3 \cdot 2^k)$. By (7) with $n - d \leq n$ and $2^k \geq \log n$,

$$A_{n-d}(v) \geq \exp\left(-2 \log 2 \cdot \frac{1 + \log n - \log \log n}{\log n}\right) \geq 2^{-2(1+1/\log n)} \geq 2^{-2.44} > \frac{1}{8}.$$

The interval $(R, 2R]$ contains at least $R/2 - 1$ odd integers, each with $1/\mu \geq 1/(2R)$, and $R \geq n/(6 \log n) \geq 3.6$, so $\sum_{\mu} 1/\mu \geq \frac{1}{4} - \frac{1}{2R} \geq 0.111$. Hence $P(G) \geq \frac{1}{2} \cdot \frac{1}{8} \cdot 2^{-k} \cdot 0.111 \geq \frac{0.111}{32 \log n} > \frac{1}{300 \log n}$. The same holds for $\mathrm{GU}_n(q)$. $\square$

These witnesses are those of [LNP09, Section 5], one cycle of length $2^k \mu$ in the middle third; the better constant comes from (7), which allows other active cycles of lower level.

4.3 Proof of Theorem 1.2

For $\mathrm{Sp}_{2m}(q)$ and $\mathrm{SO}_{2m+1}(q)$, Proposition 2.4(b), Proposition 3.1 and Lemma 3.2(b), (c) give the hypotheses of Lemma 4.1 with $N = m$, $c = \frac{1}{2}$, $k_0 = 1$, $w_0 = a_1 - 1$ and $(\theta_0, \theta_1, \theta_2) = (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$ for I_1 ; for I_2 and $m \geq 4$, every $d \in I_2$ has $d > (2m+1)/6 > m/3$, $d \leq (2m+1)/3 \leq 3m/4$ and $m-d \geq (m-1)/3 \geq m/4$, so $(\theta_0, \theta_1, \theta_2) = (\frac{1}{3}, \frac{3}{4}, \frac{1}{4})$. For $\mathrm{SO}_{2m}^\varepsilon(q)$, Proposition 2.4(c) gives $P \leq 2\mathbb{P}(\mathcal{Q}_{I_1})$ in the model of $\mathrm{Sp}_{2m}(q)$. For $\mathrm{GSp}_{2m}(q)$ and $\mathrm{GO}_{2m+1}(q)$, Lemma 2.5 gives (U1) with $w_0 = a_1 - 1$, $w_0 + k_0 = v_*$, and (U2) at the levels $v \geq v_*$ with the weights of Lemma 3.2(b), where $k_0 = 2$ for GSp_{2m} and $k_0 = 1$ for GO_{2m+1} . For $\mathrm{GO}_{2m}^\varepsilon(q)^0$ the tori of a given signed type are those of $\mathrm{GSp}_{2m}(q)$, so $P(X) \leq 2\mathbb{P}_Y(t \in Q)$, where $\mathbb{P}_Y$ is the experiment of $Y = \mathrm{GSp}_{2m}(q)$.

Let $f = 2$ for $\mathrm{SO}_{2m}^\pm$ and $\mathrm{GO}_{2m}^\pm(q)^0$ and $f = 1$ otherwise, and $\rho = 6$ for $n = 2m$, $\rho = 9$ for $n = 2m+1$. Then $\theta_2 N = n/6$ for $n = 2m$, $\theta_2 N = (n-1)/8 \geq n/9$ for $n = 2m+1 \geq 9$, and $\lfloor N/2^{k_0} \rfloor + 1 > (n-1)/2^{k_0+1} \geq n/(3 \cdot 2^{k_0})$ for $n \geq 3$. So (8) gives, for $n > \rho$ (and $n \geq 9$ if n is odd),

$$P(X) \log n \leq \mathcal{M}(n) := f \left[\left(\frac{n}{3 \cdot 2^{k_0}} \right)^{-c/2^{k_0+1}} + \frac{K}{\log(n/\rho)} \right] \log n.$$

Both terms of $\mathcal{M}$ decrease in n once $\log n > 2^{k_0+1}/c$ (at most 16) and $n > \rho$, because $\frac{d}{dL}(e^{-sL}L) = e^{-sL}(1-sL)$ and $L/(L-\log \rho)$ decreases. Let $C = 45$ ($f = 1$) or $C = 80$ ($f = 2$). If $\log n < C$, then $P(X) \leq 1 < C/\log n$; if $\log n \geq C$, then $P(X) \log n \leq \mathcal{M}(e^C)$, and $\mathcal{M}(e^{45}) = 39.25$ (Sp), 42.21 (GSp), 42.90 ($\mathrm{SO}_{2m+1}, \mathrm{GO}_{2m+1}$), $\mathcal{M}(e^{80}) = 76.72$ ($\mathrm{SO}^\pm$), 77.97 ($\mathrm{GO}^\pm(q)^0$), all below C . $\square$

The same computation with $c = 1$ gives the constant 50 in place of 100 in Theorem 1.1.

5 Groups between S and X

Proof of Corollary 1.3. If $H \leq Y$ act on the same natural module, then $|g|$ and $\text{Fix}_V(g^{|g|/2})$ do not depend on whether g is viewed in H or in Y , so $Q(H) = Q(Y) \cap H$ and $P(H) \leq [Y : H]P(Y)$. Take $Y = X$. The indices are $[\text{GL}_n(q) : \text{SL}_n(q)] = q - 1$, $[\text{GU}_n(q) : \text{SU}_n(q)] = q + 1$, $[\text{GSp}_{2m}(q) : \text{Sp}_{2m}(q)] = q - 1$ (the similitude $\text{diag}(1, \dots, 1, w, \dots, w)$ has multiplier w), $[\text{GO}_{2m+1}(q) : \text{SO}_{2m+1}(q)] = q - 1$ (Section 2), and $[\text{GO}_{2m}^\pm(q)^0 : \text{SO}_{2m}^\pm(q)] = q - 1$: the kernel of the multiplier on GO_{2m}^0 is the connected group SO_{2m} , so by Lang's theorem the multiplier maps $\text{GO}_{2m}^\pm(q)^0$ onto $\mathbb{F}_q^\times$, with kernel $\text{SO}_{2m}^\pm(q)$. So $[X : H] \leq q + 1$, and $C_X \leq 100$ gives the upper bound. The lower bound is [LNP09, Theorem 1.1], which is stated for exactly these H . (In the orthogonal case the proof of [LNP09, Lemma 4.4], in the version we read, works with a group containing the full orthogonal group, larger than $\text{GO}_{2m}^\pm(q)^0$; the bound then holds a fortiori for the groups between S and $\text{GO}_{2m}^\pm(q)^0$.) Finally $[\text{SO} : \Omega] = 2$ in odd characteristic, and $Y = \text{SO}$ gives the last statement. $\square$

No lower bound is claimed for $H = \Omega$, which lies below S in [LNP09, Table 1], and the full orthogonal groups $\text{O}_{2m}^\pm(q)$ are not covered: they are the groups of points of the disconnected group O_{2m} , and Proposition 2.2 is not available for them.

6 The asymptotic value

6.1 The function Φ

Proof of Theorem 1.4(a). Let $f(t) = \frac{\log 2}{3} 2^t e^{-2^t/3}$. The series $\sum_k f(u - k)$ converges uniformly on compact sets with all derivatives, so Φ is smooth and 1-periodic, and $c_j = \int_0^1 \Phi(u) e^{-2\pi i j u} du = \int_{\mathbb{R}} f(t) e^{-2\pi i j t} dt$. With $y = 2^t$ this is $\frac{1}{3} \int_0^\infty e^{-y/3} y^{-i\tau_j} dy = 3^{-i\tau_j} \Gamma(1 - i\tau_j)$; in particular $c_0 = 1$. Since Γ has no zeros, $c_j \neq 0$, and $|c_j|^2 = \pi \tau_j / \sinh(\pi \tau_j)$. As Φ is smooth, its Fourier series converges absolutely, so $|\Phi(u) - 1| \leq 2 \sum_{j \geq 1} |c_j|$. Since $\pi \tau_j = 2\pi^2 j / \log 2 > 28j$, $|c_j|^2 = 2\pi \tau_j e^{-\pi \tau_j} / (1 - e^{-2\pi \tau_j}) < 2.0001 \pi \tau_j e^{-\pi \tau_j}$; with $|c_1| = 4.9422 \cdot 10^{-6}$ and $|c_2| = 4.58 \cdot 10^{-12}$ from the formula, this bound gives $\sum_{j \geq 2} |c_j| < 5 \cdot 10^{-12}$, so $2 \sum_{j \geq 1} |c_j| = 9.8845 \cdot 10^{-6}$. $\square$

6.2 Proof of Theorem 1.4(b) for $\text{GL}_n(q)$ and $\text{GU}_n(q)$

Let $L = \log n$, $\omega = \log L$, $y_k = L/2^k$ and $v_k = a_1 - 1 + k$ for $k \in \mathbb{Z}$, and $C_v = \mathbb{P}(M = v, D \in I(n)) = \sum_{d \in I(n)} E_d(v) A_{n-d}(v)$, so that $P(G) = \sum_{v < a_1} C_v + \sum_{k \geq 1} C_{v_k}$. Let $\mathcal{K} = \{k \geq 1 : y_k \leq \omega, 2^k \leq \sqrt{n}\}$. Below, $O(\cdot)$, $o(\cdot)$ and “for n large” are uniform in q : by Lemma 3.2(a) the weights at the levels v_k , $k \geq 1$, do not depend on q .

Step 1: the levels outside $\mathcal{K}$. We use the proof of Lemma 4.1 with the parameters of Theorem 1.1, where $c\kappa/2 = c_0 := (2 + \log 2)/6$. First, $\sum_{v < a_1} C_v \leq ([n/2] + 1)^{-1/4}$. Second, $C_{v_k} \leq T_k \leq \frac{e^{c_0}}{2} 2^{-k} g(0)$ for every $k \geq 1$, so $\sum_{2^k > \sqrt{n}} C_{v_k} \leq e^{c_0} g(0) n^{-1/2}$. Third, if $y_k > \omega$ then $2^k < L/\omega < L' := \log(n/3)$ for n large, so k belongs to the second group of part (d) of that proof (with L' in place of L), and $T_k \leq \frac{e^{c_0}}{2L'} H(y'_k)$ with $y'_k = L'/2^k > \omega/2$; as the y'_k lie in distinct dyadic intervals, $\sum_{y_k > \omega} C_{v_k} \leq \frac{e^{c_0}}{2L'} \sum_{j: 2^{j+1} > \omega/2} \sup_{(2^j, 2^{j+1}]} H$, a tail of the convergent series Σ_H divided by $2L'$. Hence $LP(G) = L \sum_{k \in \mathcal{K}} C_{v_k} + o(1)$.

Two facts. (P) If $b_j \geq 0$ and $F(x) = \sum_{j \leq J} b_j x^j / j$, then $e^{F(x) - F(1)}$ is the generating function of $W = \sum_j j Z_j$ with independent $Z_j \sim \text{Poisson}(b_j/j)$, so by Markov's inequality $\sum_{r > r_0} [x^r] e^F = e^{F(1)} \mathbb{P}(W > r_0) \leq e^{F(1)} F'(1)/r_0$; and $|[x^r] e^{-F}| \leq [x^r] e^F$ for all r , by expanding both exponentials.

(H) For integers $M \geq 1$ and reals $Y \geq 4$,

$$\left| \sum_{\mu \leq M} \frac{1 - 2^{-v_2(\mu)}}{\mu} - \frac{\log M}{3} \right| \leq 2, \quad \left| \sum_{\mu \in (Y, 2Y]} \frac{2^{-(v_2(\mu)+1)}}{\mu} - \frac{\log 2}{3} \right| \leq \frac{5}{Y}. \quad (9)$$

For the proof let $O(Z) = \sum_{o \text{ odd} \leq Z} 1/o$ and $O'(Z) = \sum_{o \text{ odd} \in (Z, 2Z]} 1/o$. Comparing $1/o$ with $\frac{1}{2} \int_o^{o+2} dt/t$ and $\frac{1}{2} \int_{o-2}^o dt/t$ gives $\frac{1}{2} \log Z \leq O(Z) \leq 1 + \frac{1}{2} \log Z$ for $Z \geq 1$; and $0 \leq \frac{1}{2} \log \frac{o+1}{o-1} - \frac{1}{o} \leq \frac{1}{2o^3}$ for $o \geq 3$, with the end pieces of $\bigcup_o [o-1, o+1]$ against $[Z, 2Z]$, gives $|O'(Z) - \frac{1}{2} \log 2| \leq 3/Z$ for $Z \geq 2$, while $0 \leq O'(Z) \leq 2$ always. Group $\mu = 2^e o$ with o odd. The first sum is $\sum_e (2^{-e} - 4^{-e}) O(M/2^e)$, and $|O(M/2^e) - \frac{1}{2} \log M| \leq 1 + \frac{e}{2} \log 2$ for every e (if $2^e > M$ the sum is empty); as $\sum_e (2^{-e} - 4^{-e}) = \frac{2}{3}$ and $\sum_e e(2^{-e} - 4^{-e}) = \frac{14}{9}$, the error is at most $\frac{2}{3} + \frac{7}{9} \log 2 < 2$. The second sum is $\sum_e 2^{-(2e+1)} O'(Y/2^e)$, with $\sum_e 2^{-(2e+1)} = \frac{2}{3}$; the terms with $Y/2^e \geq 2$ contribute an error at most $\sum_e 2^{-(2e+1)} 3 \cdot 2^e / Y \leq 3/Y$, and those with $2^e > Y/2$ at most $2 \cdot \frac{2}{3} (2/Y)^2 \leq \frac{4}{3} Y^{-1}$.

Step 2: the levels in $\mathcal{K}$. We show that there are $\varepsilon_n \rightarrow 0$, independent of q and k , with

$$\left| C_{v_k} - \frac{\log 2}{3} 2^{-k} e^{-y_k/3} \right| \leq 2^{-k} \varepsilon_n \quad (k \in \mathcal{K}). \quad (10)$$

Fix $k \in \mathcal{K}$, $v = v_k$, $y = y_k$, with the weights of Lemma 3.2(a). Then $2^{-k} = y/L \leq \omega/L$, $2^k \leq \sqrt{n}$, $R := n/(3 \cdot 2^k) \geq \sqrt{n}/3$ and $e^{y/3} \leq L^{1/3}$. Let $h = \lfloor n/L^3 \rfloor$ and $R_0 = hL \leq n/L^2$; for n large, $1 + \log(n/h) \leq 4 \log L$.

Decomposition. Let $B(x) = \sum_{j \leq n} \beta_j x^j / j$, the series cut off at degree n (this does not change $E_d(v)$ for $d \leq n$), and write $B = B_s + B_b$, where B_s collects the terms with $j \leq h$. Then $e^B = e^{B_s} (1 + B_b + \mathcal{R})$ with $\mathcal{R} = e^{B_b} - 1 - B_b$, which has non-negative coefficients. By the computation of γ in the proof of Lemma 4.1 and the counting used there for ρ ,

$$\begin{aligned} B_s(1) &\leq B(1) \leq \frac{2+L}{3 \cdot 2^k} \leq \frac{y}{3} + 1, & B'_s(1) &\leq \frac{h}{2^{k+1}}, \\ B_b(1) &\leq \frac{1 + \log(n/h)}{2^{k+1}} \leq \frac{3 \log L}{2^k}, & \mathcal{R}(1) &\leq \frac{B_b(1)^2}{2} e^{B_b(1)}. \end{aligned}$$

Accordingly $C_v = X_1 + X_2 + X_3$, where X_1, X_2, X_3 are $\sum_{d \in I(n)} A_{n-d}(v)$ times the coefficient of x^d in $e^{B_s}, B_b e^{B_s}, \mathcal{R} e^{B_s}$. As $0 \leq A_m \leq 1$, (P) gives

$$X_1 \leq \sum_{d > n/3} [x^d] e^{B_s} \leq e^{B_s(1)} \frac{B'_s(1)}{n/3} \leq \frac{3}{2} 2^{-k} e^{y/3+1} L^{-3}, \quad X_3 \leq \mathcal{R}(1) e^{B_s(1)} \leq \frac{9}{2} 2^{-k} \frac{y \log^2 L}{L} e^{y/3+2},$$

using $2^{-k} = y/L$ and $B_b(1) \leq 1$ for n large. With $e_r^s = [x^r] e^{B_s}$,

$$X_2 = \sum_{\ell > h} \frac{\beta_\ell}{\ell} \sum_{r \geq 0: \ell+r \in I(n)} e_r^s A_{n-\ell-r}(v).$$

Put $G_{N'} = \sum_{r=0}^{N'} e_r^s A_{N'-r}(v)$ and $\tau = \sum_{r > R_0} e_r^s \leq e^{B_s(1)} B'_s(1)/R_0 \leq 2^{-k} e^{y/3+1}/(2L)$. Split the sum over $\ell > h$. (a) For $\ell \in I^\circ = (n/3, 2n/3 - R_0]$ every $r \leq R_0$ has $\ell + r \in I(n)$, so the inner sum lies in $[G_{n-\ell} - \tau, G_{n-\ell}]$. (b) For $h < \ell \leq n/3 - R_0$ only $r > R_0$ occur, so the inner sum is at most τ , and $\sum \beta_\ell / \ell \leq B(1) \leq y/3 + 1$. (c) For ℓ in $\mathcal{W} = (n/3 - R_0, n/3] \cup (2n/3 - R_0, 2n/3]$ the inner sum is at most $e^{B_s(1)}$, and the active $\ell \in \mathcal{W}$ number at most $2(R_0/2^k + 1)$, each with $\beta_\ell / \ell \leq \frac{1}{2} \cdot \frac{4}{n}$, so their sum is at most $5 \cdot 2^{-k} L^{-2}$ for n large (as $R_0 \leq n/L^2$ and $2^k \leq \sqrt{n}$). (d) For $\ell > 2n/3$ there is no term. With $S^\circ = \sum_{\ell \in I^\circ} \beta_\ell / \ell$,

$$\left| X_2 - \sum_{\ell \in I^\circ} \frac{\beta_\ell}{\ell} G_{n-\ell} \right| \leq (S^\circ + y/3 + 1) \tau + 5 \cdot 2^{-k} L^{-2} e^{y/3+1}.$$

The active $\ell \in I(n)$ are the $2^k \mu$ with $\mu \in (R, 2R]$, so by (9) and (c), $|S^\circ - \frac{\log 2}{3} 2^{-k}| \leq 2^{-k}(5/R + 5L^{-2})$.

The factor $G_{N'}$ for $n/3 \leq N' < 2n/3$. By (4), $G_{N'} = \mathbb{E}_{S_{N'}}[\prod_c \tilde{w}_{|c|}]$ with $\tilde{w}_j = \alpha_j + \beta_j \mathbf{1}[j \leq h]$, that is, $\tilde{w}_j = 1$ for inactive j , $\tilde{w}_j = 2^{-v_2(j/2^k)}$ for active $j \leq h$, and $\tilde{w}_j = \alpha_j \in [0, 1]$ for active $j > h$. Let $\hat{G}_{N'}$ be the same expectation with $\tilde{w}_j$ replaced by 1 for $j > h$. Then $\hat{G}_{N'} - \pi \leq G_{N'} \leq \hat{G}_{N'}$, where π , the probability that some active cycle is longer than h , is at most $\sum_{h < j \leq N', 2^k | j} 1/j \leq 4 \cdot 2^{-k} \log L$. By (4), $\hat{G}_{N'} = [x^{N'}](1-x)^{-1} e^{-\mathcal{Q}(x)} = \sum_{i \leq N'} [x^i] e^{-\mathcal{Q}}$ with $\mathcal{Q}(x) = \sum_{j \leq h \text{ active}} (1 - 2^{-v_2(j/2^k)}) x^j / j$, so by (P), with $\Lambda = \mathcal{Q}(1)$,

$$|\hat{G}_{N'} - e^{-\Lambda}| \leq \sum_{i > N'} [x^i] e^{\mathcal{Q}} \leq e^{\Lambda} \frac{\mathcal{Q}'(1)}{N'} \leq e^{\Lambda} \frac{3h}{2^k n} \leq 3 \cdot 2^{-k} e^{\Lambda} L^{-3}.$$

By (9), $\Lambda = 2^{-k} \sum_{\mu \leq h/2^k} (1 - 2^{-v_2(\mu)}) / \mu = 2^{-k} (\frac{1}{3} \log(h/2^k) + \theta)$ with $|\theta| \leq 3$, and $\log(h/2^k) = L - 4 \log L + \log y + \theta'$ with $\theta' \in [-\log 2, 0]$. Hence $|\Lambda - y/3| \leq \frac{y}{L} (\frac{4}{3} \log L + \frac{1}{3} |\log y| + 4) \leq 3 \log^2 L / L$ for n large, because $y \leq \omega$ and $y |\log y| \leq \max(1/e, \omega \log \omega)$. In particular $\Lambda \leq y/3 + 1$ and $|e^{-\Lambda} - e^{-y/3}| \leq |\Lambda - y/3|$.

Conclusion of Step 2. Write $\sum_{\ell \in I^\circ} \frac{\beta_\ell}{\ell} G_{n-\ell} = S^\circ e^{-y/3} + \sum_{\ell \in I^\circ} \frac{\beta_\ell}{\ell} (G_{n-\ell} - e^{-y/3})$. Then $|C_v - \frac{\log 2}{3} 2^{-k} e^{-y/3}|$ is at most 2^{-k} times the sum of the following quantities, each of which tends to 0 uniformly in $k \in \mathcal{K}$ and in q (recall $y \leq \omega = \log L$, $e^{y/3} \leq L^{1/3}$, $2^{-k} \leq \omega/L$, $R \geq \sqrt{n}/3$, $S^\circ \leq 2^{-k}$ for n large): $X_1/2^{-k} \leq \frac{3}{2} e L^{1/3-3}$; $X_3/2^{-k} \leq \frac{9}{2} e^2 L^{1/3-1} \log^3 L$; the terms with τ , at most $(\omega/3+2)e L^{1/3-1}$; the windows, $5e L^{1/3-2}$; $5/R + 5L^{-2}$; and $\sup_{N'} |G_{N'} - e^{-y/3}| \leq \pi + 3 \cdot 2^{-k} e^{y/3+1} L^{-3} + 3 \log^2 L / L$. This proves (10) with $\varepsilon_n = O(L^{-2/3} \log^3 L)$.

Step 3: summation. Since $\sum_{k \in \mathcal{K}} 2^{-k} \leq 2 \cdot 2^{-\min \mathcal{K}} \leq 2\omega/L$, (10) gives $L \sum_{k \in \mathcal{K}} C_{v_k} = \frac{\log 2}{3} \sum_{k \in \mathcal{K}} y_k e^{-y_k/3} + O(\omega \varepsilon_n)$. Now $\Phi(\log_2 L) = \frac{\log 2}{3} \sum_{k \in \mathbb{Z}} y_k e^{-y_k/3}$. The terms with $k \notin \mathcal{K}$ are those with $y_k > \omega$ (which include all $k \leq 0$), forming a geometric progression of ratio 2 with smallest term in $(\omega, 2\omega]$, on which $y e^{-y/3}$ decreases, so that their sum is at most $\sum_{i \geq 0} 2^{i+1} \omega e^{-2^i \omega/3} \rightarrow 0$; and those with $2^k > \sqrt{n}$, whose sum is at most $2L/\sqrt{n}$. With Step 1, $LP(G) = \Phi(\log_2 L) + o(1)$. For $\text{GU}_n(q)$ the proof is the same, because Lemma 3.2(a), (c) hold for a^U . $\square$

6.3 The other families

(a) $\text{Sp}_{2m}(q)$ and $\text{SO}_{2m+1}(q)$. Here $N = m$; put $L_m = \log m$, $\omega = \log L_m$, $y_k = L_m/2^k$, $\mathcal{K} = \{k \geq 1 : y_k \leq \omega, 2^k \leq \sqrt{m}\}$, and use the model of Proposition 2.4(b) with the weights of Lemma 3.2(b). Step 1 is the same, with the parameters of Theorem 1.2 (the low levels contribute at most $(\lfloor m/2 \rfloor + 1)^{-1/8}$). In Step 2 every bound used still holds ($\beta_j \leq \frac{1}{2}$, $\tilde{w}_j \in [0, 1]$, (P)), and two quantities are halved: for active $j = 2^k \mu$, $\beta_j = 2^{-(v_2(\mu)+2)}$ and $1 - \tilde{w}_j = \frac{1}{2}(1 - 2^{-v_2(\mu)})$, so $S^\circ = \frac{\log 2}{6} 2^{-k} + 2^{-k} O(1/R + L_m^{-2})$ and $\Lambda = y/6 + O(\log^2 L_m/L_m)$. For I_2 the proof is run with I_2 in place of I_1 . Since $I_2 \setminus I_1$ consists of at most the integer $\ell^* = (2m+1)/3$, the sum $\sum_{\ell \in I} \beta_\ell / \ell$ changes by at most $\beta_{\ell^*} / \ell^* \leq \frac{1}{4} \cdot \frac{3}{2m} = \frac{3}{8m} \leq 2^{-k} \frac{3}{8} m^{-1/2}$ (as $2^k \leq \sqrt{m}$), which is $2^{-k} o(1)$; and $N' = m - \ell \geq (m-1)/3$ in the bound for $\hat{G}_{N'}$. Hence $L_m P = \frac{\log 2}{6} \sum_k y_k e^{-y_k/6} + o(1) = \frac{\log 2}{3} \sum_k y_{k+1} e^{-y_{k+1}/3} + o(1) = \Phi(\log_2 L_m) + o(1)$. As $L_m / \log n \rightarrow 1$, $\log_2 \log n - \log_2 L_m \rightarrow 0$ and Φ is bounded and uniformly continuous, $(\log n)P = \Phi(\log_2 \log n) + o(1)$.

(b) $\text{SO}_{2m}^\varepsilon(q)$. By Proposition 2.4(c), $P = 2\mathbb{E}[\mathbf{1}_{\mathcal{Q}} \mathbf{1}\{\nu \equiv \varepsilon\}]$ in the model of (a) with I_1 . Let $\mathcal{Q}_{\text{low}} = \mathcal{Q} \cap \{M < a_1\}$ and $\mathcal{Q}_k = \mathcal{Q} \cap \{M = v_k\}$. First, $2\mathbb{P}(\mathcal{Q}_{\text{low}}) \leq 2(\lfloor m/2 \rfloor + 1)^{-1/8} = o(1/L_m)$. For $k \geq 1$ the cycles of odd length have levels at most $\max(a_0, b_0) < a_1$ by (2), so $\mathcal{Q}_k$ is determined by σ and by the signs and levels of the cycles of even length. Given these, if σ has a cycle of odd length, the parity of ν is a fair coin. Hence $|2\mathbb{E}[\mathbf{1}_{\mathcal{Q}_k} \mathbf{1}\{\nu \equiv \varepsilon\}] - \mathbb{P}(\mathcal{Q}_k)| \leq \mathbb{P}(\mathcal{Q}_k \cap \{\text{no odd cycle}\})$, and the sum over k of these errors is at most $\mathbb{P}_{S_m}(\text{no odd cycle}) = [x^m](1-x^2)^{-1/2} = \binom{m}{m/2} 2^{-m} \leq m^{-1/2}$ for even m (and 0 for odd m). So $P = \mathbb{P}(\mathcal{Q}) + o(1/L_m)$, and (a) applies.

family, q	$N = 10^2$	10^3	10^4	10^5	10^6	10^7
GL, $q = 3$ (= GU, $q = 5$)	0.8877	1.0152	1.0624	1.0870	1.0986	1.1033
GL, $q = 5$ (= GU, $q = 3$)	0.8988	1.0326	1.0779	1.0984	1.1061	1.1079
Sp, $q = 3$ and 5	0.9456	1.0482	1.0856	1.1015	1.1071	1.1077
SO $_{2m+1}$, $q = 3$ and 5	0.9601	1.0494	1.0857	1.1015	1.1071	—
SO $_{2m}^+$, $q = 3$	0.9421	1.0478	1.0855	1.1015	1.1071	1.1077
SO $_{2m}^-$, $q = 3$	0.9491	1.0487	1.0856	1.1015	1.1071	—

Table 1: Values of $(\log n)P(X)$, with $N = n$ for GL, GU and $N = m$ otherwise, from the witness formula in floating point. The coincidences $\text{GL}(q) \leftrightarrow \text{GU}(q')$ for $v_2(q-1) = v_2(q'+1)$ and $v_2(q+1) = v_2(q'-1)$ follow from (2), and those between $q = 3$ and $q = 5$ for Sp and SO $_{2m+1}$ from the symmetry of the signs (exchanging a_0 and b_0 exchanges the levels of the positive and the negative cycles of odd length). These values are evidence only.

(c) *The groups of similitudes.* For $\text{GSp}_{2m}(q)$, Lemma 2.5 shows that the levels $v \geq a_1 + 1$ contribute exactly the C_{v_k} , $k \geq 2$, of the model of (a), and the lower levels at most $\mathbb{E}[(3/4)^{N_2}] \leq (\lfloor m/4 \rfloor + 1)^{-1/16} = o(1/L_m)$; the level $v_1 = a_1$ of the model of (a) has $y_1 = L_m/2 > \omega$ and contributes $o(1/L_m)$ by Step 1. So $P(\text{GSp}_{2m}(q)) = P(\text{Sp}_{2m}(q)) + o(1/L_m)$. For $\text{GO}_{2m}^\varepsilon(q)^0$ the same comparison holds with the factor 2 and the parity restriction on both sides, and these are removed as in (b). For $\text{GO}_{2m+1}(q)$ the levels $v \geq a_1$ coincide with those of $\text{SO}_{2m+1}(q)$, and the lower ones contribute at most $(\lfloor m/2 \rfloor + 1)^{-1/8}$. All bounds are uniform in q . $\square$

7 Computations and reproducibility

No proof in this note depends on a computation, except the finite evaluations of the explicit constants Σ_H , $g(0)$, $g'(0)$, K and $\mathcal{M}(e^C)$ in Section 4, and of $\sum_j |c_j|$, which is used only for the numerical bound in Theorem 1.4(a). The other computations test the torus model against the groups and produce Table 1. The reproducibility package has four parts.

The exact program of this note (`check_note.py`, standard library, about 20 seconds) recomputes the constants of Section 4 (the dyadic suprema of H also by dense sampling), the majorants $\mathcal{M}(e^C)$, the constants of the lower bound, the range and Fourier coefficients of Φ , and (9) on a finite range. In exact arithmetic it checks (2), Lemma 2.3, (6), the sum of Proposition 2.4 over (signed) cycle types against the witness formula, level by level (130 cases), and the level identities of Lemma 2.5 in a model built from the 2-parts of all eigenvalues (40 cases). In floating point it checks $C_v \leq T_k$ for every level and the lower bound of Theorem 1.1 for $\text{GL}_n(q)$, $n \in \{100, 1000\}$. It computes $P(G)$ by brute force for 29 groups, among them $\text{Sp}_4(3)$, $\text{SO}_5(3)$, $\text{GSp}_4(3)$, $\text{GO}_5(3)$ and $\text{GO}_4^\pm(5)^0$, with the order of the generated group compared with the order formula, and compares the model with all 35 brute-force values recorded in the package; all agree.

The programs with which the results were first obtained (`original/`): exact evaluations of (1), brute force over 16 groups, the witness formula in floating point (programs in C) for Table 1, and the constants. *The programs of the second verification run* (`verification_run_B/`): its brute force, which includes $\text{GO}_5(3)$ ($|\text{GO}_5(3)| = 103680$, $P(\text{GO}_5(3)) = 21/64$), and its enumeration of semisimple classes described below, with their recorded outputs. *The programs of the third verification run* (`independent_run_2/`), described below, with their recorded outputs.

For $n = 1000$, Table 1 gives $P(\text{GL}_{1000}(3)) = 0.1470$, in line with the sampled proportions of about 15% in dimension 1000 reported in [LNP09]. The values near 1.10 at $N = 10^6$ – 10^7 , about 10% above the mean 1 of Φ , reflect the slow convergence in Theorem 1.4; the level-by-level ratios

$C_{v_k}/(\frac{\log 2}{3}2^{-k}e^{-y_k/3})$ for $\text{GL}_{10^6}(3)$ are 1.16, 1.18, 1.13, 1.08, 1.05, 1.03, 1.02, 1.01 for $k = 1, \dots, 8$. We do not claim any rate of convergence.

Verification. The results were first obtained with written proofs and the programs in `original/`. Two independent verification runs, both AI-assisted, examined this first written version; each re-derived the proofs line by line and wrote its own programs, and neither read the other's report. Neither found an error in a statement.

The first run compared every statement quoted from [LNP09] with the authors' preprint and [NPP14, Theorem 2.2] with its arXiv text, recomputed all constants of Section 4 and the range of Φ , and computed $P(X)$ by its own brute force over 15 groups (orders checked against the order formula), by its own exact torus model, including an independently derived model for the groups of similitudes, and by its own witness recursion, which reproduces Table 1 for $N \leq 10^4$. The second run wrote a brute force over 35 groups with the orders checked, which includes $\text{GO}_5(3)$, the one group whose brute force had not been finished in the first computations; an exact method that does not use (1), the enumeration of semisimple classes with Steinberg's count of the unipotent elements of their centralisers [Car85], which agrees with its own implementation of the torus model at all 326 values compared ($\text{GL}_n(q)$ and $\text{GU}_n(q)$ for $n \leq 16$ to 22, the symplectic and orthogonal groups for $m \leq 10$, $q \leq 17$); the values of Table 1 by class enumeration for $N \leq 10^4$, and for GL at $N = 10^5$; and a Monte Carlo simulation of the model at $N = 10^5, 10^6, 10^7$, which agrees with Table 1 within 2.2 standard errors (about ± 0.0015). It re-ran the programs of `original/` for the constants, the exact sums and Table 1, and reproduced their recorded outputs.

The brute-force values used in this note come from runs in which the order of the generated group was compared with the order formula: automatically in all three verification runs, in `check_note.py` and in the second brute-force run of `original/`, and through the printed orders in its first run, of which one line (a proper subgroup of order 192 of $\text{SO}_4^+(3)$) was discarded. The corrections required by the first two runs were made: the bound in Section 6.3(a) for I_2 ; the cut-off of the series B in Step 2; the reading of [DPS18] (see "Scope and priority"); the remark on [LNP09, Lemma 4.4] in Section 5; the bibliographic data; the record of the order checks; and the constant $\frac{9}{2}$ for X_3 .

A third verification run, also AI-assisted, examined the written note. It checked first the parts written or changed when the note was written: Lemma 4.1 and the derivation of Theorems 1.1 and 1.2 from it, the bound for I_2 in Section 6.3(a), the cut-off of B and the bound on π in Step 2, the inequalities (9), the bound by $2 \sum_j |c_j|$, Lemma 2.5 with the hypotheses (U1), (U2) for each family, Corollary 1.3 with the remark on [LNP09, Lemma 4.4], and Steps 1–3 of Section 6; then it re-derived all other proofs line by line. It compared the quoted statements with [EHHO11], with the preprint of [LNP09], with the arXiv text of [NPP14] and with the publisher's version of [DPS18], re-checked the references in Crossref and repeated the literature search. Its own programs computed the constants of Section 4 with 50 digits (the suprema of H from its mode and by dense sampling), Φ and its Fourier coefficients, and (9) for $M \leq 2 \cdot 10^6$ and $4 \leq Y \leq 2 \cdot 10^5$; a brute force over 26 groups with the orders checked, among them $\text{GU}_3(3)$, $\text{GO}_3(5)$, $\text{GO}_5(3)$ and $\text{GO}_4^\pm(5)^0$; an exact torus model in which the tori of the groups of similitudes are built from the 2-parts of their elements, without the level identities of Lemma 2.5, which agrees with all 26 brute-force values and with the 63 exact values of `original/` that it also computed; and the witness formula in floating point, with the weights computed from the valuations, which agrees with the exact model for small N , reproduces Table 1 to all printed digits for $N \leq 10^6$ and the entry for GL, $q = 3$, at $N = 10^7$, confirms $C_v \leq T_k$ at every level in the cases it computed, and gives the level ratios quoted above. It re-ran `check_note.py`, the constant programs and the grid of `fastcl.c` from the archive and reproduced their recorded outputs. It found no error in a theorem, proposition, lemma or proof. The eight corrections it required were made: the scope statements for $\text{SL}_n(q)$ and $\text{SU}_n(q)$, for which only the order of magnitude is proved; the statement

on [DPS18]; the condition in the caption of Table 1; the last of the level ratios above (printed as 1.00 before); an explicit bound for the tail of $\sum_j |c_j|$; the description of the Frobenius action for $\mathrm{GU}_n(q)$; the reference to [GP12, Theorem 1.2]; and this record.

Scope and priority. The note answers the second part of the question of [EHHO11, pp. 2154–2155], the order $\Theta(1/\log n)$, for every group H with $S \leq H \leq X$ in a line of [LNP09, Table 1] (Theorems 1.1 and 1.2, Corollary 1.3), and the first part, the asymptotic value, for the groups of Theorems 1.1 and 1.2: all groups X , and the groups S of the symplectic and orthogonal lines (Theorem 1.4). Its limits: the full orthogonal groups $\mathrm{O}_{2m}^\pm(q)$ are not covered; for $\mathrm{SL}_n(q)$, $\mathrm{SU}_n(q)$ and for $S < H < X$ only the order of magnitude is proved, with an upper constant carrying the index, and Theorem 1.4 is not claimed for these groups; for Ω only the upper bound is proved; the constants are not optimised and no rate of convergence is given; even q is outside the question. The question is due to Praeger, jointly with Lübeck and Niemeyer [EHHO11]. The lower bound, the list of groups and the description of the tori and of their involutions are from [LNP09]; the torus formula is the theorem of Niemeyer and Praeger on quokka sets [NP10, Theorem 1.3], used in the form of [LNP09, Lemma 2.3] and [NPP14, Theorem 2.2]; the facts (4) and (5) are standard [ABT03]; the classes of the Weyl groups [Car72] and the correspondence between classes of tori and F -classes [Car85] are used as in [LNP09]. The range of 2-parts that carries the main term of Theorem 1.4 is the one of [GP12, Theorem 1.2].

What was read: the problem session of [EHHO11], in a text extraction of the report; [LNP09] completely, in the authors’ preprint of 20 May 2008 (the publisher’s file refused automated access, and the journal version was not compared); [NPP14], [GP12], [GPRD20] and [NP14] in their arXiv versions; [DPS18] completely, in the publisher’s version. Seen only through their metadata, abstracts or restatements: [NP10], through the restatements in [LNP09, NPP14]; [NPP10]; [PS12] and the survey [NPS13], through their abstracts; [Car72, Car85], as cited in [LNP09]. The main result of [DPS18], [DPS18, Theorem 1.1], is a lower bound for a different event: for a strong involution t in $\mathrm{GL}_n(q)$ and a uniform $g \in \mathrm{GL}_n(q)$, the restriction of the involution $z(g)$ of $\langle tt^g \rangle$ to an eigenspace of t is a strong involution with probability at least $\kappa/\log n$ for $n \geq n_0$, where one can take $n_0 = 700$ and $\kappa = 0.0002$ [DPS18, Remark 1.2]. The proportion $|Q|/|G|$ occurs there only through the lower bound of [LNP09, Theorem 1.1] ([DPS18, Section 1.1]); [DPS18] contains no upper bound and no asymptotic value for it.

The searches were made on 18 and 19 October 2026: arXiv (strong involutions; proportions of elements powering to involutions in classical groups; the authors of the question; quokka sets and maximal tori), the lists of citing works in OpenAlex (24 works citing [LNP09], 18 citing [NP10], 10 citing [NPP10]), Semantic Scholar (in part refused), zbMATH, Crossref, and web searches (one in the first work, three by the second verification run and one by the third); the third run repeated the queries to arXiv, OpenAlex (also the 2 works citing [DPS18]), zbMATH and Crossref. No printed answer to the question was found in these sources: [LNP09] states that its method gives no upper bound, and the citing works, by title and available abstract, concern recognition algorithms, other proportions, or lower bounds for other events, such as [GP12, DPS18, GPRD20]. A search that finds nothing is not a proof of novelty, and no priority is claimed.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

References

- [ABT03] Richard Arratia, A. D. Barbour, and Simon Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach*. EMS Monographs in Mathematics. European Mathematical Society, Zürich, 2003. doi:10.4171/000.
- [Car72] Roger W. Carter. Conjugacy classes in the Weyl group. *Compositio Mathematica*, 25(1):1–59, 1972.
- [Car85] Roger W. Carter. *Finite Groups of Lie Type: Conjugacy Classes and Complex Characters*. John Wiley & Sons, Chichester, 1985. Reprinted 1993 in the Wiley Classics Library.
- [DPS18] John D. Dixon, Cheryl E. Praeger, and Ákos Seress. Strong involutions in finite special linear groups of odd characteristic. *Journal of Algebra*, 498:413–447, 2018. doi:10.1016/j.jalgebra.2017.11.047.
- [EHHO11] Bettina Eick, Gerhard Hiß, Derek F. Holt, and Eamonn A. O’Brien. Computational group theory. *Oberwolfach Reports*, 8(3):2113–2161, 2011. Report No. 37/2011; the problem session is on pp. 2154–2155. doi:10.4171/OWR/2011/37.
- [GP12] Simon Guest and Cheryl E. Praeger. Proportions of elements with given 2-part order in finite classical groups of odd characteristic. *Journal of Algebra*, 372:637–660, 2012. arXiv:1007.2983, doi:10.1016/j.jalgebra.2012.09.028.
- [GPRD20] S. P. Glasby, Cheryl E. Praeger, and Colva M. Roney-Dougal. Involution centralisers in finite unitary groups of odd characteristic. *Journal of Algebra*, 545:245–299, 2020. arXiv:1812.01362, doi:10.1016/j.jalgebra.2019.09.009.
- [LNP09] Frank Lübeck, Alice C. Niemeyer, and Cheryl E. Praeger. Finding involutions in finite Lie type groups of odd characteristic. *Journal of Algebra*, 321(11):3397–3417, 2009. doi:10.1016/j.jalgebra.2008.05.009.
- [NP10] Alice C. Niemeyer and Cheryl E. Praeger. Estimating proportions of elements in finite groups of Lie type. *Journal of Algebra*, 324(1):122–145, 2010. doi:10.1016/j.jalgebra.2010.03.018.
- [NP14] Alice C. Niemeyer and Cheryl E. Praeger. Elements in finite classical groups whose powers have large 1-eigenspaces. *Discrete Mathematics & Theoretical Computer Science*, 16(1), 2014. arXiv:1405.2385, doi:10.46298/dmtcs.3908.
- [NPP10] Alice C. Niemeyer, Tomasz Popiel, and Cheryl E. Praeger. On proportions of pre-involutions in finite classical groups. *Journal of Algebra*, 324(5):1016–1043, 2010. doi:10.1016/j.jalgebra.2010.03.026.
- [NPP14] Alice C. Niemeyer, Tomasz Popiel, and Cheryl E. Praeger. Abundant p -singular elements in finite classical groups. *Journal of Algebra*, 408:189–204, 2014. arXiv:1205.1454, doi:10.1016/j.jalgebra.2013.09.021.
- [NPS13] Alice C. Niemeyer, Cheryl E. Praeger, and Ákos Seress. Estimation problems and randomised group algorithms. In *Probabilistic Group Theory, Combinatorics, and Computing*, volume 2070 of *Lecture Notes in Mathematics*, pages 35–82. Springer, London, 2013. doi:10.1007/978-1-4471-4814-2_2.

- [PS12] Cheryl E. Praeger and Ákos Seress. Regular semisimple elements and involutions in finite general linear groups of odd characteristic. *Proceedings of the American Mathematical Society*, 140(9):3003–3015, 2012. doi:10.1090/S0002-9939-2012-11491-1.