SIAM J. Discrete Math.
Vol. 19, No. 3, pp. 691–699

THE ASYMPTOTIC NUMBER OF BINARY CODES AND BINARY MATROIDS*

MARCEL WILD†

Abstract. The asymptotic number of nonequivalent binary n-codes is determined. This is also the asymptotic number of nonisomorphic binary matroids on n elements.

Key words. asymptotic enumeration, binary codes, binary matroids, lattice of invariant subspaces

AMS subject classifications. Primary, 94A10, 05A16; Secondary, 05B35, 05A30

DOI. 10.1137/S0895480104445538

1. Introduction.

Recall that a *binary n-code* is a subspace X of the $GF(2)$-vector space $V := GF(2)^n$. Two binary n-codes $X, X' \subseteq V$ are *equivalent* if for some permutation σ of the symmetric group S_n on $\{1, 2, \ldots, n\}$ we have

$$X' = X_\sigma := \{(x_{1\sigma}, \ldots, x_{n\sigma}) \mid (x_1, \ldots, x_n) \in X\},$$

where $i\sigma$ is the image of i under σ. Let $b(n)$ be the number of equivalence classes of binary n-codes. It is well known that $b(n)$ is also the number of nonisomorphic binary matroids on an n-set. The asymptotic behavior of $b(n)$ was posed as open problem 14.5.4 in [O].

Here the setting of binary codes suits us better. For a field K let $G(n, K)$ be the (possibly infinite) number of K-linear subspaces of K^n. Mostly, K will be $GF(q)$, in which case we write $G(n, q)$ instead of $G(n, K)$. Because each equivalence class of binary n-codes has cardinality at most $n!$ it follows that $b(n) \geq G(n, 2)/n!$ for all n. It will be a corollary of our main theorem that for $n \to \infty$ asymptotically

$$b(n) \sim G(n, 2)/n!. \tag{1}$$

For $\sigma \in S_n$ let $T_\sigma : V \to V$ be the vector space automorphism defined on the canonical base by $T_\sigma(e_i) := e_{i\sigma}$. Let $\mathcal{L}(T_\sigma)$ be the lattice of all T_σ-*invariant* subspaces in the sense of linear algebra, meaning the lattice of all subspaces U with $T_\sigma(U) \subseteq U$. Since here T_σ is bijective, $T_\sigma(U) \subseteq U$ is equivalent to $T_\sigma(U) = U$, i.e., to U being a "fixed point." This allows us to apply the Cauchy–Frobenius lemma (erroneously called Burnside's lemma):

$$b(n) \quad = \quad \frac{G(n,2)}{n!} + \frac{1}{n!} \sum_{\sigma \in S_n - \{id\}} |\mathcal{L}(T_\sigma)|, \tag{2}$$

hence proving (1) is equivalent to showing

$$\sum_{\sigma \in S_n - \{id\}} |\mathcal{L}(T_\sigma)| \quad = \quad o(G(n,2)). \tag{3}$$

*Received by the editors August 6, 2004; accepted for publication (in revised form) April 4, 2005; published electronically November 23, 2005.
http://www.siam.org/journals/sidma/19-3/44553.html
†Department of Mathematics, University of Stellenbosch, 7602 Matieland, South Africa (mwild@sun.ac.za).

There are $\binom{n}{2}$ permutations $\tau \in S_n$ with one 2-cycle and $n-2$ cycles of length 1. Any such transposition τ yields a T_τ with at least $G(n-1,2)$ invariant subspaces. Indeed, say T_τ switches e_1 and e_2. Then the $n-1$ vectors $e_1+e_2,\ e_3,\dots,e_n$ are fixed by T_τ. Hence

$$\text{(4)} \qquad \binom{n}{2} G(n-1,2) \quad \text{is a } \textit{lower} \text{ bound for} \sum_{\sigma \in S_n - \{id\}} |\mathcal{L}(T_\sigma)|.$$

This shows that (3) can only be true if $G(n,2)$ grows superexponentially with n. Proving (3) was undertaken in [W1] but, as pointed out by Lax [L], there is an error in the proof of [W1, Lemma 6]. The error is fixed in the present article, which also improves upon style and organization. In fact, we shall wind up with a stronger result but as a golden thread it may be helpful, at least in section 2, to think of (3) as our target. The stronger result consists of rather sharp lower and upper bounds for $b(n)$ when n is large enough. These bounds are derived in sections 3 and 4, respectively, and the pieces are put together in section 5.

2. Four lemmata.

The first lemma reduces our preliminary target (3) to the statement that the left-hand side of (3) is $o(2^{n^2/4})$.

LEMMA 1. *For all prime powers q there are positive constants $d_1(q)$ and $d_2(q)$ such that*

$$\lim_{m\to\infty} \frac{G(2m+1,q)}{q^{(2m+1)^2/4}} = d_1(q) \quad \text{and} \quad \lim_{m\to\infty} \frac{G(2m,q)}{q^{(2m)^2/4}} = d_2(q).$$

Furthermore, all $d_i(q)$ are less than 32 *and rounded to six decimals, $d_1(2)$ is* 7.371949, *and $d_2(2)$ is* 7.371969.

Proof. Let q be fixed and put $G_n := G(n,q)$. Note that $G_0 = 1,\ G_1 = 2$. By [A, p. 94] one has

$$\text{(5)} \qquad G_{n+1} \;=\; 2G_n + (q^n - 1)G_{n-1} \quad (n \geq 1).$$

Letting $u_n := q^{-n^2/4} G_n$ $(n \geq 0)$, it follows from (5) that

$$\text{(6)} \qquad u_n \;=\; 2q^{-n/2+1/4} u_{n-1} + (1 - q^{-n+1}) u_{n-2} \quad (n \geq 2).$$

Letting $\tau_n = \tau_n(q) := 2q^{-n/2+1/4} + 1 - q^{-n+1},\ a_n := 2q^{-n/2+1/4}\tau_n^{-1}$, and $b_n := (1-q^{-n+1})\tau_n^{-1}$, we have $a_n + b_n = 1$ and

$$\text{(7)} \qquad u_n \;=\; \tau_n(a_n u_{n-1} + b_n u_{n-2}) \quad (n \geq 2).$$

From $u_0 = 1,\ u_1 = 2q^{-1/4},\ \tau_n > 1$ $(n \geq 2)$, and (7) it follows that

$$\text{(8)} \qquad u_n \;\geq\; \min\{u_0, u_1\} \;>\; 0 \quad (n \geq 0).$$

As to an upper bound, from $u_0 = 1$ and $u_1 \leq 2 \cdot 2^{-1/4} < 1.7$ we get $a_2 u_1 + b_2 u_0 \leq 1.7$, so (7) yields $u_2 \leq (1.7)\tau_2,\ u_3 \leq (1.7)\tau_2\tau_3$, and so forth. One checks that $\tau_n(q) \leq \tau_n(2)$ for $n \geq 2$ and $\tau_n(2) \leq 1 + 2^{-n/3}$ for $n \geq 7$, whence

$$\text{(9)} \quad u_n \;\leq\; (1.7)\tau_2(2)\cdots\tau_6(2) \cdot \prod_{k\geq 7}(1+2^{-k/3}) \;<\; (1.7)\cdot(6.8)\cdot e^{0.97} \;<\; 32 \quad (n \geq 0).$$

The convergence of the latter infinite product follows by taking natural logarithms and noticing that $\sum_{k\geq 7} ln(1+2^{-k/3})$ is bounded by $\sum_{k\geq 7} 2^{-k/3} < 0.97$. From (6) and (9) it follows that

$$|u_{n+2} - u_n| \;=\; |-q^{-n-1}u_n + 2q^{-\frac{n}{2}-\frac{3}{4}}u_{n+1}| \;\leq\; 32q^{-\frac{n}{3}} \quad (n \geq 0).$$

Iterating and applying the triangle inequality yields

$$(10)\quad |u_{n+2k} - u_n| \;\leq\; 32(q^{-\frac{n}{3}} + q^{-\frac{(n+2)}{3}} + \cdots + q^{-\frac{(n+2k-2)}{3}}) \;\leq\; 32\frac{q^{-\frac{n}{3}}}{1-q^{-\frac{2}{3}}} \quad (n \geq 0).$$

Cauchy's criterion therefore guarantees that both $d_1(q) := \lim_{m\to\infty} u_{2m+1}$ and $d_2(q) := \lim_{m\to\infty} u_{2m}$ exist. They are nonzero by (8). Clearly, (10) implies

$$(11)\qquad |d_2(q) - u_{2m}| \;\leq\; 32\;\frac{q^{-\frac{2m}{3}}}{1-q^{-\frac{2}{3}}} \quad (m \geq 0).$$

Combining (7) and (11), one can compute $d_2(q)$ to any desired accuracy. Ditto for $d_1(q)$. ▯

In order to get a handle on $\mathcal{L}(T_\sigma)$ we need the minimal polynomial

$$\min(T_\sigma, t) = \prod_{i=1}^{s} p_i(t)^{\mu_i},$$

where the $p_i(t) \in GF(2)[t]$ are irreducible and $\mu_i \geq 1$ $(1 \leq i \leq s)$. We seek an upper bound for $s = s(\sigma)$. Since $\min(T_\sigma, t)$ has degree at most n and since there are only finitely many irreducible polynomials in $GF(2)[t]$ of any given degree, it is clear that for any fixed $\epsilon > 0$ one can force $s \leq \epsilon n$ for all $\sigma \in S_n$, provided n is large enough. For us it will suffice that

(12) for all large enough n one has $s \leq (0.06)n$ for all $\sigma \in S_n$.

It is well known that if

$$V_i := \ker(p_i(T_\sigma)^{\mu_i}), \quad n_i := \dim(V_i) \quad (1 \leq i \leq s),$$

then $V = V_1 \oplus \cdots \oplus V_s$; and if $T_i := (T_\sigma \restriction V_i)$, then $T_i : V_i \to V_i$ has minimal polynomial $\min(T_i, t) = p_i(t)^{\mu_i}$. Furthermore, by [BF, p. 812]

$$(13)\qquad \mathcal{L}(T_\sigma) \quad \simeq \quad \mathcal{L}(T_1) \times \mathcal{L}(T_2) \quad \times \cdots \times \quad \mathcal{L}(T_s).$$

Assume that our σ is a product of r disjoint cycles $C_1, \ldots, C_r$ of lengths $\lambda_j = 2^{\alpha_j} \cdot u_j$, where $\alpha_j \geq 0$ and $u_j \geq 1$ is odd. The upcoming (14) and (15) will be the only facts for which we refer to [W1]. Namely, if we standardize $p_1(t) := t + 1$, then its corresponding parameters μ_1 and n_1 satisfy [W1, Lemma 4]

$$(14)\qquad \mu_1 = \max\{2^{\alpha_j} |\; 1 \leq j \leq r\}$$

and [W1, Lemma 5]

$$(15)\qquad r \quad \leq \quad n_1 = 2^{\alpha_1} + 2^{\alpha_2} + \cdots + 2^{\alpha_r} \quad \leq \quad n.$$

For instance, $\sigma := (1,2,\ldots,11,12)(13,14,15)(16,17)$ has $n_1 = 2^2+2^0+2^1 = 7$ and a base of V_1 is

$$e_1+e_5+e_9,\ e_2+e_6+e_{10},\ e_3+e_7+e_{11},\ e_4+e_8+e_{12},\ e_{13}+e_{14}+e_{15},\ e_{16},\ e_{17}.$$

Observe that while $\min(T_\sigma,t)$ is just the least common multiple of the polynomials $t^{\lambda_j}+1$ $(1 \le j \le r)$, the prime factors of $\min(T_\sigma,t)$ are unpredictable, and hence there is no general connection between the number r of disjoint cycles of σ and the number s of direct factors of $\mathcal{L}(T_\sigma)$. It is well known that the expected value of $r(\sigma)$ asymptotically is $ln(n)$ as $n \to \infty$. Question: What is the expected value of $s(\sigma)$ as $n \to \infty$?

LEMMA 2. *For large enough n all $\sigma \in S_n$ have $|\mathcal{L}(T_\sigma)| \le |\mathcal{L}(T_1)| \cdot 2^{\frac{(n-n_1)^2}{8}+(0.3)n}$.*

Proof. Since T_i is bijective we have $T_i^{\mu_i} \neq 0$, so $p_i(t)=t$ is impossible, so each $p_i(t)$ $(2 \le i \le s)$ has degree $d_i \ge 2$. Fix $T_i : V_i \to V_i$ with $2 \le i \le s$. According to [BF, Thm. 6] one can write $T_i = Q+S$, where $S : V_i \to V_i$ is semisimple and $Q : V_i \to V_i$ is nilpotent. Moreover, putting $K := GF(2)[t]/p_i(t)$, the map Q is K-linear in a natural sense and $\mathcal{L}(T_i) = \mathcal{L}_K(Q)$. Since $K \simeq GF(2^{d_i})$ and $\dim_K(V_i) = n_i/d_i$, it follows from Lemma 1 that

$$|\mathcal{L}(T_i)| \quad \le \quad G\left(\frac{n_i}{d_i}, 2^{d_i}\right) \quad \le \quad 2^5 \cdot (2^{d_i})^{(n_i/d_i)^2/4} \quad = \quad 2^5 \cdot 2^{n_i^2/4d_i}.$$

Using (12) and $d_i \ge 2$ $(2 \le i \le s)$ we get

$$\begin{aligned}|\mathcal{L}(T_\sigma)| &\le |\mathcal{L}(T_1)|(2^5 \cdot 2^{n_2^2/8}) \cdots (2^5 \cdot 2^{n_s^2/8})\\ &\le |\mathcal{L}(T_1)| \cdot (2^5)^{(0.06)n} \cdot 2^{n_2^2/8+\cdots+n_s^2/8}\\ &\le |\mathcal{L}(T_1)| \cdot 2^{(0.3)n+(n_2+\cdots+n_s)^2/8}. \qquad \square\end{aligned}$$

The trick to decompose T_i as $S+Q$ with Q nilpotent and $\mathcal{L}(T_i) = \mathcal{L}_K(Q)$ also works for $T_i = T_1$. In fact one verifies at once that $T_1 = I + (T_1+I)$ with $(T_1+I)^{\mu_1} = 0$ and $\mathcal{L}(T_1) = \mathcal{L}(T_1+I)$. However, $d_i \ge 2$ is essential in the proof of Lemma 2; for $d_1 = 1$ one only gets the triviality (in view of Lemma 1) $|\mathcal{L}(T_1)| = O(2^{n_1^2/4})$. On the other hand, information about $\ker(Q)$ is only available when $i=1$, and that is what makes the next lemma tick.

LEMMA 3. *Let $\sigma \in S_n$ have r disjoint cycles. With T_1, n_1, μ_1 derived from T_σ as above, one has*

(a) $\quad |\mathcal{L}(T_1)| \quad \le \quad G(r,2) \cdot G(n_1-r,2)$,

(b) $\quad |\mathcal{L}(T_1)| \quad \le \quad G(r,2)^{\mu_1}$.

Proof. Let W be any K-vector space with $\dim(W) = \overline{n}$ and $Q : W \to W$ a linear nilpotent map, say $Q^{m-1} \neq Q^m = 0$. Let $Q_2 := Q \restriction \text{ im } (Q)$. Note that $Q_2 \neq Q^2$ but $\text{im } (Q_2) = \text{ im } (Q^2)$. It is easy to see [BF, Thm. 7] that

$$\mathcal{L}(Q) = \bigcup_{X \in \mathcal{L}(Q_2)} [X, Q^{-1}(X)], \tag{16}$$

where $Q^{-1}(X) := \{w \in W |\ Q(w) \in X\}$ and $[X, Q^{-1}(X)] := \{Y \in \mathcal{L}(W) |\ X \subseteq Y \subseteq Q^{-1}(X)\}$ is an interval of the lattice $\mathcal{L}(W)$ of all subspaces of W. Its length is

$$\dim(Q^{-1}(X)) - \dim(X) \quad = \quad \dim(\ker Q) \quad =: \quad \kappa_1. \tag{17}$$

Since $Q_2 : \ \mathrm{im}(Q) \rightarrow \ \mathrm{im}(Q)$ and $\dim(\mathrm{im}Q) = \overline{n} - \kappa_1$ it follows from (16) and (17) that

$$|\mathcal{L}(Q)| \quad \leq \quad |\mathcal{L}(Q_2)| \cdot G(\kappa_1, K) \quad \leq \quad G(\overline{n} - \kappa_1, K) \cdot G(\kappa_1, K). \tag{18}$$

Iterating this idea, observe that $\ker(Q_2) = \ker(Q) \cap \mathrm{im}(Q)$, hence $\kappa_2 := \dim(\ker Q_2) \leq \kappa_1$. Putting $Q_3 := Q_2 \restriction \ \mathrm{im}(Q_2)$ one deduces, as above,

$$|\mathcal{L}(Q_2)| \quad \leq \quad |\mathcal{L}(Q_3)| \cdot G(\kappa_2, K),$$

which, when substituted into (18), yields

$$|\mathcal{L}(Q)| \quad \leq \quad |\mathcal{L}(Q_3)| \cdot G(\kappa_2, K) \cdot G(\kappa_1, K).$$

By induction and because of $|\mathcal{L}(Q_{m+1})| = 1$, one gets

$$|\mathcal{L}(Q)| \quad \leq \quad G(\kappa_m, K) \cdots G(\kappa_2, K) \cdot G(\kappa_1, K),$$

where $\kappa_m \leq \kappa_{m-1} \leq \cdots \leq \kappa_2 \leq \kappa_1$ are defined in the obvious way. Therefore

$$|\mathcal{L}(Q)| \quad \leq \quad G(\dim(\ker Q), K)^m. \tag{19}$$

We are interested, for fixed $\sigma \in S_n$, in the case $K = GF(2), Q = T_1 + I,\ W = V_1,\ \overline{n} = n_1,\ m = \mu_1$. To fix ideas suppose that $(2, 5, 7, 9)$ is one of the cycles of σ. It gives rise to exactly one nonzero $v \in V$ with $T_\sigma(v) = v$; namely $v := e_2 + e_5 + e_7 + e_9$. Therefore $Q(v) = 0$. Thus, clearly $\dim(\ker Q) = r$. See (15) for the relation between r and n_1. Claim (a) now follows from (18) in view of $\mathcal{L}(T_1) = \mathcal{L}(T_1 + I)$. Claim (b) follows from (19). $\square$

Notice that more than $\lfloor n/2 \rfloor ! 2^n$ permutations $\sigma \in S_n$ have $T_\sigma = T_1$ or, what amounts to the same, $n_1(\sigma) = n$. This is most easily seen when $n = 2^{\alpha_1}$ happens to be a power of 2. Then even $(n-1)!$ permutations $\sigma \in S_n$ have $n_1(\sigma) = n$, namely by (15) all the n-cycles.

In what follows $r = r(\sigma),\ n_1 = n_1(\sigma)$, and log is the logarithm to base 2. Putting

$\mathcal{D}_1 := \{\sigma \in S_n \ | n_1 \leq n - 6 \log n\}$,

$\mathcal{D}_2 := \{\sigma \in S_n \setminus \mathcal{D}_1 \ | 1 \leq r \leq 8 \log n_1\}$,

$\mathcal{D}_3 := \{\sigma \in S_n \setminus \mathcal{D}_1 \ | 8 \log n_1 < r < n_1 - 8 \log n_1\}$,

$\mathcal{D}_4 := \{\sigma \in S_n \setminus \mathcal{D}_1 \ | n_1 - 8 \log n_1 \leq r \leq n - 1\}$,

it is clear that $S_n - \{id\}$ is the disjoint union of the sets $\mathcal{D}_i$ $(1 \leq i \leq 4)$. The remainder of the article essentially amounts to giving upper bounds for each of the four sums $\sum_{\sigma \in \mathcal{D}_i} |\mathcal{L}(T_\sigma)|$. For $i = 4$ a lower bound will be needed as well.

Lemma 4.

$$\sum_{\sigma \in \mathcal{D}_1} |\mathcal{L}(T_\sigma)| = O(2^{(n^2/4) - n \log n}), \tag{20}$$

$$\sum_{\sigma \in \mathcal{D}_2} |\mathcal{L}(T_\sigma)| = O(2^{17 n \log^2 n}), \tag{21}$$

$$\sum_{\sigma \in \mathcal{D}_3} |\mathcal{L}(T_\sigma)| = O(2^{(n^2/4) - n \log n}). \tag{22}$$

Proof. Without always mentioning it, Lemma 1 will be used throughout the proof. As to (20), fix n and consider the maximum of the function

$$\frac{x^2}{4} + \frac{(n-x)^2}{8} + (0.3)n \quad (0 \leq x \leq n - 6\log n).$$

Since for big enough n this maximum is obtained at $x = n - 6\log n$, it follows from Lemma 2 (and Lemma 1) that for all $\sigma \in \mathcal{D}_1$

$$|\mathcal{L}(T_\sigma)| = O(2^{\frac{n_1^2}{4} + \frac{(n-n_1)^2}{8} + (0.3)n}) = O(2^{\frac{(n-6\log n)^2}{4} + \frac{(6\log n)^2}{8} + (0.3)n}) = O(2^{\frac{n^2}{4} - 2n\log n}),$$

which, in view of $|\mathcal{D}_1| \leq n! \leq n^n = 2^{n\log n}$, yields

$$\sum_{\sigma \in \mathcal{D}_1} |\mathcal{L}(T_\sigma)| \quad = \quad 2^{n\log n} \cdot O(2^{\frac{n^2}{4} - 2n\log n}) \quad = \quad O(2^{\frac{n^2}{4} - n\log n}).$$

As to (21), from $r \leq 8\log n_1 \leq 8\log n$ and $\mu_1 \leq n$ and Lemma 3(b) one deduces

$$|\mathcal{L}(T_1)| \quad \leq \quad G(8\log n, 2)^n \;\leq\; \left(8 \cdot 2^{(8\log n)^2/4}\right)^n \quad = \; O(2^{16n\log^2 n \; + \; 3n}).$$

Since $\sigma \in \mathcal{D}_2$ implies $\sigma \notin \mathcal{D}_1$, whence $n_1 > n - 6\log n$, Lemma 2 yields

$$\sum_{\sigma \in \mathcal{D}_2} |\mathcal{L}(T_\sigma)| \quad = \quad 2^{n\log n} \cdot O(2^{16n\log^2 n + \frac{36\log^2 n}{8} + 3.3n}) \quad = \quad O(2^{17n\log^2 n}).$$

As to (22), for all $\sigma \in \mathcal{D}_3$ one derives from Lemma 3(a) that

$$\begin{aligned} |\mathcal{L}(T_1)| \quad &\leq \quad G(r,2) \cdot G(n_1 - r, 2) \quad = \quad O(2^{\frac{r^2}{4} + \frac{(n_1-r)^2}{4}}) \\ &= \quad O(2^{\frac{(8\log n_1)^2}{4} + \frac{(n_1 - 8\log n_1)^2}{4}}) \quad = \quad O(2^{\frac{n_1^2}{4} - 3n_1\log n_1}), \end{aligned}$$

so by Lemma 2

$$|\mathcal{L}(T_\sigma)| = O(2^{\frac{n_1^2}{4} + \frac{(n-n_1)^2}{8} - 3n_1\log n_1 + (0.3)n}) = O(2^{\frac{n^2}{4} - 3n_1\log n_1 + (0.3)n}) = O(2^{\frac{n^2}{4} - 2n\log n}).$$

Here the last equality holds since $n_1 > n - 6\log n$. As previously, one now argues that

$$\sum_{\sigma \in \mathcal{D}_3} |\mathcal{L}(T_\sigma)| \quad = \quad 2^{n\log n} \cdot O(2^{\frac{n^2}{4} - 2n\log n}) \quad = \quad O(2^{\frac{n^2}{4} - n\log n}). \qquad \square$$

The asymptotic behavior of $b(n)$ will depend on the size of

$$Z(n) \;:=\; \sum_{\sigma \in \mathcal{D}_4} |\mathcal{L}(T_\sigma)|.$$

Lemma 4 guarantees that the sum of the other $|\mathcal{L}(T_\sigma)|$ is negligible in comparison. By Lemmata 1 and 4 it would suffice to show that $Z(n) = o(2^{n^2/4})$ in order to prove (1). But we strive for more than (1). This requires a *sharper upper* bound for $Z(n)$ (section 4), as well as a *lower* bound for $Z(n)$ (section 3).

3. A lower bound for $Z(n)$. Consider a transposition $\tau \in S_n$. As seen in the introduction, $\mathcal{L}(T_\tau)$ has size at least $G(n-1,2)$. Here is the precise value:

$$\text{(23)} \qquad \text{If } r(\tau) = n-1, \text{ then } |\mathcal{L}(T_\tau)| = 2G(n-1,2) - G(n-2,2).$$

To see (23) consider without loss of generality the transposition $\tau = (1,2)$. We claim that

$$\text{(24)} \qquad \mathcal{L}(T_{(1,2)}) = \{U \in \mathcal{L}(V) |\ \langle e_1 + e_2 \rangle \subseteq U \text{ or } U \subseteq \langle e_1 + e_2 \rangle^\perp\}.$$

To see (24), let $U \in \mathcal{L}(T_{(1,2)})$ be such that $e_1 + e_2 \notin U$. We have to show that $U \subseteq \langle e_1 + e_2 \rangle^\perp$. Assume to the contrary some $x = \sum_{i=1}^n \lambda_i e_i$ in U has scalar product $(e_1 + e_2) \cdot x \neq 0$. Then $x = e_1 + \sum_{i=3}^n \lambda_i e_i$ or $x = e_2 + \sum_{i=3}^n \lambda_i e_i$, say the former. From $T_{(1,2)}(x) = e_2 + \sum_{i=3}^n \lambda_i e_i$ being in U we get the contradiction $e_1 + e_2 = x + T_{(1,2)}(x) \in U$. This establishes one inclusion in (24). The reverse inclusion is similar and left to the reader.

By (24), $\mathcal{L}(T_{(1,2)})$ is the union of the $G(n-1,2)$-element interval sublattices $[\langle e_1 + e_2 \rangle, V]$ and $[0, \langle e_1 + e_2 \rangle^\perp]$, whose intersection is the $G(n-2,2)$-element interval sublattice $[\langle e_1 + e_2 \rangle, \langle e_1 + e_2 \rangle^\perp]$. This gives (23).

We now double the lower bound in (4). More precisely, because $G(n-2,2) = o(G(n-1,2))$ it follows from (23) and Lemma 1 that

$$\text{(25)} \qquad \sum_{r(\sigma)=n-1} |\mathcal{L}(T_\sigma)| \quad \geq \quad \binom{n}{2} \cdot 2 \cdot 7.3719 \cdot 2^{\frac{(n-1)^2}{4}} \qquad (n \text{ large}).$$

Because $r(\sigma) = n-1$ implies $\sigma \in \mathcal{D}_4$, the right-hand side of (25) is also a lower bound for $Z(n)$.

4. An upper bound for $Z(n)$. From Lemma 1 and the proof of (25) it follows at once that upon transition from 7.3719 to 7.37197 one has

$$\text{(26)} \qquad \sum_{r(\sigma)=n-1} |\mathcal{L}(T_\sigma)| \quad \leq \quad \binom{n}{2} \cdot 2 \cdot 7.37197 \cdot 2^{\frac{(n-1)^2}{4}} \qquad (n \text{ large}).$$

In order to prove that

$$\text{(27)} \qquad \sum_{\sigma \in \mathcal{D}_4} |\mathcal{L}(T_\sigma)| \quad \leq \quad \binom{n}{2} \cdot 2 \cdot 7.37198 \cdot 2^{\frac{(n-1)^2}{4}} \qquad (n \text{ large}),$$

put

$$\mathcal{D} := \{\sigma \in S_n |\ n_1(\sigma) > n - 6\log n \text{ and } n - 14\log n \leq r(\sigma) \leq n-1\}.$$

All $\sigma \in \mathcal{D}_4$ satisfy $n_1(\sigma) > n - 6\log n$, as well as

$$r \quad \geq \quad n_1 - 8\log n_1 \quad > \quad (n - 6\log n) - 8\log n \quad = \quad n - 14\log n,$$

so $\mathcal{D}_4 \subseteq \mathcal{D}$. In view of (26) it thus suffices to show

$$\text{(28)} \qquad \sum_{\sigma \in \mathcal{D}, r(\sigma) \leq n-2} |\mathcal{L}(T_\sigma)| \quad = \quad o(2^{\frac{(n-1)^2}{4}}) \quad = \quad o(2^{\frac{n^2}{4} - \frac{n}{2}}).$$

Fix $\sigma \in \mathcal{D}$ with $r(\sigma) \leq n-2$. Consider T_σ and the associated T_1. Putting $n_1 := n_1(\sigma)$ and $r := r(\sigma)$, Lemma 3(a) yields

$$\begin{aligned}|\mathcal{L}(T_1)| &\leq G(r,2) \cdot G(n_1 - r, 2)\\ &= O(2^{\frac{r^2}{4} + \frac{(n_1-r)^2}{4}}) = O(2^{\frac{(n-2)^2}{4} + \frac{2^2}{4}}) = O(2^{\frac{n^2}{4} - n}).\end{aligned}$$

From $n_1 > n - 6 \log n$ and Lemma 2 one concludes that

$$|\mathcal{L}(T_\sigma)| \quad \leq \quad 2^{\frac{(n-n_1)^2}{8} + (0.3)n} \cdot O(2^{\frac{n^2}{4} - n}) \quad = \quad O(2^{\frac{n^2}{4} - (0.7)n + \frac{36 \log^2 n}{8}}).$$

How many elements has $\mathcal{D}$ at most? We claim that $\mathcal{D}$ is contained in the class $\mathcal{D}'$ of all $\sigma \in S_n$, which have at least $n - 28 \log n$ cycles of length 1. Indeed, if $\sigma \in \mathcal{D}$ had less than $n - 28 \log n$ of them, then σ had less than $(n - 28 \log n) + 14 \log n = n - 14 \log n$ cycles altogether, contradicting the definition of $\mathcal{D}$. Hence

$$|\mathcal{D}| \quad \leq \quad |\mathcal{D}'| \quad \leq \quad \binom{n}{n - 28 \log n} [(28 \log n)!] \quad \leq \quad n^{28 \log n},$$

which implies

$$\sum_{\sigma \in \mathcal{D},\ r(\sigma) \leq n-2} |\mathcal{L}(T_\sigma)| \quad = \quad 2^{28 \log^2 n} \cdot O(2^{\frac{n^2}{4} - (0.7)n + \frac{36 \log^2 n}{8}}) \quad = \quad o(2^{\frac{n^2}{4} - \frac{n}{2}}).$$

This proves (28) and whence (27).

5. The main theorem.

We are now in a position to prove the following.

THEOREM. *For all sufficiently large n one has*

$$\left(1 + 2^{-\frac{n}{2} + 2 \log n + 0.2499}\right) \frac{G(n,2)}{n!} \quad \leq \quad b(n) \quad \leq \quad \left(1 + 2^{-\frac{n}{2} + 2 \log n + 0.2501}\right) \frac{G(n,2)}{n!}.$$

Proof. By Lemma 1 one has $G(n,2) \leq 7.3720 \cdot 2^{n^2/4}$ for all large enough n. Together with (2) and (25) this implies that for large enough n

$$\begin{aligned}b(n) = \frac{G(n,2)}{n!} &\left(1 + \frac{1}{G(n,2)} \sum_{r(\sigma) \leq n-1} |\mathcal{L}(T_\sigma)|\right)\\ &\geq \quad \frac{G(n,2)}{n!} \left(1 + \binom{n}{2} 2^{-\frac{n}{2} + 1.25} \cdot \frac{7.3719}{7.3720}\right)\\ &\geq \quad \frac{G(n,2)}{n!} \left(1 + 2^{-\frac{n}{2} + 2 \log n + 0.2499}\right).\end{aligned}$$

The last inequality holds because

$$\binom{n}{2} \cdot \frac{7.3719}{7.3720} \;=\; \frac{n^2}{2} \left(1 - \frac{1}{n}\right) \cdot \frac{7.3719}{7.3720} \;\geq\; \frac{n^2}{2} \cdot 2^{-0.00001} \cdot 2^{-0.00002} \;=\; 2^{2 \log n - 1.00003}$$

for large n. From Lemma 4 and (27) we see that

$$\sum_{r(\sigma) \leq n-1} |\mathcal{L}(T_\sigma)| \quad \leq \quad \binom{n}{2} \cdot 2 \cdot 7.3720 \cdot 2^{\frac{(n-1)^2}{4}} \qquad (n \text{ large}). \tag{29}$$

By Lemma 1 one has $G(n,2) \geq 7.3719 \cdot 2^{n^2/4}$ for all large enough n, so (29) yields

$$\begin{aligned} b(n) = \frac{G(n,2)}{n!}\left(1 + \frac{1}{G(n,2)} \sum_{r(\sigma)\leq n-1} |\mathcal{L}(T_\sigma)|\right) \\ \leq \quad \frac{G(n,2)}{n!}\left(1 + \binom{n}{2} 2^{-\frac{n}{2}+1.25} \cdot \frac{7.3720}{7.3719}\right) \\ \leq \quad \frac{G(n,2)}{n!}\left(1 + 2^{-\frac{n}{2}+2\log n + 0.2501}\right). \qquad \square \end{aligned}$$

It should be clear from the proof that the exponents 0.2499 and 0.2501 in the theorem *cannot* be replaced by $0.25 \pm \epsilon$. However, equally clear, $0.25 \pm \epsilon$ can be introduced if one distinguishes between even and odd integers. It is also obvious that the theorem implies (1). In turn, (1) implies that the fraction $\beta(n)$ of n-codes X with nontrivial automorphism group $\mathrm{Aut}(X) := \{\sigma \in S_n | \ X_\sigma = X\}$ goes to 0 for $n \to \infty$. Namely, the total number $b(n)$ of equivalence classes satisfies

$$(30) \quad b(n) \quad \geq \quad \frac{\beta(n)G(n,2)}{n!/2} + \frac{(1-\beta(n))G(n,2)}{n!} \quad = \quad \frac{(1+\beta(n))G(n,2)}{n!}.$$

By (1) this forces $\beta(n) \to 0$ as $n \to \infty$. Notice that there is no quick argument why, conversely, (30) together with $\beta(n) \to 0$ should imply (1). This relates to results in [LPR]; see [W2] for details. A year after [W2] the mistake in [W1] was also fixed in [H]; in fact Hou extends formula (1) to prime powers $q > 2$.

REFERENCES

[A] M. Aigner, *Combinatorial Theory*, Springer-Verlag, New York, 1979.

[BF] L. Brickman and P. A. Fillmore, *The invariant subspace lattice of a linear transformation*, Canad. J. Math., 19 (1967), pp. 810–822.

[H] X. D. Hou, *On the asymptotic number of non-equivalent q-ary linear codes*, J. Combin. Theory Ser. A, 112 (2005), pp. 337–346.

[L] R. F. Lax, *On the character of S_n acting on subspaces of $\mathbb{F}_q^n$*, Finite Fields Appl., 10 (2004), pp. 315–322.

[LPR] H. Lefmann, K. T. Phelps, and V. Rödl, *Rigid linear binary codes*, J. Combin. Theory Ser. A, 63 (1993), pp. 110–128.

[O] J. G. Oxley, *Matroid Theory*, Oxford University Press, New York, 1997.

[W1] M. Wild, *The asymptotic number of inequivalent binary codes and nonisomorphic binary matroids*, Finite Fields Appl., 6 (2000), pp. 192–202.

[W2] M. Wild, *The asymptotic number of binary codes and binary matroids*. This is a previous version of the present article with connections to [LPR]. Also available online at http://arxiv.org/abs/cs.IT/0408011.