Available online at www.sciencedirect.com

SciVerse ScienceDirect

Electronic Notes in Discrete Mathematics 38 (2011) 383–387

Electronic Notes in DISCRETE MATHEMATICS

www.elsevier.com/locate/endm

Reverse-free codes and permutations

Zoltán Füredi [a,1,2], Ida Kantor [b,3] Angelo Monti [c,4]
Blerina Sinaimeri [c,5]

[a] *Rényi Institute of Mathematics of the Hungarian Academy of Sciences, Pf. 127, 1364 Budapest, Hungary*

[b] *Institute for Theoretical Computer Science, Charles University 118 00 Prague, Czech Republic*

[c] *Department of Computer Science, Sapienza University of Rome, Rome 00198, Italy*

Abstract

Two codewords $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ form a *reverse-free* pair if $(a_i, a_j) \neq (b_j, b_i)$ holds whenever $1 \leq i < j \leq k$ are indices such that $a_i \neq a_j$. In a *reverse-free code*, each pair of codewords is reverse-free. The maximum size of a reverse-free code with codewords of length k and an n-element alphabet is denoted by $\overline{F}(n, k)$. Let $F(n, k)$ denote the maximum size of a reverse-free code with all codewords consisting of distinct entries.

We determine $\overline{F}(n, 3)$ and $\overline{F}(n, 3)$ exactly whenever n is a power of 3, and asymptotically for other values of n. We prove non-trivial bounds for $F(n, k)$ and $\overline{F}(n, k)$ for general k and for other related functions as well. Using VC-dimension of a matrix, we determine the order of magnitude of $\overline{F}(n, k)$ for n fixed and k tending to infinity.

Keywords: extremal combinatorics, permutations, codes, reverse-free

doi:10.1016/j.endm.2011.09.062

1 Reverse-free permutations and codes

Let k and n be natural numbers, and X an n-element alphabet. The set of all ordered k-tuples with entries in X will be denoted by X^k. A *code* $\mathcal{C}$ is simply a subset of X^k. Its members are called *codewords*, k is its *length*, and $|\mathcal{C}|$ its *size*. A typical problem in coding theory is to determine the maximum size of a code satisfying some local condition. In this paper, the codes are required to be reverse-free.

Definition 1.1 Let a and b be two distinct integers. The pair $\{a, b\}$ is a *reversed pair* for a pair of k-tuples $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, \ldots, y_k)$ if there are two indices $i, j \in [k]$ such that $(x_i, x_j) = (y_j, y_i) = (a, b)$. If $\mathbf{x}$ and $\mathbf{y}$ have no reversed pair, they are *reverse-free.*

A code $\mathcal{F}$ is called *reverse-free* if any two of its members are reverse-free. Let $\overline{F}(n, k)$ be the maximum size of a reverse-free code $\mathcal{C} \subseteq X^k$. If we also require that each codeword consist of k distinct symbols, the maximum size of such code is denoted by $F(n, k)$. A natural companion notion is that of a *flip-full* code, where every two codewords are required to have a reversed pair.

Given a family $\mathcal{F}$, define an $|\mathcal{F}| \times k$ matrix $M(\mathcal{F})$ in a natural way, by listing the the k-tuples in $\mathcal{F}$ as its rows. The family $\mathcal{F}$ is then reverse-free if $M(\mathcal{F})$ has distinct rows and does not contain a 2×2 submatrix with rows (a, b) and (b, a), with $a \neq b$. Similarly, $\mathcal{F}$ is flip-full if every two rows contain such submatrix. Many coding theory problems can be formulated in this way, as determining the maximum number of rows of a matrix with a forbidden submatrix, or with certain submatrices required for every pair (triple, quadruple...) of rows (see, e.g., Körner [4], Vapnik and Chervonenkis [6], or a survey by Anstee [1]).

Similar problems have also been investigated in the information-theoretic setting, giving rise to the notions of *robust capacity* [5] and *forbiddance* [3].

For $n = k$, we have an extremal problem on permutations. Let us mention that a similar problem with a different local condition, determining the maximum size of a set of t-intersecting permutations for n large compared to t, was solved recently by Ellis, Friedgut, and Pilpel [2], proving an analogue of Erdős-Ko-Rado Theorem.

[1] Research supported in part by the Hungarian National Science Foundation OTKA, and by the National Science Foundation under grant NFS DMS 09-01276.
[2] Email: z-furedi@illinois.edu
[3] Email: ida@kam.mff.cuni.cz
[4] Email: monti@di.uniroma1.it
[5] Email: sinaimeri@di.uniroma1.it

2 Triples

For every n and k, the set of all *increasing* k-tuples is reverse-free, hence $F(n,k) \geq \binom{n}{k}$. Is this bound tight? For $k = 3$, the answer is negative, we have $F(n,3) = \left(\frac{5}{4} + o(1)\right)\binom{n}{3}$.

Theorem 2.1 *For any $n \in \mathbb{N}$, we have*

$$\frac{5}{24}n^3 - \frac{1}{2}n^2 - O(n \log n) \leq F(n,3) \leq \frac{5}{24}n^3 - \frac{1}{2}n^2 + \frac{5}{8}n.$$

If n is a power of 3, the upper bound holds with equality.

We also determined $\overline{F}(n,3)$ with a small error term $O(n \log n)$, and exactly whenever n is a power of 3.

Proof (Sketch). We provide a recursive construction of a large reverse-free set $\mathcal{F}$. Partition $[n]$ into three parts, A, B, and C. Find maximum reverse-free families of triples $\mathcal{F}_A$, $\mathcal{F}_B$, and $\mathcal{F}_C$, on the alphabets A, B, and C respectively. Fix a maximum reverse-free family L of triples from the alphabet $\{A, B, C\}$ with allowed repetition of symbols. We have $|L| = \overline{F}(n,3) = 11$. If $X_1, X_2, X_3 \in \{A, B, C\}$ and $(X_1, X_2, X_3) \in L$, we include in $\mathcal{F}$ all the triples (x_1, x_2, x_3) such that $x_i \in X_i$ and the triple (x_1, x_2, x_3) has no reversed pairs with triples in $\mathcal{F}_A \cup \mathcal{F}_B \cup \mathcal{F}_C$, or with any triples already added. Splitting $[n]$ as equally as possible in each step, we obtain the lower bound.

On the other hand, let $\mathcal{F}$ be a reverse-free family of triples. Each 3-element set can appear up to three times as a triple in $\mathcal{F}$. For $0 \leq i \leq 3$, let T_i be the number of 3-element sets that appear i times as triples in $\mathcal{F}$. Define a directed graph $G_{1,2}$ by putting $(x_i, x_j) \in E(G_{ij})$ whenever there is a u such that the triple (x_1, x_2, u) belongs to $\mathcal{F}$. Define graphs $G_{2,3}$ and $G_{3,1}$ analogously. Define two additional directed graphs, $\mathcal{D}$ and $\mathcal{M}$ by putting

$$E(\mathcal{M}) = \{uv : uv \in E(G_i) \cap E(G_j) \text{ for some pair } i,j \in \{1,2,3\}\}$$
$$E(\mathcal{D}) = \{uv : uv \in E(G_1) \cap E(G_2) \cap E(G_3)\}.$$

If a 3-element set appears at least twice as a triple in $\mathcal{F}$, the three vertices form a directed triangle in $\mathcal{M}$. Hence $|T_2| + |T_3|$ is bounded by the number of directed triangles in $\mathcal{M}$, which is at most $(n^3 - n)/24$.

The set T_0 contains all sets $\{u, v, w\}$ such that $uv \in E(\mathcal{D})$ and $uw \in E(\mathcal{D})$. It also contains all sets $\{u, v, w\}$ such that $vu \in E(\mathcal{D})$ and $wu \in E(\mathcal{D})$. If $d^+(u)$ and $d^-(u)$ are the outdegree and indegree of a vertex u in $\mathcal{D}$, we have $|T_0| \geq \frac{1}{2}\sum_{u\in[n]}\left[\binom{d^+(u)}{2} + \binom{d^-(u)}{2}\right]$.

Also, each set in T_3 corresponds to a directed triangle in $\mathcal{D}$, so $|T_3| \leq \frac{1}{3}\sum_{u\in[n]} d^+(u)\cdot d^-(u)$. It follows that

$$|T_3| - |T_0| \leq \sum_{u\in[n]} \left[\frac{1}{3}d^+(u)\cdot d^-(u) - \frac{1}{2}\binom{d^+(u)}{2} - \frac{1}{2}\binom{d^-(u)}{2}\right] \leq \frac{n}{3}.$$

The upper bound then follows from the above by observing that $|\mathcal{F}| = |T_1| + 2|T_2| + 3|T_3|$, and $|T_0| + |T_1| + |T_2| + |T_3| = \binom{n}{3}$. □

3 Small alphabets

If n is fixed and the length of the codewords k tends to ∞, the true order of magnitude of the maximum size is polynomial in k.

Theorem 3.1 *Let $n \geq 2$, $k \geq 2$. Then*

$$\left(\frac{k}{\binom{n}{2}}\right)^{\binom{n}{2}} \leq \overline{F}(n,k) \leq \binom{k}{\leq n-1}\binom{k}{\leq n-2}\cdots\binom{k}{\leq 1} = O\left(k^{\binom{n}{2}}\right),$$

where $\binom{k}{\leq \ell}$ stands for $\sum_{0\leq i\leq \ell}\binom{k}{i}$.

Proof of the upper bound (Sketch). We use induction on n and k. Let $\mathcal{F}$ be a reverse-free code of length k, with n-element alphabet. For $\mathbf{x} \in \mathcal{F}$, define its *$i$-support* $\mathrm{supp}_i(\mathbf{x})$ as the set of indices ℓ such that $x_\ell = i$. Any set A can appear at most $\overline{F}(n-1, k-|A|)$ times as an i-support. Also, there are not too many sets that appear as i-supports. Let $\mathcal{F}_i$ be the family of i-supports. Then

$$|\mathcal{F}_i| \leq \binom{k}{n-1} + \binom{k}{n-2} + \cdots + \binom{k}{0}.$$

To prove this, let M be the matrix whose rows are the indicator vectors of the sets in $\mathcal{F}_i$. We say that the *VC-dimension* of a $(0,1)$-matrix is at least s if one can find s columns such that the matrix restricted to these columns contains all the 2^s possible rows. It is known that if the VC-dimension is at most s, then

$$m \leq \sum_{0\leq i\leq s}\binom{k}{i}. \tag{1}$$

We claim that the VC-dimension of $\mathcal{F}_1$ is at most $n-1$. If this is not the case, then there are indices $s_1, \ldots, s_n$ such that M restricted to these columns contains the complement of the identity matrix. Then, by pigeonhole principle, we find a reversed pair in $\mathcal{F}$. □

4 Further results and open problems

When k is fixed and n tends to infinity, we do not even know whether the trivial lower bound $F(n,k) \geq \binom{n}{k}$ is asymptotically tight. Our best upper bound is $F(n,k) \leq k!\binom{n}{k}/(1.686\cdots + o(1))^k$. Similar results can be proved for $\overline{F}(n,k)$ as well. Another interesting open problem for which we only have partial results is to determine $F(n,n)$.

Define a graph $P_{n,k}$ on the vertex set consisting of all the k-tuples with non-repeating entries in $[n]$ by making two k-tuples adjacent if and only if they have a reversed pair. The independence number of $P_{n,k}$ is equal to $F(n,k)$. Similarly, the clique number of $P_{n,k}$ is equal to the cardinality of the maximum flip-full code of non-repeating k-tuples, denoted by $G(n,k)$. The quantity $G(n,k)$ is interesting for its own sake. But moreover, via the clique-coclique bound

$$F(n,k)G(n,k) \leq k!\binom{n}{k},$$

a lower bound on $G(n,k)$ provides an upper bound on $F(n,k)$. The case $k=n$, i.e., permutations, is especially interesting. We have

$$\frac{1}{8}(1.515\dots)^n \leq G(n,n) \leq \frac{n^2 n!}{(1.898\dots)^n}.$$

References

[1] R. P. Anstee. A survey of forbidden configuration results (online draft, last updated may 25, 2010). http://www.math.ubc.ca/~anstee/FCsurvey10.pdf.

[2] D. Ellis, E. Friedgut, and H. Pilpel. Intersection theorems for permutations. *to appear.*

[3] Emanuela Fachini and János Körner. Forbiddance and capacity. *Graphs and Combinatorics*, pages 1–9, 2010. 10.1007/s00373-010-0987-9.

[4] János Körner. On the extremal combinatorics of the Hamming space. *J. Combin. Theory Ser. A*, 71(1):112–126, 1995.

[5] János Körner and Gábor Simonyi. A Sperner-type theorem and qualitative independence. *J. Comb. Theory Ser. A*, 59(1):90–103, 1992.

[6] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probab. and its Applications*, 16(2):264–280, 1971.