Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory
June 15-21, 2012, Pomorie, Bulgaria *pp. 29–34*

Strong isometries of codes [1]

S. V. Avgustinovich avgust@math.nsc.ru
Sobolev Institute of Mathematics, Novosibirsk State University
E. V. Gorkunov evgumin@gmail.com
Sobolev Institute of Mathematics, Novosibirsk State University

Abstract. For q-ary codes we generalize the theorem on reconstructing a binary code by dimensions of its subcodes. The new reconstructing theorem involves arbitrary positive integer q. The term of a correlation coefficient for a set of subcodes of a code is proposed. Here correlation coefficient of subcodes is a q-ary analog for dimension of a subcode of a binary code.

1 Introduction

Let E_q^n be a q-ary cube, that is the set of words of length n over the alphabet $\{0, 1, 2, \ldots, q-1\}$ with q symbols. The *Hamming distance* $d(x, y)$ between two words $x, y \in E_q^n$ is determined by the number of positions where x and y differ. The *Hamming weight* $w(x)$ of a word $x \in E_q^n$ is the number of its nonzero symbols, i.e. $w(x) = d(x, 0)$. The cube E_q^n equipped with the Hamming distance is a metric space. Any subset C of the space E_q^n is called a *code*. The elements of a code C are called *codewords*. Two codes are *equivalent* if there exists an isometry of the space E_q^n mapping one of the codes into the other one.

In this paper we investigate metric invariants of codes. We propose a generalization for the notion of a dimension of a binary code introduced and investigated in [1]. We show that a bijection between two arbitrary q-ary codes preserving every correlation coefficient of their subcodes is extendable to an isometry of E_q^n. As a consequence of this fact, one can conclude that knowing the multiset of correlation coefficients of subcodes of a q-ary code is sufficient to reconstruct the code up to equivalence.

It is well known that metric invariants of two q-ary codes can not always determine if the codes are equivalent or not. For example, all Hadamard codes are pairwise isometric, but there exist many nonequivalent Hadamard codes. Some results concerning research what kinds of metric information are sufficient to reconstruct a code can be found, e.g., in [1–5].

In case $q = 2$ it is turned out [1] that a bijection between two binary codes preserving dimensions of their subcodes can be extended to an isometry of the

[1] This research is partially supported by the Russian Foundation for Basic Research (project No. 10-01-00424) and the Federal Target Program "Scientific and Scientific-Pedagogical Personnel of Innovative Russia" for 2009–2013 (government contract No. 02.740.11.0368).

space E_2^n. In other words, the multiset of dimensions of subcodes of a binary code determines the code up to equivalence. Here *dimension* of a code means dimension of the minimum face of E_2^n containing the code. Further research showed [6] that taking all dimensions of subcodes of a code is redundant. Namely, a bijection between two binary codes can be extended to an isometry of E_2^n if it preserves dimensions of their subcodes with even cardinality.

Direct generalization of the dimension method to reconstruct q-ary codes is useless. There exist bijective mappings between nonequivalent ternary codes preserving dimensions of their subcodes. The approach proposed in this paper is developed in detail for ternary codes only. In conclusion we give some remarks on correlation coefficients of subcodes of q-ary codes in general.

2 Necessary notations and definitions

Consider a ternary code C of length n. If all of its codewords have the same symbol at the i-th position, we call the position *unessential* for the code C. By $N(C)$ we denote the set of all unessential positions of C. For disjoint subcodes C_1 and C_2 of the code C, let $K(C_1, C_2)$ denote the number of positions from $N(C_1) \cap N(C_2)$ at which codewords from different codes are distinct. Formally,

$$K(C_1, C_2) = |\{i \in N(C_1) \cap N(C_2) \colon x_i \neq y_i \text{ for any } x \in C_1 \text{ and } y \in C_2\}|,$$

We refer to $K(C_1, C_2)$ as *correlation coefficient* of the codes C_1 and C_2. Let us outline some evident equalities concerning this notion.

1. $K(x, y) = d(x, y)$ for any $x, y \in E_3^n$;
2. $K(\{x, y\}, \varnothing) = n - d(x, y)$ for any $x, y \in E_3^n$;
3. $K(C, \varnothing) = n - \mathrm{Dim}(C)$, where $\mathrm{Dim}(C)$ is dimension of the code $C \subseteq E_3^n$ in the sense mentioned above.

By this means, one can easily note that correlation coefficients contain some metric information about the code structure.

Following [6], we say a bijection $I \colon C_1 \to C_2$ between two codes $C_1, C_2 \subseteq E_3^n$ to be a *strong isometry* if it preserves any correlation coefficient of subcodes of C_1, i.e. $K(A, B) = K(I(A), I(B))$ for any $A, B \subseteq C_1$. For a code matrix M_1 of the code C_1, let $M_2 = I(M_1)$ denote the code matrix of the code C_2 obtained from M_1 by applying the map I to every row of M_1.

Consider a code $C \subseteq E_3^n$ of cardinality m. Let M be its code matrix. The ternary alphabet generates a correspondence between columns of the matrix M and partitions of the set $\{1, \ldots, m\}$ into three subsets. Each subset of such an *alphabet partition* includes indices of the matrix rows with the same symbol in the corresponding column. For example, the column $(0, 1, 0, 1, 2)^{\mathsf{T}}$ has $(\{1, 3\}, \{2, 4\}, \{5\})$ as its alphabet partition. Thus every code matrix has its

own multiset of alphabet partitions. At the same time, it is not difficult to see that matrices with equal multisets of alphabet partitions represent equivalent codes.

Proposition 1. *Suppose $C_1, C_2 \subseteq E_3^n$ are ternary codes, and M_1 is a code matrix of C_1. A bijection $I\colon C_1 \to C_2$ is extendable to an isometry of the space E_3^n if and only if the multisets of alphabet partitions of the matrices M_1 and $M_2 = I(M_1)$ are equal.*

In the sequel we will keep usage of the following notation:

- $\mathcal{M} = \{1, \ldots, m\}$ is a set of indices for rows of an $m \times n$ matrix M.
- $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ are alphabet partitions.
- $P, Q, R \subseteq \mathcal{M}$ are subsets of $\mathcal{M}$.
- $a_0(\mathcal{P}) = P$, $a_1(\mathcal{P}) = Q$, $a_2(\mathcal{P}) = R$ denote three parts of an alphabet partition $\mathcal{P} = (P, Q, R)$.

Taking into account Proposition 1, it can be noticed that the key step in our approach to reconstruct a code $C \subseteq E_3^n$ is to determine how many columns with a certain alphabet partition a code matrix M of the code contains. Let $k(\mathcal{P})$ be the number of columns with the alphabet partition $\mathcal{P}$ in the matrix M. For any subset $S \subseteq \mathcal{M}$, denote by $C(S)$ the subcode of the code C formed by those rows of M that have indices in S. If the code is clear from the context, then $K(Q, R)$ is used instead of $K(C(Q), C(R))$.

At the end of the section we define a partial order $\preccurlyeq$ over the set of alphabet partitions. For two partitions of $\mathcal{M}$, put

$$(P_1, Q_1, R_1) \preccurlyeq (P_2, Q_2, R_2) \text{ if an only if } P_1 \subseteq P_2,\ Q_1 \supseteq Q_2,\ \text{and } R_1 \supseteq R_2.$$

This partial order helps us to get a relation between correlation coefficients and the value $k(\mathcal{P})$. Namely, for any code $C \subseteq E_3^n$, its code matrix M, and an alphabet partition $\mathcal{P} = (P, Q, R)$, the following equality can easily be verified:

$$K(Q, R) = \sum_{\mathcal{Q} \preccurlyeq \mathcal{P}} k(\mathcal{Q}). \tag{1}$$

3 Reconstructing codes by correlation coefficients

In this section we prove the main result:

Theorem 1. *Any strong isometry between ternary codes can be extended to an isometry of the whole space E_3^n.*

As indicated above, Proposition 1 allows us to focus our attention on a code C whose codewords are enumerated and form an unknown code matrix M. The aim is to obtain the multiset of alphabet partitions of M using the known correlation coefficients of subcodes of C. In order to show that this multiset is identified by correlation coefficients, we invert the relation (1) by means of the Möbius function for the partial order $\preccurlyeq$. Some information on posets, their Möbius functions and the Möbius inversion theorem can be found, e.g., in [7, Chapter 2].

The correlation coefficient is a well-defined function on the set of alphabet partitions. Indeed, put $K(\mathcal{P}) = K(a_1(\mathcal{P}), a_2(\mathcal{P}))$ for any alphabet partition $\mathcal{P}$ and, conversely, $K(Q, R) = K(\mathcal{P})$ for any disjoint subsets $Q, R \subseteq \mathcal{M}$ and the alphabet partition $\mathcal{P} = (\mathcal{M} \setminus (Q \cup R), Q, R)$.

By the Möbius inversion theorem we get a reversion of (1):

$$k(\mathcal{P}) = \sum_{\mathcal{Q} \preccurlyeq \mathcal{P}} \mu(\mathcal{Q}, \mathcal{P}) K(a_1(\mathcal{Q}), a_2(\mathcal{Q})), \tag{2}$$

where $\mu(\mathcal{Q}, \mathcal{P})$ is the Möbius function of the poset of alphabet partitions with the partial order $\preccurlyeq$. According to the theory of Möbius functions, we have the following axioms for μ:

$$\begin{aligned} &\mu(\mathcal{Q}, \mathcal{P}) = 0 && \text{if } \mathcal{Q} \not\preccurlyeq \mathcal{P}, \\ &\mu(\mathcal{P}, \mathcal{P}) = 1 && \text{for all } \mathcal{P}, \\ &\mu(\mathcal{Q}, \mathcal{P}) = -\sum_{\mathcal{Q} \preccurlyeq \mathcal{R} \prec \mathcal{P}} \mu(\mathcal{Q}, \mathcal{R}) && \text{if } \mathcal{Q} \prec \mathcal{P}. \end{aligned} \tag{3}$$

On the base of these axioms we derive the Möbius function for the partially ordered set of alphabet partitions with the partial order introduced in Section 2.

Proposition 2. *Given a partial order $\preccurlyeq$ by the rule*

$$(P_1, Q_1, R_1) \preccurlyeq (P_2, Q_2, R_2) \text{ if and only if } P_1 \subseteq P_2,\ Q_1 \supseteq Q_2, \text{ and } R_1 \supseteq R_2,$$

the poset of alphabet partitions has the Möbius function of the form

$$\mu(\mathcal{Q}, \mathcal{P}) = (-1)^{|a_0(\mathcal{P})| - |a_0(\mathcal{Q})|} \quad \text{for } \mathcal{Q} \preccurlyeq \mathcal{P}. \tag{4}$$

Proof. The assertion is certainly true when $|a_0(\mathcal{P})| - |a_0(\mathcal{Q})| = 0$ or 1. By induction assume (4) to be true for $|a_0(\mathcal{P})| - |a_0(\mathcal{Q})| < s$ and consider a case with $|a_0(\mathcal{P})| - |a_0(\mathcal{Q})| = s$. Then (3) becomes

$$\mu(\mathcal{Q}, \mathcal{P}) = -1 + \binom{s}{1} - \binom{s}{2} + \ldots - \binom{s}{j}(-1)^j + \ldots - \binom{s}{s-1}(-1)^{s-1}, \tag{5}$$

since there are $\binom{s}{j}$ partitions $\mathcal{R}$ with $\mathcal{Q} \preccurlyeq \mathcal{R} \prec \mathcal{P}$ and $|a_0(\mathcal{R})| - |a_0(\mathcal{Q})| = j$. Namely, such alphabet partitions are obtained from $\mathcal{Q}$ by adjoining to $a_0(\mathcal{Q})$ any j of the s elements of $a_0(\mathcal{P})$ not in $a_0(\mathcal{Q})$.

Comparison of (5) with the binomial expansion of $(1-1)^s$ gives $\mu(\mathcal{Q}, \mathcal{P}) = (-1)^s$. This proves the proposition. □

Combining (2) and (4), we obtain an explicit formula for $k(\mathcal{P})$.

Proposition 3. *The number of columns with an alphabet partition $\mathcal{P}$ in a code matrix of any ternary code is equal to*

$$k(\mathcal{P}) = \sum_{\mathcal{Q} \preccurlyeq \mathcal{P}} (-1)^{|a_0(\mathcal{P})| - |a_0(\mathcal{Q})|} K(a_1(\mathcal{Q}), a_2(\mathcal{Q})).$$

Now let us return to Theorem 1 and finalize its proof. Proposition 3 implies that the matrices M_1 and M_2 have the same multisets of alphabet partitions. It remains to apply Proposition 1.

If there exists a strong isometry between two codes, then the codes are called *strongly isometric*. Theorem 1 yields the following.

Corollary 1. *Strongly isometric ternary codes are equivalent.*

4 Conclusion

In this paper, for the ternary case, we demonstrate how to reconstruct a code having the correlation coefficients of its subcodes. To get a description of this approach for q-ary codes in general, one should slightly change the notion of a correlation coefficient of subcodes. Some remarks on that are made below.

Given a q-ary code $C \subseteq E_q^n$, we should consider correlation coefficient of a set consisting of $q-1$ mutually disjoint subcodes of C. By analogy with the definition for $q = 3$, given a set of mutually disjoint subcodes $C_1, \ldots, C_{q-1} \subseteq C$, the correlation coefficient $K(C_1, \ldots, C_{q-1})$ enumerates every position from $N(C_1) \cap \ldots \cap N(C_{q-1})$ such that codewords of each of the subcodes have their unique symbol at this position. All notation used above should be changed in an obvious way to take the new definition into account. It is needed to omit the word "ternary" and replace E_3^n by E_q^n in Theorem 1 and Corollary 1 to formulate the final result.

In order to reconstruct a code, we begin with a code matrix of the code. Concerning the matrix, we only know correlation coefficient of each set of $q-1$ mutually disjoint subsets of its rows, because such a row subset corresponds to a subcode of the code. The principal point is to determine the multiset of alphabet partitions of the code matrix. An $m \times n$ matrix can contain all of q^m different columns, while n is large enough. Therefore, generally speaking, to construct the multiset of alphabet partitions of the code matrix it is necessary

to obtain q^m values, which show the number of occurrences in the matrix for each m column. On the other hand, it is not difficult to see that the set of all correlation coefficients of subcodes is excessive for that. Thus a natural question is arises: For a q-ary code, what is the minimal set of correlation coefficients of its subcodes that is sufficient to reconstruct the code up to equivalence?

References

[1] S. V. Avgustinovich, On a Strong Isometry of Binary Codes, *Diskretn. Anal. Issled. Oper., Ser. 1,* **7** (3), 3–5, 2000 (in Russian).

[2] Zh. K. Abdurakhmanov, *On the Geometric Structure of Error-Correcting Codes,* Doctoral thesis, Tashkent, 1991 (in Russian).

[3] S. V. Avgustinovich, F. I. Solov'eva, To Metric Rigidity of Binary Codes, *Probl. Inform. Transm.,* **39** (2), 23–28, 2003.

[4] V. Yu. Krasin, On Weak Isometries of the Boolean Cube, *Diskretn. Anal. Issled. Oper., Ser. 1,* **13** (4), 26–32, 2006 (in Russian).

[5] F. I. Solov'eva, S. V. Avgustinovich, T. Honold, and W. Heise, On the Extendability of Code Isometries, *J. Geom.,* **61** (1–2), 3–16, 1998.

[6] E. V. Gorkunov, S. V Avgustinovich, On the Reconstruction of Binary Codes from the Dimensions of Their Subcodes, *J. Appl. Ind. Math.,* **5** (3), 348–351, 2011; transl. from *Diskretn. Anal. Issled. Oper.,* **17** (5), 15–21, 2010 (in Russian).

[7] M. Hall, Jr., *Combinatorial Theory,* John Wiley & Sons, New York, 1998.