0336_e

DISTANCE REGULARITY OF KERDOCK CODES

F. I. Solov′eva and N. N. Tokareva UDC 519.725

Abstract: A code is called *distance regular*, if for every two codewords $\mathbf{x}$, $\mathbf{y}$ and integers i, j the number of codewords $\mathbf{z}$ such that $d(\mathbf{x},\mathbf{z}) = i$ and $d(\mathbf{y},\mathbf{z}) = j$, with d the Hamming distance, does not depend on the choice of $\mathbf{x}$, $\mathbf{y}$ and depends only on $d(\mathbf{x},\mathbf{y})$ and i, j. Using some properties of the discrete Fourier transform we give a new combinatorial proof of the distance regularity of an arbitrary Kerdock code. We also calculate the parameters of the distance regularity of a Kerdock code.

Keywords: distance regular code, Kerdock code, Reed–Muller code, discrete Fourier transform, bent function, distance regular graph, association scheme

§ 1. Introduction

Essentially using some properties of the discrete Fourier transform, in the paper we give a new combinatorial proof of the distance regularity of an arbitrary Kerdock code. The proof is shorter and simpler than the complicated algebraic proof by Delsarte (see [1]) in studying regular Hamming association schemes. Moreover, we calculate the parameters δ_{ij}^k of distance regularity of a Kerdock code which were unknown before. The Kerdock codes are asymptotically optimal codes closely connected with so good codes as Reed–Muller codes, Preparata codes, and also with other combinatorial and algebraic objects. These codes are the first infinite class of distance regular codes with weight spectrum having more than three nonzero values.

Let us remind that a code is called *distance regular*, if for every two codewords $\mathbf{x}$, $\mathbf{y}$ and integers i, j the number of codewords $\mathbf{z}$ such that $d(\mathbf{x},\mathbf{z}) = i$ and $d(\mathbf{y},\mathbf{z}) = j$, with d the Hamming distance, does not depend on the choice of $\mathbf{x}$, $\mathbf{y}$ and depends only on $d(\mathbf{x},\mathbf{y})$ and i, j.

Distance regularity is a strong structural property of codes. Informally speaking distance regularity demonstrates a high symmetry of the code structure: between every two codewords at some fixed distance from one another all other codewords are replaced in the same way independently of the choice of this pair. The property of distance regularity is closely connected with other regular properties of the characteristic graphs of the codes, their metrical properties, with metrical Hamming association schemes (see [2, Chapter 21]) and distance regular graphs (see [3]). Studying this property of a code was initiated by Delsarte [1], Bannai and Ito [4], and Levenshtein [5].

By now the distance regularity of codes is not studied sufficiently yet. We know only a few codes with this property. All binary Hadamard codes are distance regular and among them the first order Reed–Muller codes; the binary and ternary Golay codes and also the series of the codes obtained from them by extending, shortening, and puncturing are distance regular (see [5] and [6]). In [7] it is established that all perfect binary codes with code distance 3 are not distance regular with the exception of the Hamming codes of lengths 3 and 7. An analogous result takes place for extended perfect binary codes with code

The first author was partially supported by the Royal Swedish Academy of Sciences. The second author was supported by the Russian Science Support Foundation, the Integration Project of the Siberian Branch of the Russian Academy of Sciences “A Tree-Like Catalog of Internet Mathematical Resources” (Grant No. 35), and the Russian Foundation for Basic Research (Grants 07–01–00248 and 08–01–00671). Both authors were partially supported by Novosibirsk State University.

Novosibirsk. Translated from *Sibirskiĭ Matematicheskiĭ Zhurnal*, Vol. 1, No. 0, pp. 669–682, *******–*******, 1960. Original article submitted May 30, 2006.

distance 4 (see [8]). It is shown in [9] that all Preparata codes are not distance regular with the exception (up to equivalence) of the Preparata code of length 16 called the Nordstrom–Robinson code.

Consider the structure of the paper. In § 2 we give the necessary definitions and properties of Reed–Muller codes, bent functions, Kerdock codes, and the discrete Fourier transform. In § 3 we prove the auxiliary Lemmas 1–3 describing some properties of bent functions and cosets of the first order Reed–Muller code. In § 4 we show that to check the distance regularity property of Kerdock codes it suffices to consider (see the definition of distance regularity) a pair of codewords $\mathbf{x}$, $\mathbf{y}$ such that one codeword is the all zero vector (see Lemma 4). Using the lemma of § 3, we prove for these codewords that the number of codewords $\mathbf{z}$ (see above the definition of distance regularity) is a constant depending on the parameters i, j, k and code length (see Lemmas 5–8). From here we infer the main result of this paper (see Theorem 1) that every Kerdock code is distance regular. In § 5 we make some comments on the distance regularity property of a code.

§ 2. Necessary Definitions and Notions

Consider the metric space E^n of all binary vectors of length n equipped with the Hamming metric. The *Hamming distance* $d(\mathbf{x}, \mathbf{y})$ between vectors $\mathbf{x}$ and $\mathbf{y}$ is the number of coordinates where the vectors differ. The *Hamming weight* $w(\mathbf{x})$ of a vector $\mathbf{x}$ is the number of the nonzero coordinates of $\mathbf{x}$. An arbitrary subset of E^n is called a *binary code*. The *code distance* d is the minimum distance between its distinct vectors called *codewords*. All zero and all one vectors are denoted by $\mathbf{0}$ and $\mathbf{1}$. The family of the numbers A_i, $i = 0, \ldots, n$, of all codewords of weight i for a code of length n is called its *weight spectrum*.

We will give other definitions and notions on mostly following [2].

2.1. Reed–Muller codes. Consider a set of Boolean functions of m variables $v_1, \ldots, v_m$. Each of these functions can uniquely be represented as a polynomial of degree at most r called the *Zhegalkin polynomial*:

$$\sum_{j=1}^{m} \sum_{1 \le i_1, \ldots, i_j \le m} a_{i_1 \ldots i_j} v_{i_1} \ldots v_{i_j} + a,$$

where $a_{i_1 \ldots i_j}$, a are either 0 or 1, the numbers $i_1, \ldots, i_j$ are pairwise distinct and the sum is given modulo 2. Let $\mathbf{v} = (v_1, \ldots, v_m)$ be a vector ranging over the space E^m and let $\mathbf{f}$ be the binary vector of length 2^m obtained from a Boolean function $f(v_1, \ldots, v_m)$.

The *rth order binary Reed–Muller code* $R(r, m)$ of length $n = 2^m$, $0 \le r \le m$, is the set of all binary vectors $\mathbf{f}$ of length n, where f is a Boolean function presenting a polynomial of degree at most r. The Reed–Muller code $R(r, m)$ has size $2^{1+\binom{m}{1}+\cdots+\binom{m}{r}}$ and code distance 2^{m-r}, where $\binom{m}{i}$ stands for the relevant binomial coefficient.

Given two binary vectors $\mathbf{a} = (a_1, \ldots, a_m)$ and $\mathbf{v} = (v_1, \ldots, v_m)$ of length m, we denote by $\mathbf{a} \cdot \mathbf{v} = a_1 v_1 + \cdots + a_m v_m$ their inner product (here the sum is modulo 2). The first order Reed–Muller code $R(1, m)$ can be described by all linear functions $\mathbf{a} \cdot \mathbf{v} + b$ of a variable $\mathbf{v}$, where $\mathbf{a}$ is a fixed vector of length m, the constant b is either 0 or 1. In other words every codeword of $R(1, m)$ can be represented as

$$\sum_{i=1}^{m} a_i \mathbf{v}^i + b \cdot \mathbf{1},$$

where $\mathbf{v}^i$ denotes the binary vector of length $n = 2^m$ corresponding to the Boolean function equal to v_i. The size of the code $R(1, m)$ is $2n$, its code distance is $n/2$. It is not difficult to see that nonzero components of the weight spectrum of the code are $A_0 = A_n = 1$ and $A_{n/2} = 2n - 2$.

Almost all codewords of the second order Reed–Muller code $R(2, m)$ correspond to the so-called *quadratic* Boolean functions with Zhegalkin polynomials of degree 2.

2.2. Bent functions. Further we will use maximal nonlinear Boolean functions that have maximal possible Hamming distances from the set of all linear functions. In the sequel we suppose that

$$n = 2^m \text{ for even } m \geq 4, \quad d = (n - \sqrt{n})/2.$$

A Boolean function f of m variables $v_1, \ldots, v_m$ is called *bent*, if the vector $\mathbf{f}$ has the Hamming distance d or $n - d$ from every codeword of $R(1, m)$. From this definition it follows that the weight of $\mathbf{f}$ is either d or $n - d$.

Proposition 1. *A coset $\mathbf{f} + R(1, m)$ corresponding to a bent function f of m variables has n vectors of weight d and n vectors of weight $n - d$.*

It is easy to see that every vector of the coset $\mathbf{f} + R(1, m)$ is associated with some bent function.

2.3. Kerdock codes. The *Kerdock code* K of length $n = 2^m$ for even $m \geq 4$ consists of the code $R(1, m)$ together with $(n - 2)/2$ cosets of $R(1, m)$ in $R(2, m)$ such that the Boolean functions associated with these cosets are quadratic bent functions with the property that the sum of every two of them is again a bent function.

It follows from the properties of bent functions of m variables (see [2]) that each Kerdock code K of length n has size n^2 and is a maximal subcode of the code $R(2, m)$ with code distance $(n - \sqrt{n})/2$. Moreover, Sidel′nikov (see [10]) established that an upper bound on the cardinality of any binary code of length n with code distance $(n - \sqrt{n})/2$ is equivalent to the function n^2 as $n \to \infty$. Therefore, a Kerdock code is asymptotically optimal in the class of all codes of length n with this distance. The nonzero values of the Kerdock code weight distribution are

$$A_0 = A_n = 1, \quad A_d = A_{n-d} = n(n - 2)/2, \quad A_{n/2} = 2n - 2.$$

Note that every Kerdock code possesses the *antipodal property*; i.e., for an arbitrary codeword $\mathbf{x}$ the word $\mathbf{x} + \mathbf{1}$ belongs to the code.

A code is *distance invariant* if the number of codewords at distance i from a given codeword $\mathbf{x}$ does not depend on the choice of $\mathbf{x}$ and depends only on i. According to [2, Chapter 15], we have

Proposition 2. *Every Kerdock code is distance invariant.*

The first code of this sort with each admissible length n was constructed by Kerdock in 1972 (see [11]). For $n = 2^m$ such that $m - 1$ is a composite number Kantor [12] constructed the set of $2^{\sqrt{m}/2}$ pairwise nonequivalent Kerdock codes of length n in 1982. In 1989 Nechaev [13] constructed a Kerdock code of length n for every admissible n which can be represented in terms of linear recurrent sequences over the ring $\mathbb{Z}_4$ and is equivalent to the original Kerdock code [11]. Later in 1994 Hammons et al. [14] suggested another construction of Kerdock codes for every admissible length. They proved that under the Gray map the original Kerdock code [11] is the image of an extended linear cyclic code over $\mathbb{Z}_4$ (see [14] and also [15]).

Kerdock codes are closely correlated with the Preparata codes; i.e., the binary codes of maximal size with length $n = 2^m$, even $m \geq 4$, having code distance 6 (see [2, Chapter 15]). These codes are formally dual in the sense that their weight spectra satisfy the MacWilliams identity (see [2, Chapter 5]). The Preparata and Kerdock codes of length 16 coincide. This code is unique up to equivalence and called the *Nordstrom–Robinson code* [16].

Every Kerdock code of length n can be represented by the so-called Kerdock set composed of $n/2$ skew-symmetric binary matrices of order $m \times m$ containing the all zero matrix such that a difference of every two of them is a nonsingular matrix. Every matrix from the Kerdock set corresponds to a quadratic form which uniquely defines the assigned coset of the code $R(1, m)$ in the code $R(2, m)$ that belongs to the Kerdock code. This approach to describing Kerdock codes is linked with some special problems in finite geometries and quadratic forms. Many Kerdock codes of length $n = 2^m$ are constructed using projective planes and special nonassociative algebras over $GF(2^{m-1})$ (see [17, 18]).

Let $\rho(m - 1)$ be the number of prime divisors of $m - 1$ with multiplicity counted. In [19] it is established that for every even $m \geq 4$ there exist at least

$$\frac{(2^{m-1} - 1)^{\rho(m-1)-3}}{(m - 1)^2}$$

pairwise nonequivalent Kerdock codes of length $n = 2^m$.

To investigate the distance regularity of Kerdock codes of length $n = 2^m$ we will essentially use the linear and bent functions of m variables and the properties of the discrete Fourier transform.

2.4. The discrete Fourier transform. A detailed description of the properties of the discrete Fourier transform (also called the *Walsh–Hadamard transform*) given in this section can be found in [2, Chapter 14].

Let $\mathbf{v} = (v_1, \ldots, v_m)$ be a vector ranging over the space of binary vectors E^m. A Boolean function f of m variables $v_1, \ldots, v_m$ corresponds uniquely to the binary vector $\mathbf{f}$ composed of the values of f and having length $n = 2^m$. With a Boolean function f we associate a real function F of m variables $v_1, \ldots, v_m$ as follows:

$$F(\mathbf{v}) = (-1)^{f(\mathbf{v})}. \tag{1}$$

Using (1) and replacing in every coordinate of a vector $\mathbf{f}$ the value 0 by 1 and 1 by -1, we obtain from a binary vector $\mathbf{f}$ the real vector $\mathbf{F}$ of the same length corresponding to F. In the sequel unless otherwise stated we will also use the formula (1) for other Boolean functions. For example, to the Boolean functions $f_{\mathbf{a}}$ and h by (1) we associate the real functions $F_{\mathbf{a}}$ and H respectively.

The *discrete Fourier transform* of a real vector $\mathbf{F}$ is defined as

$$\widehat{F}(\mathbf{u}) = \sum_{\mathbf{v} \in E^m} (-1)^{\mathbf{u}\cdot\mathbf{v}} F(\mathbf{v}) \quad \text{for all } \mathbf{u} \in E^m. \tag{2}$$

For every vector $\mathbf{u}$ of length m we get from (2):

$$\widehat{F}(\mathbf{u}) = n - 2d\biggl(\mathbf{f}, \sum_{i=1}^{m} u_i \mathbf{v}^i\biggr), \tag{3}$$

$$\widehat{F}(\mathbf{u}) = -n + 2d\biggl(\mathbf{f}, \sum_{i=1}^{m} u_i \mathbf{v}^i + \mathbf{1}\biggr) \tag{4}$$

(see details in [2, Chapter 14, Section 3]). Observe the following *inversion formula* of the discrete Fourier transform

$$F(\mathbf{v}) = \frac{1}{n} \sum_{\mathbf{u} \in E^m} (-1)^{\mathbf{u}\cdot\mathbf{v}} \widehat{F}(\mathbf{u}) \quad \text{for all } \mathbf{v} \in E^m. \tag{5}$$

§ 3. Auxiliary Lemmas

To prove the main theorem we will need some auxiliary propositions on the cosets of the first order Reed–Muller code, bent functions, and the discrete Fourier transform.

Let f be a bent function of m variables (throughout what follows f will denote a function of this sort). By Proposition 1 in the coset $\mathbf{f} + R(1, m)$ of the first order Reed–Muller code there are exactly n vectors of weight d and n vectors of weight $n - d$ antipodal to them. To every vector $\mathbf{a}$ in E^m we put in one-to-one correspondence the vector $\mathbf{f}_{\mathbf{a}}$ of length $n = 2^m$ and weight d from this coset. The vector can be described by the Boolean function

$$f_{\mathbf{a}}(\mathbf{v}) = f(\mathbf{v}) + \mathbf{a} \cdot \mathbf{v} + b_{\mathbf{a}} \tag{6}$$

of m variables; here $b_{\mathbf{a}}$ equal to either 0 or 1 is uniquely determined from the condition $d(\mathbf{f}_{\mathbf{a}}, \mathbf{0}) = d$. It is evident that the different vectors $\mathbf{a}$ and $\mathbf{a}'$ of E^m correspond to the different vectors $\mathbf{f}_{\mathbf{a}}$ and $\mathbf{f}_{\mathbf{a}'}$ of this coset. It should be noted that according to (3) if the all zero vector is substituted for a variable of the function $\widehat{F}_{\mathbf{a}}$ then this condition is equivalent to the condition

$$\widehat{F}_{\mathbf{a}}(\mathbf{0}) = n - 2d \tag{7}$$

independently of the choice of $\mathbf{a}$ in E^m.

From (1) and (6) we get

$$F_{\mathbf{a}}(\mathbf{v}) = (-1)^{\mathbf{a}\cdot\mathbf{v} + b_{\mathbf{a}}} F(\mathbf{v}). \tag{8}$$

Proposition 3. *If* $\mathbf{u}$ *is a vector of length* m *then for the discrete Fourier transform of* $\mathbf{F_a}$ *we have*

$$\widehat{F}_{\mathbf{a}}(\mathbf{u}) = (-1)^{b_{\mathbf{a}}}\widehat{F}(\mathbf{u}+\mathbf{a}). \tag{9}$$

The proof of this fact follows from the chain of equalities:

$$\widehat{F}_{\mathbf{a}}(\mathbf{u}) \stackrel{(2)}{=} \sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{u}\cdot\mathbf{v}}F_{\mathbf{a}}(\mathbf{v}) \stackrel{(8)}{=} \sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{u}\cdot\mathbf{v}+\mathbf{a}\cdot\mathbf{v}+b_{\mathbf{a}}}F(\mathbf{v}) \stackrel{(2)}{=} (-1)^{b_{\mathbf{a}}}\widehat{F}(\mathbf{u}+\mathbf{a}).$$

Lemma 1. *Let* $f_{\mathbf{a}}(\mathbf{v})$ *be a bent function of* m *variables as defined in* (6)*. Then for every vector* $\mathbf{v}$ *in* E^m *we have*

$$\sum_{\mathbf{a}\in E^m} F_{\mathbf{a}}(\mathbf{v}) = \frac{n}{n-2d}.$$

Proof. Given a binary vector $\mathbf{v}$ of length m, consider the inversion formula (5) for the discrete Fourier transform

$$F(\mathbf{v}) = \frac{1}{n}\sum_{\mathbf{a}\in E^m}(-1)^{\mathbf{a}\cdot\mathbf{v}}\widehat{F}(\mathbf{a}).$$

Using (9), we infer

$$F(\mathbf{v}) = \frac{1}{n}\sum_{\mathbf{a}\in E^m}(-1)^{\mathbf{a}\cdot\mathbf{v}+b_{\mathbf{a}}}\widehat{F}_{\mathbf{a}}(\mathbf{0}).$$

Involving (7), (8) and multiplying both sides of this equality by $F(\mathbf{v})$, we obtain

$$F^2(\mathbf{v}) = \frac{n-2d}{n}\sum_{\mathbf{a}\in E^m} F_{\mathbf{a}}(\mathbf{v}).$$

Since $F^2(\mathbf{v}) = 1$ for every binary vector $\mathbf{v}$, we come to the required equality. □

Lemma 2. *Let* $f_{\mathbf{a}}(\mathbf{v})$ *be a bent function of* m *variables as defined in* (6)*. Then for every nonzero vector* $\mathbf{u}$ *in* E^m *we have*

$$\sum_{\mathbf{a}\in E^m}\widehat{F}_{\mathbf{a}}(\mathbf{u}) = 0.$$

Proof. Using (9), represent the required equality as

$$\sum_{\mathbf{a}\in E^m}(-1)^{b_{\mathbf{a}}}\widehat{F}(\mathbf{u}+\mathbf{a}) = 0.$$

Inserting the formula (2) for the discrete Fourier transform, on the left-hand side of the equality, we get

$$\sum_{\mathbf{a}\in E^m}(-1)^{b_{\mathbf{a}}}\sum_{\mathbf{v}\in E^m}(-1)^{(\mathbf{u}+\mathbf{a})\cdot\mathbf{v}}F(\mathbf{v}).$$

Changing the order of sums and using (8), we obtain

$$\sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{u}\cdot\mathbf{v}}\sum_{\mathbf{a}\in E^m}F_{\mathbf{a}}(\mathbf{v}).$$

From Lemma 1 we derive

$$\frac{n}{n-2d}\sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{u}\cdot\mathbf{v}}.$$

Note that if $\mathbf{v}$ ranges over E^m for every fixed nonzero vector $\mathbf{u}$, the inner product $\mathbf{u}\cdot\mathbf{v}$ takes values 0 and 1 equally often. Hence,

$$\sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{u}\cdot\mathbf{v}} = 0.$$

Therefore, we arrive at the required equality. □

Lemma 3. *Let $f_{\mathbf{a}}(\mathbf{v})$ be a bent function of m variables as defined in* (6). *Then for a binary vector* $\mathbf{g}$ *of length n and weight d the sum of all values of real functions $(-1)^{g(\mathbf{v})}F_{\mathbf{a}}(\mathbf{v})$, where* $\mathbf{a}$ *belongs to E^m, does not depend on the choice of* $\mathbf{g}$ *and equals a constant; i.e.,*

$$\sum_{\mathbf{a}\in E^m,\,\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}F_{\mathbf{a}}(\mathbf{v})=n.$$

Proof. Changing the order of sums, we obtain

$$\sum_{\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}\sum_{\mathbf{a}\in E^m}F_{\mathbf{a}}(\mathbf{v}).$$

Using Lemma 1, we get

$$\frac{n}{n-2d}\sum_{\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}.$$

Note that the expression $\sum_{\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}$ equals the difference between the number of zeros and units in the binary vector $\mathbf{g}$. Since $\mathbf{g}$ has weight d, this difference is $n-2d$. Whence we get the required equality. □

§ 4. Distance Regularity of Kerdock Codes

Consider an arbitrary Kerdock code K of length $n=2^m$ for an even $m\geq 4$. The set of all codewords of weight i of K we denote by K_i. Recall that A_i denote the corresponding value of the weight spectrum of a code; i.e., $|K_i|=A_i$. Fix nonnegative integers i,j,k. Following [7], for a code C we denote by $\delta_{ij}^k(\mathbf{x})$ the number of weight j codewords at distance k from the codeword $\mathbf{x}$ of weight i. To prove the distance regularity of a Kerdock code we will use the following

Lemma 4. *If for a Kerdock code of length n for all admissible i,j,k the functions δ_{ij}^k are constants depending only on i, j, k, and n then every Kerdock code is distance regular.*

Proof. Let K be a Kerdock code. For all codewords $\mathbf{x}$ and $\mathbf{y}$ in K at distance i from each other we denote by $\Delta_{jk}^i(\mathbf{x},\mathbf{y})$ the number of codewords $\mathbf{z}$ such that $d(\mathbf{x},\mathbf{z})=j$ and $d(\mathbf{y},\mathbf{z})=k$. Let us prove that the values of Δ_{jk}^i do not depend on the choice of $\mathbf{x}$ and $\mathbf{y}$. Since the switching of a Kerdock code K by any codeword $\mathbf{x}$ is an isometry; i.e., preserves the Hamming distance, it is easy to see that the code $\mathbf{x}+K$ is again a Kerdock code. Whence $\Delta_{jk}^i(\mathbf{x},\mathbf{y})=\delta_{ij}^k(\mathbf{x}+\mathbf{y})$, where by the condition of this lemma δ_{ij}^k are constants depending only on i,j,k, and n. Therefore, K is distance regular. □

Below we prove (see Lemmas 5–8) that for a Kerdock code all functions δ_{ij}^k for all admissible i,j,k are constants depending on i,j,k,n and not depending on the choice of the code. From this and Lemma 4, we will get the distance regularity of a Kerdock code (see below Theorem 1), the constants δ_{ij}^k play a role of the parameters of distance regularity.

To prove this we consider some properties of the functions δ_{ij}^k for all admissible i,j,k from the spectrum of all possible distances $\{0,d,n/2,n-d,n\}$ between codewords of a Kerdock code (see the propositions below).

By the antipodality of a Kerdok code we have

Proposition 4. *For all i,j,k and a codeword* $\mathbf{x}$ *of weight i we have*

$$\delta_{ij}^k(\mathbf{x})=\delta_{i,n-j}^{n-k}(\mathbf{x})=\delta_{n-i,n-j}^k(\mathbf{x}+1)=\delta_{n-i,j}^{n-k}(\mathbf{x}+1). \tag{10}$$

Given i,j,k and calculating by two different ways the numbers of ordered codeword pairs $\mathbf{x}\in K_i$ and $\mathbf{y}\in K_j$ such that $d(\mathbf{x},\mathbf{y})=k$, we obtain the following well-known equality

$$\sum_{\mathbf{x}\in K_i}\delta_{ij}^k(\mathbf{x})=\sum_{\mathbf{y}\in K_j}\delta_{ji}^k(\mathbf{y}). \tag{11}$$

By Proposition 3 every Kerdock code is distance invariant, and so for each codeword $\mathbf{x}$ of weight i the number of codewords at distance k from $\mathbf{x}$ equals A_k. Every codeword among all A_j codewords of weight j has one of the five possible distances from an arbitrary codeword $\mathbf{x}$ of weight i. This property follows from the weight spectrum properties of a Kerdock code. Hence, we have

Proposition 5. *For all i, j, k and an arbitrary codeword $\mathbf{x}$ of weight i, the following hold:*

$$A_k = \delta_{i0}^k(\mathbf{x}) + \delta_{id}^k(\mathbf{x}) + \delta_{i,n/2}^k(\mathbf{x}) + \delta_{i,n-d}^k(\mathbf{x}) + \delta_{in}^k(\mathbf{x}), \tag{12}$$

$$A_j = \delta_{ij}^0(\mathbf{x}) + \delta_{ij}^d(\mathbf{x}) + \delta_{ij}^{n/2}(\mathbf{x}) + \delta_{ij}^{n-d}(\mathbf{x}) + \delta_{ij}^n(\mathbf{x}). \tag{13}$$

In the sequel we denote the identical equality by $\equiv$.

Lemma 5. *The functions δ_{ij}^k such that at least one of the parameters i, j, k is either 0 or n are constants depending only on i, j, k, and n.*

Proof. By Proposition 3 every Kerdock code is distance invariant. From this we easily get

$$\delta_{ij}^0 \equiv \delta_{i0}^j \equiv \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j; \end{cases} \quad \delta_{ij}^n \equiv \delta_{in}^j \equiv \begin{cases} 0 & \text{if } i + j \neq n, \\ 1 & \text{if } i + j = n. \end{cases} \tag{14}$$

Using the Kerdock code weight distribution, we obtain

$$\delta_{0j}^k \equiv \begin{cases} 0 & \text{if } j \neq k, \\ A_j & \text{if } j = k; \end{cases} \quad \delta_{nj}^k \equiv \begin{cases} 0 & \text{if } j + k \neq n, \\ A_j & \text{if } j + k = n, \end{cases} \tag{15}$$

which completes the proof. □

We suppose further that $i, j, k \in \{d, n/2, n-d\}$. To prove the lemmas below we will use the properties (10)–(15) of δ_{ij}^k.

Lemma 6. *$\delta_{n/2,n/2}^{n/2} \equiv 2n - 4$; $\delta_{d,n/2}^{n/2} \equiv 0$; and $\delta_{dd}^{n/2} \equiv n - 1$.*

Proof. Since $K_0 \cup K_{n/2} \cup K_n$ is the first order Reed–Muller code $R(1, m)$, every codeword $\mathbf{x}$ of weight $n/2$ in the Kerdock code is at distance $n/2$ from exactly $A_{n/2} - 2$ codewords of weight $n/2$. Then

$$\delta_{n/2,n/2}^{n/2} \equiv 2n - 4.$$

Using Proposition 5 (see the property (12) for $i = k = n/2$) and (14) for every codeword $\mathbf{x}$ of weight $n/2$, we get

$$2n - 2 = 1 + \delta_{n/2,d}^{n/2}(\mathbf{x}) + 2n - 4 + \delta_{n/2,n-d}^{n/2}(\mathbf{x}) + 1.$$

Considering that δ_{ij}^k assume nonnegative values, from the last equality we obtain $\delta_{n/2,d}^{n/2} \equiv \delta_{n/2,n-d}^{n/2} \equiv 0$. Therefore by (11) for $i = k = n/2$, $j = d$ we get $\delta_{d,n/2}^{n/2} \equiv 0$. By Proposition 5 for $i = d$, $k = n/2$ and every weight d codeword $\mathbf{x}$ the equality (12) turns into equality $2n - 2 = \delta_{dd}^{n/2}(\mathbf{x}) + \delta_{d,n-d}^{n/2}(\mathbf{x})$. Hence, by Proposition 4 we have $\delta_{dd}^{n/2} \equiv \delta_{d,n-d}^{n/2} \equiv n - 1$. □

Lemma 7. *$\delta_{n/2,d}^d \equiv n(n-2)/4$.*

Proof. Consider a codeword $\mathbf{g}$ of weight $n/2$ from a Kerdock code. Since $\mathbf{g}$ belongs to $R(1, m)$, this vector $\mathbf{g}$ corresponds to the linear Boolean function g of m variables. Put

$$g(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_m v_m$$

for some nonzero vector $\mathbf{u} = (u_1, \ldots, u_m)$ of length m (the case $g(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} + 1$ is analogous). The vectors of weight d in cosets of the code $R(1, m)$ contained in the Kerdock code define K_d. The number of the cosets is $(n-2)/2$. Consider the vectors of weight d in an arbitrary coset $\mathbf{f} + R(1, m)$ where f is some bent function. The number of these vectors is equal to n. Each of these vectors can be described by the bent function $f_{\mathbf{a}}(\mathbf{v}) = f(\mathbf{v}) + \mathbf{a} \cdot \mathbf{v} + b_{\mathbf{a}}$ satisfying (7), i.e. $\widehat{F}_{\mathbf{a}}(\mathbf{0}) = n - 2d$, where $\mathbf{a}$ is the vector of length m and $b_{\mathbf{a}}$ is either 0 or 1 (see (6)). From the definition of a bent function and (3) we derive

$$\widehat{F}_{\mathbf{a}}(\mathbf{u}) = \begin{cases} n - 2d & \text{if } d(\mathbf{f_a}, \mathbf{g}) = d, \\ -n + 2d & \text{if } d(\mathbf{f_a}, \mathbf{g}) = n - d. \end{cases}$$

By Lemma 2 for every nonzero vector $\mathbf{u}$ of length m we have $\sum_{\mathbf{a} \in E^m} \widehat{F}_{\mathbf{a}}(\mathbf{u}) = 0$. Therefore the distances d and $n - d$ between $\mathbf{g}$ and weight d vectors of the coset $\mathbf{f} + R(1, m)$ are met equally often, namely $n/2$ times. Multiplying this number by the number $(n-2)/2$ of nontrivial cosets of the code $R(1, m)$, we obtain $\delta_{n/2,d}^d \equiv n(n-2)/4$. □

Lemma 8. *$\delta_{dd}^{d} \equiv (n-d)(n-4)/2$ and $\delta_{dd}^{n-d} \equiv d(n-4)/2$.*

PROOF. Let $\mathbf{g}$ be a weight d codeword of a Kerdock code. Consider a bent function f of m variables such that its corresponding vector $\mathbf{f}$ belongs to the Kerdock code but $\mathbf{g}$ does not belong to $\mathbf{f}+R(1,m)$. The set of weight d vectors of $\mathbf{f}+R(1,m)$ can be represented by the bent functions $f_{\mathbf{a}}(\mathbf{v}) = f(\mathbf{v})+\mathbf{a}\cdot\mathbf{v}+b_{\mathbf{a}}$ satisfying (7), i.e., $\widehat{F}_{\mathbf{a}}(\mathbf{0}) = n-2d$. Recall that here $\mathbf{a}$ is (as above) the vector of length m and $b_{\mathbf{a}}$ is either 0 or 1 (see (6)). Let $h(\mathbf{v}) = g(\mathbf{v})+f(\mathbf{v})$. By the definition of a Kerdock code the function h is bent. Suppose that

$$H(\mathbf{v}) = (-1)^{h(\mathbf{v})}.$$

Prove that

$$(-1)^{b_{\mathbf{a}}}\widehat{H}(\mathbf{a}) = \begin{cases} n-2d & \text{if } d(\mathbf{g},\mathbf{f}_{\mathbf{a}}) = d, \\ -n+2d & \text{if } d(\mathbf{g},\mathbf{f}_{\mathbf{a}}) = n-d. \end{cases} \tag{16}$$

In fact we have

$$d(\mathbf{g},\mathbf{f}_{\mathbf{a}}) = d\biggl(\mathbf{g},\mathbf{f}+\sum_{i=1}^{m} a_i\mathbf{v}^i + b_{\mathbf{a}}\cdot\mathbf{1}\biggr) = d\biggl(\mathbf{h},\sum_{i=1}^{m} a_i\mathbf{v}^i + b_{\mathbf{a}}\cdot\mathbf{1}\biggr).$$

For $\widehat{H}(\mathbf{a})$ we use (3) in the case $b_{\mathbf{a}}=0$ and (4) if $b_{\mathbf{a}}=1$. Since h is bent, the distance $d\bigl(\mathbf{h},\sum_{i=1}^{m} a_i\mathbf{v}^i+b_{\mathbf{a}}\cdot\mathbf{1}\bigr)$ can only be equal to d or $n-d$. Hence, we have (16).

Using (2) for the discrete Fourier transform, we can convert the left-hand side of (16) to the form

$$(-1)^{b_{\mathbf{a}}}\sum_{\mathbf{v}\in E^m}(-1)^{\mathbf{a}\cdot\mathbf{v}}H(\mathbf{v}).$$

Inserting $H(\mathbf{v}) = (-1)^{g(\mathbf{v})}F(\mathbf{v})$ and using $F(\mathbf{v}) = (-1)^{\mathbf{a}\cdot\mathbf{v}+b_{\mathbf{a}}}F_{\mathbf{a}}(\mathbf{v})$ (see(8)), we obtain

$$\sum_{\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}F_{\mathbf{a}}(\mathbf{v}).$$

Therefore from (16) we get

$$\frac{1}{n-2d}\sum_{\mathbf{v}\in E^m}(-1)^{g(\mathbf{v})}F_{\mathbf{a}}(\mathbf{v}) = \begin{cases} 1 & \text{if } d(\mathbf{g},\mathbf{f}_{\mathbf{a}}) = d, \\ -1 & \text{if } d(\mathbf{g},\mathbf{f}_{\mathbf{a}}) = n-d. \end{cases} \tag{17}$$

By the definition of a Kerdock code the function $g+f_{\mathbf{a}}$ is bent for every vector $\mathbf{a}$. Then the weight of every vector $\mathbf{g}+\mathbf{f}_{\mathbf{a}}$ is either d or $n-d$. Therefore the distance between $\mathbf{g}$ and $\mathbf{f}_{\mathbf{a}}$ can only be d or $n-d$. Denote by $\mu_{d,f}(\mathbf{g})$ and $\mu_{n-d,f}(\mathbf{g})$ the number of the vectors $\mathbf{f}_{\mathbf{a}}$ at distance d and $n-d$ from the vector $\mathbf{g}$ respectively. Since the number of all vectors $\mathbf{f}_{\mathbf{a}}$ is n, we have

$$\mu_{d,f}(\mathbf{g}) + \mu_{n-d,f}(\mathbf{g}) = n.$$

Summing both sides of (17) over all vectors $\mathbf{a}$ of length m and applying Lemma 3, we get

$$\mu_{d,f}(\mathbf{g}) - \mu_{n-d,f}(\mathbf{g}) = \frac{n}{n-2d}.$$

Solving the system of the last two equations and inserting $d = (n-\sqrt{n})/2$, we have

$$\mu_{d,f}(\mathbf{g}) = n-d, \quad \mu_{n-d,f}(\mathbf{g}) = d.$$

Note that the so-obtained values do not depend on the choice of the codeword $\mathbf{g}$ of weight d and the bent function f such that the coset $\mathbf{f}+R(1,m)$ does not contain $\mathbf{g}$. For the fixed weight d codeword $\mathbf{g}$ the number of cosets of the Kerdock code by the first order Reed–Muller code such that the cosets differ from $R(1,m)$ and do not contain the vector $\mathbf{g}$ is equal to $(n-4)/2$. Since the distance between the codeword $\mathbf{g}$ and every vector of the coset $\mathbf{g}+R(1,m)$ containing the word $\mathbf{g}$ is not equal to d and $n-d$, we conclude that

$$\delta_{dd}^{d}(\mathbf{g}) = \mu_{d,f}(\mathbf{g})(n-4)/2, \quad \delta_{dd}^{n-d}(\mathbf{g}) = \mu_{n-d,f}(\mathbf{g})(n-4)/2.$$

Inserting $\mu_{d,f}$ and $\mu_{n-d,f}$, we obtain the required equalities for δ_{dd}^{d} and δ_{dd}^{n-d}. $\square$

Theorem 1. *Every Kerdock code of length $n = 2^m$ for an even $m \geq 4$ is distance regular.*

PROOF. From Lemmas 6–8, the properties (10)–(13) of δ_{ij}^k, and the formulas (14), (15) we obtain that every function δ_{ij}^k with parameters i, j, k from $\{d, n/2, n-d\}$ is a constant, i.e.,

$$\delta_{dd}^d \equiv \delta_{n-d,n-d}^d \equiv \delta_{d,n-d}^{n-d} \equiv \delta_{n-d,d}^{n-d} \equiv (n-d)(n-4)/2,$$

$$\delta_{d,n-d}^d \equiv \delta_{n-d,d}^d \equiv \delta_{d,d}^{n-d} \equiv \delta_{n-d,n-d}^{n-d} \equiv d(n-4)/2,$$

$$\delta_{n/2,d}^d \equiv \delta_{n/2,n-d}^d \equiv \delta_{n/2,d}^{n-d} \equiv \delta_{n/2,n-d}^{n-d} \equiv n(n-2)/4,$$

$$\delta_{n/2,n/2}^d \equiv \delta_{d,n/2}^{n/2} \equiv \delta_{n/2,d}^{n/2} \equiv \delta_{n/2,n-d}^{n/2} \equiv \delta_{n-d,n/2}^{n/2} \equiv \delta_{n/2,n/2}^{n-d} \equiv 0, \quad \delta_{n/2,n/2}^{n/2} = 2n-4,$$

$$\delta_{d,n/2}^d \equiv \delta_{n-d,n/2}^d \equiv \delta_{dd}^{n/2} \equiv \delta_{d,n-d}^{n/2} \equiv \delta_{n-d,d}^{n/2} \equiv \delta_{n-d,n-d}^{n/2} \equiv \delta_{d,n/2}^{n-d} \equiv \delta_{n-d,n/2}^{n-d} \equiv n-1.$$

These constants do not depend on the choice of the Kerdock code of length n. From here by Lemmas 4 and 5 we derive the distance regularity of every Kerdock code. □

Kerdock codes form an infinite class of distance regular codes with weight spectrum containing more than three (namely five) nonzero values.

A code is *strong distance invariant* if the number of codeword pairs $\mathbf{x}$ and $\mathbf{y}$ such that $d(\mathbf{x}, \mathbf{y}) = k$, $d(\mathbf{x}, \mathbf{z}) = i$, and $d(\mathbf{y}, \mathbf{z}) = j$ for any codeword $\mathbf{z}$ depends on the numbers i, j, k and does not depend on the choice of $\mathbf{z}$. This property is weaker than the property of distance regularity.

Corollary 1. *Every Kerdock code of length $n = 2^m$ for every even $m \geq 4$ is strong distance invariant.*

By [20] all perfect binary codes with code distance 3 are strong distance invariant but according to [7] they are not distance regular.

§ 5. Comments

Here we discuss some approaches to studying the distance regularity (nonregularity) of arbitrary codes.

1. To investigate the distance regularity of a reduced code (i.e. containing the all zero vector) sometimes it is helpful to use information about the different group transformations mapping the code into itself. The *automorphism group* $\mathrm{Aut}(C)$ of a code C of length n consists of all isometries of E^n (the combinations of the permutations on n coordinate positions and switchings by vectors from E^n) that transform the code into itself. A code C is *transitive* if $\mathrm{Aut}(C)$ acts transitively on the set of its codewords. The subgroup $\mathrm{Sym}(C)$ of $\mathrm{Aut}(C)$ corresponding to all permutations on n coordinates with the trivial switching by all zero vector is called the *symmetry group* of C. According to [9] we have

Proposition 6. *If the group $\mathrm{Sym}(C)$ of a transitive code C acts transitively on every set of all codewords of a fixed weight then C is distance regular.*

2. Consider a generalization of Lemma 4. Let for a reduced code C of length n the numbers $\delta_{ij}^k(\mathbf{x})$ be (as in § 4) the numbers of codewords of weight j at distance k from the codeword $\mathbf{x}$ of weight i.

Proposition 7. *Assume that for every codeword $\mathbf{x}$ of the code C of length n for all admissible i, j, k all functions δ_{ij}^k for the code $\mathbf{x}+C$ are constants c_{ij}^k depending on i, j, k, n and not depending on $\mathbf{x}$. Then C is distance regular.*

3. To prove the distance nonregularity of a code we can use the property (11) of δ_{ij}^k. Assuming that δ_{ij}^k are constants it is possible sometimes to prove (for example, see [7–9]) that some of the equalities $A_i \delta_{ij}^k = A_j \delta_{ji}^k$ do not hold for any integer value of δ_{ij}^k, where A_i and A_j are the corresponding values of the weight spectrum of the reduced code.

The results of this paper are announced in [21]. The authors are very grateful to P. Sharpin, G. Kerigyan, and K. Bay for useful discussions.

References

1. *Delsarte P.,* An Algebraic Approach to the Association Schemes of Coding Theory, Philips Research Reports Supplements 10, Historical Jrl., Ann Arbor (1973).
2. *MacWilliams F. J. and Sloane N. J. A.*, The Theory of Error-Correcting Codes, North-Holland, Amsterdam (1977).
3. *Brouwer A. E., Cohen A. M., and Neumaier A.*, Distance-Regular Graphs, Springer-Verlag, Berlin (1989).
4. *Bannai E. and Ito T.*, Algebraic Combinatorics. I: Association Schemes, Benjamin, London (1984).
5. *Levenshtein V. I.*, "Universal bounds for codes and designs," in: Handbook of Coding Theory. Vol. 1, Elsevier, Amsterdam, 1998, pp. 499–648.
6. *Topalova S. T.*, "Distance regularity of some linear codes," in: Abstracts of the Annual Workshop on Algebraic and Combinatorial Coding Theory, St. Zagora, Bulgaria, 2000, p. 18.
7. *Avgustinovich S. V. and Solov′eva F. I.*, "On distance regularity of perfect binary codes," Problems Inform. Transmission, **34**, No. 3, 247–249 (1998).
8. *Avgustinovich S. V. and Solov′eva F. I.*, "New constructions and properties of perfect codes," in: Proc. Intern. Workshop "Discrete Analysis and Operation Research" [in Russian], Novosibirsk, 2000, pp. 5–10.
9. *Solov′eva F. I. and Tokareva N. N.*, "On distance nonregularity of Preparata codes," Siberian Math. J., **48**, No. 2, 327–333 (2007).
10. *Sidel′nikov V. M.*, "Extremal polynomials used in bounds of code volume," Problems Inform. Transmission, **16**, No. 3, 174–186 (1981).
11. *Kerdock A. M.*, "A class of low-rate non-linear binary codes," Inform. Control, **20**, No. 2, 182–187 (1972).
12. *Kantor W. M.*, "An exponential number of generalized Kerdock codes," Inform. Control, **53**, No. 1–2, 74–80 (1982).
13. *Nechaev A. A.*, "The Kerdock code in cyclic form," Diskret. Mat., **1**, No. 4, 123–139 (1989).
14. *Hammons A. R., Kumar P. V., Calderbank A. R., Sloane N. J. A., and Solé P.*, "The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes," IEEE Trans. Inform. Theory, **40**, No. 2, 301–319 (1994).
15. *Wan Zhe-Xian*, Quaternary Codes, World Scientific, Singapore (1997).
16. *Nordstrom A. W. and Robinson J. P.*, "An optimum nonlinear code," Inform. Control, **11**, No. 5–6, 613–616 (1967).
17. *Kantor W. M.*, "Codes, quadratic forms and finite geometries," Proc. Sympos. Appl. Math., **50**, 153–177 (1995).
18. *Calderbank A. R., Cameron P. J., Kantor W. M., and Seidel J. J.*, "$\mathbb{Z}_4$-Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets," Proc. London Math. Soc., **75**, 436–480 (1997).
19. *Kantor W. M. and Williams M. E.*, "Symplectic semifield planes and $\mathbb{Z}_4$-linear codes," Trans. Amer. Math. Soc., **356**, No. 3, 895–938 (2004).
20. *Vasil′eva A. Yu.*, "Strong distance invariance of perfect binary codes," Diskret. Anal. Issled. Oper. Ser. 1, **9**, No. 4, 33–40 (2002).
21. *Solov′eva F. I. and Tokareva N. N.*, "On the property of distance regularity of Kerdock and Preparata codes," in: Proc. Tenth Intern. Workshop "Algebraic and Combinatorial Coding Theory," 3–9 September 2006, Zvenigorod, Russia; IITP, Moscow, 2006, pp. 248–251.

F. I. Solov′eva; N. N. Tokareva
Sobolev Institute of Mathematics, Novosibirsk, Russia
E-mail address: sol@math.nsc.ru; tokareva@math.nsc.ru