Crandall's problem

Jürgen Bierbrauer
Department of Mathematical Sciences
Michigan Technological University
HOUGHTON (MI) 49931 (USA)

November 13, 2001

1 Crandall's problem - covering functions

Ron Crandall (xrc@aptix.com) posed a problem, which arose in a cryptographic context (steganography). He suspected a link to coding theory. Such a link does indeed exist. It will turn out to be a version of the covering radius problem. I start with the original problem, the way I understand it.

We want to device methods to hide a certain small picture inside a given big picture. The aim of the design is to arrange things such that the existence of the hidden image will never be suspected. The big picture is given to us. It consists of a lot of pixels. One bit is extracted from each pixel. We use a natural method, block coding. The pixels are divided into blocks of N. Let $x = (x_1, x_2, \ldots, x_N) \in I\!F_2^N$ be the N-tuple of bits extracted from the given block of N pixels. We want to use this block to build n bits of the hidden image. Let $y = (y_1, y_2, \ldots, y_n)$ be the part of the hidden image to be constructed from x. We have no control on x. The tuple y is given to us. We have to construct a function $f : I\!F_2^N \longrightarrow I\!F_2^n$. Assume f is given. The ideal situation is when $f(x) = y$. Imagine this is not the case. We have to change x, that is to replace x by x' such that $f(x') = y$. This means that we have to change the pixels. As the changes made within the picture should not be evident we are interested in changing a minimal amount of pixels. The number of pixels that have to be changed is the Hamming distance $d(x, x')$. This leads to the following conflicting aims in the construction of f : when

N is given, then we want n to be big (our block of pixels should give a non negligible part of the hidden image) and we want the maximum possible $d(x, x')$ to be small (the number of pixels that have to be changed within each block should be small), say $d(x, x') \leq \rho$

2 Translation in coding terms

The discussion in the preceding section leads us to the binary case of the following definition. There is no reason not to consider arbitrary finite fields as ground fields.

Definition 1 *A function $f : \mathbb{F}_q^N \longrightarrow \mathbb{F}_q^n$ is a* **covering function** $COV(\rho, N, n)_q$ *if for every $y \in \mathbb{F}_q^n$ and $x \in \mathbb{F}_q^N$ there is some $x' \in \mathbb{F}_q^N$ such that $f(x') = y$ and $d(x, x') \leq \rho$.*

For the application sketched in Section 1 we are interested in constructing binary covering functions $(\rho, N, n) = (\rho, N, n)_2$ such that the **change rate** ρ/N is small and the **rate** n/N is large. Clearly, both rates are bounded by 1 and the aims are in conflict. Fix $y \in \mathbb{F}_q^n$. The defining property of Definition 1 says that every vector $x \in \mathbb{F}_q^N$ is at Hamming distance at most ρ from some word from $f^{-1}(y)$. In other words $f^{-1}(y)$ is a code with **covering radius** $\leq \rho$ for every $y \in \mathbb{F}_q^n$. Equivalently we may say that $\mathbb{F}_q^N$ is partitioned into q^n codes, where each of these codes has covering radius $\leq \rho$.

Definition 2 *A* **large set** *of m covering codes $LCOV[m](N, \rho)_q$ is a partition of $\mathbb{F}_q^N$ into m subcodes, where each subcode has covering radius $\leq \rho$.*

Theorem 1 *The following are equivalent:*

- *A covering function $COV(\rho, N, n)_q$.*
- *A large set $LCOV[q^n](N, \rho)_q$.*

We see that the idea of using **large sets** emerges once again. This idea of partitioning the universe (in our case $\mathbb{F}_q^N$) into parts all of which have the same structure was popularized in design theory by the work of Teirlinck. Stinson proved in [26] that **resilient functions** may equivalently be described as large sets of orthogonal arrays (see also [3]). In our context we

have shown that the structures meeting the requirements of Section 1 are equivalent with large sets of covering codes.

It is natural to consider the special case of **linear** covering functions. In that case $f^{-1}(0)$ will be a linear space, and hence a linear covering code. As $f^{-1}(y)$ is a coset of $f^{-1}(0)$ for every y and these cosets have the same properties as the linear code we see that the large set will exist automatically.

Theorem 2 *The following are equivalent:*

- *A linear covering function* $COV(\rho, N, n)_q$.
- *A linear code* $[N, N-n]_q$ *(q-ary, length* N*, codimension* n*) with covering radius* $\leq \rho$.

Let H be a check matrix of the linear code in Theorem 2. Then H is an $(n, N)-$matrix. We may describe the corresponding linear covering function f by $f(x) = Hx$. Here we write our vectors as column vectors. The function $l(n, \rho; q)$ is defined as the smallest number N such that there exists a linear q-ary code $\mathcal{C}$ of length N, dimension $N - n$ and covering radius $\leq \rho$. This function has been studied by coding theorists for a long time, especially in the binary case $q = 2$. Theorem 2 shows that $l(n, \rho; q)$ is the smallest N such that a linear covering function $COV(\rho, N, n)_q$ exists. Let us denote by $l^*(n, \rho; q)$ the smallest N such that a covering function $COV(\rho, N, n)_q$ exists.

3 Bounds

We start from a trivial bound, the **sphere packing bound.**

Definition 3 *The number of q-ary n-tupels of* **weight** *(= number of nonzero entries) at most i is*

$$V_q(i, n) = \sum_{j=0}^{i} \binom{n}{j} (q-1)^j \tag{1}$$

Let $\mathcal{C} \subseteq I\!F_q^n$ be a (not necessarily linear) code of covering radius ρ. The union of the balls of radius ρ centered at the codewords must be the whole space. This leads to the following lower bound on the number of codewords:

Theorem 3 (sphere packing bound) *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a (not necessarily linear) code of covering radius ρ. Then*

$$| \mathcal{C} | \geq \frac{q^N}{V_q(\rho, N)}$$

When will the bound of Theorem 3 be achieved with equality? This means that the balls of radius ρ centered at the codewords partition $\mathbb{F}_q^N$. Such codes are known as **perfect codes.** The minimum distance of such a perfect code is $2\rho+1$. The parameters of perfect codes have been completely determined. In each case there are linear codes with these parameters. The Hamming codes have parameters $[\frac{q^k-1}{q-1}, \frac{q^k-1}{q-1} - k, 3]_q$ and covering radius 1. Aside of this infinite family only two more linear perfect codes exist: the binary Golay code $[23, 11, 7]$ with $\rho = 3$ and the ternary Golay code $[11, 6, 5]_3$ with $\rho = 2$. This yields the following precise values of our l-functions:

Theorem 4

$$l^*(k, 1; q) = l(k, 1; q) = \frac{q^k - 1}{q - 1}$$

$$l^*(11, 3; 2) = l(11, 3; 2) = 23 \text{ and } l^*(5, 2; 3) = l(5, 2; 3) = 11.$$

The situation is easy for large ρ. Recall that $l(n, \rho; q)$ is the minimum N such that a q-linear code of length N and codimension n exists, which has covering radius $\leq \rho$. Clearly $N \geq n$, and $N = n$ means that the code is the 0-code, of covering radius n. Let $\rho < n$. Then $N > n$.

Lemma 1 *The repetition code $[N, 1, N]_q$ has covering radius $N - \lceil N/q \rceil$.*

Proof: Let $x \in \mathbb{F}_q^N$ and λ_i the frequency of entry $i \in \mathbb{F}_q$ as an entry of x. There is an $i \in \mathbb{F}_q$ such that $\lambda_i \geq \lceil N/q \rceil$. It follows that the distance of x from the repetition code is $leqN - \lceil N/q \rceil$. Clearly we have equality.■
We see that the repetition code is optimal as long as $n + 1 - \lceil (n+1)/q \rceil \leq \rho < n$.

Theorem 5 *We have $l(n, n; q) = n$ and $l(n, \rho; q) = n + 1$ if $n + 1 - \lceil (n + 1)/q \rceil \leq \rho < n$.*

4 Asymptotics

Let f be a covering function $(\rho, N, n)_q$. The steganographic application motivates the following asymptotic question: Let $\overline{\rho} = \rho/N$ (the change rate or **relative covering radius**) and $R = n/N$ (the information rate). $\overline{\rho}$ is an upper bound (worst case) on the fraction of the number of pixels that have to be changed. We want $\overline{\rho}$ to be small and we want the rate to be as big as possible. It may be expected that we can achieve better values by admitting codes of larger length, in analogy with the case of the (relative) minimum distance.

Let $\alpha(x)$ be the limsup (for $N \longrightarrow \infty$) of the rate of covering functions (ρ, N, n) having $\overline{\rho} \leq x$. It is obvious (using the direct sum) that whenever a covering function (ρ, N, n) exists, then for every natural number c there is a covering function $COV(c\rho, cN, cn)$. Because of the results of the preceding section we have $\alpha(x) = 1$ if $\overline{\rho} \geq (q-1)/q$. We can therefore restrict to $\overline{\rho} < (q-1)/q$. The sphere packing bound Theorem 3 says

$$\sum_{i=0}^{\rho} \binom{N}{i} (q-1)^i \geq q^n.$$

Take base 2 logarithms and divide by N. It is an often used fact in coding theory that for $\overline{\rho} \leq (q-1)/q$ the asymptotically dominating term on the left is $i = \rho$. Moreover $\binom{N}{\rho}$ goes to infinity like $2^{N \cdot H(\overline{\rho})}$. Here $H(x)$ is the **entropy function:**

$$H(x) = -xlog_2(x) - (1-x)log_2(1-x).$$

We obtain the following asymptotic bound:

Theorem 6 *Let $\alpha(x)$ be the limsup (for $N \longrightarrow \infty$) of the rate of covering functions $COV(\rho, N, n)$ having $\overline{\rho} \leq x$. Then $\alpha(x) = 1$ for $x \geq (q-1)/q$. If $x \leq (q-1)/q$, then*

$$\alpha(x) \leq H_q(x) = \frac{1}{log_2(q)} \{H(x) + log_2(q-1) \cdot x\}$$

$H_q(x)$ is known as the q-ary entropy function. In particular we see that for every covering function $COV(\rho, N, n)_q$ we must have $n/N \leq \alpha(\rho/N) \leq H_q(x)$. It follows from a probability argument based on linear codes that we have in fact equality in Theorem 6 (see [29]). This answers the question what is asymptotically possible. Remains the problem to describe covering functions of small length with a good rate.

5 Code extensions, complete caps and covering codes

Let $\mathcal{C}$ be a code $[n+1, k+1, d]_q$. It is easy to construct a code $[n, k, d]_q$ out of $\mathcal{C}$. In fact, choose a coordinate and consider the subcode of $\mathcal{C}$ consisting of the codewords that vanish at that coordinate. This subcode does the job. The other way around is much harder. Let $\mathcal{D}$ be a code $[n, k, d]_q$. If there is a code $\mathcal{C}$ with parameters $[n+1, k+1, d]_q$ such that $\mathcal{D}$ can be obtained from $\mathcal{C}$ in the way described above, then $\mathcal{C}$ is called an **extension** of $\mathcal{D}$ and we say that $\mathcal{D}$ can be **extended.**

Back to our problem. Let H be a q-ary $(n, N)-$matrix of **strength** $s+1$ (meaning that any $s+1$ columns of H are linearly independent). In particular H is a check matrix of a code $\mathcal{C}$ with parameters $[N, N-n, s+2]_q$. Assume there is some vector in $x \in \mathbb{F}_q^n$, which cannot be written as a linear combination of s columns of H. This is equivalent with the statement that the lengthening of H obtained by adding vector x still has strength $s+1$ and is a check matrix of a code $[N+1, N-n+1, s+2]$, which is an extension of $\mathcal{C}$. Observe in particular that we really have equivalence. We have seen the following:

Theorem 7 *Let $\mathcal{C}$ be a code $[n, k, d]_q$, which cannot be extended to a code $[n+1, k+1, d]_q$. Then there is a linear covering function $COV(d-2, n, n-k)_q$, in other words $l(n-k, d-2; q) \leq n$.*

This theorem provides a first link between error-correcting codes (and hence the minimum distance) and covering codes. A relatively brutal way of making sure that a certain code $[n, k, d]_q$ cannot be extended is to prove there is no code $[n+1, k+1, d]_q$ at all. We can bring the known bounds on codes in play. Assume we know for some reason that there is no code $[n+1, k+1, d]_q$. Let j be minimal such that $[n-j, k-j, d]_q$ exists. Clearly such a j exists ($j = k$, for example) and therefore also a minimal such j. We have $j \geq 0$. As code $[n-j, k-j, d]_q$ is not extendable we conclude $l(n-k, d-2; q) \leq n-j \leq n$.

Theorem 8 *If a linear code $[n+1, k+1, d]_q$ does not exist, then*

$$l(n-k, d-2; q) \leq n.$$

In case $\rho = 2$ Theorem 7 can naturally be formulated in a geometrical setting. Let H be a q-ary $(n, N)-$matrix of **strength** 3. Consider the columns of H as points of $PG(n-1, q)$ (recall that the points of the projective geometry $PG(n-1, q)$ are the 1-dimensional subspaces of $I\!F_q^n$). Strength 3 means that no three of our points are on a line (lines correspond to 2-dimensional subspaces). Sets of points no three of which are on a line are called **caps**. We see that our matrix H describes precisely a cap of n points (an n-cap) in $PG(n-1, q)$. Also, the code $\mathcal{C}$ whose check matrix is H will be extendable if and only if the cap can be embedded in an $(n+1)-$cap. Caps, which cannot be embedded in larger caps, are known as **complete caps.** This shows that Theorem 7 can be written as follows in case $s = 2$:

Theorem 9 *If there is a complete N-cap in projective geometry $PG(n-1, q)$, then*

$$l(n, 2; q) \leq N.$$

This raises the problem of determining the smallest cardinality of complete caps. Fortunately this problem has been studied by geometers for a long time.

6 Binary linear covering codes

We collect what is known about the function $l(n, \rho) = l(n, \rho; 2)$. The straightforward lower bound given in Section **??** can be marginally improved. This improvement is known as the **van Wee bound.** In the binary case it is as follows:

$$M\{\sum_{i=0}^{\rho} \binom{N}{\rho} - \frac{\binom{N}{\rho}}{\lceil (N-\rho)/(\rho+1) \rceil} (\lceil (N+1)/(\rho+1) \rceil - (N+1)/(\rho+1)\} \geq 2^N.$$

René Struik succeeded in improving on the van Wee bound [27, 28]. This led to improved lower bounds, which have been incorporated in the following table. We mention here

$$l(2m-1, 2) \geq 2^m + 1 \text{ for } m \geq 3$$

[14] contains a list of the best known upper bounds for small values of ρ and n. Whenever I know a lower bound I include it. If the value is known to equal $l(n, \rho)$ I write an equality sign.

$n\backslash\rho$	1	2	3	4	5
2	=3	=2			
3	=7	=4	=3		
4	=15	=5	=5	=4	
5	=31	=9	6	6	
6	=63	=13	7	7	
7	=127	=19	11	8	
8		25-26	14	9	
9		34-39	17-18	13	
10		53	21-22	16	
11		65-79	=23	20	15
12		92-107	31-38	24	18
13		129-159	38-53	25	20
14		182-215	63	29	26
15		257-319	75	32-37	28
16		363-431	75-95	49	27-31
17		513-639	126	44-62	
18		725-863	153	53-77	
19		1025-1279	148-205	62-84	
20		1449-1727	187-255	73-93	
21		2049-2559	235-308	86-125	51-75
22		2897-3455	295-383	103-150	88
23					98
24					76-107
25					123
26					147
27					173
28					204

7 The output

Crandall believes the range from .05 to .10 is most likely to be generally useful. Let us see how close the entries of the table bring us to the asymptotically optimal values

$$\alpha(.1) = H(.1) = 0.468995 \text{ and } \alpha(.05) = H(.05) = 0.286396$$

Here we use the trivial fact that the existence of a linear covering function $COV(\rho, N, n)$ implies the existence of a linear $COV(\rho, N', n)$ for every $N' \geq N$. Let us first consider relative covering radius .1 The first column (Hamming codes) yields a rate of .3 The second column ($\rho = 2$) gives us $COV(2, 20, 7)$ implying a rate of $7/20 = .35$ The third column improves this (via the Golay code) to $11/30$, and column $\rho = 4$ yields a rate of $15/40 = .375$ This system is already in the old table of [18]. An entry of the table in [9] shows that a linear $COV(5, 47, 19)$ exists, and hence $COV(5, 50, 19)$, of rate 0.38 Finally there is a linear $COV(6, 59, 23)$ in [9], implying $COV(6, 60, 23)$ of rate 0.383 In the case of $\overline{\rho} = 0.05$ the situation is similar. The best rate visible from the table is given by the $(\rho = 5)-$column. Covering function $COV(5, 98, 23)$ implies $COV(5, 100, 23)$, of rate .23

8 The football pool problem

The football pool problem is the problem of finding the best possible betting system. As there are three possible results for each game (win-loss-draw) the problem is ternary. If n games are involved, then σ_n is the minimum size of a code in $I\!F_3^n$ of covering radius 1. Most of thw work has indeed been done in this setting: ternary, covering radius 1. There is however a nice construction, which works over arbitrary finite fields. It was restricted to covering radius 1 in its original formulation in [5]. The obvious generalization is given in [28].

Theorem 10 *Let $M = \{m_1, m_2, \ldots, m_N\}$ be a set of N elements in $I\!F_q^n$. Let $S \subset I\!F_q^n$ be such that every element of $I\!F_q^n$ can be written as a sum of some element of S and a linear combination of at most ρ elements of M. Let $W = \{(w_1, w_2, \ldots, w_N) \in I\!F_q^N \mid \sum_{i=1}^N w_i m_i \in S\}$.*
Then $W \subset I\!F_q^N$ has covering radius $\leq \rho$. If M generates the whole space $I\!F_q^n$, then $\mid W \mid = \mid S \mid q^{N-n}$.

Proof: This is almost trivial. Let $x = (x_1, x_2, \ldots, x_N) \in I\!F_q^N$ be given. Consider $\lambda = \sum_{i=1}^N x_i m_i \in I\!F_q^n$. We can write λ as a sum of some $s \in S$ and a linear combination of most ρ elements of M. This means that we can change x in at most ρ coordinates and obtain an element of W. The statement concerning the cardinality of W is a trivial fact from linear algebra.■

Let us check what this does. Let H be the (n, N)-matrix whose columns are the elements of M. Assume also that H has maximum rank n. Let L be the set of elements in $I\!F_q^n$ which can be written as linear combinations of $\leq \rho$ columns of H. We know from Section 1 that a linear $COV(\rho, N, n)_q$ exists if and only if H can be chosen such that $L = I\!F_q^n$. In that case we obtain a linear covering code of length N and dimension $N - n$. Assume this is not the case or we cannot find the corresponding matrix. Choose H such that $\mid L \mid$ is as big as possible. Let S be such that $S + L = I\!F_q^n$. Theorem 10 gives us a covering code (same length, same covering radius) of size $\mid S \mid q^{N-n}$. We see that this is interesting only if $COV(\rho, N, n)_q$ cannot be constructed. Moreover the covering code is by construction a union of $\mid S \mid$ cosets of the code $\mathcal{C}$ whose check matrix is H. So here is what Theorem 10 says: even if $\mathcal{C}$ is not a covering code, then the union of certain cosets of $\mathcal{C}$ is. Most importantly we get a recipe which cosets to choose. Let $\mid S \mid \geq q^i$. Then the parameters of the resulting covering code are not better than those of a linear $COV(\rho, N, n-i)_q$. The construction is of interest only when this linear function cannot be constructed. Our real aim are systems (large sets), not individual covering codes. When will Theorem 10 yield a (nonlinear) system? Observe that when S satisfies the requirements of the theorem, then every coset $S + x$ also does. We need that there are a certain number of mutually exclusive cosets of S. The most natural choice is to use a linear subspace as S.

Theorem 11 *Let H be a q-ary $(n, N)-$matrix of rank n and L the set of elements from $I\!F_q^n$, which can be written as linear combinations of at most ρ columns of H. Let S be a linear subspace of dimension i such that $S+L = I\!F_q^n$ (equivalently: every coset of S contains an element of L). Then we can construct a $COV(\rho, N, n-i)_q$, equivalently a family of q^{n-i} pairwise disjoint covering codes (length N, covering radius ρ). Each member of the family is a union of cosets of the code $\mathcal{C}$ whose check matrix is H.*

This has a big disadvantage. If S is a linear space of dimension i, then the union of the cosets of the linear code whose check matrix is H with S as set of representatives is itself a linear code. In order to get a nonlinear construction we need to partition $I\!F_q^n$ into cosets of sets S, where S is not a linear subspace. The following example shows, that this can happen. We work in $I\!F_2^8$. Let S consist of the 0-vector and the seven words of weight 2,

which have a 1 in the first coordinate. We have $\mid S \mid = 8$. Every coset $S + x$ consists of 8 elements. When will two different cosets $S + x$ and $S + y$ have a vector in common? This means $s_1 + x = s_2 + y$ or $x + y = s_1 + s_2 \in S + S$. As $S + S$ consists of the 0-vector and of all vectors of weight 2 we need to choose representatives x_i such that $x_i + x_j (i \neq j)$ never has weight 2. Here are examples for such representatives:

i	x_i
1	00000000
2	11111111
3	11110000
4	00001111
5	11001100
6	00110011
7	10101010
8	01010101
9	01100110
10	10011001
11	10010110
12	01101001
13	11000011
14	00111100
15	10100101
16	01011010

The union of these cosets is the set of all codewords of even weight, the all-even code $[8, 7, 2]$.

Large sets of covering codes have appeared in the literature already. A name that appeared in the literature is **covering by coverings CBC** (see [16]).

Here is the first part of a table from [16] with bounds on the number $K(n, \rho)$, the minimum number of binary words of length n, such that the balls of radius ρ covers the space. Intervals are supplied. If only one value is given this is the minimum

$n\backslash\rho$	2	3
3	2	1
4	2	2
5	2	2
6	4	2
7	7	2
8	11-12	4
9	14-16	7
10	23-30	9-12
11	36-44	12-16
12	61-80	18-28
13	97-128	27-42
14	157-256	43-64
15	308-480	69-112
16	512-896	114-224
17	859-1536	186-352
18	1702-3056	316-640
19	2897-2^{12}	511-1024
20	5328-2^{13}	889-2^{11}
21	9893-$7 \cdot 2^{11}$	1475-2^{12}
22	17,316-$3 \cdot 2^{13}$	2536-2^{12}
23	30,667-2^{15}	2^{12}
24	60,350-2^{16}	8123-2^{13}
25	107,203-2^{17}	13,896-2^{14}
26	190,765-2^{18}	23,718-2^{15}

9 Factors and the Preparata codes

A large set of codes is a partition of the space of all tuples into codes, where all the participating codes have the same parameters. Such partitions are often used in many mathematical disciplines. A natural refinement of this idea applies to pairs of codes contained in each other, as follows:

Definition 4 *Let $\mathcal{C} \supset \mathcal{D}$ be a chain of (not necessarily linear) q-ary codes. We say that $\mathcal{C}/\mathcal{D}$ is a* **factorization** *if $\mathcal{C}$ can be written as the disjoint union*

of cosets of $\mathcal{D}$. The number of participating cosets is of course

$$[\mathcal{C} : \mathcal{D}] = \frac{\mid \mathcal{C} \mid}{\mid \mathcal{D} \mid},$$

the **index** *of $\mathcal{D}$ in $\mathcal{C}$. In most cases the index will be a power of q. We define the* **codimension** *of $\mathcal{D}$ in $\mathcal{C}$ as*

$$dim(\mathcal{C}/\mathcal{D}) = log_q([\mathcal{C} : \mathcal{D}]) \text{ in these cases}$$

We will also use the term **dimension** *in this sense.*

Observe that if $v_i, i = 1, 2, \ldots, [\mathcal{C} : \mathcal{D}]$ is a system of representatives of $\mathcal{D}$ in $\mathcal{C}$, then it is also a system of representatives with respect to $\mathcal{D} + v_j$ for any fixed j. If $\mathcal{C}$ and $\mathcal{D}$ are linear codes, then clearly $\mathcal{C}/\mathcal{D}$ will always be a factorization and the codimension has its usual meaning. For instance, the binary Hamming code $H(n)$ of length $2^n - 1$ has codimension n in the ambient space $I\!F_2^{2^n-1}$ and the extended Hamming code $\overline{H(n)}$ has codimension $n+1$ in its ambient space $I\!F_2^{2^n}$ and codimension n in the all-even code $\mathbf{1}^\perp$. The reason for not restricting the definition of a factorization to linear codes is the fact that an important class of nonlinear codes, the **Preparata codes,** lead to good factorizations. We collect the most important properties of the Preparata codes here.

Let $n \geq 4$ be even. The Preparata code $Pr(n)$ has length $2^n - 1$ and dimension $2^n - 2n$ (although it is not linear). Its minimum distance is 5. We have $Pr(n) \subset H(n)$. The number of vectors at distance 1 or 2 from $Pr(n)$ is $\mid Pr(n) \mid (2^n - 1 + \binom{2^n-1}{2}) = \mid I\!F_2^{2^n-1} \mid - \mid Pr(n) \mid$. As no word from $H(n)$ has this property it follows that we have counted every vector outside $H(n)$ precisely once. It is also known that $H(n)$ can be written as a disjoint union of cosets of $Pr(n)$. As $Pr(n)$ has minimum distance 5 and is not perfect it has $\rho \geq 3$. We have seen that equality holds.

Theorem 12 *Let $n \geq 4$ be even. We have $dim H(n) = 2^n - (n+1)$. The Preparata code $Pr(n)$ has codimension $n-1$ in $H(n)$ and $H(n)/Pr(n)$ is a factorization. Moreover every vector $x \notin H(n)$ has distance either 1 or 2 from precisely one word of $Pr(n)$. $\rho(Pr(n)) = 3$.*

The same kind of counting can be applied to the lengthened codes (after adding a parity check bit) of length 2^n.

Theorem 13 *Let $n \geq 4$ be even. We have $dim(\mathbf{1}^{\perp})/\overline{H(n)}) = n$ and $dim(\overline{H}(n)/\overline{Pr}(n)) = n-1$. Here $\overline{Pr}(n)$ has minimum distance 6 and $\overline{H}(n)$ has minimum distance 4. Moreover every vector of even weight outside $\overline{H}(n)$ is at distance 2 from precisely one word of $\overline{Pr}(n)$ and every vector outside $\overline{H}(n)$ is at distance ≤ 3 from $\overline{Pr}(n)$. $\rho(\overline{Pr}(n)) = 4$. Observe $dim(\overline{H}(n)/\overline{Pr}(n)) = n-1$.*

An important parameter of a factorization $\mathcal{C}/\mathcal{D}$ is its norm.

Definition 5 *Let $\mathcal{C}/\mathcal{D}$ be a factorization of the q-ary code $\mathcal{C}$ into cosets $\mathcal{D}_i$ of its subcode $\mathcal{D} = \mathcal{D}_1$. The* **norm** *$N(\mathcal{C}/\mathcal{D})$ is the maximum over all vectors x in the ambient space of*

$$Min_i\{d(x, \mathcal{D}_i)\} + Max_j\{d(x, \mathcal{D}_j)\}$$

The norm will be used in the blockwise direct sum construction of covering codes. Consider the special case when $\mathcal{C} = U$ is the ambient space. The factorization is then a large set consisting of cosets of $\mathcal{D}$ and we have $N(U/\mathcal{D}) = \rho(\mathcal{D})$. An obvious general bound is $N(\mathcal{C}/\mathcal{D}) \leq \rho(\mathcal{C}) + \rho(\mathcal{D}$.

Theorem 14

$$N(I\!F_q^{(q^n-1)/(q-1)}/H(n)) = 1.$$

$$N(I\!F_2^{2^n}/\overline{H}(n)) = 2$$

Let $n \geq 4$ be even. Then

$$N(H(n)/Pr(n)) = 3 \text{ and } N(\overline{H}(n)/\overline{Pr}(n)) = 4.$$

Proof: The first two statements are clear as $H(n)$ has covering radius 1. Let $q = 2, x \notin H(n)$. The minimum in Definition 5 is 1, the maximum is 2 as we see from Theorem 12. Let $x \in H(n)$. This time the minimum is 0, the maximum is 3.■

10 The blockwise direct sum

We start from construction X4 for linear codes.

Theorem 15 (blockwise direct sum, linear case) *Let $\mathcal{C}_1 \supset \mathcal{D}_1$ be codes $[n_1, k_1, d_1] \supset [n_1, l_1, D_1]$ and $\mathcal{C}_2 \supset \mathcal{D}_2$ codes with parameters $[n_2, k_2, d_2] \supset [n_2, l_2, D_2]$, where $k_1 - l_1 = k_2 - l_2 = \kappa$. Fix complements U_i of $\mathcal{D}_i$ in $\mathcal{C}_i$ and an isomorphism $\alpha : U_1 \longrightarrow U_2$. We construct a code $\mathcal{C}$ of length $n_1 + n_2$ and dimension $l_1 + l_2 + \kappa$ as the image of*

$$\phi : \mathcal{D}_1 \oplus \mathcal{D}_2 \oplus U_1 \longrightarrow I\!F_q^{n_1+n_2}.$$

Here ϕ is defined by $\phi(x, 0, 0) = (x, 0^{n_2}), \phi(0, y, 0) = (0^{n_1}, y)$ and $\phi(0, 0, z) = (z, \alpha(z))$. Then $\mathcal{C} = \phi(\mathcal{D}_1 \oplus \mathcal{D}_2 \oplus U_1)$ has minimum distance $\geq min\{D_1, D_2, d_1 + d_2\}$.

Proof: As ϕ is an injective mapping the dimension is obvious. In the determination of the minimum distance we distinguish two cases: if $z = 0$, then $wt(\phi(x, y, 0)) = wt(x) + wt(y) \geq min\{D_1, D_2\}$ provided $(x, y) \neq (0, 0)$. If $z \neq 0$, then $wt(\phi(x, y, z)) = wt(x + z) + wt(y + \alpha(z)) \geq d_1 + d_2$.■
What does this mean from a combinatorial point of view? The mapping α is a bijection between the cosets of $\mathcal{D}_1$ in $\mathcal{C}_1$ and the cosets of $\mathcal{D}_2$ in $\mathcal{C}_2$. Pair (a, b) is in the blockwise direct sum if and only if $a \in \mathcal{C}_1, b \in \mathcal{C}_2$ and $\alpha(a+\mathcal{D}_1) = b+\mathcal{D}_2$. We see how this can be generalized to cover also nonlinear codes:

Definition 6 (blockwise direct sum) *Let $\mathcal{C}_1/\mathcal{D}_1$ (q-ary, of length n_1) and $\mathcal{C}_2/\mathcal{D}_2$ (q-ary, of length n_2) be factorizations with the same index $[\mathcal{C}_1 : \mathcal{D}_1] = [\mathcal{C}_2 : \mathcal{D}_2]$. Choose a bijection between the cosets of the two factorizations. The blockwise direct sum has length $n_1 + n_2$. It is defined by*

$$\mathcal{C} = \cup_{r=1}^{k} \mathcal{D}_1(r) \times \mathcal{D}_2(r),$$

where $\mathcal{D}_1(r)$ and $\mathcal{D}_2(r)$ are the cosets of the factorizations.

The number of codewords is

$$| \mathcal{C} | = | \mathcal{C}_1 | \cdot | \mathcal{D}_2 | = | \mathcal{C}_2 | \cdot | \mathcal{D}_1 |$$

Theorem 16 *Let $\mathcal{C}$ be the blockwise direct sum of two factorizations with identical index, as in Definition 6. The minimum distance of $\mathcal{C}$ is $min\{d(\mathcal{D}_1), d(\mathcal{D}_2), d(\mathcal{C}_1)+ d(\mathcal{C}_2)\}$. The covering radius of $\mathcal{C}$ satisfies*

$$\rho(\mathcal{C}) \leq \lfloor (N(\mathcal{C}_1/\mathcal{D}_1) + N(\mathcal{C}_2/\mathcal{D}_2))/2 \rfloor.$$

Proof: The minimum distance is obvious. Let $(x, y) \in \mathcal{C}$, where $x \in \mathcal{D}_1(i), y \in \mathcal{D}_2(i)$. Let $N_j = N(\mathcal{C}_j/\mathcal{D}_j), j = 1, 2$. Choose j, k such that $d(x, \mathcal{D}_1(j))$ and $d(y, \mathcal{D}_2(k))$ are minimal. It follows from the definition of the norm that the sum of the distances from (x, y) to $\mathcal{D}_1(j) \times \mathcal{D}_2(j)$ and to $\mathcal{D}_1(k) \times \mathcal{D}_2(k)$ is it most $N_1 + N_2$. One of the two distances must be $\leq (N_1 + N_2)/2$.■

Theorem 17 *Let $\mathcal{C}/\mathcal{D}$ a factorization of binary codes, with norm $N(\mathcal{C}/\mathcal{D}) = N$. Then the lengthened codes $\overline{\mathcal{C}}/\overline{\mathcal{D}}$ also form a factorization. Its norm is the even number among $\{N + 1, N + 2\}$.*

Proof : The first statement is clear. In fact, the indices are the same. The norm of the lengthened codes is even. This follows from Definition 5. The norm is defined as a sum of two numbers. If x is in the all even code, then both these numbers are even, in the contrary case both are odd. If x is in the ambient space of $\mathcal{C}$, and v in the ambient space of the lengthened code, then $d((x, 0), v) + d((x, 1), v) = 2d(x, v) + 1$. It follows that either $d((x, 0), v) = d(x, v) + 1$ or $d((x, 1), v) = d(x, v) + 1$. It is clear that the norm is bigger then N, but at most $N + 2$.■

As we are primarily interested in large sets the following result is important:

Theorem 18 *Let $\mathcal{C}_1/\mathcal{D}_1$ and $\mathcal{C}_2/\mathcal{D}_2$ be factorizations (lengths n_1 and n_2, respectively) of identical index, so that we can form the blockwise direct sum $\mathcal{C} = \mathcal{C}_1/\mathcal{D}_1 \times \mathcal{C}_2/\mathcal{D}_2$. Then the following hold:*

1. *$\mathcal{C}_1 \times \mathcal{C}_2/\mathcal{C}$ is a factorization, with index $[\mathcal{C}_i : \mathcal{D}_i], i = 1, 2$ and norm $N \leq \rho(\mathcal{C}_1) + \rho(\mathcal{C}_2) + \rho(\mathcal{C})$.*
2. *If $\mathbb{F}_q^{n_1}/\mathcal{C}_1$ and $\mathbb{F}_q^{n_2}/\mathcal{C}_2$ both are factorizations, then also $\mathbb{F}_q^{n_1+n_2}/\mathcal{C}$ is a factorization.*

Proof: Statement 1. is easy to prove. Let w_k be a system of representatives of $\mathcal{D}_2$ in $\mathcal{C}_2$. As the numbers are right it suffices to prove that the corresponding cosets of $\mathcal{C}_1 \times \mathcal{C}_2$ in $\mathcal{C}$ are disjoint. So assume $w_k + (x_1, y_1) = w_{k'} + (x_2, y_2)$, where $(x_1, y_1) \in D_1(a) \times D_2(a)$ and $(x_2, y_2) \in D_1(b) \times D_2(b)$.

The second coordinate shows $k = k'$. The statement concerning the norm is obvious.

Consider 2. By assumption we have representatives u_i, v_j, w_k such that the $u_i + \mathcal{C}_1$ partition $I\!F_q^{n_1}$, the $v_j + \mathcal{C}_1$ partition $I\!F_q^{n_2}$ and the $w_k + \mathcal{D}_2$ partition $\mathcal{C}_2$. As the numbers are right it remains to prove that the cosets $\mathcal{C} + (u_i, v_j + w_k)$ are pairwise disjoint. Assume there are elements $(x_1, y_1) \in D_1(a) \times D_2(a)$ and $(x_2, y_2) \in D_1(b) \times D_2(b)$ such that

$$(x_1, y_1) + (u_i, v_j + w_k) = (x_2, y_2) + (u_{i'}, v_{j'} + w_{k'}).$$

The first coordinate shows $i = i'$, whence $a = b$. The second coordinate shows $y_1 + v_j + w_k = y_2 + v_{j'} + w_{k'}$. As $y_1, y_2 \in D_2(a)$ and $D_2(a)$ is a coset of D_2 we can find $z_1, z_2 \in D_2$ such that $z_1 + v_j + w_k = z_2 + v_{j'} + w_{k'}$. It follows from our assumptions that $j = j', k = k'$.■

Armed with these basic structural facts it is easy now to apply the blockwise direct sum in a variety of situations. Here is a list of factorizations. The basic parameters are readily deduced from what has been said before. We write U for the ambient space. Its dimension is the length. d is the minimum distance of the larger of the codes, D the minmum distance of the smaller code. Observe that all members of the factorization are cosets of $\mathcal{D}$ and therefore have the same distance D. In column *dim* we note the dimension of the larger code. $\mathbf{1}^\perp$ is the all-even code.

factorization	length	dim	codim	(d, D)	norm	condition
$U/H(n)$	$\frac{q^n-1}{q-1}$	$\frac{q^n-1}{q-1}$	n	(1,3)	1	
$U/\overline{H}(n)$	2^n	2^n	$n+1$	(1,4)	2	$q = 2$
$H(n)/Pr(n)$	$2^n - 1$	$2^n - 1 - n$	$n-1$	(3,5)	3	$q = 2, n$ even
$\overline{H}(n)/\overline{Pr}(n)$	2^n	$2^n - 1 - n$	$n-1$	(4,6)	4	$q = 2, n$ even
$\mathbf{1}^\perp/\overline{Pr}(n)$	2^n	$2^n - 1$	$2n-1$	(2,6)	4	$q = 2, n$ even
$\mathbf{1}^\perp/\overline{H}(n)$	2^n	$2^n - 1$	n	(2,4)	2	$q = 2$

For example, consider the factorization $H(n)/Pr(n)$. We know the length, the dimensions and the minimum distances $d = 3, D = 5$. Let $x \notin H(n)$. As $H(n)$ is a perfect code of covering radius 1 it follows that $d(x, H(n)) = 1$. It was noted in Theorem 12 that $d(x, Pr(n)) \leq 2$. Let $x \in H(n)$. Clearly the

first summand, the minimum of Definition 5, is 0, and the maximum is ≤ 3. We conclude that the norm is 3.

All we need to do is to use two factorizations with the same index and apply the blockwise direct sum. Minimum distance and covering radius of the resulting code $\mathcal{C}$ can be read off from Theorem 16. If the corresponding condition of Theorem 18 is met we obtain a large set and hence a covering function. We collect some binary examples in the following table:

first fact.	length	second fact.	length	d	ρ	COV	cond.
$H(n)/Pr(n)$	2^n-1	$\mathbf{1}^\perp/\overline{Pr}(n/2)$	$2^{n/2}$	5	3	$(3, 2^n+2^{n/2}-1, 2n+1)$	$4 \mid n$
$H(n)/Pr(n)$	2^n-1	$\mathbf{1}^\perp/\overline{H}(n-1)$	2^{n-1}	4	2	$(2, 3\cdot 2^{n-1}-1, 2n)$	$2 \mid n$
$H(n)/Pr(n)$	2^n-1	$\overline{H}(n)/\overline{Pr}(n)$	2^n	5	3	$(3, 2^{n+1}-1, 3n)$	$2 \mid n$

An example from the table is a (non-linear) $COV(3, 31, 12)$. Its relative covering radius is < 0.1 and it reaches a rate of $12/31 \sim 0.387$ This is better than the linear functions from Section 7 and the length is only 31.

11 Covering codes of large radius

Rodemich [24] studies covering codes of large radius. His main result is that a q-ary covering code of length n and radius $n-2$ has $\geq q^2/(n-1)$ words, and he characterizes the case of equality in terms of OA of index unity and strength 2 (hence mutually orthogonal latin squares). Baartmans-Sane [1] give a more accessible proof for a more complete result in this case. The cases of equality are again characterized by the existence of mutually orthogonal latin squares. They start from a somewhat different motivation: a q-ary combination lock with n coordinates. The secret number therefore is a q-ary n-tuple. The lock is defective. It is possible to open it when the word agrees with the secret key in just 2 positions. It will suffice to try out the elements of a covering code to open the lock.

The construction is easy to understand, whereas the bound requires some work: Write $q = l(n-1)+r, 0 \leq r < n-1$ and let $Q = Q_1 \cup \ldots \cup Q_{n-1}$ be a partition of the alphabet such that $|Q_i| \in \{l, l+1\}$ (if $r = 0$ we choose $|Q_i| = l$ for all i). Assume an $OA_1(n, l)$ and an $OA_1(2, n, l+1)$ exist, equivalently $n-2$ MOLs of orders l and $l+1$ (the latter is not needed when $r = 0$). Let

O_i be such an OA with Q_i as alphabet. Then the union of the rows of all the $O_i, i = 1, \ldots, n-1$ form a covering code of radius $n-2$.

The proof is trivial: given an aribtrary n-tuple $x = (x_1, \ldots, x_n)$. There must be two coordinates, say 1 and 2, such that $x_1, x_2 \in Q_i$ for the same i. As O_i is an OA, x agrees with a row from O_i in coordinates 1 and 2.

References

[1] A. Baartmans and S.Sane: *Combination locks and orthogonal arrays,* manuscript.

[2] J.Bierbrauer and Y.Edel: *Lengthening and the Gilbert-Varshamov bound, IEEE Transactions on Information Theory* **43** (1997),991-992.

[3] J.Bierbrauer, K.Gopalakrishnan,D.R.Stinson : *Orthogonal Arrays, Resilient Functions, Error Correcting Codes and Linear Programming Bounds, SIAM Journal on Discrete Mathematics* **6** (1996),424-452.

[4] Bhandari and Duraijaran: *A note on bounds for q-ary covering codes, IEEE Transactions on Information Theory* **42** (1996),1640-1642.

[5] A.Blokhuis and C.Lam: *More coverings by rook domains, Journal of Combinatorial Theory A* **36** (1984),240-244.

[6] Brualdi and V.Pless: *On the length of codes with a given covering radius, Coding Theory and Design Theory I,* New York, Springer 1990.

[7] Brualdi,V.Pless, R.M.Wilson: *Short codes with a given covering radius, IEEE Transactions on Information Theory* **35** (1989),99-109.

[8] J.C.Cock and P.Östergård: *Ternary covering codes derived from BCH-codes, Journal of Combinatorial Theory A* **43** (1997),283-289.

[9] G.Cohen,I.Honkala,S.Litsyn,A.Lobstein: *Covering Codes,* North Holland, Amsterdam 1997. ISBN 0-444-82511-8.

[10] Cohen, Honkala, Litsyn, Solé: *Long packing and covering codes, IEEE Transactions on Information Theory* **43** (1997),1617-1619.

[11] G. Cohen, Karpovsky, Mattson, Schatz: *Covering radius - survey and recent results, IEEE Transactions on Information Theory* **31** (1985),328-343.

[12] Davydov: *Construction of linear covering codes, Problems in Information Transmission* **26** (1990),317-331.

[13] A.A.Davydov: *Construction of nonlinear covering codes, IEEE Transactions on Information Theory* **43** (1997),1639-1647.

[14] A.A.Davydov and A.Labinskaya: *Constructions, families and tables of binary linear covering codes, IEEE Transactions on Information Theory* **40** (1994),1270-1279.

[15] Dougherty and Janwa: *Covering radius computations for binary cyclic codes, Math Comput* **57** (1991),415-434.

[16] T.Etzion and Greenberg: *Constructions for perfect mixed codes and other covering codes, IEEE Transactions on Information Theory* **39** (1993),209-214.

[17] Gabidulin, Davydov, Tombak: *Codes with covering radius 2 and other new covering codes, IEEE Transactions on Information Theory* **37** (1991),219-224.

[18] Graham and N.J.A. Sloane: *On the covering radius of codes, IEEE Transactions on Information Theory* **31** (1985),385-401.

[19] X.Hou: *On the norm and covering radius of the first-order Reed-Muller codes, IEEE Transactions on Information Theory* **43** (1997),1025-1027.

[20] A.Klapper: *The multicovering radii of codes, IEEE Transactions on Information Theory* **43** (1997),1372-1377.

[21] Li and Chen: *New lower bounds for binary covering codes, IEEE Transactions on Information Theory* **40** (1994),1122-1129.

[22] F.Levy-dit-Vehel and S.Litsyn: *More on the covering radius of BCH-codes, IEEE Transactions on Information Theory* **42** (1996),1023-1028.

[23] F. J. MacWilliams, N. J. A. Sloane. *The Theory of Error-Correcting Codes,* North-Holland, 1977.

[24] E.R.Rodemich: *Coverings by Rook domains, Journal of Combinatorial Theory A* **9**(1970), 117-128.

[25] N.J.A.Sloane, S.M.Reddy, C.L.Chen: *New binary codes, IEEE Transactions on Information Theory* **18** (1972),503-510.

[26] D.R.Stinson: *Resilient functions and large sets of orthogonal arrays, Congressus Numerantium* **92**(1993),105-110.

[27] R.Struik: *An improvement of the van Wee bound for binary linear covering codes, IEEE Transactions on Information Theory* **40** (1994),1280-1284.

[28] R.Struik: *On the structure of linear codes with covering radius 2 and 3, IEEE Transactions on Information Theory* **40** (1994),1406-1416.

[29] R.Struik: *Covering Codes,* Ph.D. dissertation, Eindhoven 1994.

[30] Tsai: *The covering radius of extremal self-dual code D11 and its application, IEEE Transactions on Information Theory* **43** (1997),316-319.

[31] Wille: *New binary covering codes obtained by simulated annealing, IEEE Transactions on Information Theory* **42** (1996),300-302.