Error Correcting Codes for the Asymmetric Channel

Torleiv Kløve,
Department of Informatics,
University of Bergen
HIB,
N-5020 Bergen,
Norway

This report was first made in 1981 and revised in 1983.
The bibliography was updated in 1995.

Contents

Chapter 1

Introduction

In many binary communication systems, the probabilities of the crossovers $1 \to 0$ and $0 \to 1$ are approximately the same, and the systems are well modeled by the binary symmetric channel (BSC). Error correcting codes for BSCs have been studied extensively, see e.g. [61].

In other communication systems, the probability of a $1 \to 0$ crossover is much larger than the probability of a $0 \to 1$ crossover. This applies, for instance, to some data storing systems, see Constantin and Rao [14] and optical communication, see McEliece and Rodemich [64]. Neglecting the low probability $0 \to 1$ crossover, the communication system is modeled by the Z-channel. Error correcting codes for the Z-channel have been much less studied than the codes for the BSC.

In the following notes, I give a unified account of error correcting codes for the Z-channel. The notes are based on lectures given at the University of Bergen, in the autumn term 1980, and more resent results (up to 1983) have been included. At the end of each chapter, I have given references for the results of that chapter.

Chapter 2

Definitions and basic results

2.1 The Z-channel

Definition 2.1 *The* ***binary*** *(completely)* ***asymmetric channel*** *(the* ***Z-channel****) is the channel with* $\{0,1\}$ *as input and output alphabets, where the crossover* $1 \to 0$ *occurs with positive probability* p*, whereas the crossover* $0 \to 1$ *never occurs, cfr. Fig 2.1.*

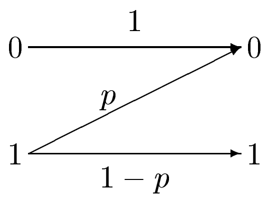

Figure 2.1: The binary asymmetric channel

Interchanging the role of "0" and "1" (*complemenatation*) we get a "complementary Z-channel". Any code for the Z-channel will by complementation give a code with the same properties for the complementary channel. However, it turns out that a code for the Z-channel will be a code with the same error correcting capabilities for the complementary Z-channel also without complementation.

2.2 Codes

Definition 2.2 *A **code** of **length** n is a subset of* $\{0,1\}^n$.

Definition 2.3 *A code C is a t-**code** (i.e. t asymmetric error correcting code) if it can correct up to t errors, that is, there exists a rule (a decoder) such that if $\mathbf{x} \in C$ and $\mathbf{v}$ is obtained from $\mathbf{x}$ by changing at most t 1s in $\mathbf{x}$ in 0s, then the rule will recover $\mathbf{x}$ from $\mathbf{v}$.*

Definition 2.4 *The set of all t-codes of length n will be denoted by $\mathcal{A}(n,t)$.*

Definition 2.5 *The maximal size of a t-code of length n will be denoted by $\alpha(n,t)$.*

2.3 Asymmetric distance

Definition 2.6 *For $\mathbf{x} = (x_1, x_2, \ldots, x_n), \mathbf{y} = (y_1, y_2, \ldots, y_n) \in \{0,1\}^n$ let*

$$\begin{array}{rl}
(i) & N(\mathbf{x},\mathbf{y}) := \#\{i \mid x_i = 0 \textit{ and } y_i = 1\}, \\
(ii) & \Delta(\mathbf{x},\mathbf{y}) := \max\{N(\mathbf{x},\mathbf{y}), N(\mathbf{y},\mathbf{x})\}, \\
(iii) & d(\mathbf{x},\mathbf{y}) := N(\mathbf{x},\mathbf{y}) + N(\mathbf{y},\mathbf{x}), \\
(iv) & \mathbf{x} \leq \mathbf{y} \textit{ if and only if } N(\mathbf{y},\mathbf{x}) = 0.
\end{array}$$

Here and in the following $\#X$ denotes the cardinality of the set X.

Both Δ and d are metrics on $\{0,1\}^n$, we leave the easy verification to the reader. d is the **Hamming metric** and Δ the **asymmetric metric** (the name may be confusing, the metric is of course symmetric, i.e. $\Delta(\mathbf{x},\mathbf{y}) = \Delta(\mathbf{y},\mathbf{x})$. "Asymmetric" refers to the metric's importance in the study of codes for the asymmetric channel).

Definition 2.7 *For $\mathbf{x} \in \{0,1\}^n$ let*

$$w(\mathbf{x}) := \#\{i \mid x_i = 1\}.$$

$w(\mathbf{x})$ is known as the **(Hamming) weight** of $\mathbf{x}$. Note that

$$w(\mathbf{x}) = d(\mathbf{x},\mathbf{0}) = \Delta(\mathbf{x},\mathbf{0}).$$

Here and in the following we use the notations

$$\mathbf{0} = (0,0,\ldots,0) \text{ and } \mathbf{1} = (1,1,\ldots,1).$$

Sometimes we find it convenient to illustrate parts of proofs by figures. This will be done as in Figure 2.2. The figure has the following interpretation: $\mathbf{x}$ and $\mathbf{y}$ are binary vectors of the same length, say n. There are a positions i such that $x_i = y_i = 0$ (these positions need not be adjacent), b positions i such that $x_i = 0$ and $y_i = 1$, etc. Hence $b = N(\mathbf{x}, \mathbf{y})$, $c = N(\mathbf{y}, \mathbf{x})$, $b + c = d(\mathbf{x}, \mathbf{y})$, $c + d = w(\mathbf{x})$, $a + b + c + d = n$, etc.

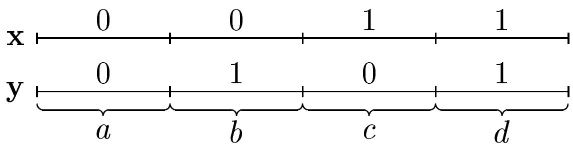

Figure 2.2:

The two metrics d and Δ are related as shown by the following lemma.

Lemma 2.1 *For* $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$ *we have*

$$2\Delta(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y}) + |w(\mathbf{x}) - w(\mathbf{y})|.$$

Proof: First we note that (cfr. Figure 2.3)

$$N(\mathbf{x}, \mathbf{y}) + w(\mathbf{x}) = N(\mathbf{y}, \mathbf{x}) + w(\mathbf{y}).$$

By symmetry, we may assume that $w(\mathbf{x}) \geq w(\mathbf{y})$. Then $N(\mathbf{x}, \mathbf{y}) \leq N(\mathbf{y}, \mathbf{x})$.

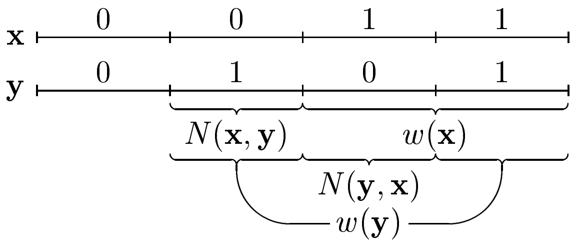

Figure 2.3:

Hence

$$\begin{aligned} 2\Delta(\mathbf{x}, \mathbf{y}) &= 2N(\mathbf{y}, \mathbf{x}) \\ &= N(\mathbf{x}, \mathbf{y}) + N(\mathbf{y}, \mathbf{x}) + (N(\mathbf{y}, \mathbf{x}) - N(\mathbf{x}, \mathbf{y})) \\ &= d(\mathbf{x}, \mathbf{y}) + w(\mathbf{x}) - w(\mathbf{y}). \end{aligned}$$

Definition 2.8 *For* $\mathbf{x} \in \{0,1\}^n$*, let*

$$S_t(\mathbf{x}) = \left\{ \mathbf{v} \in \{0,1\}^n \;\middle|\; \mathbf{v} \leq \mathbf{x} \textit{ and } N(\mathbf{v},\mathbf{x}) \leq t \right\}.$$

$S_t(\mathbf{x})$ is the set of vectors obtained by changing t or less 1s in $\mathbf{x}$ into 0s.

Lemma 2.2 $C \in \mathcal{A}(n,t)$ *if and only if* $S_t(\mathbf{x}) \cap S_t(\mathbf{y}) = \emptyset$ *for all* $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$.

Proof: If $S_t(\mathbf{x}) \cap S_t(\mathbf{y}) = \emptyset$ for all $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$, then a decoding rule is to decode $\mathbf{v}$ into the unique $\mathbf{x}$ such that $\mathbf{v} \in S_t(\mathbf{x})$. Hence $C \in \mathcal{A}(n,t)$. On the other hand, if $\mathbf{v} \in S_t(\mathbf{x}) \cap S_t(\mathbf{y})$ for some $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$, then there is no way to tell if $\mathbf{v}$ is obtained from $\mathbf{x}$ or from $\mathbf{y}$. Hence, $C \notin \mathcal{A}(n,t)$.

Theorem 2.1 $C \in \mathcal{A}(n,t)$ *if and only if* $\Delta(\mathbf{x},\mathbf{y}) > t$ *for all* $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$.

Proof: If $C \in \mathcal{A}(n,t)$ and $\mathbf{x} \in C$, let $\mathbf{u}$ be any vector such that $\mathbf{u} \neq \mathbf{x}$ and $\Delta(\mathbf{u},\mathbf{x}) \leq t$. Define $\mathbf{v}$ by

$$v_i = \begin{cases} 1 & \text{if } x_i = u_i = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbf{v} \in S_t(\mathbf{x}) \cap S_t(\mathbf{u})$ (cfr. Figure 2.4). By Lemma 2.2, $\mathbf{u} \notin C$. Hence $\Delta(\mathbf{x},\mathbf{y}) > t$ for all $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$. On the other hand, if $C \notin \mathcal{A}(n,t)$,

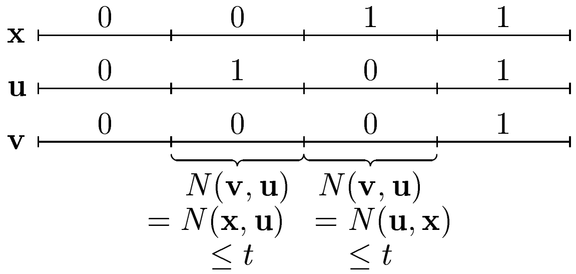

Figure 2.4:

then there exist $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$, and a $\mathbf{v} \in S_t(\mathbf{x}) \cap S_t(\mathbf{y})$. Then (cfr. Figure 2.5) $N(\mathbf{x},\mathbf{y}) \leq N(\mathbf{v},\mathbf{y}) \leq t$ and $N(\mathbf{y},\mathbf{x}) \leq N(\mathbf{v},\mathbf{x}) \leq t$. Hence $\Delta(\mathbf{x},\mathbf{y}) \leq t$. $\square$

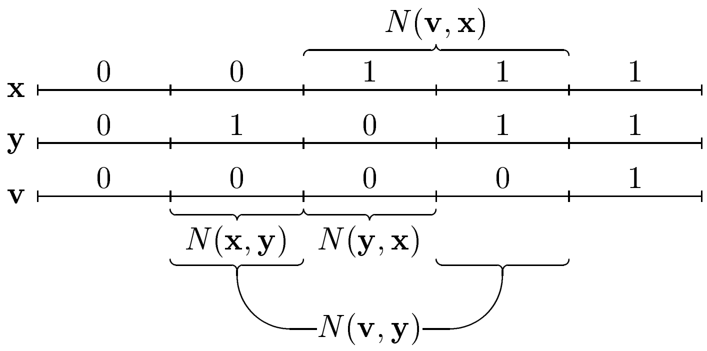

Figure 2.5:

If we use Lemma 2.1, we see that C is a t-code if and only if $d(\mathbf{x},\mathbf{y}) + |w(\mathbf{x}) - w(\mathbf{y})| > 2t$ for all $\mathbf{x},\mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$. This may be compared with the fact that a code C corrects t errors on the binary symmetric channel if and only if $d(\mathbf{x},\mathbf{y}) > 2t$ for all $\mathbf{x},\mathbf{y} \in C$.

Finally, we prove another lemma which generalizes Lemma 2.2 and which will be applied in the next chapter.

Definition 2.9 *For $s \geq 0$, $s' \geq 0$, and $\mathbf{x} \in \{0,1\}^n$, let*

$$\begin{aligned} S_{s',s}(\mathbf{x}) &= \left\{\mathbf{v} \in \{0,1\}^n \;\middle|\; \mathbf{v} \leq \mathbf{x} \textit{ and } N(\mathbf{v},\mathbf{x}) \leq s\right\} \\ &\quad \cup \left\{\mathbf{v} \in \{0,1\}^n \;\middle|\; \mathbf{x} \leq \mathbf{v} \textit{ and } N(\mathbf{x},\mathbf{v}) \leq s'\right\}. \end{aligned}$$

Note that $S_t(\mathbf{x}) = S_{0,t}(\mathbf{x})$.

Lemma 2.3 *Let $C \in \mathcal{A}(n,t)$ and $0 \leq s \leq t$. If $\mathbf{x},\mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$, then $S_{t-s,s}(\mathbf{x}) \cap S_{t-s,s}(\mathbf{y}) = \emptyset$.*

Proof: Suppose that $\mathbf{v} \in S_{t-s,s}(\mathbf{x}) \cap S_{t-s,s}(\mathbf{y})$. Without loss of generality, we may assume that $\mathbf{y} \leq \mathbf{x}$, i.e. $N(\mathbf{x},\mathbf{y}) = 0$, and so $\Delta(\mathbf{x},\mathbf{y}) = N(\mathbf{y},\mathbf{x})$. We consider three cases, which also are illustrated by figures.

Case I, $\mathbf{v} \leq \mathbf{y} \leq \mathbf{x}$. Then

$$N(\mathbf{y},\mathbf{x}) = N(\mathbf{v},\mathbf{x}) - N(\mathbf{v},\mathbf{y}) \leq N(\mathbf{v},\mathbf{x}) \leq s \leq t.$$

Hence $\Delta(\mathbf{x},\mathbf{y}) \leq t$, contradicting Theorem 2.1.

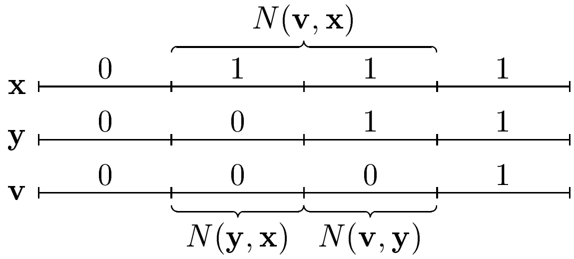

Figure 2.6: Case I, $\mathbf{v} \leq \mathbf{y} \leq \mathbf{x}$.

Case II, $\mathbf{y} \leq \mathbf{v} \leq \mathbf{x}$. In this case

$$\Delta(\mathbf{x}, \mathbf{y}) = N(\mathbf{y}, \mathbf{x}) = N(\mathbf{v}, \mathbf{x}) + N(\mathbf{y}, \mathbf{v}) \leq s + (t - s) = t,$$

again contradicting Theorem 2.1.

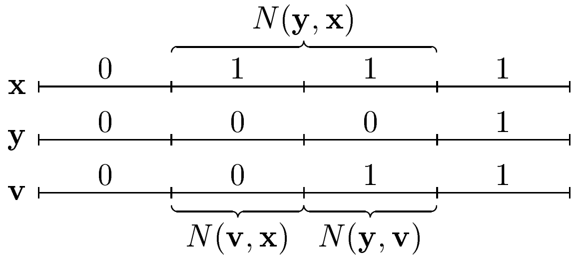

Figure 2.7: Case II, $\mathbf{y} \leq \mathbf{v} \leq \mathbf{x}$.

Case III, $\mathbf{y} \leq \mathbf{x} \leq \mathbf{v}$. In this case

$$\Delta(\mathbf{x}, \mathbf{y}) = N(\mathbf{y}, \mathbf{x}) \leq N(\mathbf{y}, \mathbf{v}) \leq t - s \leq t,$$

again contradicting Theorem 2.1.

2.4 Notes

2.1. The capacity of the Z-channel was determined by Silverman [74].

2.3. Our definition of asymmetric metric is due to Rao and Chawla [69]. Varshamov [82] introduced a metric $\overline{\rho}$ defined by

$$\overline{\rho}(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y}) + |w(\mathbf{x}) - w(\mathbf{y})|.$$

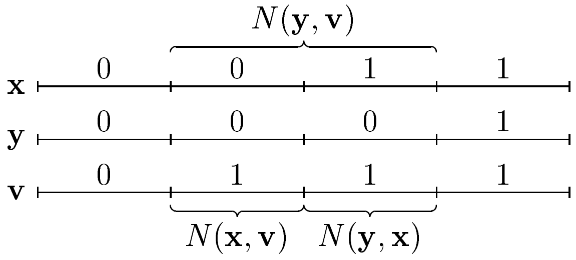

Figure 2.8: Case III, $\mathbf{y} \leq \mathbf{x} \leq \mathbf{v}$.

By Lemma 2.1, $\overline{\rho}(\mathbf{x},\mathbf{y}) = 2\Delta(\mathbf{x},\mathbf{y})$.

Theorem 2.1 is essentially due to Kim and Freiman [50]. They proved that $C \in \mathcal{A}(n, y)$ if and only if for all $\mathbf{x},\mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$ we have

$$\begin{array}{rl} (i) & |w(\mathbf{x}) - w(\mathbf{y})| \geq t+1 \\ \text{or } (ii) & |w(\mathbf{x}) - w(\mathbf{y})| \leq t \text{ and } d(\mathbf{x},\mathbf{y}) \geq 2(t+1) - |w(\mathbf{x}) - w(\mathbf{y})|. \end{array}$$

Varshamov [82] formulated the same result in terms of the metric $\overline{\rho}$.

Lemma 2.3 is due to Delsarte and Piret [18].

Chapter 3

Upper bounds

3.1 The Varshamov bound

In this chapter we give some upper bounds on $\alpha(n,t)$. We first give a bound due to Varshamov which is easy to formulate and prove. We use the following notations.

Definition 3.1 *For* $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ *let*

$$\overline{\mathbf{x}} := (1 - x_1, 1 - x_2, \ldots, 1 - x_n).$$

Definition 3.2 *For* $C \subseteq \{0,1\}^n$ *let*

$$\overline{C} := \{\overline{\mathbf{x}} \mid \mathbf{x} \in C\}.$$

Definition 3.3 *For* $C \subseteq \{0,1\}^n$ *let*

$$C_r := \#\{\mathbf{x} \in C \mid w(\mathbf{x}) = r\}.$$

Lemma 3.1 *If* $C \in \mathcal{A}(n,t)$*, then* $\overline{C} \in \mathcal{A}(n,t)$*.*

Proof: We have

$$\begin{aligned} N(\overline{\mathbf{x}}, \overline{\mathbf{y}}) &= \#\{i \mid 1 - x_i = 0 \text{ and } 1 - y_i = 1\} \\ &= \#\{i \mid x_i = 1 \text{ and } y_i = 0\} = N(\mathbf{y}, \mathbf{x}). \end{aligned}$$

Hence $\Delta(\overline{\mathbf{x}}, \overline{\mathbf{y}}) = \max\{N(\overline{\mathbf{x}}, \overline{\mathbf{y}}), N(\overline{\mathbf{y}}, \overline{\mathbf{x}})\} = \Delta(\mathbf{y}, \mathbf{x})$. By Theorem 2.1

$$\begin{aligned} C \in \mathcal{A}(n,t) &\Rightarrow \Delta(\mathbf{y}, \mathbf{x}) > t \text{ for all } \mathbf{y}, \mathbf{x} \in C,\ \mathbf{y} \neq \mathbf{x} \\ &\Rightarrow \Delta(\overline{\mathbf{x}}, \overline{\mathbf{y}}) > t \text{ for all } \overline{\mathbf{x}}, \overline{\mathbf{y}} \in \overline{C},\ \overline{\mathbf{x}} \neq \overline{\mathbf{y}} \\ &\Rightarrow \overline{C} \in \mathcal{A}(n,t). \end{aligned}$$

Theorem 3.1 *For $n \geq 1$ and $t \geq 1$ we have*

$$\alpha(n,t) \leq \frac{2^{n+1}}{\sum_{j=0}^{t} \left\{ \binom{\lfloor n/2 \rfloor}{j} + \binom{\lceil n/2 \rceil}{j} \right\}}.$$

Proof: Let $C \in \mathcal{A}(n,t)$. By Lemma 2.2, $\bigcup_{\mathbf{x}\in C} S_t(\mathbf{x})$ is a disjoint union. Further, if $w(\mathbf{x}) = r$, then $\#S_t(\mathbf{x}) = \sum_{j=0}^{t} \binom{r}{j}$. Hence

$$2^n \geq \# \bigcup_{\mathbf{x}\in C} S_t(\mathbf{x}) = \sum_{\mathbf{x}\in C} \#S_t(\mathbf{x}) = \sum_{r=0}^{n} C_r \sum_{j=0}^{t} \binom{r}{j}.$$

By Lemma 3.1, $\overline{C} \in \mathcal{A}(n,t)$. Since $\overline{C}_r = C_{n-r}$, we similarly get

$$2^n \geq \sum_{r=0}^{n} C_{n-r} \sum_{j=0}^{t} \binom{r}{j} = \sum_{r=0}^{n} C_r \sum_{j=0}^{t} \binom{n-r}{j}.$$

Adding the two inequalities, we get

$$2^{n+1} \geq \sum_{r=0}^{n} C_r \sum_{j=0}^{t} \left\{ \binom{r}{j} + \binom{n-r}{j} \right\}.$$

Since

$$\binom{r}{j} + \binom{n-r}{j} \geq \binom{\lfloor n/2 \rfloor}{j} + \binom{\lceil n/2 \rceil}{j}$$

for all r and j, we get

$$2^{n+1} \geq \sum_{r=0}^{n} C_r \sum_{j=0}^{t} \left\{ \binom{\lfloor n/2 \rfloor}{j} + \binom{\lceil n/2 \rceil}{j} \right\} = \#C \sum_{j=0}^{t} \left\{ \binom{\lfloor n/2 \rfloor}{j} + \binom{\lceil n/2 \rceil}{j} \right\}.$$

Hence

$$\#C \leq \frac{2^{n+1}}{\sum_{j=0}^{t} \left\{ \binom{\lfloor n/2 \rfloor}{j} + \binom{\lceil n/2 \rceil}{j} \right\}}$$

for all $C \in \mathcal{A}(n,t)$. $\square$

For $t = 1$ we get $\alpha(n,1) \leq 2^{n+1}/(n+2)$.

3.2 The programming bound

Lemma 3.2 *Let $n > t \geq 1$. If $C \in \mathcal{A}(n,t)$, then there exists a code $C' \in \mathcal{A}(n,t)$ such that $\mathbf{0}, \mathbf{1} \in C'$ and $\#C' \geq \#C$.*

Proof: If $w(\mathbf{x}) \leq t$, then $\mathbf{0} \in S_t(\mathbf{x})$. Hence C contains at most one codeword of weight t or less. We remove this from C (if it exists) and include $\mathbf{0}$. Similarly, there is at most one codeword of weight $n-t$ or more. This we replace with $\mathbf{1}$. The resulting code is C' which clearly has the stated properties. □

Lemma 3.3 *Let $C \in \mathcal{A}(n,t)$, $0 \leq r \leq n$, and $0 \leq s \leq t$. Then*

$$\sum_{j=0}^{s} \binom{r+j}{r} C_{r+j} + \sum_{i=1}^{t-s} \binom{n-r+1}{n-r} C_{r-i} \leq \binom{n}{r}.$$

Proof: By Lemma 2.3, $\bigcup_{\mathbf{x} \in C} S_{t-s,s}(\mathbf{x})$ is a disjoint union. If $\mathbf{x} \in C$ and $w(\mathbf{x}) = r+j$, where $0 \leq j \leq s$, then $S_{t-s,s}(\mathbf{x})$ contains $\binom{r+j}{r}$ vectors of weight r. If $\mathbf{x} \in C$ and $w(\mathbf{x}) = r-i$, where $1 \leq i \leq t-s$, then $S_{t-s,s}(\mathbf{x})$ contains $\binom{n-r+i}{n-r}$ vectors of weight r. Hence the number of vectors of weight r in $\bigcup_{\mathbf{x} \in C} S_{t-s,s}(\mathbf{x})$ is

$$\sum_{j=0}^{s} \binom{r+j}{r} C_{r+j} + \sum_{i=1}^{t-s} \binom{n-r+1}{n-r} C_{r-i}.$$

Since the total number of vectors of weight r in $\{0,1\}^n$ is $\binom{n}{r}$, the lemma follows. □

Definition 3.4 *$A(n,d,w)$ denotes the maximal number of vectors in $\{0,1\}^n$ of weight w and with Hamming distance at least d apart.*

Lemma 3.4 *Let $C \in \mathcal{A}(n,t)$ and $0 \leq r \leq n$. Then*

$$\sum_{j=s}^{r} A(r-s, 2t+2, r-j) C_j \leq A(n+r-s, 2t+2, r).$$

Proof: For $s \leq j \leq r$, let E_j be a code of length $r-s$, constant weight $r-j$, and Hamming distance at least $2t+2$ between distinct codewords, and such that $\#E_j = A(r-s, 2t+2, r-j)$. Let

$$X = \bigcup_{j=s}^{r} \left\{ (\mathbf{x}|\mathbf{v}) \;\middle|\; \mathbf{x} \in C,\ w(\mathbf{x}) = j,\ \mathbf{v} \in E_j \right\}.$$

Then X is a code of length $n+(r-s)$ and constant weight r. We shall prove that the Hamming distance between distinct codewords in X is at least $2t+2$. Let $(\mathbf{x}|\mathbf{v}), (\mathbf{x}'|\mathbf{v}') \in X$, $(\mathbf{x}|\mathbf{v}) \neq (\mathbf{x}'|\mathbf{v}')$. Then

$$d((\mathbf{x}|\mathbf{v}), (\mathbf{x}'|\mathbf{v}')) = d(\mathbf{x}, \mathbf{x}') + d(\mathbf{v}, \mathbf{v}').$$

If $\mathbf{x} = \mathbf{x}'$, then $\mathbf{v}, \mathbf{v}' \in E_{w(\mathbf{x})}$ and $\mathbf{v} \neq \mathbf{v}'$. Hence

$$d((\mathbf{x}|\mathbf{v}), (\mathbf{x}'|\mathbf{v}')) = d(\mathbf{v}, \mathbf{v}') \geq 2t+2.$$

If $\mathbf{x} \neq \mathbf{x}'$, then $\Delta(\mathbf{x}, \mathbf{x}') \geq t+1$. Further,

$$d(\mathbf{v}, \mathbf{v}') \geq |w(\mathbf{v}) - w(\mathbf{v}')| = |w(\mathbf{x}) - w(\mathbf{x}')|.$$

Hence

$$d((\mathbf{x}|\mathbf{v}), (\mathbf{x}'|\mathbf{v}')) \geq d(\mathbf{x}, \mathbf{x}') + |w(\mathbf{x}) - w(\mathbf{x}')| = 2\Delta(\mathbf{x}, \mathbf{x}') \geq 2t+2.$$

Therefore

$$\begin{aligned} A(n+r-s, 2t+2, r) \geq \#X &= \sum_{j=s}^{r} \#\{\mathbf{x} \in C \mid w(\mathbf{x}) = j\} \#E_j \\ &= \sum_{j=s}^{r} C_j A(r-s, 2t+2, r-j). \end{aligned}$$

$\square$

Theorem 3.2 *For $n \geq 2t \geq 2$, let*

$$M(n,t) := \max \sum_{r=0}^{n} z_r$$

where the maximum is taken over all $(z_0, z_1, \ldots, z_n)$ satisfying the following constraints.

(i) z_r *are non-negative integers,*
(ii) $z_0 = z_n = 1$, $z_r = z_{n-r} = 0$ *for* $1 \leq r \leq t$,
(iii) $\sum_{j=0}^{s} \binom{r+j}{r} z_{r+j} + \sum_{i=1}^{t-s} \binom{n-r+1}{n-r} z_{r-i} \leq \binom{n}{r}$ *for* $0 \leq s \leq t$, $0 \leq r \leq n$,
(iv) $\sum_{j=s}^{r} A(r-s, 2t+2, r-j) z_j \leq A(n+r-s, 2t+2, r)$ *for* $0 \leq s \leq r$,
(v) $\sum_{j=s}^{r} A(r-s, 2t+2, r-j) z_{n-j} \leq A(n+r-s, 2t+2, r)$ *for* $0 \leq s \leq r$.

Then $\alpha(n,t) \leq M(n,t)$.

Proof: Let $C \in \mathcal{A}(n,t)$ be a code of size $\alpha(n,t)$. By Lemma 3.2 we may assume that $\mathbf{0}, \mathbf{1} \in C$. Hence $C_0 = C_n = 1$ and $C_r = C_{n-r} = 0$ for $1 \leq r \leq t$. Therefore $z_r = C_r$ for $r = 0, 1, \ldots, n$ satisfies (i) and (ii). By Lemma 3.3 it satisfies (iii). By Lemma 3.4 it satisfies (iv), and by Lemma 3.4 applied to $\overline{C}$ it satisfies (v). Hence

$$M(n,t) \geq \sum_{r=0}^{n} C_r = \#C = \alpha(n,t).$$

$\square$

The bound given by Theorem 3.2 is the best upper bound known for $\alpha(n,t)$. The bound is not explicit, it is given as the solution of an integer programming problem. Other bounds which are weaker, but are simpler to compute, are given below.

3.3 The constant weight code bound

Theorem 3.3 *For* $n > 2t \geq 2$, *let* $B_t, B_{t+1}, \ldots, B_{n-t-1}$ *be defined by*

$$\begin{aligned} B_t &:= 2, \\ B_r &:= \min_{t \leq j < r} \left\{ B_j + A(n+r-j-1, 2t+2, r) \right\} \text{ for } r > t. \end{aligned}$$

Then $\alpha(n,t) \leq B_{n-t-1}$.

Proof: By the definition of the B_r, there exist $r_0, r_1, \ldots, r_m$ such that

$$t = r_0 < r_1 < \cdots < r_m = n-t-1$$

and

$$B_{r_j} = B_{r_{j-1}} + A(n + r_j - r_{j-1} - 1, 2t+2, r_j)$$

for $1 \le j \le m$. Therefore

$$B_{n-t-1} = 2 + \sum_{j=1}^{m} A(n + r_j - r_{j-1} - 1, 2t + 2, r_j).$$

Let $C \in \mathcal{A}(n,t)$ be such that $\mathbf{0}, \mathbf{1} \in C$ and $\#C = \alpha(n,t)$. By lemma 3.4, $\sum_{i=s}^{r} C_i \le A(n + r - s, 2t + 2, r)$, and so we get

$$\begin{aligned}
\alpha(n,t) &= \sum_{i=0}^{n} C_i = 2 + \sum_{i=t+1}^{n-t-1} C_i \\
&= 2 + \sum_{j=1}^{m} \sum_{i=r_{j-1}+1}^{r_j} C_i \\
&\le 2 + \sum_{j=1}^{m} A(n + r_j - r_{j-1} - 1, 2t + 2, r_j) = B_{n-t-1}.
\end{aligned}$$

3.4 An almost explicit bound

By relaxing some of the constraints in Theorem 3.2, we can obtain a linear programming problem which can be solved.

Theorem 3.4 *For $n > 2t \ge 2$, let $y_0, y_1, \ldots, y_n$ be defined by*

$$\begin{aligned}
y_0 &:= 1, \\
y_r &:= 0 \quad \textit{for } 1 \le r \le t, \\
y_{t+r} &:= \frac{1}{\binom{t+r}{t}} \left\{ \binom{n}{r} - \sum_{j=0}^{t-1} y_{r+j} \binom{r+j}{j} \right\} \quad \textit{for } 1 \le r \le \frac{n}{2} - t, \\
y_{n-r} &:= y_r \quad \textit{for } 0 \le r \le \frac{n}{2}.
\end{aligned}$$

Then $\alpha(n,t) \le \sum_{r=0}^{n} y_r$.

Proof: Let $M^*(n,t) := \max \sum_{r=0}^{n} z_r$, where the maximum is taken over the following constraints.

(i) z_r are non-negative integers,
(ii) $z_0 = 1$, $z_r = 0$ for $1 \le r \le t$,
(iii) $\sum_{j=0}^{t} \binom{r+j}{r} z_{r+j} \le \binom{n}{r}$ for $0 \le r \le \frac{n}{2} - t$,
(iv) $z_{n-r} = z_r$ for $0 \le r \le \frac{n}{2}$.

Let $\tilde{z}_0, \tilde{z}_1, \ldots, \tilde{z}_n$ be integers satisfying (i)-(v) in Theorem 3.2, and such that $\sum_{r=0}^{n} \tilde{z}_r = M(n,t)$. Let $z_r = (\tilde{z}_r + \tilde{z}_{n-r})/2$ for $0 \leq r \leq n$. Then $z_0, z_1, \ldots, z_n$ clearly satisfy (i), (ii) and (iv) in Theorem 3.4. They also satisfy (iii) since this is obtained from (iii) in Theorem 3.2 by putting $s = t$ and $s = 0$ and adding. Finally, we note that

$$M^*(n,t) \geq \sum_{r=0}^{n} z_r = M(n,t) \geq \alpha(n,t).$$

We shall prove that $y_0, y_1, \ldots, y_n$ is the unique solution giving the maximum $M^*(n,t)$.

Let $z_0, z_1, \ldots, z_n$ be real numbers satisfying (i)-(iv) and $\sum_{r=0}^{n} z_r = M^*(n,t)$. Let $Z_r = \sum_{j=0}^{t} z_{r+j}\binom{r+j}{j}$. We shall prove that $Z_r = \binom{n}{r}$ for $0 < r \leq \frac{n}{2} - t$. We split the proof in three lemmas.

Lemma 3.5 *If $Z_k < \binom{n}{k}$ for some k, $0 < k \leq \frac{n}{2} - t$, and $z_{k+t+u} = 0$ for $1 \leq u \leq s$, where $1 \leq s \leq \min\left\{t, \frac{n}{2} - t - k\right\}$, then $Z_{k+s} < \binom{n}{k+s}$.*

Proof:

$$\begin{aligned}
Z_{k+s} &= \sum_{j=0}^{t} z_{k+s+j}\binom{k+s+j}{k+s} \\
&= \sum_{j=s}^{t} z_{k+j}\binom{k+j}{k+s} \qquad \text{since } z_{k+j} = 0 \text{ for } t < j \leq t+s \\
&= \sum_{j=s}^{t} z_{k+j}\binom{k+j}{k}\frac{\binom{j}{s}}{\binom{k+s}{s}} \\
&\leq \frac{\binom{t}{s}}{\binom{k+s}{s}} \sum_{j=s}^{t} z_{k+j}\binom{k+j}{k} \\
&\leq \frac{\binom{t}{s}}{\binom{k+s}{s}} Z_k \\
&< \frac{\binom{t}{s}}{\binom{k+s}{s}} \binom{n}{k} < \binom{n}{k+s}.
\end{aligned}$$

Lemma 3.6 *If $0 < k \leq \frac{n}{2} - t$ and $z_{k+t+u} = 0$ for $1 \leq u \leq \min\left\{t, \frac{n}{2} - t - k\right\}$, then $Z_k = \binom{n}{k}$.*

Proof: Suppose that $Z_k < \binom{n}{k}$. Let

$$\Delta = \min\left\{ \frac{\binom{n}{k+s} - Z_{k+s}}{\binom{k+t}{k+s}} \;\middle|\; 0 \le s \le \min\left\{t, \frac{n}{2} - t - k\right\} \right\}.$$

By Lemma 3.5, $\Delta > 0$. Let

$$\begin{aligned} z^*_{k+t} &= z_{k+t} + \Delta, \\ z^*_r &= z_r \quad \text{for } 0 \le r \le \frac{n}{2},\ r \ne k+t, \\ z^*_{n-r} &= z^*_r \quad \text{for } 0 \le r \le \frac{n}{2}. \end{aligned}$$

Then $z^*_0, z^*_1, \ldots, z^*_n$ clearly satisfy (i), (ii) and (iv). They also satisfy (iii): if $r < k$ or $r > k + t$, then $Z^*_r = Z_r < \binom{n}{r}$, and if $0 \le s \le t$, then

$$Z^*_{k+s} = Z_{k+s} + \binom{k+t}{k+s}\Delta \le \binom{n}{k+s}.$$

On the other hand,

$$\sum_{r=0}^{n} z^*_r \ge \sum_{r=0}^{n} z_r + \Delta > M^*(n,t).$$

This contradicts the definition of $M^*(n,t)$. Hence $Z_k = \binom{n}{k}$.

Lemma 3.7 *If $0 < k \le \frac{n}{2} - t$ and there exists a J such that $1 \le J \le \min\left\{t, \frac{n}{2} - t - k\right\}$, $z_{k+t+J} > 0$, and $z_{k+t+u} = 0$ for $1 \le u \le J$, then $Z_k = \binom{n}{k}$.*

Proof: Suppose that $Z_k < \binom{n}{k}$. Let

$$\Delta = \min\left\{ \frac{\binom{k+t+J}{k+t}}{\binom{t}{J}} z_{k+t+J}, \min\left\{ \frac{\binom{n}{k+s} - Z_{k+s}}{\binom{k+t}{k+s}} \;\middle|\; 0 \le s \le J \right\} \right\}$$

and

$$\delta = \frac{\binom{t}{J}}{\binom{k+t+J}{k+t}} \Delta.$$

By Lemma 3.5, $\Delta > \delta > 0$. Let

$$\begin{aligned} z^*_{k+t} &= z_{k+t} + \Delta, \\ z^*_{k+t+J} &= z_{k+t+J} - \delta, \\ z^*_r &= z_r \quad \text{for } 0 \le r \le \frac{n}{2},\ r \ne k+t,\ r \ne k+t+J, \\ z^*_{n-r} &= z^*_r \quad \text{for } 0 \le r \le \frac{n}{2}. \end{aligned}$$

Clearly $z^*_0, z^*_1, \dots, z^*_n$ clearly satisfy (ii) and (iv). They also satisfy (i) since $\delta \le z_{k+t+J}$. To show that they satisfy (iii) we consider the various cases:

$$\begin{aligned} Z^*_r &= Z_r < \binom{n}{r} \quad \text{for } r < k \text{ or } r > k+t+J, \\ Z^*_{k+s} &= Z_{k+s} + \binom{k+t}{k+s}\Delta \le \binom{n}{k+s} \quad \text{for } 0 \le s < J, \\ Z^*_{k+s} &= Z_{k+s} - \binom{k+t+J}{k+s}\delta < \binom{n}{k+s} \quad \text{for } t < s < t+J, \\ Z^*_{k+s} &= Z_{k+s} + \binom{k+t}{k+s}\Delta - \binom{k+t+J}{k+s}\delta \le Z_{k+s} \le \binom{n}{k+s} \text{ for } t < s < t+J \end{aligned}$$

since

$$\binom{k+t}{k+s}\Delta - \binom{k+t+J}{k+s}\delta = \binom{k+t}{k+s}\Delta\left\{1 - \frac{\binom{t}{J}}{\binom{t+J-s}{J}}\right\} \le 0.$$

As in the proof of Lemma 3.6 we get $\sum_{r=0}^{n} z^*_r > M^*(n,t)$, a contradiction.

To complete the proof of Theorem 3.4, we observe that if $0 < r \le \frac{n}{2} - t$, then, by Lemmas 3.6 and 3.7,

$$\sum_{j=0}^{t} z_{r+j}\binom{r+j}{j} = Z_r = \binom{n}{r}.$$

Together with (ii) and (iv), this determines the z_r uniquely, and by induction we get $z_r = y_r$ for all r. Hence, $\alpha(n,t) \le M^*(n,t) = \sum_{r=0}^{n} y_r$. □

The bound given by Theorem 3.4 is usually weaker than $M(n,t)$. However, it is quite simple to compute. Moreover, there exists a more explicit expression for y_r.

Theorem 3.5 *Let $\beta_t(m)$ be defined by*

$$\begin{aligned}
\beta_t(m) &= 0 \quad \text{for } m<0,\\
\beta_t(0) &= 1,\\
\beta_t(m) &= -\sum_{j=0}^{t}\beta_t(m+j-t)\frac{t!}{j!} \quad \text{for } m>0.
\end{aligned}$$

Then

$$y_{k+t}=\sum_{m=1}^{k}\beta_t(k-m)\frac{t!m!}{(t+k)!}\binom{n}{m} \quad \text{for } 0\le k\le \frac{n}{2}-t, \tag{3.1}$$

and there exist complex numbers $\zeta_1,\zeta_2,\dots,\zeta_t$, $\theta_1,\theta_2,\dots,\theta_t$ such that

$$\beta_t(m)=\sum_{j=1}^{t}\theta_j\zeta_j^m \quad \text{for all } m\ge 0. \tag{3.2}$$

In particular, for $m\ge 0$ we have

$$\begin{aligned}
\beta_1(m) &= (-1)^m,\\
\beta_2(m) &= -(-1+i)^{m-1}-(-1-i)^{m-1}.
\end{aligned}$$

Proof: We prove (3.1) by induction on k. For $k=0$ we get $y_t=0$ which is true. Let $k>0$ and suppose that (3.1) is true for lower values. Then

$$\begin{aligned}
y_{t+k} &= \frac{\binom{n}{k}}{\binom{t+k}{t}}-\sum_{j=0}^{t-1}\frac{\binom{k+j}{j}}{\binom{t+k}{t}}\sum_{m=1}^{k+j-t}\beta_t(k+j-t-m)\frac{t!m!}{(k+j)!}\binom{n}{m}\\
&= \beta_t(0)\frac{t!k!}{(t+k)!}\binom{n}{k}-\sum_{m=1}^{k-1}\frac{t!m!}{(t+k)!}\binom{n}{m}\sum_{j=0}^{t-1}\frac{t!}{j!}\beta_t(k-m+j-t)\\
&= \sum_{m=1}^{k}\beta_t(k-m)\frac{t!m!}{(t+k)!}\binom{n}{m}.
\end{aligned}$$

To prove (3.2), we note that $\beta_t(m)$ is the solution of a linear recurrence whose characteristic polynomial is

$$f_t(x)=\sum_{j=0}^{t}\frac{t!}{j!}x^j.$$

Since $\frac{d}{dx}f_t(x) = f_t(x) - x^t$ and $f_t(0) \neq 0$, all the zeros of the equation $f_t(x) = 0$ are simple. Hence $\beta_t(m) = \sum_{j=1}^{t} \theta_j \zeta_j^m$ for $m \geq 0$, where $\zeta_1, \zeta_2, \ldots, \zeta_t$ are the zeros of $f_t(x)$ and $\theta_1, \theta_2, \ldots, \theta_t$ are suitable complex numbers. In particular, $f_1(x) = x + 1$ with the zero $\zeta_1 = -1$ and $f_2(x) = x^2 + 2x + 2$ with the zeros $\zeta = -1 \pm i$.

3.5 The Borden bounds

Definition 3.5 *Let* $\mathbf{u} : \mathbb{Z} \to \{0,1\}^t$ *be defined by*

$$\mathbf{u}(w) := (s(w), s(w+1), \ldots, s(w_t - 1))$$

where

$$\begin{aligned} s(w) &:= 0 \textit{ if } w \equiv 0, 1, \ldots, t \pmod{2t+2}, \\ &:= 1 \textit{ if } w \equiv t+1, t+2, \ldots, 2t+1 \pmod{2t+2}. \end{aligned}$$

Further, let

$$f(w, w') := d(\mathbf{u}(w), \mathbf{u}(w')).$$

Lemma 3.8 *We have*

$$\begin{array}{lll} (i) & f(w,w') = f(w',w), & \\ (ii) & f(w+t+1, w') = t - f(w,w'), & \\ (iii) & f(0,0) = 0, & \\ & f(0,w') = w' - 1 & \textit{for } 1 \leq w' \leq t+1, \\ & f(w,w') = w' - w & \textit{for } 1 \leq w \leq w' \leq t+1, \\ & f(w,w') = w' - w - 1 & \textit{for } 1 \leq w \leq t+1 \textit{ and } t+2 \leq w' \leq w+t+1, \\ (iv) & \textit{if } w \leq w' \leq w+t+1, \textit{ then } w' - w - 1 \leq f(w,w') \leq w' - w. & \end{array}$$

Proof: (i) is obvious and (ii) follows from the fact that $\mathbf{u}(w+t+1) = \overline{\mathbf{u}(w)}$. Combining this and the fact that $\mathbf{u}(0) = \mathbf{u}(1) = 0$ and

$$\mathbf{u}(w) = (0, 0, \ldots, 0, \overbrace{1, 1, \ldots, 1}^{w-1})$$

for $1 \leq w \leq t+1$, we get (iii). Finally, we get (iv) by combining (i), (ii) and (iii).

Definition 3.6 *$A(n,t)$ is the maximal number of codewords in a binary code of length n and Hamming distince d between distinct codewords.*

Theorem 3.6 *For $n \geq t$ we have $\alpha(n,t) \leq A(n+t, 2t+1)$.*

Proof: Let $C \in \mathcal{A}(n,t)$. Define

$$D := \left\{ (\mathbf{x}|\mathbf{u}(w(\mathbf{x})) \;\middle|\; \mathbf{x} \in C \right\}.$$

Clearly, $D \subseteq \{0,1\}^{n+t}$. We shall prove that if $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$, then

$$d((\mathbf{x}|\mathbf{u}(w(\mathbf{x}))), (\mathbf{y}|\mathbf{u}(w(\mathbf{y})))) \geq 2t+1.$$

This implies that $\#C = \#D \leq A(n+t, 2t+1)$, and the theorem follows.

Hence, let $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$. Without loss of generality assume that $N(\mathbf{y},\mathbf{x}) \geq N(\mathbf{x},\mathbf{y})$. This implies that $N(\mathbf{y},\mathbf{x}) = \Delta(\mathbf{x},\mathbf{y}) \geq t+1$ and that $w(\mathbf{x}) \geq w(\mathbf{y})$. We have

$$\begin{aligned} d((\mathbf{x}|\mathbf{u}(w(\mathbf{x}))), (\mathbf{y}|\mathbf{u}(w(\mathbf{y})))) &= d(\mathbf{x},\mathbf{y}) + d(\mathbf{u}(w(\mathbf{x})), \mathbf{u}(w(\mathbf{y}))) \\ &= N(\mathbf{x},\mathbf{y}) + N(\mathbf{y},\mathbf{x}) + f(w(\mathbf{x}), w(\mathbf{x})). \end{aligned}$$

If $N(\mathbf{x},\mathbf{y}) + N(\mathbf{y},\mathbf{x}) \geq 2t+1$, then we are finished. If, on the other hand, $N(\mathbf{x},\mathbf{y}) + N(\mathbf{y},\mathbf{x}) \leq 2t$, then $N(\mathbf{x},\mathbf{y}) \leq t-1$. We consider two cases.
Case I, $N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y}) \leq t+1$. Then $w(\mathbf{y}) \leq w(\mathbf{x}) \leq w(\mathbf{y}) + t + 1$, and hence

$$\begin{aligned} f(w(\mathbf{x}), w(\mathbf{y})) &\geq w(\mathbf{x}) - w(\mathbf{y}) - 1 \\ &= N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y}) - 1 \\ &\geq 2t + 1 - N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y}). \end{aligned}$$

Case II, $t+2 \leq N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y}) \leq 2t$. Then $w(\mathbf{y}) + t + 2 \leq w(\mathbf{x}) \leq w(\mathbf{y}) + 2t$ and so $w(\mathbf{y}) \leq w(\mathbf{x}) - t - 1(2?) \leq w(\mathbf{y}) + t - 1$. Hence

$$\begin{aligned} f(w(\mathbf{x}), w(\mathbf{y})) &= t - f(w(\mathbf{x}), w(\mathbf{y}) - t - 1) \\ &\geq t - (w(\mathbf{x}) - t - 1 - w(\mathbf{y})) \\ &= 2t + 1 - (w(\mathbf{x}) - w(\mathbf{y})) \\ &= 2t + 1 - (N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y})) \\ &\geq 2t + 1 - N(\mathbf{y},\mathbf{x}) - N(\mathbf{x},\mathbf{y}). \end{aligned}$$

Hence $N(\mathbf{x},\mathbf{y}) + N(\mathbf{y},\mathbf{x}) + f(w(\mathbf{x}), w(\mathbf{y})) \geq 2t+1$ in both cases.

Theorem 3.7 *For $n \geq t$ we have $\alpha(n,t) \leq (t+1)A(n,2t+1)$.*

Proof: Let $C \in \mathcal{A}(n,t)$ with $\#C = \alpha(n,t)$. For $r = 0, 1, \ldots, t$, let

$$S_r = \left\{ \mathbf{x} \in C \;\middle|\; w(\mathbf{x}) \equiv 2r \text{ or } 2r+1 \pmod{2t+2} \right\}.$$

We shall prove that S_r is a code with Hamming distance at least $2t+1$ between distinct codewords. Let $\mathbf{x}, \mathbf{y} \in S_r$, $\mathbf{x} \neq \mathbf{y}$. Then
$|w(\mathbf{x}) - w(\mathbf{y})| \leq 1$, in which case

$$d(\mathbf{x},\mathbf{y}) = 2\Delta(\mathbf{x},\mathbf{y}) - |w(\mathbf{x}) - w(\mathbf{y})| \geq 2(t+1) - 1 = 2t+1,$$

or $|w(\mathbf{x}) - w(\mathbf{y})| \geq 2t+1$, in which case

$$d(\mathbf{x},\mathbf{y}) \geq |w(\mathbf{x}) - w(\mathbf{y})| \geq 2t+1.$$

Hence $\#S_r \leq A(n,2t+1)$ for $r = 0, 1, \ldots, t$ and so

$$\alpha(n,t) = \#C = \sum_{r=0}^{t} \#S_r \leq (t+1)A(n,2t+1).$$

Corollary 3.1 *For $n \geq t$ we have*

$$\alpha(n,t) \leq \frac{(t+1)2^n}{\sum_{j=0}^{t} \binom{n}{j}} = \frac{(t+1)!2^n}{n^t}\Big(1 + o(n)\Big).$$

Proof: The Hamming bound, see e.g. MacWilliams and Sloane [61] states that

$$A(n,2t+1) \leq \frac{2^n}{\sum_{j=0}^{t} \binom{n}{j}}.$$

3.6 Notes

3.1. Lemma 3.1 and Theorem 3.1 are due to Varshamov [82].

3.2. The first programming bound was given by Goldbaum [41]. The present presentation follows Kløve [56].

Lemma 3.3 is due to Delsarte and Piret [18], the special case $s = t$ had been proved by Goldbaum [41].

Lemma 3.4 is due to Kløve [56]; Delsarte and Piret [18] gave the result $\sum_{j=s}^{r} C_j \leq A(n+r-s, 2t+2, r)$.

3.3 and 3.4 are from Kløve [56].

3.5. The bound $\alpha(n,1) \leq A(n+1,3)$ was proved by Stanley and Yoder [76]. The general Theorems 3.6 and 3.7 was given by Borden [6]. Bassalygo [4] gave the bound $\alpha(n,t) \leq (2t+1)A(n,2t+1)$.

It appears that for large n, Theorem 3.4 gives the best explicit bound on $\alpha(n,1)$ and Theorem 3.7 combined with known bounds on $A(n,d)$ gives the best explicit bound on $\alpha(n,t)$ for $t > 1$.

Chapter 4

Codes correcting single errors

In this chapter we describe the known 1-codes. For each code we give a decoding algorithm in a Pascal-like language. The existence of a decoding algorithm of course proves that the code is a 1-code. In the next chapter, we describe the known t-codes for $t > 1$. A code correcting t errors on the binary symmetric channel is in particular a t-code. However, we will restrict ourselves to codes which are designed to correct asymmetric errors (i.e. $1 \to 0$ errors). The decodings algorithms, may or may not be efficient, the main emphasis is to show that unique decoding is possible, i.e. that the codes are able to correct t errors.

4.1 Kim-Freiman codes

For $m \geq 1$, let H_m be a code of length m which is able to correct one *symmetric* error.

Code construction.

If $n = 2m$, then

$$C = \Big\{ (\mathbf{x}|\mathbf{x}\oplus\mathbf{h}) \;\Big|\; \mathbf{x} \in \{0,1\}^m,\ w(\mathbf{x}) \text{ even},\ \mathbf{h} \in H_m \backslash \{\mathbf{0}\} \Big\} \cup \Big\{ (\mathbf{x}|\mathbf{x}) \;\Big|\; \mathbf{x} \in \{0,1\}^m \Big\}.$$

If $n = 2m + 1$, then

$$C = \Big\{ (\mathbf{x}|(\mathbf{x}|0)\oplus\mathbf{h}) \;\Big|\; \mathbf{x} \in \{0,1\}^m,\ w(\mathbf{x}) \text{ even},\ \mathbf{h} \in H_{m+1} \backslash \{\mathbf{0}\} \Big\} \cup \Big\{ (\mathbf{x}|\mathbf{x}|0) \;\Big|\; \mathbf{x} \in \{0,1\}^m \Big\}.$$

Decoding algorithm for $n = 2m$.

Comment: The received vector is $(\mathbf{y}|\mathbf{y}')$ where $\mathbf{y}, \mathbf{y}' \in \{0,1\}^m$.

if (there exist $\mathbf{c} \in H_m$ such that $d(\mathbf{y} \oplus \mathbf{y}', \mathbf{c}) \leq 1$)
  then $\mathbf{g} := \mathbf{c}$
  else decoding has failed;
if $(\mathbf{g} = \mathbf{0})$ **then**
  if $(w(\mathbf{y}) \leq w(\mathbf{y}'))$ **then** $\mathbf{z} := (\mathbf{y}'|\mathbf{y}')$
    else $\mathbf{z} := (\mathbf{y}|\mathbf{y})$
else
  if $(w(\mathbf{y})$ is even$)$ **then** $\mathbf{z} := (\mathbf{y}|\mathbf{y} \oplus \mathbf{g})$
    else $\mathbf{z} := (\mathbf{y}' \oplus \mathbf{g}|\mathbf{y}')$;
decode into $\mathbf{z}$.

The decoding algorithm is similar when $n = 2m + 1$.

Proof of the decoding algorithm.

Case I, $(\mathbf{x}|\mathbf{x} \oplus \mathbf{h})$ is sent, where $w(\mathbf{x})$ is even, and $\mathbf{h} \neq \mathbf{0}$.
Subcase Ia, no errors occurs or one error occurs in the second part. Then $\mathbf{y} = \mathbf{x}$ and $\mathbf{y}' = (\mathbf{x} \oplus \mathbf{h}) - \mathbf{e}$ where $w(\mathbf{e}) \leq 1$. Hence $\mathbf{y} \oplus \mathbf{y}' = \mathbf{h} \oplus \mathbf{e}$ and so $\mathbf{g} = \mathbf{h} \neq \mathbf{0}$. Further, $w(\mathbf{y}) = w(\mathbf{x})$ is even, and so $\mathbf{z} = (\mathbf{y}|\mathbf{y} \oplus \mathbf{h}) = (\mathbf{x}|\mathbf{x} \oplus \mathbf{h})$.
Subcase Ib, one error occurs in the first part. In this case $\mathbf{y} = \mathbf{x} - \mathbf{e}$ and $\mathbf{y}' = \mathbf{x} \oplus \mathbf{h}$ where $w(\mathbf{e}) \leq 1$. Hence $\mathbf{g} = \mathbf{h}$, $w(\mathbf{y})$ is odd, and so $\mathbf{z} = (\mathbf{y}' \oplus \mathbf{h}|\mathbf{y}') = (\mathbf{x}|\mathbf{x} \oplus \mathbf{h})$.
Case II, $(\mathbf{x}|\mathbf{x})$ is sent.
Subcase IIa, no errors occurs or one error occurs in the first part. Then $\mathbf{y} = \mathbf{x} - \mathbf{e}$, $\mathbf{y}' = \mathbf{x}$, $\mathbf{g} = \mathbf{0}$, $w(\mathbf{y}) \leq w(\mathbf{y}')$. Hence $\mathbf{z} = (\mathbf{y}'|\mathbf{y}') = (\mathbf{x}|\mathbf{x})$.
Subcase IIb, one error occurs in the second part. This subcase is similar.

The size of the codes.

Let $h_m = \#H_m$. The size of the Kin-Freiman code is

$$2^{m-1}(1 + h_m) \quad \text{if} \quad n = 2m,$$
$$2^{m-1}(1 + h_{m+1}) \quad \text{if} \quad n = 2m + 1.$$

.

Remark. If $n = 2^r - 1$, the Kim-Freiman code of length n is smaller than the Hamming code of the same length, for all other values of n it is larger.

4.2 Stanley-Yoder codes

Code construction.

Let G be a group of order $n+1$ such that every element commutes with its conjugates (i.e. $abab^{-1} = bab^{-1}a$ for all $a, b \in G$). Let $g_1, g_2, \ldots, g_n, g_{n+1}$ be an ordering of the elements of G such that every conjugacy class appears as a set of consecutive elements, $g_m, g_{m+1}, \ldots, g_{m+k}$, in the ordering, and $g_{n+1} = e$, the identity. For every $g \in G$, let

$$C_g = \left\{ (x_1, x_2, \ldots, x_n) \in \{0,1\}^n \;\middle|\; \prod_{i=1}^{n} g_i^{x_i} = g \right\}.$$

Decoding algorithm for C_g.

Comment: The received vector is $\mathbf{y}$; $\mathbf{e}_k$ is the k'th unit vector.

$h := \prod_{i=1}^{n} g_i^{y_i}$;
$k := 0$;
while $(h \neq g$ and $k \leq n)$ **do**
begin
    $k := k+1$;
    if $(y_k = 1)$ **then begin** $h := g_k^{-1}h$; $g := g_k^{-1}g$ **end**
                else if $(g_k h = g)$ **then** $h := g$
end;
if $(k = 0)$ **then** decode into $\mathbf{y}$
else if $(k \leq n)$ **then** decode into $\mathbf{y} + \mathbf{e}_k$
else decoding has failed.

Proof of decoding algorithm.

Let $\mathbf{x} \in C_g$ be sent. If no errors has occurred, then $\mathbf{y} = \mathbf{x}$. In this case $h = g$ initially, we do not run through the while-loop at all, and we decode into $\mathbf{y} = \mathbf{x}$.

Suppose an error has occurred in position j. Then

$$\begin{aligned} y_i &= x_i \quad \text{for} \quad i \neq j, \\ y_j &= 0, \qquad\quad x_j = 1. \end{aligned}$$

We shall prove the following statement:
(*) If the while-loop has been repeated k times, then

$$\text{(i) } k < j,\ g = \prod_{i=k+1}^{n} g_i^{x_i},\ h = \prod_{i=k+1}^{n} g_i^{y_i},\ \text{and } h \neq g$$
$$\text{or (ii) } k = j \text{ and } g = h.$$

First we note that if $k < j$, $g = \prod_{i=k+1}^{n} g_i^{x_i}$, and $h = \prod_{i=k+1}^{n} g_i^{y_i}$, then $g = w_1 g_j w_2$ and $h = w_1 w_2$, and so $g \neq h$. We prove (*) by induction. Clearly (i) is true for $k = 0$. Suppose (*) for some $k \geq 0$. If $k < j$, then we repeat the loop once more. We first increase k by one to $K = k + 1$. If $y_K = 1$, then $x_K = 1$ also. Hence we change g to $\prod_{i=K+1}^{n} g_i^{x_i}$ and h to $\prod_{i=K+1}^{n} g_i^{y_i}$. Hence (i) of (*) is true for $K = k + 1$ in this case. On the other hand, if $y_K = 0$, then there are two possibilities for x_K. If $K = j$, then $x_K = 1$ and so $h = \prod_{i=K+1}^{n} g_i^{x_i}$ and $g = g_K \prod_{i=K+1}^{n} g_i^{x_i}$. Hence $g_K h = g$ and so (ii) is true for $K = k = 1$ in this case. Finally, if $K < j$, then $x_K = 0$. Hence, $g = \prod_{i=K+1}^{n} g_i^{x_i}$ and $h = \prod_{i=K+1}^{n} g_i^{y_i}$. Suppose $g_K h = g$. Then

$$g_K \prod_{i=K+1}^{j-1} g_i^{x_i} = \left(\prod_{i=K+1}^{j-1} g_i^{x_i} \right) g_j.$$

Hence g_K and g_j are conjugates. This implies that $g_K, g_{K+1}, \ldots, g_j$ all belong to the same conjugacy class and hence they commute. Therefore $g_K = g_j$ which is impossible since $K < j$. Therefore (i) is true in this case also. From (*) it follows that the while-loop is repeated until $k = j$. Then we go on and decode into $\mathbf{y} + \mathbf{e}_j$.

The size of the codes.

Since $\{C_g \mid g \in G\}$ is a partition of $\{0, 1\}$ into $n + 1$ parts,

$$\max_{g \in G} \#C_g \geq \frac{2^n}{n+1}.$$

Determination of $\#C_g$ has been done only for the codes based on an Abelian group G.

4.3 Constantin-Rao codes

These codes are the Stanley-Yoder codes based on an Abelian group G. Hence, if the group operation is written $+$, then

$$C_g = \left\{ (x_1, x_2, \ldots x_n) \;\middle|\; \sum_{i=1}^{n} x_i g_i = g \right\},$$

where $g_1, g_2, \ldots, g_n$ are the non-identity elements of G.

For these codes there is a simpler decoding algorithm than the one given for Stanley-Yoder codes in general. Let g_0 be the identity element.

Decoding algorithm.

Comment: The received vector is $\mathbf{y}$;
Comment: $\mathbf{e}_k$ is the k'th unit vector for $k > 0$; $\mathbf{e}_0 = \mathbf{0}$.

$h := g - \sum_{i=1}^{n} y_i g_i$;
decode into $\mathbf{y} + \mathbf{e}_k$ where k is the index such that $g_k = h$.

Proof of decoding algorithm.

Let $\mathbf{x} \in C_g$ be sent. If no errors occur, then $\mathbf{y} = \mathbf{x}$ and hence $h = 0 = g_0$, and we decode into $\mathbf{y} + \mathbf{e}_0 = \mathbf{x}$. If an error occurs in position j, then $y_i = x_i$ for $i \neq j$ and $y_j = 0$, $x_j = 1$. Hence $h = \sum_{i=1}^{n} (x_i - y_i) g_i = g_j$, and we decode into $\mathbf{y} + \mathbf{e}_j = \mathbf{x}$.

4.4 Ananiashvili codes

Code construction.

Let $\mathbf{u} : \{0,1\}^k \to \{0,1\}^m$, where $m = \lceil \log_2(k+1) \rceil$, be defined as follows: for $(x_1, x_2, \ldots, x_k) \in \{0,1\}^k$, let s be defined by

$$s \equiv \sum_{i=1}^{k} x_i \, i \pmod{k+1}, \quad 0 \leq s \leq k,$$

and let $\sum_{i=1}^{m-1} u_i 2^{i-1}$ be the binary expansion of s. Finally, let

$$u_m \equiv \sum_{i=1}^{m-1} u_i \pmod 2, \quad u_m \in \{0,1\}.$$

Then

$$\mathbf{u}(\mathbf{x}) := (u_1, u_2, \dots, u_m).$$

The code is

$$C := \left\{ (\mathbf{x}|\mathbf{u}(\mathbf{x})) \;\middle|\; \mathbf{x} \in \{0,1\}^k \right\}.$$

Decoding algorithm.

Comment: The received vector is $(\mathbf{y}|\mathbf{v})$ where $\mathbf{y} \in \{0,1\}^k$, $\mathbf{v} \in \{0,1\}^m$.
Comment: $\mathbf{e}_j$ is the j'th unit vector for $j > 0$, $\mathbf{e}_0 = \mathbf{0}$.

if $(\sum_{i=1}^{m} v_i \equiv 1 \pmod 2)$ **then** decode into $(\mathbf{y}|\mathbf{u}(\mathbf{y}))$
else
begin
determine b by $b \equiv \sum_{i=1}^{m-1} v_i 2^{i-1} - \sum_{i=1}^{k} y_i\, i \pmod{k+1}$, $0 \le b \le k$;
$\mathbf{y} := \mathbf{y} + \mathbf{e}_b$;
decode into $(\mathbf{y}|\mathbf{u}(\mathbf{y}))$
end.

Proof of decoding algorithm.

Let $(\mathbf{x}|\mathbf{u}(\mathbf{x}))$ be sent.

If no errors occur, then $\sum_{i=1}^{m} v_i \equiv 0 \pmod 2$, and

$$b \equiv \sum_{i=1}^{m-1} u_i 2^{i-1} - \sum_{i=1}^{k} x_i i \equiv s - s \equiv 0 \pmod{k+1}.$$

Hence we decode into $(\mathbf{y}|\mathbf{u}(\mathbf{y})) = (\mathbf{x}|\mathbf{u}(\mathbf{x}))$.

If an error occurs in $\mathbf{u}(\mathbf{x})$, then $\mathbf{y} = \mathbf{x}$ and $\sum_{i=1}^{m} v_i = (\sum_{i=1}^{m} u_i) - 1 \equiv 1 \pmod 2$, and we decode into $(\mathbf{y}|\mathbf{u}(\mathbf{y})) = (\mathbf{x}|\mathbf{u}(\mathbf{x}))$.

Finally, if an error occurs in position j of $\mathbf{x}$, then $\mathbf{v} = \mathbf{u}(\mathbf{x})$, $y_i = x_i$ for $i \ne j$, $y_j = 0$, and $x_j = 1$. Hence

$$b \equiv s - (s - j) \equiv j \mod k+1.$$

Hence $b = j$ and we decode into $(\mathbf{y} + \mathbf{e}_j|\mathbf{u}(\mathbf{y} + \mathbf{e}_j)) = (\mathbf{x}|\mathbf{u}(\mathbf{x}))$.

The size of the codes.

The length of the code C is $n = k + \lceil \log_2(k+1) \rceil$ and $\#C = 2^k$.

4.5 Delsarte-Piret Codes

The Delsarte-Piret codes are 1-codes of lengths between 7 and 11, and with related constructions.

The main idea of the constructions.

As before, let C_w denote the number of codewords in C of weight w. The main idea of Delsarte and Piret's constructions is to let $C_w = 0$ for $w = w_1, w_2, \ldots, w_s$ and use some known combinatorial construction to get codewords of weights $w_i + 1, w_i + 2, \ldots, w_{i+1} - 1$, $i = 1, 2, \ldots$. The point is that if $w(\mathbf{c}_1) < w_i$ and $w(\mathbf{c}_2) > w_i$, then $\Delta(\mathbf{c}_1, \mathbf{c}_2) \geq 2$, hence the various constructions may be done independently. For all the constructions, $\mathbf{0}, \mathbf{1} \in C$ and so $C_0 = C_n = 1$ and $C_1 = C_{n-1} = 0$.

Construction for $n = 11$.

We let $C_4 = C_7 = 0$.

We get $C_2 + C_3 = 20$ by the following construction: It is known that $A(12, 4, 3) = 20$. Let X be a code of length 12, constant weight 3, minimum distance 4, and size 20. Let

$$T_{11} := \left\{ (x_1, x_2, \ldots, x_{11}) \,\middle|\, (x_1, x_2, \ldots, x_{11}, x_{12}) \in X \text{ for } x_{12} = 0 \text{ or } x_{12} = 1 \right\}.$$

For $\mathbf{x}, \mathbf{y} \in T_{11}$, $\mathbf{x} \neq \mathbf{y}$, we have

$$w(\mathbf{x}) = w(\mathbf{y}) \text{ and } d(\mathbf{x}, \mathbf{y}) \geq 4,$$

or

$$|w(\mathbf{x}) - w(\mathbf{y})| = 1 \text{ and } d(\mathbf{x}, \mathbf{y}) \geq 3.$$

In any case $\Delta(\mathbf{x}, \mathbf{y}) \geq 2$. Moreover, $w(\mathbf{x}) \in \{2, 3\}$ for all $\mathbf{x} \in T_{11}$.

We get $C_8 + C_9 = 20$ by the choice $\{\overline{\mathbf{x}} \mid \mathbf{x} \in T_{11}\}$.

Finally, we get $C_5 + C_6 = 132$ by the following construction: Starting from a $(5, 6, 12)$ Steiner system and deleting one coordinate we get

$$R_{11} := \bigcup_{i=1}^{12} R(i),$$

where $R(i)$ is the set of cyclic shifts of $\mathbf{r}(i)$ defined by

$$\begin{aligned}
\mathbf{r}(1) &:= (11011100010),\\
\mathbf{r}(2) &:= (10110010011),\\
\mathbf{r}(3) &:= (01101011010),\\
\mathbf{r}(4) &:= (10000111110),\\
\mathbf{r}(5) &:= (11110001100),\\
\mathbf{r}(6) &:= (11001010101),\\
\mathbf{r}(i) &:= \overline{\mathbf{r}(i-6)} \text{ for } 7 \le i \le 12.
\end{aligned}$$

The size of the code is $\sum_{i=0}^{11} C_i = 174$.

Construction for $n = 10$.

We let $C_3 = C_7 = 0$.

We can get $C_2 = C_8 = 5$ since $A(10, 4, 2) = 5$.

We get $C_4 + C_5 + C_6 = 96$ by the following construction:

$$\begin{aligned}
R_{10} &:= \left\{ (x_1, x_2, \ldots, x_{10}) \;\middle|\; (x_1, x_2, \ldots, x_{10}, 0) \in R_{11} \right\}\\
&\quad \cup \left\{ (x_1, x_2, \ldots, x_{10}) \;\middle|\; (x_1, x_2, \ldots, x_{10}, 1) \in \bigcup_{i=7}^{12} R(i) \right\}.
\end{aligned}$$

The size of the code is 108.

Construction for $n = 9$.

We let $C_2 = C_7 = 0$.

We get $C_3 + C_4 + C_5 + C_6 = 60$ by the following construction:

$$\begin{aligned}
R_9 &:= \left\{ (x_1, x_2, \ldots, x_9) \;\middle|\; (x_1, x_2, \ldots, x_9, 0) \in R_{10} \right\}\\
&\quad \cup \left\{ (x_1, x_2, \ldots, x_9) \;\middle|\; \mathbf{x} = (x_1, x_2, \ldots, x_9, 1) \in R_{10},\ w(\mathbf{x}) = 4 \right\}.
\end{aligned}$$

The size of the code is 62.

Construction for $n = 8$.

We get $C_2 + C_3 + C_4 + C_5 + C_6 = 34$ by the following construction:

$$\begin{aligned} R_8 \ := \ & \left\{ (x_1, x_2, \dots, x_8) \,\middle|\, (x_1, x_2, \dots, x_8, 0) \in R_9 \right\} \\ & \cup \left\{ (x_1, x_2, \dots, x_8) \,\middle|\, \mathbf{x} = (x_1, x_2, \dots, x_8, 1) \in R_9,\ w(\mathbf{x}) = 3 \right\}. \end{aligned}$$

The size of the code is 36.

Construction for $n = 7$.

By shortening R_8 we get a code R_7 with 18 code words.

4.6 The size of 1-codes

We give a table of the size of the maximal known 1-code in each of the classes constructed above. The table also contains the best upper bounds given in Chapter 3.

4.7 Notes

In addition to correcting one single error, many of the codes are also able to *detect* many combinations of multiple errors. Decoding algorithms taking this into account would be less simple than the ones we have given. We have decided to give the simpler algorithms to make the underlying ideas clearer. In most cases it is straight-forward to rewrite the algorithms so as to detect many multiple errors.

Varshamov [82] proved that almost all linear codes which are able to correct t asymmetric errors are also able to correct t symmetric errors. Therefore, to go beyond t-symmetric-error correcting codes, non-linear constructions are needed.

4.1. The codes were defined by Kim and Freiman [50]. They used Hamming codes as the codes H_m in the construction.

4.2 and 4.3. The main idea of the construction is due to Varshamov and Tenengolts [91]. They used a cyclic group G. The general construction

n	S	KF	CR	A	DP	U
5	4	6	6	4	-	6
6	8	12	10	8	-	12
7	16	12	16	-	18	18
8	20	24	32	16	36	36
9	38	40	52	32	62	64
10	72	80	94	64	108	118
11	144	144	172	128	174	210
12	256	288	316	-	-	410
13	512	544	586	256	-	786
14	1024	1088	1096	512	-	1500
15	2048	1344	2048	1024	-	2828
16	2560	2688	3856	2048	-	5486

Key to abbreviations:

S: code correcting one symmetric error
KF: Kim-Freiman code
CR: Constantin-Rao code
A: Ananiashvili code
DP: Delsarte-Piret code
U: upper bound

was given by Stanley and Yoder [76]. The construction was rediscovered by Constantin and Rao [14] who used an Abelian group. The properties of Constantin-Rao codes will be discussed in detail in the next chapter.

4.4. The construction is due to Ananiashvili [1]. The codes are separable, the $\mathbf{x}$ part of $(\mathbf{x}|\mathbf{u}(\mathbf{x}))$ may be used for information. However, for most code lengths, the Ananiashvili codes are smaller than the Hamming codes. On the other hand, Ananiashvili codes are able to detect a large fraction of all double errors.

4.5. The constructions are due to Delsarte and Piret [18].

Chapter 5

Properties of Constantin-Rao codes

5.1 Definitions

In this chapter we shall give the main properties of the Constantin-Rao codes. It is clear that isomorphic groups define the same set of codes. Since any Abelian group is isomorphic to a unique direct sum of cyclic groups of prime power order, it is no restriction to assume that G is so defined, i.e.

$$G = \bigoplus_{j=1}^{J} G_j$$

where

$$G_j = \mathbb{Z}_{p_j^{\alpha_j}}.$$

The order of G is $N := \prod_{j=1}^{J} p_j^{\alpha_j}$. Each element of G is represented by a J-tuple $\mathbf{g} = (g_1, g_2, \ldots, g_J)$ where $0 \leq g_j < p_j^{\alpha_j}$ for $1 \leq j \leq J$, let

$$\zeta_j := \exp\left(\frac{2\pi i}{p_j^{\alpha_j}}\right).$$

For $\mathbf{g}, \mathbf{h} \in G$ let

$$\langle \mathbf{g}, \mathbf{h} \rangle := \prod_{j=1}^{J} \zeta_j^{g_j h_j}.$$

We note that $\langle \mathbf{g}, \mathbf{h} \rangle \langle \mathbf{g}, \mathbf{k} \rangle = \langle \mathbf{g}, \mathbf{h} + \mathbf{k} \rangle$ and $\langle \mathbf{g}, j\mathbf{h} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle^j$. The Constantin-Rao codes are defined by

$$C_{\mathbf{g}} := \left\{ \mathbf{x} \in \{0,1\}^{N-1} \;\middle|\; \sum_{i=1}^{N-1} x_i \mathbf{g}_i = \mathbf{g} \right\}$$

where $G = \{\mathbf{g}_0 = \mathbf{0}, \mathbf{g}_1, \ldots, \mathbf{g}_{N-1}\}$.

5.2 The weight distribution

Definition 5.1 *For $\mathbf{g} \in G$ and $0 \le w \le N-1$, let*

$$t(\mathbf{g}, w) := \# \left\{ \mathbf{x} \in C_{\mathbf{g}} \;\middle|\; w(\mathbf{x}) = w \right\},$$

$$T_{\mathbf{g}}(y) := \sum_{w=0}^{N-1} t(\mathbf{g}, w) y^w.$$

Our aim is to determine $t(\mathbf{g}, w)$ for all w, or equivalently, the polynomial $T_{\mathbf{g}}(y)$. Note that $\# C_{\mathbf{g}} = T_{\mathbf{g}}(1)$.

For $\mathbf{h} \in G$, let

$$\hat{T}_{\mathbf{h}}(y) := \sum_{\mathbf{g} \in G} \langle -\mathbf{h}, \mathbf{g} \rangle T_{\mathbf{g}}(y).$$

Lemma 5.1 *For $\mathbf{g} \in G$ we have*

$$T_{\mathbf{g}}(y) = \frac{1}{N} \sum_{\mathbf{h} \in G} \langle \mathbf{h}, \mathbf{g} \rangle \hat{T}_{\mathbf{h}}(y).$$

Proof: First we note that

$$\sum_{\mathbf{h} \in G} \langle \mathbf{h}, \mathbf{g} \rangle = \prod_{j=1}^{J} \sum_{h_j=0}^{p_j^{\alpha_j}-1} \zeta_j^{g_j h_j} = \begin{cases} N & \text{if } \mathbf{g} = \mathbf{0}, \\ 0 & \text{if } \mathbf{g} \ne \mathbf{0}. \end{cases}$$

Hence

$$\begin{aligned} \sum_{\mathbf{h} \in G} \langle \mathbf{h}, \mathbf{g} \rangle \hat{T}_{\mathbf{h}}(y) &= \sum_{\mathbf{h} \in G} \sum_{\mathbf{k} \in G} \langle \mathbf{h}, \mathbf{g} \rangle \langle -\mathbf{h}, \mathbf{k} \rangle \hat{T}_{\mathbf{k}}(y) \\ &= \sum_{\mathbf{k} \in G} \hat{T}_{\mathbf{k}}(y) \sum_{\mathbf{h} \in G} \langle \mathbf{h}, \mathbf{g} - \mathbf{k} \rangle = N\, T_{\mathbf{g}}(y). \end{aligned}$$

Lemma 5.2 *For* $\mathbf{h} \in G$ *we have*

$$\hat{T}_{\mathbf{h}}(y) = \frac{1}{1+y}\Big(1-(-y)^d\Big)^{N/d},$$

where d *is the order of* $\mathbf{h}$.

Proof:

$$\begin{aligned}
\hat{T}_{\mathbf{h}}(y) &= \sum_{\mathbf{g}\in G}\sum_{w=0}^{N-1}\langle -\mathbf{h},\mathbf{g}\rangle y^w \#\Big\{(x_1,x_2,\ldots,x_{N-1}) \Bigm| \sum_{i=1}^{N-1} x_i\mathbf{g}_i = \mathbf{g},\ \sum_{i=1}^{N-1} x_i = w\Big\} \\
&= \sum_{\mathbf{x}\in\{0,1\}^{N-1}} \langle -\mathbf{h}, \sum_{i=1}^{N-1} x_i\mathbf{g}_i\rangle\, y^{\sum_{i=1}^{N-1} x_i} \\
&= \sum_{\mathbf{x}\in\{0,1\}^{N-1}} \prod_{i=1}^{N-1}\Big(\langle -\mathbf{h},\mathbf{g}_i\rangle^{x_i} y^{x_i}\Big) \\
&= \prod_{i=1}^{N-1}\Big(1+\langle -\mathbf{h},\mathbf{g}_i\rangle y\Big).
\end{aligned}$$

Let $o(\mathbf{h})$ denote the order of $\mathbf{h}$. If $o(\mathbf{h}) = d$, then $\mathbf{g} \mapsto \langle -\mathbf{h},\mathbf{g}\rangle$ is a homomorphism from G onto the complex d'th roots of unity. Hence, if ζ is a primitive d'th root of unity, then

$$\prod_{i=0}^{N-1}\Big(1-\langle -\mathbf{h},\mathbf{g}_i\rangle z\Big) = \Big(\prod_{i=0}^{d-1}(1-\zeta^i z)\Big)^{N/d} = \Big(1-z^d\Big)^{N/d}.$$

Putting $z = -y$, the lemma follows.

Definition 5.2 *For* $\mathbf{g} \in G$, *let*

$$S_{\mathbf{g}}(d) := \sum_{\substack{\mathbf{h}\in G\\ o(\mathbf{h})=d}} \langle \mathbf{h},\mathbf{g}\rangle.$$

Theorem 5.1 *For* $\mathbf{g} \in G$ *we have*

$$T_{\mathbf{g}}(y) = \frac{1}{N(1+y)}\sum_{cd=N}\Big(1-(-y)^d\Big)^c S_{\mathbf{g}}(d)$$

and

$$\#C_{\mathbf{g}} = \frac{1}{2N}\sum_{\substack{cd=N\\ d\ \text{odd}}} 2^c S_{\mathbf{g}}(d).$$

Proof: From Lemmas 5.1 and 5.2 we get

$$\begin{aligned} T_{\mathbf{g}}(y) &= \frac{1}{N(1+y)} \sum_{\mathbf{h}\in G} \langle \mathbf{h}, \mathbf{g}\rangle \left(1-(-y)^{o(\mathbf{h})}\right)^{N/o(\mathbf{h})} \\ &= \frac{1}{N(1+y)} \sum_{cd=N} \left(1-(-y)^{d}\right)^{c} \sum_{\substack{\mathbf{h}\in G \\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g}\rangle, \end{aligned}$$

and

$$\#C_{\mathbf{g}} = T_{\mathbf{g}}(1) = \frac{1}{2N} \sum_{cd=N} \left(1-(-1)^{d}\right)^{c} S_{\mathbf{g}}(d).$$

5.3 The function $S_{\mathbf{g}}$

The usefulness of Theorem 5.1 depends on having an explicit expression for $S_{\mathbf{g}}(d)$ and we shall derive such an expression (Theorem 5.3 below).

Lemma 5.3 *Let* $\mathbf{h} \in G$. *If* $o(\mathbf{h}) = d_1 d_2$, *where* $\gcd(d_1, d_2) = 1$, *then* $\mathbf{h}$ *has a unique decomposition* $\mathbf{h} = \mathbf{h}_1 + \mathbf{h}_2$ *such that* $o(\mathbf{h}_1) = d_1$ *and* $o(\mathbf{h}_2) = d_2$.

Proof: Let a_1 and a_2 be integers such that $a_1 d_1 + a_2 d_2 = 1$. Let $\mathbf{h}_1 = a_2 d_2 \mathbf{h}$ and $\mathbf{h}_2 = a_1 d_1 \mathbf{h}$. Then $\mathbf{h}_1 + \mathbf{h}_2 = \mathbf{h}$, $o(\mathbf{h}_1) = d_1$, and $o(\mathbf{h}_2) = d_2$. Suppose that $\mathbf{h} = \mathbf{h}_1' + \mathbf{h}_2'$ where $o(\mathbf{h}_1') = d_1$ and $o(\mathbf{h}_2') = d_2$. Then $\mathbf{h}_1 - \mathbf{h}_1' = \mathbf{h}_2' - \mathbf{h}_2$. Hence $\mathbf{h}_1 - \mathbf{h}_1' = a_1 d_1 (\mathbf{h}_1 - \mathbf{h}_1') + a_2 d_2 (\mathbf{h}_2 - \mathbf{h}_2') = 0$ and so $\mathbf{h}_1 = \mathbf{h}_1'$ and $\mathbf{h}_2 = \mathbf{h}_2'$.

Theorem 5.2 *The function* $S_{\mathbf{g}}(d)$ *is multiplicative.*

Proof: If $\gcd(d_1, d_2) = 1$, then, by Lemma 5.3,

$$S_{\mathbf{g}}(d_1 d_2) = \sum_{\substack{\mathbf{h}\in G \\ o(\mathbf{h})=d_1 d_2}} \langle \mathbf{h}, \mathbf{g}\rangle = \sum_{\substack{\mathbf{h}_1,\mathbf{h}_2\in G \\ o(\mathbf{h}_1)=d_1 \\ o(\mathbf{h}_2)=d_2}} \langle \mathbf{h}_1, \mathbf{g}\rangle \langle \mathbf{h}_2, \mathbf{g}\rangle = S_{\mathbf{g}}(d_1) S_{\mathbf{g}}(d_2).$$

Definition 5.3 *For* $\mathbf{g} \in G$ *let*

$$W_{\mathbf{g}}(n) := \sum_{cd=n} S_{\mathbf{g}}(d),$$

$$W_{\mathbf{g}}^{(j)}(n) := \sum_{\substack{h_j \in G_j \\ o(h_j) | n}} \zeta_j^{g_j h_j}.$$

Lemma 5.4 *For* $\mathbf{g} \in G$ *we have*

(i) $W_{\mathbf{g}}$ *is multiplicative,*
(ii) $S_{\mathbf{g}}(d) = \sum_{mn=d} \mu(m) W_{\mathbf{g}}(n)$,
(ii) $W_{\mathbf{g}}(n) = \prod_{j=1}^{J} W_{\mathbf{g}}^{(j)}(n)$,
(iv) *if* q *is a prime, then*

$$W_{\mathbf{g}}^{(j)}(q^\delta) = 1 \text{ if } q \neq p_j,$$
$$W_{\mathbf{g}}^{(j)}(p_j^\delta) = p_j^{\min(\delta,\alpha_j)} \text{ if } \delta \leq \epsilon_j, \qquad = 0 \text{ if } \delta > \epsilon_j,$$

where $p_j^{\epsilon_j}$ *is the exact power of* p *dividing* g_j *(we put* $\epsilon_j = \infty$ *if* $g_j = 0$*).*

Proof: (i) follows from the definition and Theorem 5.2.
(ii) follows by the Moebius' inversion formula.
(iii). We have

$$\begin{aligned} W_{\mathbf{g}}(n) &= \sum_{d|n} \sum_{\substack{\mathbf{h}\in G \\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g} \rangle = \sum_{\substack{\mathbf{h}\in G \\ o(\mathbf{h})|n}} \prod_{j=1}^{J} \zeta_j^{g_j h_j} \\ &= \prod_{j=1}^{J} \sum_{\substack{h_j\in G_j \\ o(h_j)|n}} \zeta_j^{g_j h_j} = \prod_{j=1}^{J} W_{\mathbf{g}}^{(j)}(n). \end{aligned}$$

(iv). First we note that if $q \neq p_j$, then $o(h_j)|q^\delta$ if and only if $o(h_j) = 1$, i.e. $h_j = 0$. Next, suppose $q = p_j$. If $\delta \geq \alpha_j$, then $o(h_j)|p_j^\delta$ for all $h_j \in G_j$. Hence

$$W_{\mathbf{g}}^{(j)}(p_j^\delta) = \sum_{h_j=1}^{p_j^{\alpha_j}} \zeta_j^{g_j h_j} = \begin{cases} p_j^{\alpha_j} & \text{if } g_j = 0 \quad (\text{i.e. } \epsilon_j = \infty), \\ 0 & \text{if } g_j \neq 0 \quad (\text{i.e. } \epsilon_j < \infty). \end{cases}$$

If $\delta < \alpha_j$, then $o(h_j)|p_j^\delta$ if and only if $p_j^{\alpha_j-\delta}|h_j$. Hence

$$W_{\mathbf{g}}^{(j)}(p_j^\delta) = \sum_{h=1}^{p_j^\delta} \zeta_j^{g_j p_j^{\alpha_j-\delta} h} = \begin{cases} p_j^{\alpha_j} & \text{if } g_j \equiv 0 \pmod{p_j^\delta} \quad (\text{i.e. } \epsilon_j \geq \delta), \\ 0 & \text{if } g_j \not\equiv 0 \pmod{p_j^\delta} \quad (\text{i.e. } \epsilon_j < \delta). \end{cases}$$

Therefore, if $\delta > \epsilon_j$ (in which case $\epsilon_j < \infty$), then $W_{\mathbf{g}}^{(j)}(p_j^\delta) = 0$, and if $\delta \leq \epsilon_j$, then

$$W_{\mathbf{g}}^{(j)}(p_j^\delta) = \begin{cases} p_j^\delta & \text{if } \delta < \alpha_j, \\ p_j^{\alpha_j} & \text{if } \delta \geq \alpha_j. \end{cases}$$

Definition 5.4 *(i) For $p|N$ let $\mathcal{J}_p := \{j \mid p_j = p\}$,*
(ii) for $\mathbf{g} \in G$ and $p|N$, let

$$e_p(\mathbf{g}) := \begin{cases} \min_{j\in\mathcal{J}_p} \epsilon_j & \text{if } \epsilon_j \neq \infty \text{ for some } j \in \mathcal{J}_p, \\ \max_{j\in\mathcal{J}_p} \alpha_j & \text{if } \epsilon_j = \infty \text{ for all } j \in \mathcal{J}_p, \end{cases}$$

and

$$\Delta(\mathbf{g}) := \prod_{p|N} p^{e_p(\mathbf{g})}$$

where ϵ_j is defined in Lemma 5.4(iv).
(iii) Let

$$W_G\left(\prod_p p^{\delta_p}\right) := \prod_{p|N} p^{\sum_{j\in\mathcal{J}} \min(\delta_p,\alpha_j)}.$$

Theorem 5.3 *For $\mathbf{g} \in G$ we have*

$$S_{\mathbf{g}}(d) = \sum_{\substack{mn=d \\ n|\Delta(\mathbf{g})}} \mu(m) W_G(n).$$

Proof: By Theorem 5.2 and Lemma 5.4, if $n = \prod_p p^{\delta_p}$, then

$$W_{\mathbf{g}}(n) = \prod_p W_{\mathbf{g}}\left(p^{\delta_p}\right) = \prod_p \prod_{j=1}^{J} W_{\mathbf{g}}^{(j)}\left(p^{\delta_p}\right) = \prod_p \prod_{j\in\mathcal{J}} W_{\mathbf{g}}^{(j)}\left(p^{\delta_p}\right).$$

Hence

$$W_{\mathbf{g}}(n) = \prod_p \prod_{j\in\mathcal{J}} W_{\mathbf{g}}^{(j)}\left(p^{\min(\delta_p,\alpha_j)}\right)$$

if $p^{\delta_p}|g_j$ for all p and all $j \in \mathcal{J}_p$, and $W_{\mathbf{g}}(n) = 0$ otherwise, i.e.

$$W_{\mathbf{g}}(n) = \begin{cases} W_G(n) & \text{if } n|\Delta(\mathbf{g}), \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$S_{\mathbf{g}}(d) = \sum_{mn=d} \mu(m) W_{\mathbf{g}}(n) = \sum_{\substack{mn=d \\ n|\Delta(\mathbf{g})}} \mu(m) W_G(n).$$

5.4 Maximal Constantin-Rao codes

Theorem 5.4 *If* $\mathbf{g}, \mathbf{h} \in G$ *and* $\Delta(\mathbf{g})|\Delta(\mathbf{h})$, *then* $\#C_{\mathbf{g}} \leq \#C_{\mathbf{h}}$.
In particular, $\#C_{\mathbf{1}} \leq \#C_{\mathbf{g}} \leq \#C_{\mathbf{0}}$ *for all* $\mathbf{g} \in G$.

Proof: By Theorems 5.1 and 5.3 we have

$$\#C_{\mathbf{g}} = \frac{1}{2N} \sum_{\substack{cd=N \\ d \ odd}} 2^c \sum_{\substack{mn=d \\ n|\Delta(\mathbf{g})}} \mu(m) W_G(n) = \frac{1}{2N} \sum_{\substack{en=N \\ n|\Delta(\mathbf{g}) \\ n \ odd}} W_G(n) \sum_{\substack{mc=e \\ m \ odd}} \mu(m) 2^c.$$

Since

$$\sum_{\substack{mc=e \\ m \ odd}} \mu(m) 2^c > 0$$

for all e, the first part of the theorem follows. Further, $\Delta(\mathbf{1}) = 1$ and $\Delta(\mathbf{0}) = \prod_p p^{\max_{j\in\mathcal{J}} \alpha_j}$. Hence $\Delta(\mathbf{1})|\Delta(\mathbf{g})|\Delta(\mathbf{0})$ for all $\mathbf{g} \in G$.

Theorem 5.5 *If* $\mathbf{g} \in G = \mathbb{Z}_{p^{\alpha+\beta}} \oplus H$, $\mathbf{g}' \in G' = \mathbb{Z}_{p^\alpha} \oplus \mathbb{Z}_{p^\beta} \oplus H$, *and* $\Delta(\mathbf{g}) = \Delta(\mathbf{g}')$, *then* $\#C_{\mathbf{g}} \leq \#C_{\mathbf{g}'}$.

Proof: First we note that

$$\frac{W_{G'}\left(\prod_p p^{\delta_p}\right)}{W_G\left(\prod_p p^{\delta_p}\right)} = \frac{q^{\min(\delta_q,\alpha)+\min(\delta_q,\beta)}}{q^{\min(\delta_q,\alpha+\beta)}} \geq 1.$$

Hence

$$\#C_{\mathbf{g}'} - \#C_{\mathbf{g}} = \frac{1}{2N} \sum_{\substack{en=N \\ n|\Delta(\mathbf{g}) \\ n \ odd}} \left\{W_{G'}(n) - W_G(n)\right\} \sum_{\substack{mc=e \\ m \ odd}} \mu(m) 2^c \geq 0.$$

Theorem 5.6 *The largest Constantin-Rao code of length* $N-1$ *is the code* $C_{\mathbf{0}}$ *based on the group*

$$G = \bigoplus_{p|N} \bigoplus_{i=1}^{n_p} \mathbb{Z}_p$$

where $N = \prod_{p|N} p^{n_p}$.

5.5 Shortened Constantin-Rao codes

Definition 5.5 *Let G be an Abelian group. For $\mathbf{g} \in G$ and $1 \le j \le N-1$, let*

$$C_{\mathbf{g}}^{(j)} = \Big\{ (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_{N-1}) \;\Big|\; (x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_{N-1}) \in C_{\mathbf{g}} \Big\}$$

where $C_{\mathbf{g}}$ is the Constantin-Rao code.

Remark. We can in the same way construct a shortened code from all codewords in $C_{\mathbf{g}}$ having a 1 in position j. However, this is the same code as $C_{\mathbf{g}-\mathbf{g}_j}^{(j)}$, and we get nothing new.

Definition 5.6 *For $\mathbf{g} \in G$, $1 \le j \le N-1$, and $0 \le w \le N-2$, let*

$$t^{(j)}(\mathbf{g}, w) := \#\Big\{ \mathbf{x} \in C_{\mathbf{g}}^{(j)} \;\Big|\; w(\mathbf{x}) = w \Big\},$$

$$T_{\mathbf{g}}^{(j)}(y) := \sum_{w=0}^{N-2} t^{(j)}(\mathbf{g}, w) y^w.$$

We shall determine $T_{\mathbf{g}}^{(j)}(y)$. The determination is similar to the determination of $T_{\mathbf{g}}(y)$ and so parts of the proof is sketchy.

Theorem 5.7 *For $\mathbf{g} \in G$ and $1 \le j \le N-1$ we have*

$$T_{\mathbf{g}}^{(j)}(y) = \frac{1}{N(1+y)} \sum_{cd=N} \Big(1 - (-1)^d\Big)^{c-1} \sum_{m=0}^{d-1} (-y)^m S_{\mathbf{g}-m\mathbf{g}_j}(d).$$

Proof: Let

$$\hat{T}_{\mathbf{h}}^{(j)}(y) := \sum_{\mathbf{g} \in G} \langle -\mathbf{h}, \mathbf{g} \rangle T_{\mathbf{g}}^{(j)}(y).$$

Then

$$T_{\mathbf{g}}^{(j)}(y) = \frac{1}{N} \sum_{\mathbf{h} \in G} \langle \mathbf{h}, \mathbf{g} \rangle \hat{T}_{\mathbf{h}}^{(j)}(y).$$

Further

$$\hat{T}_{\mathbf{h}}^{(j)}(y) = \prod_{\substack{i=1 \\ i \ne j}}^{N-1} \Big(1 + \langle -\mathbf{h}, \mathbf{g}_i \rangle y\Big) = \frac{\Big(1 - (-y)^{o(\mathbf{h})}\Big)^{N/o(\mathbf{h})}}{(1+y)(1 + \langle -\mathbf{h}, \mathbf{g}_j \rangle y)}.$$

Hence

$$T_{\mathbf{g}}^{(j)}(y) = \frac{1}{N(1+y)} \sum_{cd=N} \left(1-(-1)^d\right)^{c-1} \sum_{\substack{\mathbf{h}\in G\\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g}\rangle \frac{1-(-y)^d}{1+\langle -\mathbf{h}, \mathbf{g}_j\rangle y}$$

and

$$\begin{aligned}
\sum_{\substack{\mathbf{h}\in G\\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g}\rangle \frac{1-(-y)^d}{1+\langle -\mathbf{h}, \mathbf{g}_j\rangle y} &= \sum_{\substack{\mathbf{h}\in G\\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g}\rangle \sum_{m=0}^{d-1} \langle -\mathbf{h}, \mathbf{g}_j\rangle^m (-y)^m \\
&= \sum_{m=0}^{d-1} (-y)^m \sum_{\substack{\mathbf{h}\in G\\ o(\mathbf{h})=d}} \langle \mathbf{h}, \mathbf{g} - m\mathbf{g}_j\rangle \\
&= \sum_{m=0}^{d-1} (-y)^m S_{\mathbf{g}-m\mathbf{g}_j}(d).
\end{aligned}$$

5.6 Notes

The exposition in 5.1-5.5 is based on Helleseth and Kløve [48].

5.2. The exposition follows closely McEliece and Rodemich [64].

5.3. The function $S_{\mathbf{g}}$ for $\mathbf{g}$ in a cyclic group is known as von Sterneck's function. For the cyclic group case, Theorem 5.1 and Theorem 5.3 were proved by von Sterneck, see Bachmann [3, Chapter 5]. The results have been rediscovered several times. Ginzburg [40] gave the expressions for $\#C_{\mathbf{g}}$ and Stanley and Yoder [76] and Mazur [62] gave the weight distribution. Dynkin and Togonidze [22] determined the size and weight distribution of $C_{\mathbf{0}}$ when G is the additive group of a finite field and $n = \#G - 1$. The expression for $S_{\mathbf{g}}$ in general was given by Helleseth and Kløve [48].

5.4. Constantin and Rao [14] proved that $\#C_{\mathbf{g}} \leq \#C_{\mathbf{0}}$, and they conjectured Theorem 5.6. The first proof of Theorem 5.6 was given by McEliece and Rodemich [64].

5.5. Theorem 5.7 has not been published before.

Chapter 6

Generalized Varshamov codes

6.1 Preliminaries

Varshamov gave several classes of codes to correct multiple errors. Varshamov's constructions have been generalized in various ways and we shall describe these generalizations below. A common feature for all these codes is that a vital step in the decoding is the solution of an equation over a finite field. Therefore we start by summarizing a few facts about finite fields and equations.

Throughout the chapter, p denotes a prime, q a power of p, and F_q is the finite field with q elements (it is unique up to isomorphism). $F_q^* := F_q \setminus \{0\}$ is a cyclic group under multiplication, i.e. there exists a (*primitive element*) $\beta \in F_q$ such that $\beta^{q-1} = 1$ and $\beta^i \neq 1$ for $0 < i < q-1$. If β is a primitive element of F_{q^t} and $m := (q^t - 1)/(q-1)$, then $\gamma := \beta^m$ is a primitive element of F_q. Moreover, $1, \beta, \beta^2, \ldots, \beta^{t-1}$ is a basis for F_{q^t} as a vector space over F_q.

Let $x, u_1, u_2, \ldots, u_r$ be variables and $\mathbf{u} := (u_1, u_2, \ldots, u_r)$. The elementary symmetric functions $\sigma_l(\mathbf{u})$, $l = 0, 1, 2, \ldots$ are defined by

$$\prod_{i=1}^{r}(x + u_i) = \sum_{l=0}^{\infty} \sigma_l(\mathbf{u}) x^{r-l}.$$

Alternatively,

$$
\begin{aligned}
&\sigma_0(\mathbf{u}) = 1,\\
&\sigma_l(\mathbf{u}) = \sum_{1\le i_1<i_2<\cdots<i_l\le r} u_{i_1}u_{i_2}\cdots u_{i_l} \text{ for } 1 \le l \le r,\\
&\sigma_l(\mathbf{u}) = 0 \text{ for } l > r.
\end{aligned}
$$

If $I = \{i_1, i_2, \ldots, i_s\}$, where $1 \le i_1 < i_2 < \cdots < i_s \le r$, then $\mathbf{u}_I := (u_{i_1}, u_{i_2}, \ldots, u_{i_s})$.

Lemma 6.1 *(i) If π is a permutation of $\{1, 2, \ldots, r\}$, then*

$$\sigma_l(u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(r)}) = \sigma_l(\mathbf{u}).$$

(ii) We have

$$\sigma_l\Big((\mathbf{u}|\mathbf{0})\Big) = \sigma_l(\mathbf{u}).$$

(iii) If $\{I, J\}$ is a partition of $\{1, 2, \ldots, r\}$, then

$$\sigma_l(\mathbf{u}) = \sum_{j=0}^{l} \sigma_j(\mathbf{u}_I)\sigma_{l-j}(\mathbf{u}_J).$$

(iv) We have

$$\sigma_l(\mathbf{u}+\mathbf{1}) = \sum_{j=0}^{l} \binom{r-j}{r-l}\sigma_j(\mathbf{u}).$$

Proof: (i).

$$\sum_{l=0}^{\infty} \sigma_l(u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(r)})x^{r-l} = \prod_{i=1}^{r}(x+u_{\pi(i)}) = \prod_{i=1}^{r}(x+u_i) = \sum_{l=0}^{\infty}\sigma_l(\mathbf{u})x^{r-l}.$$

(ii) If $\mathbf{0} \in \{0\}^s$, then

$$\sum_{l=0}^{\infty}\sigma_l\Big((\mathbf{u}|\mathbf{0})\Big)x^{r+s-l} = \Big(\prod_{i=1}^{r}(x+u_i)\Big)(x+0)^s = \sum_{l=0}^{\infty}\sigma_l(\mathbf{u})x^{r-l+s}.$$

(iii) If $\#I = s$ and $\#J = r-s$, then

$$
\begin{aligned}
\sum_{l=0}^{\infty}\sigma_l(\mathbf{u})x^{r-l} &= \prod_{i\in I}(x+u_i)\prod_{i\in J}(x+u_i)\\
&= \sum_{j=0}^{\infty}\sigma_j(\mathbf{u}_I)x^{s-j}\sum_{k=0}^{\infty}\sigma_k(\mathbf{u}_J)x^{r-s-k}\\
&= \sum_{l=0}^{\infty}x^{r-l}\sum_{j+k=l}\sigma_j(\mathbf{u}_I)\sigma_k(\mathbf{u}_J).
\end{aligned}
$$

$$
\begin{aligned}
(iv) \quad \sum_{l=0}^{\infty}\sigma_l(\mathbf{u}+\mathbf{1})x^{r-l} &= \prod_{i=1}^{r}(x+u_i+1) = \prod_{i=1}^{r}((x+1)+u_i) \\
&= \sum_{j=0}^{r}\sigma_j(\mathbf{u})(x+1)^{r-j} = \sum_{j=0}^{r}\sigma_j(\mathbf{u})\sum_{k=0}^{r-j}\binom{r-j}{k}x^k \\
&= \sum_{l=0}^{r}x^{r-l}\sum_{j=0}^{l}\sigma_j(\mathbf{u})\binom{r-j}{r-l}.
\end{aligned}
$$

The lemma follows by equating coefficients in each case. □

Let $\alpha_1, \alpha_2, \ldots, \alpha_r \in F_q$. If $a_l := \sigma_l(\alpha_1, \alpha_2, \ldots, \alpha_r)$ for $l = 0, 1, \ldots, r$, then the equation

$$\sum_{l=0}^{r}(-1)^l a_{r-l}x^l = 0$$

has as its roots exactly the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$. Hence we recover this set by solving the equation. Moreover, to find the roots of an equation over F_q is a finite job since there are exactly q possibilities.

6.2 First construction

Let G be a finite Abelian group.

Definition 6.1 *A V_t-set in G is a subset $H = \{h_1, h_2, \ldots, h_n\}$ of G containing n elements such that all the sums*

$$h_{i_1} + h_{i_2} + \cdots + h_{i_r},$$

where $1 \le i_1 < i_2 < \cdots < i_r \le n$, $0 \le r \le t$, are distinct.

Code construction.

Let H be a V_t-set in G. For $g \in G$, let

$$C_g := \left\{(x_1, x_2, \ldots, x_n) \in \{0,1\}^n \;\middle|\; \sum_{i=1}^{n} x_i h_i = g\right\}.$$

Remark. The codes based on the V_1-set $G \setminus \{0\}$ are the Constantin-Rao codes.

Decoding algorithm.

Comment: The received vector is $\mathbf{y}$.
Comment: $\mathbf{e}_i$ is the i'th unit vector for $i > 0$, $\mathbf{e}_0 = \mathbf{0}$.

$c := g - \sum_{i=1}^{n} y_i h_i$;
if $(\exists i_1 < i_2 < \cdots < i_r$ such that $h_{i_1} + h_{i_2} + \cdots + h_{i_r} = c)$
  then decode into $\mathbf{y} + \sum_{j=1}^{r} \mathbf{e}_{i_j}$
  else decoding has failed.

Proof of decoding algorithm.

Let $\mathbf{x}$ be sent. Suppose errors occur in positions $j_1, j_2, \ldots, j_s$ where $0 \leq s \leq t$, $1 \leq j_1 < j_2 < \cdots < j_s \leq n$. Then

$$y_j = x_j \quad \text{for} \quad j \notin \{j_1, j_2, \ldots, j_s\},$$
$$y_j = 0 \text{ and } x_j = 1 \quad \text{for} \quad j \in \{j_1, j_2, \ldots, j_s\}.$$

Hence $c = h_{j_1} + h_{j_2} + \cdots + h_{j_s}$, and we decode into $\mathbf{y} + \sum_{l=1}^{s} \mathbf{e}_{j_l} = \mathbf{x}$.

Remarks. (i) The actual design of the if-part of the decoding algorithm will in each case depend on the structure of the set H. Below we give a couple of constructions of V_t-sets, these have the stronger property that the sums $h_{i_1} + h_{i_2} + \cdots + h_{i_r}$ where $1 \leq i_1 \leq i_2 \leq \cdots \leq i_r \leq n$ (i.e. repetitions are allowed), $0 \leq r \leq t$, are distinct.
(ii) If H is a V_t-set, then so is any subset H' of H. Starting with H', we may construct a class of codes C'_g of length $n' := \#H'$. The properties of these codes will depend on H' as well as g. In particular, $\max_{g \in G} C'_g$ will depend, not only on n' and H, but also on H'. Almost nothing is known on how to choose H' so as to maximize $\max_{g \in G} C'_g$.

Bound on the size of C_g

Since $\{C_g \mid g \in G\}$ is a partition of $\{0,1\}^n$, we get

$$\max_{g \in G} C_g \geq \frac{2^n}{\#G}.$$

6.3 Two V_t-sets

First construction

Let $G := \mathbb{Z}_{q^t-1}$. Let β be a primitive element in F_{q^t}. Let

$$H = \left\{ h \in G \;\middle|\; \beta^{h+1} - \beta \in F_q^* \right\}.$$

Then $n := \#H = q - 1$.

We use the notation

$$\alpha_g := \beta^{g+1} - \beta \text{ for all } g \in G.$$

We define the function $L : F_q \to F_q$ by

$$\beta^{L(\alpha)+1} = \beta + \alpha.$$

Decoding algorithm.

Comment: c is the sum whose addends we shall determine.

determine $(\gamma_0, \gamma_1, \ldots, \gamma_{t-1}) \in F_q^t$ by $\beta^{c+t} - \beta^t = \sum_{j=0}^{t-1} \gamma_j \beta^j$;
$\{\delta_1, \delta_2, \ldots, \delta_t\} := \left\{ \delta \;\middle|\; \delta^t + \sum_{j=0}^{t-1} (-1)^j \gamma_j \delta^j = 0 \right\}$;
$z_i := L(\delta_i)$ for $i = 1, 2, \ldots, t$;
$\{h_{i_1}, h_{i_2}, \ldots, h_{i_r}\} := \{z_i \mid 1 \le i \le t,\ z_i \ne 0\}$.

Proof of the decoding algorithm.

Let $\mathcal{I} := \{j_1, j_2, \ldots, j_s\}$ where $1 \le j_1 \le j_2 \le \cdots \le j_s \le n$, $0 \le s \le t$, and $c := \sum_{j \in \mathcal{I}} h_j$. Let

$$(z'_0, z'_1, \ldots, z'_t) := (h_{j_1}, h_{j_2}, \ldots, h_{j_s}, 0, 0, \ldots, 0).$$

Then $\sum_{k=1}^t (z'_k + 1) = c + t$. Hence

$$\beta^{c+t} = \prod_{k=1}^t \beta^{z'_k+1} = \prod_{k=1}^t (\beta + \alpha_{z'_k}) = \beta^t + \sum_{l=0}^{t-1} \sigma_{t-l}(\alpha) \beta^l.$$

Therefore, $\gamma_j = \sigma_{t-l}(\alpha)$ for $j = 0, 1, \ldots, t-1$, where $\alpha = (\alpha_{z'_1}, \alpha_{z'_2}, \ldots, \alpha_{z'_t})$. Hence $\delta_1, \delta_2, \ldots, \delta_t$ are the same as $\alpha_{z'_1}, \alpha_{z'_2}, \ldots, \alpha_{z'_t}$ is some order and so $\{z_1, z_2, \ldots, z_t\} = \{z'_1, z'_2, \ldots, z'_t\}$ and $\{h_{i_1}, h_{i_2}, \ldots, h_{i_r}\} = \{h_{j_1}, h_{j_2}, \ldots, h_{j_s}\}$.

Second construction

Let $m := (q^{t+1}-1)/(q-1)$ and $G := \mathbb{Z}_m$. Let β be a primitive root in $F_{q^{t+1}}$. Let

$$H := \left\{ h \;\middle|\; 0 < h < m \text{ and } \exists \alpha_h^{(0)}, \alpha_h^{(1)} \in F_q \text{ such that } \beta^h = \alpha_h^{(0)} + \alpha_h^{(1)}\beta \right\}.$$

Then $n := \#H = q$.

We define the function Λ by

$$\Lambda(\alpha_h^{(0)}/\alpha_h^{(1)}) = h.$$

Decoding algorithm.

Comment: c is the sum whose addends we shall determine.

determine $(\gamma_0, \gamma_1, \dots, \gamma_t) \in F_q^t$ by $\beta^c = \sum_{j=0}^t \gamma_j \beta^j$;
$\{\delta_1, \delta_2, \dots, \delta_r\} := \left\{ \delta \;\middle|\; \sum_{j=0}^t (-1)^j \gamma_j \delta^j = 0 \right\}$;
$h_{i_k} := \Lambda(\delta_k)$ for $k = 1, 2, \dots, r$.

Proof of the decoding algorithm.

Let $\mathcal{I} := \{j_1, j_2, \dots, j_s\}$ where $1 \le j_1 \le j_2 \le \dots \le j_s \le n$, $0 \le s \le t$, and $c := \sum_{j \in \mathcal{I}} h_j$. Since $\beta^h \notin F_q$ when $0 < h < m$, $\alpha_h^{(1)} \neq 0$. Hence

$$\begin{aligned}
\beta^c = \prod_{j \in \mathcal{I}} \beta^{h_j} &= \prod_{j \in \mathcal{I}} \left(\alpha_{h_j}^{(0)} + \alpha_{h_j}^{(1)} \beta \right) \\
&= \prod_{j \in \mathcal{I}} \alpha_{h_j}^{(1)} \prod_{j \in \mathcal{I}} \left(\beta + \alpha_{h_j}^{(0)}/\alpha_{h_j}^{(1)} \right) \\
&= \prod_{j \in \mathcal{I}} \alpha_{h_j}^{(1)} \sum_{l=0}^{s} \sigma_{s-l}(\alpha) \beta^l,
\end{aligned}$$

where $\alpha := (\alpha_{h_j}^{(0)}/\alpha_{h_j}^{(1)})_{j \in \mathcal{I}}$. Hence

$$\gamma_l = \begin{cases} \sigma_{s-l}(\alpha) \prod_{j \in \mathcal{I}} \alpha_{h_j}^{(1)} & \text{for } 0 \le l \le s, \\ 0 & \text{for } s < l \le t. \end{cases}$$

Therefore,

$$\sum_{l=0}^{t} \gamma_l z^l = \prod_{j\in\mathcal{I}} \alpha_{h_j}^{(1)} \Big(\sum_{l=0}^{s} \sigma_{s-l}(\alpha) z^l\Big),$$

and so the roots of $\sum_{l=0}^{t}(-1)^l \gamma_l z^l = 0$ are $\Big\{\alpha_{h_j}^{(0)}/\alpha_{h_j}^{(1)} \ \Big|\ j \in \mathcal{I}\Big\}$. Hence

$$\{\Lambda(\delta_k) \mid k = 1, 2, \ldots, r\} = \{h_j \mid j \in \mathcal{I}\}.$$

6.4 Second construction

Code construction

Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct non-zero elements of F_q and $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_n)$. For $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$, let $\mathbf{x}\alpha = (x_1\alpha_1, x_2\alpha_2, \ldots, x_n\alpha_n)$. For $g_1, g_2, \ldots, g_t \in F_q$, let

$$C_{g_1,g_2,\ldots,g_t} := \Big\{\mathbf{x} \in \{0,1\}^n \ \Big|\ \sigma_l(\mathbf{x}\alpha) = g_l \text{ for } 1 \le l \le t\Big\}.$$

Decoding algorithm.

Comment: The received vector is $\mathbf{y}$.
Comment: $\mathbf{e}_j$ is the j'th unit vector.

$h_l := \sigma_l(\mathbf{y}\alpha)$ for $l = 0, 1, \ldots, t$;
$A_0 := 1$;
for $l := 1$ **to** t **do** $A_l := g_l - \sum_{j=1}^{l} h_j A_{l-j}$;
$\{\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_r}\} := \Big\{z \neq 0 \ \Big|\ \sum_{l=0}^{t}(-1)^l A_{t-l} z^l = 0\Big\}$;
decode into $\mathbf{y} + \sum_{k=1}^{r} \mathbf{e}_{i_k}$.

Proof of decoding algorithm.

Let $\mathbf{x} \in C_{g_1,g_2,\ldots,g_t}$ be sent. Suppose errors occurs in positions $j_1, j_2, \ldots, j_s$ where $0 \le s \le t$ and $1 \le j_1 < j_2 < \cdots < j_s \le n$. Let $\mathcal{J} := \{j_1, j_2, \ldots, j_s\}$ and $\mathcal{I} := \{1, 2, \ldots, n\} \setminus \mathcal{J}$. Then $\mathbf{x}\alpha_{\mathcal{J}} = \alpha_{\mathcal{J}}$,

$$g_l = \sigma_l(\mathbf{x}\alpha) = \sigma_l(\mathbf{x}\alpha_{\mathcal{I}}|\mathbf{x}\alpha_{\mathcal{J}}) \text{ for } 1 \le l \le t,$$

and

$$h_l = \sigma_l(\mathbf{y}\alpha) = \sigma_l(\mathbf{x}\alpha_{\mathcal{I}}|\mathbf{0}) = \sigma_l(\mathbf{x}\alpha_{\mathcal{I}}) \text{ for } 1 \leq l \leq t.$$

By Lemma 6.1(iii), $g_l = \sum_{j=0}^{l} h_j \sigma_{l-j}(\alpha_{\mathcal{J}})$. Hence $A_l = \sigma_l(\alpha_{\mathcal{J}})$ for $0 \leq l \leq t$. Therefore $\{\alpha_{j_1}, \alpha_{j_2}, \ldots, \alpha_{j_s}\}$ are the non-zero roots of $\sum_{l=0}^{t}(-1)^l A_{t-l} z^l = 0$ and we decode into $\mathbf{y} + \sum_{k=1}^{s} \mathbf{e}_{j_k} = \mathbf{x}$.

Bound on the size of the code.

Since $\left\{ C_{g_1,g_2,\ldots,g_t} \;\middle|\; g_1, g_2, \ldots, g_t \in F_q \right\}$ is a partition of $\{0,1\}^n$, we have

$$\max \#C \geq \frac{2^n}{q^t}.$$

6.5 Third construction

Code construction

Let β be a primitive element of F_q. Let $a_1, a_2, \ldots, a_n$ be distinct non-zero elements of $\mathbb{Z}_{q-1}$, and let $\alpha_i := \beta^{a_i} - 1$ for $i = 1, 2, \ldots, n$, $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\mathbf{a} := (a_1, a_2, \ldots, a_n)$. For $g_1, g_2, \ldots, g_{t-1} \in F_q$ and $m \in \mathbb{Z}_{q-1}$, let

$$C_{g_1,g_2,\ldots,g_{t-1},m} := \left\{ \mathbf{x} \in \{0,1\}^n \;\middle|\; \sigma_l(\mathbf{x}\alpha) = g_l \text{ for } 1 \leq l \leq t-1,\ \sigma_1(\mathbf{x}\mathbf{a}) = m \right\}.$$

Decoding algorithm.

Comment: The received vector is $\mathbf{y}$.
Comment: $\mathbf{e}_j$ is the j'th unit vector.

$h_l := \sigma_l(\mathbf{y}\alpha)$ for $l = 0, 1, \ldots, t-1$;
$M := \sigma_1(\mathbf{y}\mathbf{a})$;
$A_0 := 1$;
for $l := 1$ **to** $t-1$ **do** $A_l := g_l - \sum_{j=1}^{l} h_j A_{l-j}$;
for $l := 0$ **to** $t-1$ **do** $B_l := \sum_{j=0}^{l} \binom{t-j}{t-l} A_j$;
$B_t := \beta^{m-M}$;
$\{a_{i_1}, a_{i_2}, \ldots, a_{i_r}\} := \left\{ z \neq 0 \;\middle|\; \sum_{l=0}^{t}(-1)^l B_{t-l} \beta^{zl} = 0 \right\}$;
decode into $\mathbf{y} + \sum_{k=1}^{r} \mathbf{e}_{i_k}$.

Proof of decoding algorithm.

Let $\mathbf{x} \in C_{g_1,g_2,\dots,g_{t-1},m}$ be sent. Suppose errors occurs in positions $j_1, j_2, \dots, j_s$ where $0 \le s \le t$ and $1 \le j_1 < j_2 < \dots < j_s \le n$. Let $\mathcal{J} := \{j_1, j_2, \dots, j_s\}$ and $\mathcal{I} := \{1, 2, \dots, n\} \setminus \mathcal{J}$. Then $\mathbf{x}\alpha_{\mathcal{J}} = \alpha_{\mathcal{J}}$,

$$g_l = \sigma_l(\mathbf{x}\alpha) = \sigma_l(\mathbf{x}\alpha_{\mathcal{I}}|\mathbf{x}\alpha_{\mathcal{J}}) \text{ for } 1 \le l \le t-1,$$

and

$$h_l = \sigma_l(\mathbf{y}\alpha) = \sigma_l(\mathbf{x}\alpha_{\mathcal{I}}|\mathbf{0}) \text{ for } 1 \le l \le t-1.$$

By Lemma 6.1(iii), $g_l = \sum_{j=0}^{l} h_j \sigma_{l-j}(\alpha_{\mathcal{J}})$. Hence $A_l = \sigma_l(\alpha_{\mathcal{J}}|\mathbf{0})$ for $0 \le l \le t-1$, where $\mathbf{0} \in \{0\}^{t-s}$. By Lemma 6.1(iv)

$$\begin{aligned} B_l &= \sum_{j=0}^{t} \binom{t-j}{t-l} \sigma_j(\alpha_{\mathcal{J}}|\mathbf{0}) = \sigma_l(\alpha_{\mathcal{J}} + \mathbf{1}|\mathbf{1}) \\ &= \sigma_l\left(\beta^{a_{j_1}}, \beta^{a_{j_2}}, \dots, \beta^{a_{j_s}}, \beta^0, \beta^0, \dots, \beta^0\right) \end{aligned}$$

for $0 \le l \le t-1$. Further

$$B_t = \beta^{a_{j_1}+a_{j_2}+\dots+a_{j_s}} = \sigma_t\left(\beta^{a_{j_1}}, \beta^{a_{j_2}}, \dots, \beta^{a_{j_s}}, \beta^0, \beta^0, \dots, \beta^0\right).$$

Hence $\{i_1, i_2, \dots, i_r\} = \{j_1, j_2, \dots, j_s\}$ and we decode into $\mathbf{y} + \sum_{k=1}^{s} \mathbf{e}_{j_k} = \mathbf{x}$.

Bound on the size of the code.

Since $\left\{C_{g_1,g_2,\dots,g_{t-1},m} \;\middle|\; g_1, g_2, \dots, g_{t-1} \in F_q \text{ and } m \in \mathbb{Z}_{q-1}\right\}$ is a partition of $\{0,1\}^n$, we have

$$\max \#C \ge \frac{2^n}{(q-1)q^{t-1}}.$$

6.6 Lower bounds on $\alpha(n,t)$

Theorem 6.1 *(i) If n is a power of a prime, then*

$$\alpha(n,t) \ge \frac{2^n}{n^t + n^{t-1} + \dots + 1}.$$

(ii) If $n+1$ is a power of a prime, then

$$\alpha(n,t) \ge \frac{2^n}{(n+1)^t - 1}.$$

(iii) If q is the least prime power $\geq n+2$, then, for $t \geq 2$,

$$\alpha(n,t) \geq \frac{2^n}{q^t - q^{t-1}}.$$

Proof: (i) and (ii) follows from the code construction in 6.2, using the second and first V_t-set of 6.3 respectively, and (iii) follows from the code construction in 6.5.

Corollary 6.1 *For fixed t,*

$$\alpha(n,t) \geq \frac{2^n}{n^t}\Big(1 + o(1)\Big)$$

when $n \to \infty$.

Proof: from the theory of primes it is well known that there exists a $\theta < 1$ such that for all large n there exists a prime p such that $n < p < n + n^\theta$. Hence, if q is the least prime power $\geq n$, then $q < n + n^\theta$, and we get

$$\alpha(n,t) \geq \frac{2^n}{(n+n^\theta)^t - (n+n^\theta)^{t-1}} = \frac{2^n}{n^t}\Big(1 + o(1)\Big).$$

6.7 Notes

The basic idea of the code constructions in this chapter is due to Varshamov and Zograbjan [92] and Varshamov [88],[89], generalizing the 1-code construction of Varshamov and Tenengolts [91].

6.2. The general construction is due to Delsarte and Piret [17]. They also gave a general formula for the weight distribution of these codes.

6.3. The two V_t-sets are due to Bose and Chowla [9]. They were first applied to code construction by Graham and Sloane [45]. For $t = 2$, the first V_t-set was given by Singer [75] and applied to code construction by Varshamov [88].

6.4. If $t < p$, then $\left\{(g, g^2, \ldots, g^t) \,\middle|\, g \in F_q^*\right\}$ is a V_t set in F_q^t and the codes constructed from this V_t-set coincide with our construction in this case. This may be proved using Newton's equations connecting power sums and elementary symmetric functions. For $q = p$, this construction is due to Varshamov and Zograbjan [92] (for $t = 2$) and Varshamov [88] (for general t). Mazur [62] proved that if $q = p$ and $n = p - 1$, then

$$\left| \#C_{g_1, g_2, \ldots, g_t} - \frac{2^{p-1}}{p^t} \right| \leq \frac{2^{(p-1)/2}(p-1)\left(p^t e^{t\sqrt{p/2}} - 1\right)}{\sqrt{e}\, p^t \left(p\, e^{\sqrt{p/2}} - 1\right)}$$

for any choice of $g_1, g_2, \ldots, g_t$. For general q, Dynkin and Togonidze [21] gave a special case of the construction. For $t < p$ the general construction is due to Delsarte and Piret [17]. Graham and Sloane [45] used a closely related construction to construct constant weight codes.

6.5. If $t \leq p$, then $\left\{ \left(\beta^a - 1, \beta^{2a} - 1, \ldots, \beta^{(t-1)a} - 1, a\right) \,\middle|\, a \in \mathbb{Z}_{q-1} \right\}$ is a V_t-set in $F_q^{t-1} \times \mathbb{Z}_{q-1}$ and the codes constructed from this V_t-set coincides with our construction. For $q = p$, the construction is due to Varshamov and Zograbjan [92] (for $t = 2$) and Varshamov [88] (for general t). Decoding was discussed by Nalbandjan [65],[66].

Chapter 7

Other multiple error correcting codes

7.1 Modified Kim-Freiman codes

Code construction

Suppose $n = (t+1)m$ and let H be a code of length m having Hamming distance at least $2t+1$ between distinct code words, and $\mathbf{0} \in H$.

$$\begin{aligned} C \;\; := \;\; & \Big\{ (\mathbf{x}|\mathbf{x} \oplus \mathbf{h}_1| \cdots |\mathbf{x} \oplus \mathbf{h}_t) \;\Big|\; \mathbf{x} \in \{0,1\}^m,\ w(\mathbf{x}) \text{ even},\ \mathbf{h}_1, \mathbf{h}_1, \ldots, \mathbf{h}_1 \in H \Big\} \\ & \cup \Big\{ (\mathbf{x}|\mathbf{x}| \cdots |\mathbf{x}) \;\Big|\; \mathbf{x} \in \{0,1\}^m,\ w(\mathbf{x}) \text{ odd} \Big\}. \end{aligned}$$

Proof that C is a t-code

Let $\mathbf{u} = (\mathbf{x}|\mathbf{x} \oplus \mathbf{h}_1| \cdots |\mathbf{x} \oplus \mathbf{h}_t)$ and $\mathbf{u}' = (\mathbf{x}'|\mathbf{x}' \oplus \mathbf{h}'_1| \cdots |\mathbf{x}' \oplus \mathbf{h}'_t)$, where $\mathbf{u} \neq \mathbf{u}'$.
Case I, $\mathbf{h}_i \neq \mathbf{h}'_i$ for some i. Then

$$\begin{aligned} 2\Delta(\mathbf{u}, \mathbf{u}') & \geq d(\mathbf{u}, \mathbf{u}') \geq d(\mathbf{x}, \mathbf{x}') + d(\mathbf{x} \oplus \mathbf{h}_i, \mathbf{x}' \oplus \mathbf{h}'_i) \\ & = d(\mathbf{x}' \oplus \mathbf{h}_i, \mathbf{x} \oplus \mathbf{h}_i) + d(\mathbf{x} \oplus \mathbf{h}_i, \mathbf{x}' \oplus \mathbf{h}'_i) \\ & \geq d(\mathbf{x}' \oplus \mathbf{h}_i, \mathbf{x}' \oplus \mathbf{h}'_i) = d(\mathbf{h}_i, \mathbf{h}'_i) \geq 2t+1. \end{aligned}$$

Hence $\Delta(\mathbf{u}, \mathbf{u}') \geq t+1$.
Case II, $\mathbf{h}_i = \mathbf{h}'_i$ for all i, $\mathbf{h}_i \neq \mathbf{0}$ for some i. Then $\mathbf{x} \neq \mathbf{x}'$, and $w(\mathbf{x})$ and

$w(\mathbf{x}')$ are both even. Hence, putting $\mathbf{h}_0 = \mathbf{0}$, we get

$$2\Delta(\mathbf{u}, \mathbf{u}') \geq d(\mathbf{u}, \mathbf{u}') = \sum_{i=0}^{t} d(\mathbf{x} \oplus \mathbf{h}_i, \mathbf{x}' \oplus \mathbf{h}_i') = (t+1)\, d(\mathbf{x}, \mathbf{x}') \geq 2(t+1).$$

Case III, $\mathbf{h}_i = \mathbf{h}_i' = \mathbf{0}$ for all i. Then $\mathbf{x} \neq \mathbf{x}'$ and so

$$\Delta(\mathbf{u}, \mathbf{u}') = (t+1)\, \Delta(\mathbf{x}, \mathbf{x}') \geq t+1.$$

Size of the code.

Choosing H as large as possible, we get

$$\#C = 2^{m-1}\Big(A(m, 2t+1) + 1\Big).$$

Remark. Kim and Freiman proposed a larger code than the one we have given. E.g. for $t = 2$ and $n = 3m$, $m \geq 5$, they gave

$$C := \Big\{(\mathbf{x}|\mathbf{x} \oplus \mathbf{h}_1|\mathbf{x} \oplus \mathbf{h}_2) \;\Big|\; \mathbf{x} \in \{0,1\}^m,\ \mathbf{h}_1, \mathbf{h}_2 \in H\Big\}.$$

However, this is not a 2-code. To show this, let $\mathbf{h}_1 \neq \mathbf{h}_2$, and suppose they differ in position j, say. Then

$$\Delta\Big((\mathbf{0}|\mathbf{h}_1|\mathbf{h}_2), (\mathbf{e}_j|\mathbf{e}_j \oplus \mathbf{h}_1|\mathbf{e}_j \oplus \mathbf{h}_2)\Big) = 2.$$

7.2 Delsarte-Piret 2-codes

Delsarte and Piret gave constructions for 2-codes for $8 \leq n \leq 14$ and $n = 16$, and these are the largest known codes for theses parameters.

$n = 16$.

Let C' be the cyclic $[17, 8]$ code generated by $x^3 + x^5 + x^6 + x^7 + x^{10} + x^{11} + x^{12} + x^{14}$. Let

$$\begin{aligned} T \;:=\; \{ & x^8 + x^{10} + x^{11} + x^{14},\ x + x^8 + x^{10} + x^{13},\ x^2 + x^4 + x^{11} + x^{13}, \\ & x^2 + x^5 + x^7 + x^{14},\ x + x^4 + x^7 + x^9,\ x^5 + x^{11} + x^{12} + x^{15}, \\ & 1 + x^6 + x^{12} + x^{13},\ x^2 + x^3 + x^9 + x^{15},\ 1 + x^3 + x^4 + x^{10}, \\ & x^5 + x^6 + x^9 + x^{10}\}. \end{aligned}$$

Let
$$C := \left\{ \mathbf{x} \in \{0,1\}^{16} \;\middle|\; (\mathbf{x}|0) \in C' \cup T \text{ or } (\mathbf{x}|1) \in C' \cup T \right\}.$$
Then C is a 2-code and $\#C = 266$.

$n = 14$.

Let C' be the Nordstrom-Robinson code of length 16. Let $S_w(abc)$ denote the set of codewords in C' of weight$w$, having a, b and c as their last three coordinates. Similarly, $S_w(bc) := S_w(0bc) \cup S_w(1bc)$. Let

$$C'' := \{\mathbf{0},\mathbf{1}\} \cup S_6(00) \cup S_6(10) \cup S_6(11) \cup S_8(01) \cup S_8(10) \cup S_{10}(00) \cup S_{10}(01) \cup S_{10}(11).$$

Let
$$C := \left\{ \mathbf{x} \in \{0,1\}^{14} \;\middle|\; (\mathbf{x}|bc) \in C'' \text{ for some } b, c \in \{0,1\} \right\}.$$
Then C is a 2-code and $\#C = 186$.

$n = 13$.

Using the same notations as for $n = 14$, let

$$\begin{aligned} C''' \;:=\; & \{\mathbf{0},\mathbf{1}\} \cup S_6(000) \cup S_6(010) \cup S_6(011) \cup S_6(111) \cup S_8(001) \cup S_8(010) \\ & \cup S_{10}(000) \cup S_{10}(001) \cup S_{10}(011). \end{aligned}$$

Let

$$C^* := \left\{ \mathbf{x} \in \{0,1\}^{13} \;\middle|\; (\mathbf{x}|abc) \in C''' \text{ for some } a, b, c \in \{0,1\} \right\}.$$

Then C is a 2-code and $\#C = 98$.

$n = 12$.

There exists a position such that zero appears in this position for 50 of the code words of C^*. Deleting this coordinate in these codewords we get a 2-code of length 12 having 50 codewords.

$n = 11$.

Let

$$\begin{aligned} C' &:= \Big\{\mathbf{0}, (11000100000), (00110001000), (00001010001)\Big\} \\ &\cup \Big\{\text{all cyclic shifts of } (11101001000)\Big\}, \end{aligned}$$

and let

$$C := C' \cup \overline{C'}.$$

Then C is a 2-code having 30 codewords.

$n = 8, 9, 10$.

The columns of the following matrices are 2-codes of lengths 8,9, and 10 and sizes 7, 12, and 18 respectively.

0100101	010100011011	000011100101111001
0100111	010010101011	000101110010110101
0101011	010001110011	001010111000101011
0001011	001100101101	010001011001011101
0001101	001010110101	010100101100010111
0011101	001001011101	010010010110011011
0010111	000110010111	001001001110011011
0010011	000101100111	011000100011001111
	000011001111	001100010101100111
		000110001011110011

7.3 Notes

7.1. The code of Kim and Freiman [51] were the first codes constructed for correction of multiple asymmetric errors, and are included here for this reason.

7.2. The constructions appear in Delsarte and Piret [17],[18]. For a discussion of the $[17, 8]$ code and the Nordstrom-Robinson code, see e.g. MacWilliams and Sloane [61].

Dynkin and Togonidze [21] discussed the use of a certain cyclic code for correcting multiple asymmetric errors.

Codes to correct t or less adjacent errors have been constructed by Varshamov et al. [90], Oganesyan and Yagdzhyan [67], Klimiashvili [54], and Tenengolts [80].

Chapter 8

Error burst correction

8.1 Preliminaries

Definition 8.1 *If* $\mathbf{x} \in \{0,1\}^n$ *is transmitted and errors occur in positions* $i_1, i_2, \ldots, i_r$ *where* $i_1 < i_2 < \cdots i_r$*, then we say that an (*error*)* burst *of* length $i_r - i_1 + 1$ *has occurred.*

Remark. Any error pattern is a burst as defined above. The codes we construct are able to correct a burst of length less than or equal to some predefined bound, i.e. the length of the burst, not the number of errors is the focus of attention.

8.2 Generalized Oganesyan-Yagdzhyan codes

Code construction.

In this construction

b is a positive integer (maximal burst length),
$c := 2b - 1$,
m is a positive integer such that $\gcd(m, c) = 1$
and such that all prime factors of m exceed b,
$n := cm$.

For $a_0 \in \mathbb{Z}_m$ and $a_j \in \mathbb{Z}_2$ for $1 \le j \le c$, let

$$C_{a_0,a_1,\ldots,a_c} := \Big\{ \mathbf{x} \in \{0,1\}^n \,\Big|\, \sum_{i=1}^{n} i x_i \equiv a_0 \pmod{m}, \\ \sum_{k=0}^{m-1} x_{j+kc} \equiv a_j \pmod{2} \text{ for } 1 \le j \le c \Big\}.$$

These codes are able to correct a burst of length b or less. We give a slightly less formal description of the decoding algorithm than usual.

Decoding algorithm.

Comment: $\mathbf{y}$ is the received vector.
Comment: $\mathbf{e}_i$ is the i'th unit vector.

determine $a, b_1, b_2, \ldots, b_c$ by
$a \equiv \left(\sum_{i=1}^{n} i y_i\right) - a_0 \pmod{m}$, $0 \le a < m$;
$b_j \equiv \left(\sum_{k=0}^{m-1} y_{j+kc}\right) - a_j \pmod{2}$, $b_j <\in \{0,1\}$, for $j = 1, 2, \ldots, c$;
$s := \sum_{j=1}^{c} b_j$;
if $s = 0$ **then** decode into $\mathbf{y}$
else
begin
    determine j_1 by $b_{j_1} = 1$ and $b_j = 0$ for $j_1 - b + 1 \le j \le j_1 - 1$;
        Comment: let $b_j := b_{j+c}$ if $j < 1$
    determine k by $kcs \equiv a - \sum_{j=j_1}^{j_1+b-1} j b_j \pmod{m}$, $0 \le k < m$;
        Comment: let $b_j := b_{j-c}$ if $j > c$
    decode into $\mathbf{y} + \sum_{j=j_1}^{j_1+b-1} b_j \mathbf{e}_{j+kc}$
end.

Remark. If no j_1 or no k exist, then decoding has failed.

Proof of decoding algorithm.

Let $\mathbf{x}$ be sent. If no error occurs, then $b_j = 0$ for $j = 1, 2, \ldots, c$, $s = 0$, and we decode into $\mathbf{y} = \mathbf{x}$. Suppose errors occur in positions $i_1, i_2, \ldots, i_r$, where $i_1 < i_2 < \cdots i_r \le i_1 + b$. Let $i_l = \iota_l + \lambda c$ where $0 \le \iota_l < c$. Then $b_j = 1$ for

$j \equiv \iota_l \pmod{c}$, $1 \le l \le r$, and $b_j = 0$ otherwise. Hence $s = r$ and $j_1 = i_1$. Further,

$$a \equiv \sum_{l=1}^{r} i_l \equiv \sum_{l=1}^{r} (\iota_l + \lambda c) \equiv \sum_{j=i_1}^{i_1+b-1} j b_j + \lambda c r \pmod{m}.$$

Hence $k = \lambda$ (we have $\gcd(cr, m) = 1$ since $\gcd(c, m) = 1$ and if p is a prime factor of m, then $p > b \ge r$) and we decode into $\mathbf{y} + \sum_{l=1}^{r} \mathbf{e}_{\iota_l + \lambda c}$.

Bound on the size of the codes.

If $N_{n,g}$ is the size of the Varshamov-Tenengolts code C_g of length n, then

$$\max_{a_1, a_2, \ldots, a_c} \#C_{a_0, a_1, \ldots, a_c} \ge \frac{N_{n,a_0}}{2^c}$$

since $\left\{ C_{a_0, a_1, \ldots, a_c} \;\middle|\; a_i \in \{0, 1\} \text{ for } i = 1, 2, \ldots, c \right\}$ is a partition of C_{a_0}. In particular

$$\max_{a_0, a_1, \ldots, a_c} \#C_{a_0, a_1, \ldots, a_c} \ge \frac{N_{n,0}}{2^c} = \frac{1}{2^c(n+1)} \sum_{\substack{d | n+1 \\ d \ odd}} 2^{(n+1)/d-1} \phi(d).$$

8.3 Davydov-Dzodzuashivili-Tenengolts codes

Code construction

In this construction

k and b are positive integers,
$\kappa := \lceil k/b \rceil$,
$m := \lceil \log_2 k \rceil$.

For $\mathbf{x} \in \{0, 1\}^k$, let

$$\mathbf{x}^{(i)} := (x_{k-ib+1}, x_{k-ib+2}, \ldots, x_{k-ib+b})$$

for $i = 1, 2, \ldots, \kappa$ (where $x_j = 0$ for $j \le 0$) and

$$\mathbf{x}^{(0)} := \bigoplus_{i=1}^{\kappa} \mathbf{x}^{(i)}.$$

Let $s(\mathbf{x})$ be defined by

$$s(\mathbf{x}) \equiv \sum_{i=1}^{\kappa} i\, w(\mathbf{x}^{(i)}) \pmod{2^{m+1}},\ 0 \le s(\mathbf{x}) < 2^{m+1}.$$

Let

$$s(\mathbf{x}) = \sum_{j=0}^{m} s_j 2^j$$

be the binary expansion of $s(\mathbf{x})$, and define $\mathbf{u}(\mathbf{x})$ by

$$\mathbf{u}(\mathbf{x}) := (\overline{s_m}, s_0, s_1, \ldots, s_m).$$

The code is

$$C := \left\{ \left(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x})\right) \,\middle|\, \mathbf{x} \in \{0,1\}^k \right\}.$$

This code is able to correct a burst of length b or less.

Decoding algorithm.

Comment: $(\mathbf{y}|\mathbf{y}_0|\mathbf{v})$ is the received vector,
Comment: $\mathbf{y} \in \{0,1\}^k$, $\mathbf{y}_0 \in \{0,1\}^b$, $\mathbf{v} \in \{0,1\}^{m+2}$.

if $v_1 = v_{m+2}$ or $\mathbf{y}^{(0)} = \mathbf{y}_0$, **then** decode into $(\mathbf{y}|\mathbf{y}^{(0)}|\mathbf{u}(\mathbf{y}))$
else
begin
  determine σ by $\sigma \equiv \left(\sum_{i=1}^{m+2} v_i 2^{i-2}\right) - s(\mathbf{y}) \pmod{2^{m+1}}$, $-2^m < \sigma \le 2^m$;
  if $\sigma < 0$ **then** decode into $(\mathbf{y}|\mathbf{y}^{(0)}|\mathbf{u}(\mathbf{y}))$
  else
  begin
    $\mathbf{z} := \mathbf{y}_0 \oplus \mathbf{y}^{(0)}$;
    $r := w(\mathbf{z})$;
    $i := \lfloor \sigma / r \rfloor$;
    $\rho := \sigma - ir$;
    determine $\mathbf{z}_1$ and $\mathbf{z}_2$ such that $\mathbf{z} = (\mathbf{z}_1|\mathbf{z}_2)$, $w(\mathbf{z}_1) = \rho$,
      and such that $\mathbf{z}_2$ is as short as possible;
    determine $\tilde{\mathbf{y}} \in \{0,1\}^k$ by $\tilde{\mathbf{y}}^{(i)} := \mathbf{y}^{(i)} \oplus (\mathbf{z}_2|\mathbf{0})$,
      $\tilde{\mathbf{y}}^{(i+1)} := \mathbf{y}^{(i+1)} \oplus (\mathbf{0}|\mathbf{z}_1)$
      $\tilde{\mathbf{y}}^{(j)} := \mathbf{y}^{(j)}$ for $j \notin \{i, i+1\}$;

decode into $(\tilde{\mathbf{y}}|\tilde{\mathbf{y}}^{(0)}|\mathbf{u}(\tilde{\mathbf{y}}))$

end;

end.

Proof of the decoding algorithm.

Let $(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x}))$ be sent. If no error occurs, then $\mathbf{y}^{(0)} = \mathbf{y}_0$ and we decode into $(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x}))$. Suppose a burst of length b or less has occurred. Then either $\mathbf{x}$ or $\mathbf{u}(\mathbf{x})$ are received without errors.

Case I, all the errors are in the $\mathbf{x}|\mathbf{x}^{(0)}$ part; say the right-most error occurs in $\mathbf{x}^{(L)}$. Then $\mathbf{x}^{(L)} - \mathbf{y}^{(L)} = (\mathbf{f}|\mathbf{0})$ where we assume that the rightmost element in $\mathbf{f}$ is a 1, let the length of $\mathbf{f}$ be λ. Then $\mathbf{y}^{(L+1)} = \mathbf{x}^{(L+1)} - (\mathbf{0}|\mathbf{g})$, for some $\mathbf{g}$ of length $b-\lambda$, and $\mathbf{y}^{(j)} = \mathbf{x}^{(j)}$ for $j \notin \{L, L+1\}$. Therefore, since $\mathbf{u}(\mathbf{x})$ is received error free, we get

$$\sigma \equiv s(\mathbf{x}) - \Big(s(\mathbf{x}) - Lw(\mathbf{f}) - (L+1)w(\mathbf{g})\Big) \equiv Lw((\mathbf{f}|\mathbf{g})) + w(\mathbf{g}) \pmod{2^{m+1}}.$$

However, $0 \le L\,w(\mathbf{f}) + (L+1)w(\mathbf{g}) \le k \le 2^m$, and so

$$\sigma = L\,w((\mathbf{f}|\mathbf{g})) + w(\mathbf{g}).$$

Further $\mathbf{z} = (\mathbf{f}|\mathbf{0}) + (\mathbf{0}|\mathbf{g})$. Therefore, $r = w((\mathbf{f}|\mathbf{g}))$, $i = L$, and $\rho = w(\mathbf{g})$. Hence $\mathbf{z}_1 = \mathbf{g}$ and $\mathbf{z}_2 = \mathbf{f}$, and so $\tilde{\mathbf{y}} = \mathbf{x}$ and we decode in $(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x}))$.

Case II, there are one or more errors in the $\mathbf{u}(\mathbf{x})$ part. Then $\mathbf{x}$ is received without errors. Suppose $v_j = 0$ and $u_j = 1$ for $j = j_1, j_2, \ldots, j_r$ and $v_j = u_j$ otherwise. If $v_1 \ne u_1$ or $v_{m+2} \ne u_{m+2}$, then $v_1 = v_{m+2} = 0$ (since $u_1 = \overline{u_{m+2}}$), and we decode into $(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x}))$. Otherwise

$$\sigma \equiv \sum_{i=2}^{m+1} (v_i - u_i)2^{i-2} \pmod{2^{m+1}}$$

and so, since

$$-(2^m - 1) \le \sum_{i=2}^{m+1} (v_i - u_i)2^{i-2} < 0,$$

we have $-2^m < \sigma < 0$, and we decode into $(\mathbf{x}|\mathbf{x}^{(0)}|\mathbf{u}(\mathbf{x}))$.

Size of the codes.

The length of the code is $n = k + b + \lceil \log_2 k \rceil + 2$ and $\#C = 2^k$.

8.4 Notes

8.2. The construction given is essentially due to Oganesyan and Yagdzhyan [67]. However, they only considered the case where m is a prime.

8.3. The construction is due to Davydov, Dzodzuashivili, and Tenengolts [16].

Chapter 9

Codes for non-binary alphabets

9.1 Preliminaries

Many of constructions given in the previous chapters carry over to the non-binary case. We shall give the necessary definitions and state the main results without proofs. The (input and output) alphabet is $A = \{0, 1, \ldots, a-1\}$ of size a.

Definition 9.1 *An a-ary channel is* asymmetric *if it has the property that if symbol b is sent, then only symbols from the set $\{0, 1, \ldots, b\}$ can be received.*

Definition 9.2 *For $\mathbf{x}, \mathbf{y} \in A^n$, let*

(i) $w(\mathbf{x}) := \sum_{i=1}^{n} x_i$,
(ii) $N(\mathbf{x}, \mathbf{y}) := \sum_{i=1}^{n} \max\{y_i - x_i, 0\}$,
(iii) $\Delta(\mathbf{x}, \mathbf{y}) := \max\{N(\mathbf{x}, \mathbf{y}), N(\mathbf{y}, \mathbf{x})\}$.

If $\mathbf{x}$ is sent and $\mathbf{y}$ is received, we say that $w(\mathbf{x} - \mathbf{y})$ errors have occurred. A code correcting t errors is called a t-code.

Theorem 9.1 *$C \subset A^n$ is a t-code if and only if $\Delta(\mathbf{x}, \mathbf{y}) > t$ for all $\mathbf{x}, \mathbf{y} \in C$, $\mathbf{x} \neq \mathbf{y}$.*

9.2 1-codes

The Stanley-Yoder construction carries over to arbitrary alphabet size: if G is a group with the properties given in 4.1 and $g \in G$, then

$$C_g := \left\{ \mathbf{x} \in A^n \;\middle|\; g_1^{x_1} g_2^{x_2} \cdots g_n^{x_n} = g \right\}.$$

The weight distribution of the corresponding generalization of Constantin-Rao codes is given by the following theorem.

Theorem 9.2 *Let G be an Abelian group of order N, and let $g \in G$. Let*

$$t(g, w) := \#\left\{ \mathbf{x} \in C_g \;\middle|\; w(\mathbf{x}) = w \right\}$$

and

$$T_g(y) := \sum_{w=0}^{N-1} t(g, w) y^w.$$

Then

$$T_g(y) = \frac{1}{N} \frac{1-y}{1-y^a} \sum_{cd=N} \frac{\left(1 - y^{ad/\gcd(a,d)}\right)^{c \cdot \gcd(a,d)}}{\left(1 - y^d\right)^c} S_{\mathbf{g}}(d).$$

9.3 t-codes

The Varshamov constructions carry over to arbitrary alphabet size. The definition of a V_t-set has to be modified to require that the sums $h_{i_1} + h_{i_2} + \cdots + h_{i_r}$, where $1 \le i_1 \le i_2 \le \cdots \le i_r \le n$, $0 \le r \le t$, and where each subscript appears at most $a - 1$ times, are distinct. We note that the two V_t-sets given in 6.3 satisfy this condition. If H is a modified V_t-set in G and $g \in G$, then

$$C_g := \left\{ \mathbf{x} \in A^n \;\middle|\; \sum_{i=1}^{n} x_i h_i = g \right\},$$

and

$$\max_{g \in G} \#C_g \ge \frac{a^n}{\#G}.$$

The constructions in 6.4 and 6.5 have the following generalizations (using the notations of 6.4 and 6.5):

$$C_{g_1,g_2,\dots,g_t} := \Big\{ \mathbf{x} \in A^n \;\Big|\; \sigma_l(\mathbf{x}\alpha) = g_l \text{ for } 1 \le l \le t \Big\},$$

and

$$C_{g_1,g_2,\dots,g_{t-1},m} := \Big\{ \mathbf{x} \in A^n \;\Big|\; \sigma_l(\mathbf{x}\alpha) = g_l \text{ for } 1 \le l \le t-1,\ \sigma_1(\mathbf{x}\mathbf{a}) = m \Big\}.$$

9.4 Notes

All the results stated in this chapter are generalizations of results previously given in the binary case. The results may be proved by a modification of the proofs of the binary results.

9.1. The definition of Δ and Theorem 9.1 are due to Delsarte and Piret [17],[19].

9.2. Theorem 9.2 is due to Helleseth and Kløve [48].

9.3. Varshamov [88] considered codes over arbitrary alphabet size. Generalizations were given by Delsarte and Piret [17].

Kipshidze et al. [52] gave a code for the asymmetric channel with seven symbols.

Robinson [71] discussed codes for a ternary asymmetric channel with some additional restrictions. This coding problem was further discussed by Kløve [59].

Codes for an asymmetric channel of a different kind was given by Arakelov and Tenengolts [2]

Bibliography

[1] G. G. Ananiashvili, "On a class of asymmetric single-error-correcting non-linear codes" (in Russian), *Doklady Akad. Nauk. Georgian SSR* 53 no. 3, 1969, 549-552.

[2] V. A. Arakelov and G. M. Tenengolts, "Certain classes of correcting codes" (in Russian), *Trudy Vychisl. Centra Akad. Nauk. Armjan SSR, Erevan* 6 (1970) 64-77.

[3] P. Backmann, *Niedere Zahlentheorie*, Teuber Verlag, Leipzig 1910.

[4] L. A. Bassalygo, "New upper bounds for error correcting codes" (in Russian), *Problemy Peredachi Informacij* 2 no. 4, 1965, 41-44 (trans: Problems of Information Transmission 2 no. 4, 32-34).

[5] J. M. Berger, "A note on error detection codes for asymmetric channels", *Information and Contr.* 4 (1961) 68-73.

[6] J. M. Borden, "Bounds and constructions for error correcting/detecting codes on the Z-channel", *Abstracts of papers, 1981 IEEE International Symposium on Information Theory*, Feb. 9-12, 1981, 94-95.

[7] J. M. Borden, "A low-rate bound for asymmetric error-correcting codes", Report, Worchester Polytech. Inst. 1982.

[8] J. M. Borden, "A low-rate bound for asymmetric error-correcting codes", *IEEE Trans. on Information Th.* IT-29 (1983) xx.

[9] R. C. Bose and S. Chowla, "Theorems in additive theory of numbers", *Comm. Math. Helvetici* 37 (1962) 141-147.

[10] B. Bose and R. S. Cunningham, "Systematic and multiple-error-correcting asymmetric codes", manuscript.

[11] A. E. Brouwer, J. B. Shearer, N. J. A. Sloane, and W. D. Smith, "A new table of constant weight codes", *IEEE Trans. on Information Th.* IT-36 (1990) 1334-1380.

[12] S. D. Constantin, "Characterization and design of error correcting/detecting codes for asymmetric memories", Ph.D. Thesis, Dept. Comp. Sci, Southers Methodist Univ., Dallas, Texas (1977).

[13] S. D. Constantin and T. R. M. Rao, "On the theory of binary asymmetric error correcting codes", *Information and Contr.* 40 (1979) 20-36.

[14] S. D. Constantin and T. R. M. Rao, "Concatenated group theoretic codes for binary asymmetric channels", *AFIPS Conf. Proc.* 46 (1979) 837-842.

[15] N. Darwish and B. Bose, "New lower bounds for single asymmetric error correcting codes", preprint 1991.

[16] A. A. Davydov, A. G. Dzodzuashivili, and G. M. Tenengolts, "Code correcting nonsymmetric bursts of errors during data exchange between computers" (in Russian), *Avtomatika i Telemekhanika* 33 no. 2 (1972) 178-184 (trans: Automation and Remote Control 33, 1220-1225).

[17] Ph. Delsarte and Ph. Piret, "Spectral enumerators for certain unidirectional error correcting codes", *Report, Phillips Research Lab.*, Brussels 1979.

[18] Ph. Delsarte and Ph. Piret, "Bounds and constructions for binary asymmetric error-correcting-codes", *IEEE Trans. on Information Th.* IT-27 (1981) 125-128; correction in: IT-36 (1990) 954.

[19] Ph. Delsarte and Ph. Piret, "Spectral enumerators for certain additive-error-correcting codes over integer alphabets", *Information and Contr.* 48 (1981) 193-210.

[20] V. N. Dynkin and M. Kirov, "Synthesis of binary codes correcting asymmetric errors" (in Bulgarian), *Avtom. Izchislitelna Tekh.* 14 no.1 (1980) 28-35.

[21] V. N. Dynkin and V. A. Togonidze, "Cyclic codes correcting multiple asymmetric errors" (in Russian), *Doklady Akad. Nauk. Georgian SSR* 77 no. 1 (1974) 33-55.

[22] V. N. Dynkin and V. A. Togonidze, "On asymmetric error correcting codes" (in Russian), *Doklady Akad. Nauk. Georgian SSR* 77 no. 3 (1975) 561-564.

[23] T. Etzion, "Lower bounds for asymmetric and unidirectional codes", *IEEE Trans. on Information Th.* IT-37 (1911) 1696-1704.

[24] S. B. Fajn, "A certain polyadic code that corrects errors" (in Russian), *Sakharth. SSR mech. Akad. Gamothvl. Centr. Shrom.* 9 no. 1 (1969) 98-102.

[25] S. B. Fajn, Z. Sh. Kipshidze, and M. A. Klimiashvili, "A polynomial code that corrects multiple non-symmetric errors of variable degree" *Doklady Akad. Nauk. Georgian SSR* 61 (1971) 561-563.

[26] S. B. Fajn, Z. Sh. Kipshidze, and M. A. Klimiashvili, "On a class of polynomial codes" (in Russian) *Doklady Akad. Nauk. Georgian SSR* 62 (1971) 29-31.

[27] S. B. Fajn, Z. Sh. Kipshidze, and M. A. Klimiashvili, "Error correcting by the polynomial methods" (in Russian) *Doklady Akad. Nauk. Georgian SSR* 63 (1971) 397-399.

[28] S. B. Fajn, Z. Sh. Kipshidze, and M. A. Klimiashvili, "On error correcting polynomial codes" (in Russian) *Sakharth. SSR mech. Akad. Gamothvl. Centr. Shrom.* 11 no. 1 (1972) 80-86.

[29] G. Fang, "Binary block codes for correcting asymmetric or uindirectional errors", PhD thesis, Dept. of Mathematics and Comp. Sci., Eindhoven Univ. of Tech. (1993).

[30] G. Fang and I. Honkala, "On perfectness of binary codes for correcting asymmetric errors", manuscript 1993.

[31] G. Fang and H. C. A. van Tilborg, "Some new results on binary asymmetric error-correcting codes", *Proc. of IEEE Intern. Symposium on Information Th.*, Budapest, Hungary, (1991).

[32] G. Fang and H. C. A. van Tilborg, "New tables of AsEC and UEC codes", *Report 91-WSK-02, Eindhoven Univ. of Tech.*, August 1991.

[33] G. Fang and H. C. A. van Tilborg, "Bounds and constructions of asymmetric or unidirectional error-correcting codes", *Applicable Algebra in Engineering, Communication and Computing* 3 no. 4 (1992) 269-300.

[34] G. Fang, H. C. A. van Tilborg, and F. W. Sun, "Weakly perfect binary block codes for correcting asymmetric errors", *Proc. of Intern. Symposium on Communication Theory*, Tainan, Taiwan, Dec. 1991, 57-60.

[35] G. Fang, H. C. A. van Tilborg, and F. W. Sun, "On uniformly weakly perfect codes for correcting asymmetric errors; some bounds and constructions", *Collection of Papers Dedicated to the Memory of David Gevorkian, Armenia*, 1992.

[36] G. Fang, H. C. A. van Tilborg, F. W. Sun, and I. Honkala, "Some features of binary block codes for correcting asymmetric errors", *Proc. AAECC*, Springer Lecture Notes in Computer Science 1993?.

[37] C. V. Freiman, "Optimal error detection codes for completely asymmetric binary channels", *Information and Contr.* 5 (1962) 64-71.

[38] C. V. Freiman, "On the use of coset codes in asymmetric channels", *IEEE Trans. on Information Theory* IT-9 (1963) 118.

[39] D. N. Gevorkjan and A. G. Mhitarjan, "A class of codes that correct single great asymmetric errors" (in Russian), *Dokl. Akad. Nauk. Armjan. SSR* 70 no. 4 (1980) 216-218.

[40] B. D. Ginzburg, "Determination of the size of Varshamov-Tenebgolts codes correcting one asymmetric error" (in Russian), *Proc. Second Conf. on the Theory of Codes and their Applications*, sec.1, part 1 (1965) 25-27.

[41] I. Ya. Goldbaum, "Estimate for the number of signals in codes correcting non-symmetric errors" (in Russian), *Avtomatika i Telemekhanika* 32 no. 11 (1971) 94-97 (trans: Automat. Rem. Control 32, 1783-1785).

[42] I. Ya. Goldbaum, "A new bound on the number of signals in binary codes correcting asymmetric errors" (in Russian), *Problemy Peredachi Informacij* 13 no. 1 (1976) 102-104 (trans: Problems of Information Transmission 13, 74-76).

[43] S. W. Golomb, "The limiting behavior of the Z-channel", *IEEE Trans. on Information Theory* IT-26 (1980) 372.

[44] T. H. Gordon, "Error coding bounds for the binary asymmetric channel", *IEEE Trans. on Information Theory* IT-9 (1963) 206-208.

[45] R. L. Graham and N. J. A. Sloane, "Lower bounds for constant weight codes", *IEEE Trans. on Information Theory* IT-27 (1980) 37-43.

[46] C. Helgesen, "Asymmetric codes with minimal block length, and the weight enumirator for the Ananiashvili code" (in Norwegian), Cand. real. thesis, Dept. of Math., Univ. of Bergen (1983).

[47] C. Helgesen, "Asymmetric error-correcting codes with minimal block length", unpublished manuscript (1983).

[48] T. Helleseth and T. Kløve, "The size and weight distribution of 1-error-correcting group-theoretic codes for asymmetric channels", Report, Dept. of Math., Univ. of Bergen (1979).

[49] T. Helleseth and T. Kløve, "On group-theoretic codes for asymmetric channels", *Information and Contr.* 49 (1981) 1-9.

[50] W. H. Kim and C. V. Freiman, "Single error-correcting-codes for asymmetric binary channels", *IRE Trans. on Information Theory* IT-5 (1959) 62-66.

[51] W. H. Kim and C. V. Freiman, "Multiple error correcting codes for a binary asymmetric channel", *IEEE Trans. on Circuit Theory* CT-6, Special supplement on International Symposium on Circuit and Information Theory (1959) 71-78.

[52] Z. Sh. Kipshidze, M. A. Klimiashvili, T. V. Kobashvili, I. O. Urumov, and G I. Habalashvili, "Some weighted error-correcting-codes" (in Russian), *Izdat. Tibilis. Univ. Tiblisi* 45 (1978) (review in: Math. Reviews 58 no. 20788).

[53] Z. Sh. Kipshidze, M. A. Klimiashvili, T. V. Kobashvili, and I. O. Urumov, "Weighted codes with basis $q = 7$" (in Russian), *Trudy Vychisl. Centra Akad. Nauk. Gruzin SSR* 19 no. 2, Math. i Tekhn. Kibern. (1979) 12-16.

[54] M. A. Klimiashvili, "Classes of asymmetric error correcting codes" (in Russian), *Doklady Akad. Nauk. Georgian SSR* 79 no. 1 (1975) 141-144.

[55] T. Kløve, "A class of constant weight codes", Report, Dept. of Math., Univ. of Bergen (1979).

[56] T. Kløve, "Upper bounds on codes correcting asymmetric errors", *IEEE Trans. on Information Theory* IT-27 (1981) 128-131.

[57] T. Kløve, "A lower bound for $A(n, 4, w)$", *IEEE Trans. on Information Theory* IT-27 (1981) 257-258.

[58] T. Kløve, "Error correcting codes for the asymmetric channel", *Report 18-09-07-81, Dept. of Pure Mathematics, Univ. Bergen* 1981; revised and extended 1983.

[59] T. Kløve, "On Robinson's coding problem", *IEEE Trans. on Information Theory* IT-29 (1983) 450-454.

[60] R. R. Kuzyurin, "Minimal covering and maximal packings of $(k-1)$-subsets by k-subsets" (in Russian), *Matem. Zametki* 21 no. 4 (1977) 565-571 (trans: Mathematical Notes 21, 316-320).

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

[62] L. E. Mazur, "Certain codes that correct non-symmetric errors" (in Russian), *Problemy Peredachi Informacij* 10 no. 4 (1974) 40-46 (trans: Problems of Information Transmission 10, 308-312).

[63] R. J. McEliece, "Comment on 'A class of codes for asymmetric channels and a problem from the additive theory of numbers'", *IEEE Trans. on Information Theory* IT-19 (1973) 137.

[64] R. J. McEliece and E. R. Rodemich, "The Constantin-Rao construction for binary asymmetric error-correcting-codes", *Information and Contr.* 44 (1980) 187-196.

[65] M. N. Nalbandjan, "Note on two classes of non-linear codes" (in Russian), *Problemy Peredachi Informacij* 10 no. 2 (1974) 61-63 (trans: Problems of Information Transmission 10, 139-141).

[66] M. N. Nalbandjan, "A class of codes correcting multiple asymmetric errors" (in Russian), *Doklady Akad. Nauk. Georgian SSR* 77 no. 2 (1975) 405-408.

[67] S. Sh. Oganesyan and V. G. Yagdzhyan, "Classes of codes which correct error bursts in an asymmetric channel" (in Russian), *Problemy Peredachi Informacij* 6 no. 4 (1970) 27-34 (trans: Problems of Information Transmission 6, 303-309).

[68] T. R. N. Rao and S. D. Constantin, "Group theoretic codes for binary asymmetric channels", Tech. rep. CS 76014, Dept. Comp. Sci., Southern Methodist Univ., Dallas, Texas 1976.

[69] T. R. N. Rao and A. S. Chawla, "Asymmetric error codes for some LSI semi-conductor memories", *Proc. 7'th Annual Southeastern Symp. on Systems Theory* (1975) 170-171.

[70] T. R. N. Rao and E. Fujiwara, *Error-control coding for computer systems*, Prentice Hall (1989).

[71] J. P. Robinson, "An asymmetric error-correcting ternary code", *IEEE Trans. on Information Theory* IT-24 (1978) 258-261.

[72] Y. Saitoh, K. Yamaguchi, and H. Imai, "Some new binary codes correcting asymmetric/unidirectional errors", *IEEE Trans. on Information Theory* IT-36 (1990) 645-647.

[73] A. Shiozaki, "Construction for binary asymmetric error-correcting codes", *IEEE Trans. on Information Theory* IT-28 (1982) 787-789.

[74] R. A. Silverman, "On binary channels and their cascades", *IRE Trans. on Information Theory* IT-1 (1955) 19-27.

[75] J. Singer, "A theorem in finite projective geometry, and some applications to number theory", *Trans. Amer. math. Soc.* 43 (1938) 377-385.

[76] R. P. Stanley and M. F. Yoder, "A study of Varshamov codes for asymmetric channels", *Jet Prop. Lab. Tech. Rep.* 32-1526 vol. 14 (1982) 117-122.

[77] G. M. Tenengolts, "On a class of codes for the asymmetric binary channel" (in Russian), in: *Cybernetics*, Nauka Publ., Moscow (1967) 120-129.

[78] G. M. Tenengolts, "A code correcting double non-symmetric errors" (in Russian), referred to in [67].

[79] G. M. Tenengolts, "Some properties of codes correcting asymmetric errors" (in Russian), in: *Theory and Applications of Automata*, Nauka Publ. Moscow (1968) 243-244 (Review in Zbl. Math. 184, p. 441).

[80] G. M. Tenengolts, "Classes of codes correcting bit loss and errors in the preceding bit" (in Russian), *Avtomatika i Telemekhanika* 37 no. 5 (1976) 174-179 (trans: Automation and Remote Contr. 37, 797-802).

[81] J. H. van Lint and H. H. Weber, "Some combinatorial codes for the binary asymmetric channel", preprint 1993.

[82] R. R. Varshamov, "Some features of linear codes that correct asymmetric errors" (in Russian), *Doklady Akad. Nauk. SSSR* 157 no. 3 (1964) 546-548 (trans: Soviet Physics-Doklady 9, 1965, 538-540).

[83] R. R. Varshamov, "Estimates of the number of signals in codes with correction of nonsymmetric errors" (in Russian), *Avtomatika i Telemekhanika* 25 no. 11 (1964) 1628-1629 (trans: Automation and Remote Contr. 25, 1468-1469).

[84] R. R. Varshamov, "An arithmetic function applicable to coding theory" (in Russian), *Doklady Akad. Nauk. SSSR* 161 no. 3 (1965) 540-543 (trans: Soviet Physics-Doklady 10, 1965, 185-187).

[85] R. R. Varshamov, "On the theory of asymmetric codes" (in Russian), *Doklady Akad. Nauk. SSSR* 164 no. 4 (1965) 757-760 (trans: Soviet Physics-Doklady 10, 1965, 901-903).

[86] R. R. Varshamov, "On the mathematical theory of asymmetric coding systems" (in Russian), *Revue Roumaine de Mathematique pures et appliquees* 10 no. 2 (1965) 165-169.

[87] R. R. Varshamov, "Mathematical methods for increasing reliability when transmitting information on non-symmetric channels", in: *On the theory of relay organization*, Nauka Publ. Moscow (1966) 117-133 (review in RZM 1971 no. 2 V448).

[88] R. R. Varshamov, "A general method of constructing asymmetric coding systems, related to the solution of a combinatorial problem proposed by Dixon" (in Russian), *Doklady Akad. Nauk. SSSR* 194 no. 2 (1970) 284-287 (trans: Soviet Physics-Doklady 15,1970,811-813).

[89] R. R. Varshamov, "A class of codes for asymmetric channels and a problem from the additive theory of numbers", *IEEE Trans. on Information Theory* IT-19 (1973) 92-95.

[90] R. R. Varshamov, S. Sh. Oganesyan, and V. G. Yagdzhyan, "Non-linear binary codes which correct one and two adjacent errors for asymmetric channels" (in Russian), *Proc. First Conf. of Young Specialists at Computer Centers*, Erevan, vol. 2 (1969).

[91] R. R. Varshamov and G. M. Tenengolts, "Correcting code for single asymmetric errors" (in Russian), *Avtomatika i Telemekhanika* 26 no. 2 (1965) 288-292 (trans: Automation and Remote Contr. 26, 286-290).

[92] R. R. Varshamov and E. P. Zograbjan, "A class of codes correcting two asymmetric errors" (in Russian), *Trudy Vychisl. Centra Akad. Nauk. Armjan. SSR i Erevan* 6 (1970) 54-58.

[93] R. R. Varshamov and E. P. Zograbjan, "Codes correcting packets of non-symmetric errors" (in Russian), *Proc. 4'th Symposium on Problems in Information Systems* vol. 1 (1970) 87-96 (Review in RZM 1971 no. 2 V448).

[94] J. H. Weber, "Bounds and constructions for binary block codes correcting asymmetric or unidirectional errors", *PhD thesis, Dept. of Electrical Eng., Delft Univ. of Tech.* (1989).

[95] J. H. Weber, C. de Vroedt, and D. E. Boekee, "New upper bounds on the size of codes correcting asymmetric errors", *IEEE Trans. on Information Theory* IT-33 (1987) 434-437.

[96] J. H. Weber, C. de Vroedt, and D. E. Boekee, "Bounds and constructions for binary codes of length less than 24 and asymmetric distance less that 6", *IEEE Trans. on Information Theory* IT-34 (1988) 1321-1331.

[97] Z. Zhang and X.-G. Xia, "New lower bounds for binary codes of asymmetric distance two", *IEEE Trans. on Information Theory* IT-38 (1992) 1592-1597.