# Constructing codes

# CONSTRUCTING CODES

H.J. TIERSMA

# CONSTRUCTING CODES

# CONSTRUCTING CODES

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Technische Universiteit Eindhoven, op gezag
van de Rector Magnificus, prof.ir. M. Tels,
voor een commissie aangewezen door het College
van Dekanen in het openbaar te verdedigen
op dinsdag 24 oktober 1989 te 14.00 uur.

door
Herman Jacobus Tiersma
geboren te Eindhoven

Dit proefschrift is goedgekeurd door
de promotoren
prof.dr. J.H. van Lint
prof.dr.ir. J.P.M. Schalkwijk

ABSTRACT

This thesis concerns constructions and bounds for codes for three different kinds of situations. The first chapter treats bounds and constructions for codes for a so-called three-way channel. The bounds are derived by applying random-coding techniques. The construction of the codes is more or less and hoc. Furthermore, a generalization of Schalkwijk's idea for the two-way channel is investigated. The second chapter is on bounds and constructions of error correcting codes for the binary adder two-access channel. The lower bounds are derived by techniques similar to the Gilbert-Varshamov argument. The construction of the codes uses concatenation techniques. The chapter is also partly concerned with the construction of codes over a ternary alphabet supplied with the somewhat unusual Manhattan-distance. These codes then are used as building blocks in the concatenation method. The third chapter is concerned with codes constructed from Algebraic Geometry. Especially the codes from Hermitian curves are studied thoroughly.

**Contents**

# Chapter 0

# General Introduction

## 0.0. Introduction

This thesis contains the results of my research during the period '83-'88 when I was working for the Netherlands Organization for Pure Research Z.W.O. The investigated subjects have in common that they deal with constructions and bounds for codes. However, the techniques involved are quite different. The first problem is about a three-way channel, which is a generalization of the two-way channel introduced by Shannon and studied thorougly by Schalkwijk. The second problem is about coding for the noisy-access binary adder channel. This problem was already studied by among others H.C.A. van Tilborg and T. Kasami. It appeared that there is a close relation between ternary codes in the Manhattan metric (i.e. the $L$-distance, which will be defined in Chapter 2) and error correcting code pairs. Therefore we will give bounds on this kind of codes.
The third subject is becoming more and more important these days and is concerned with codes from Algebraic Geometry. In the chapter dealing with this subject we give an overview of general results. My own contribution was a detailed study of the codes from Hermitian Curves.

## 0.1. Notations, Definitions and Remarks

In this thesis are some notations and concepts that are used a lot of times. In this section we discuss the most important ones.

Convention 1. All logarithms in this thesis are log to base 2 so by $\log x$ we mean $\log_2 x$.

Definition 2. The binary entropy function $h(x)$ is defined to be:

$$h(x) := -x \log x - (1-x) \log(1-x).$$

Definition 3. If $a,b,c$ are events with probabilities $P(a), P(b), P(c), P(a|b)$ etc., then the mutual information of $a$ and $b$ conditional to $c$ is defined as:

$$I(a;b|c) := \log(P(a,b|c)/(P(a|c)P(b|c))).$$

(Cf. [1], Chapter 2).

Definition 4. If $A,B$ and $C$ are sets with probability measures $P(a), P(b), P(a|b), P(c)$ etc. then the average mutual information of $A$ and $B$ conditional to $C$ is defined to be:

$$I\ (A;B|C):=\sum_{(a,b,c)\in A*B*C} I(a;b|c)P(a,b,c).$$

(cf. [1], Chapter 2).

DEFINITION 5. If $X$ is a stochastic variable with probability distribution $P(X)$, then the entropy $H(X)$ is defined to be:

$$H(X):=\sum_{X} P(X)\log P(X).$$

REMARK 6. For more details about entropy functions, average mutual information etc. we refer to Gallagers excellent book [1].

DEFINITION 7. A $q$-ary code of length $n$ and distance $d$ will be called a $q$-ary $(n,d)$-code ($n,d$ and $q$ are integers).

REFERENCES

1. R.G. GALLAGER. (1968). *Information Theory and Reliable Communication,* John Wiley and Sons Inc..

# Chapter 1

# The BMC as a Three-way Channel

## 1.0. Introduction

In 1961 Shannon presented one of the first papers ([1]) on two-way communication channels at the fourth Berkeley Symposium. In this paper he introduced a new kind of channel, a so-called two-way channel, and he studied the capacity region for this channel, giving inner bounds and outer bounds for it. For some two-way channels the inner bound is equal to the outer bound but in other (more interesting) cases, they differ. As an example of such a channel (for which the inner bound differs from the outer bound) Shannon mentioned the binary multiplying channel. He also discusses some coding methods for this channel.

For a long time, it was unknown whether the capacity region for this channel coincides with the inner bound region or not. Then in 1980, Schalkwijk ([3]) managed to prove that the capacity region is larger than the inner bound region by exhibiting a coding strategy which operates beyond the Shannon inner bound.

At first sight, this coding strategy has nothing to do with the problem of finding "good" codes for this channel (either block codes or variable length codes). However, at the sixth symposium on Information theory in the Benelux held in Mierlo the Netherlands, Tolhuizen ([5]) showed that Schalkwijks "continuous" strategy leads to "good" block codes as well as too "good" variable length codes.

The aim of the coming chapter is to generalize parts of these results to a three-way channel (which will be defined in Section 2).

Before we do this, we will give a more detailed description of the two-way situation so that the reader gets used to the concepts involved.

## 1.1. The binary multiplying channel as a two-way channel

From Shannon [1] we recall the definition of a two-way channel. A typical two-way channel is shown in Figure 1.

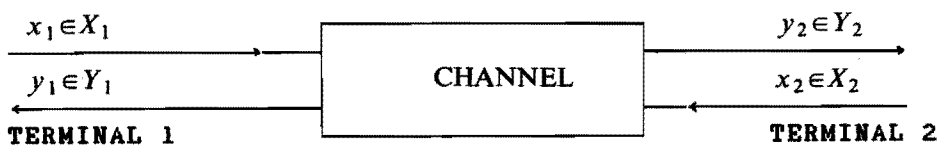



FIGURE 1. A typical two-way channel.

$x_i$ and $y_i$ are the input letter, respectively output letter, at terminal i. They are elements of the input alphabet $X_i$, respectively output alphabet $Y_i$, $(i=1,2)$. Every channel use consists of feeding an input letter at both terminals simultaneously and obtaining an output letter at both terminals. The outputs are related statistically to the inputs and may be to previous inputs and outputs (if the channel has memory). The problem is to communicate as effectively as possible (in both directions).

The case we wish to study is when the channel is discrete and memoryless. This means that the sets $X_1, X_2, Y_1$, and $Y_2$ are finite, and that $y_1$ and $y_2$ are only statistically related to $x_1$ and $x_2$ and not to previous inputs or outputs. A block code pair of length $n$ for such a channel, with $M_1$ messages at terminal 1 and $M_2$ messages at terminal 2 consists of two sets of $n$ functions:

$$f_0(m_1),\ f_1(m_1,y_{11}),\ldots,f_{n-1}(m_1,y_{11},\ldots,y_{1\,n-1})$$

$$g_0(m_2),\ g_1(m_2,y_{21}),\ldots,g_{n-1}(m_2,y_{21},\ldots,y_{2\,n-1}).$$

The functions $f$ take values in $X_1$ and the functions $g$ take values in $X_2$. The $m_i$ take values from 1 to $M_i, (i=1,2)$. the $y_{ji}$ take values in $Y_j, (i=1,\ldots,n, j=1,2)$. The functions $f_i$ prescribe how the input symbol $x_{1i}$ at terminal 1 will be chosen dependent on $m_1$ and $y_{1k}\, k=1,\ldots,i-1$, and the same for $g$ at terminal 2. A decoding system for a block code pair of length $n$ consists of a pair of functions $\phi(m_1,y_{11},\ldots,y_{1n})$ and $\psi(m_2,y_{21},\ldots,y_{2n})$ which take values from 1 to $M_2$ respectively 1 to $M_1$. The decoding functions $\phi$ and $\psi$ give a rule to decide on the message that was sent at the other terminal considering the message at the own terminal together with the $n$ channel outputs. In the sequel we will assume that all messages are equiprobable (at both terminals) having probabilities $1/M_1$ respectively $1/M_2$. The signalling rates for the block code pair defined above are defined to be: $R_2^1 := (\log M_1)/n$, $R_1^2 := (\log M_2)/n$. They represent the amount of information passing through the channel in the 1-2 direction, respectively the $2-1$ direction. From a code pair and a decoding system together with the conditional probabilities which define the channel (i.e. the probabilities of the output symbols given the input symbols), one could compute the probability of incorrect decoding. Averaging over all messages (for each direction), we get the error probabilities $P_{e1}$ and $P_{e2}$ at the two terminals.

The capacity region of the channel is defined to be the set of points $(R_1,R_2)$ such that for every $\epsilon$ there exists a block code pair and a decoding system, having rates $R_1^2$ and $R_2^1$ such that $|R_1^2 - R_1| < \epsilon$ and $|R_2^1 - R_2| < \epsilon$ and $P_{e1}$ and $P_{e2}$ are both less than $\epsilon$. Sometimes it happens that at the terminals 1 and 2 the message of the other terminal is known (with error probability zero), but the terminals keep on communicating on the actual message pair, because they have not used the full block length. In this case it is better to modify the encoding/decoding procedure somewhat and obtain a variable length code pair.

The procedure then works as follows. The encoding procedure makes use of the functions $f_i(m_1,y_{11},\ldots,y_{1i})$ and $g_i(m_2,y_{21},\ldots,y_{2i})$, taking values from $X_1$,

respectively $X_2$. The decoding procedure makes use of the functions: $\phi_i(m_1,y_{1\,1},\ldots,y_{1\,i})$, $\psi_i(m_2,y_{2\,1},\ldots,y_{2\,i})$ which take values from the sets $\{1,\ldots,M_2,\text{ not}\}$, respectively $\{1,\ldots,M_1,\text{ not}\}$, $i=0,\ldots,n$.
If it is possible to decode (without error), the functions take as values the decoded message; otherwise they take as value "not". (Note that in this case it is not allowed to make decoding errors for otherwise the coding procedure does not work.) The procedure is described by the following diagram.

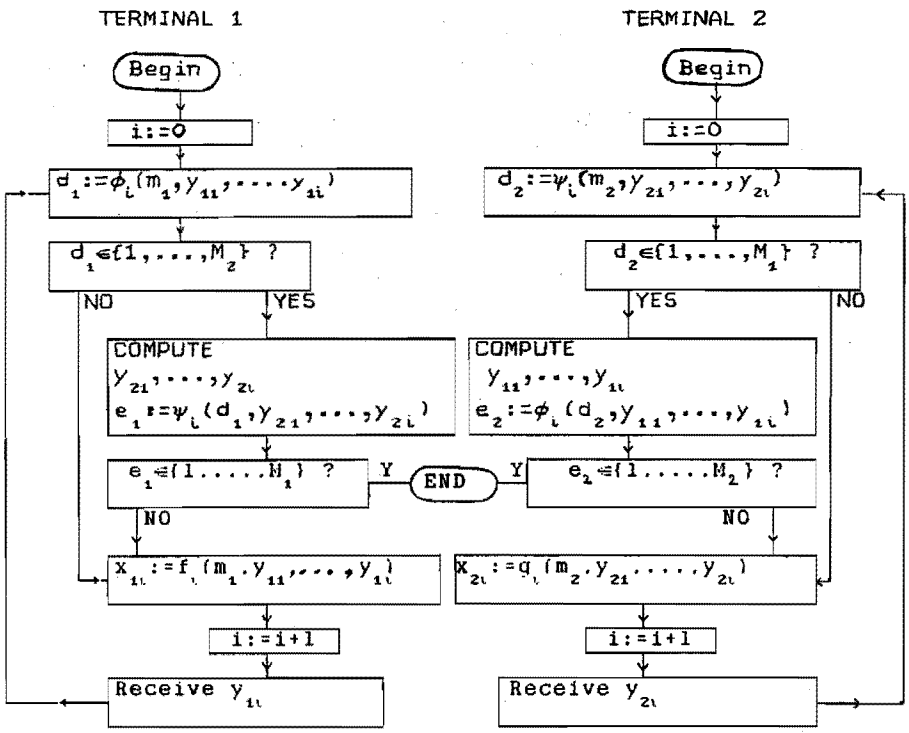



FIGURE 2. Scheme for the encoding/decoding procedure for a variable length code.

Now with each message pair $(m_1,m_2)$ there corresponds a number of transmissions $N(m_1,m_2)$. The total number of transmissions for a variable length coding procedure is defined as:

$$N:=\sum_{(m_1,m_2)\in\{1,\ldots,M_1\}\times\{1,\ldots,M_2\}} N(m_1,m_2).$$

The average length (the average number of transmissions) is equal to $n=N/(M_1M_2)$. The signalling rates for a variable length code pair are defined to be:

$$R_2^1=(\log M_1)/n \quad \text{and} \quad R_1^2=(\log M_2)/n.$$

In his paper [1], Shannon managed to prove the following remarkable result

for the capacity region $G$ of a two-way channel.

THEOREM 1. *Let* $I_2^1$ *and* $I_1^2$ *be defined as:*

$$I_2^1 := I(X_1;Y_2|X_2) \quad \textit{and} \quad I_1^2 := (X_2;Y_1|X_1).$$

*Let furthermore* $S$ *and* $T$ *be defined as:*

$$S := \{(I_1^2, I_2^1) \mid \textit{the probabilitydistribution on the inputs is a product distribution} \},$$

$$T := \{(I_1^2, I_2^1) \mid \textit{the probability distribution on the inputs is arbitrary}\}.$$

*Then the convex hull of* $S$ *and the convex hull of* $T$ *can serve as an inner bound region respectively an outer bound region for the capacity region* $G$.

That the convex hull of $S$ is an inner bound region for the capacity region $G$, Shannon proves by using a random coding argument. The outer bound region is proved by using some obvious inequalities (Cf. Sections 4-9 from [1]).
The next thing we wish to do is to specify the channel. The channel we wish to consider is called the binary multiplying channel due to Blackwell (cf. [1]) and it is depicted in Figure 3.

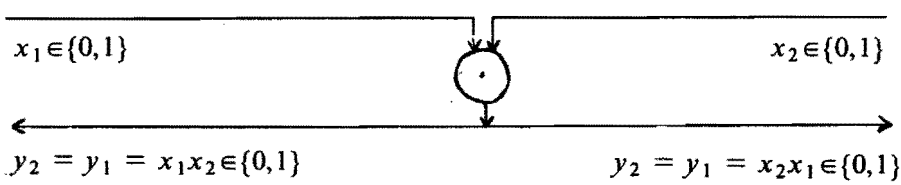



FIGURE 3. The binary multiplying channel for two users.

The inputs of the users are multiplied and the result is the output for both users. First we shall give an example of a block code for this channel, operating without error.

EXAMPLE 1.

$M_1 = M_2 = 2; n = 2.$

$f_0(1)=0,\ f_0(2)=1; g_0(1)=0,\ g_0(2)=1$

$f_1(1;0)=1,\ f_1(2;0)=0,\ f_1(2,1)=0; g_1(1;0) = 1,\ g_1(2;0)=0, g_1(2;1)=0$

$\phi(00)=2,\ \phi(01)=1,\ \phi(10)=2;\ \psi(00)=2,\ \psi(01)=1,\ \psi(10) = 2.$

$R_1^2 = 0.5,\ R_2^1 = 0.5,\ R_{\text{sum}} = R_1^2 + R_2^1 = 1.$

We can represent this code by the following diagram in which, for each message pair, the output sequence is given (in this case the output for both users is the same).

| $m_1/m_2$ | 1 | 2 |
|---|---|---|
| 1 | 01 | 00 |
| 2 | 00 | 10 |

FIGURE 4. Output table for the given code.

Next we shall give an example of a variable length code pair.

EXAMPLE 2. (Hagelbarger, Cf. Shannon [1].)

$M_1 = M_2 = 2.$

$f_0(1) = 0\ ;\ f_0(2) = 1\ ;\ f_1(1;0) = 1\ ;\ f_1(2;0) = 0.$

$g_0(1) = 0\ ;\ g_0(2) = 1\ ;\ g_1(1;0) = 1\ ;\ g_1(2;0) = 0.$

$\phi_0(1) = 2\ ;\ \phi_0(0) = \text{not}\ ;\ \phi_1(00) = 2\ ;\ \phi_1(01) = 1.$

$\psi_0(1) = 2\ ;\ \psi_0(0) = \text{not}\ ;\ \psi_1(00) = 2\ ;\ \psi_1(01) = 1.$

$N(1,1) = 2,\ N(1,2) = 2,\ N(2,1) = 2,\ N(2,2) = 1.$

$N = 7\ ;\ n = 7/4.$

$R_1^2 = 4/7,\ R_2^1 = 4/7,\ R_{\text{sum}} = 8/7 = 1.14285\ldots.$

Again we can represent the code pair in a diagram in which, for every message pair, the output is given.

| $m_1/m_2$ | 1 | 2 |
|---|---|---|
| 1 | 01 | 00 |
| 2 | 00 | 1 |

FIGURE 5. Output table for the given code.

In [3] Schalkwijk represents the information to be sent over the channel as points in the unit square. According to him, we can represent coding strategies for the binary multiplying channel as strategies for subdividing the unit square. Following his approach, we shall now calculate the best possible sumrate we can get using a Hagelbarger type of strategy.

At beginning of transmission, the situation is clearly that of Figure 6, where the message point $(m_1,m_2)$ is uniformly distributed over the unit square. The initial thresholds for $m_1$ and $m_2$ divide the unit square into four subrectangles.

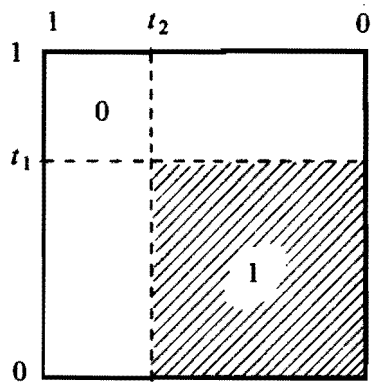



FIGURE 6. Unit square with initial $t$-thresholds.

On the first transmission, encoder $j$ produces an input $x_j=1$ if $m_j\in[0,t_j]$, otherwise sends $x_j=0$, where $j=1,2$. Hence, after receiving a 1, the message point $(m_1,m_2)$ is in the shaded subrectangle.
(i.e. $(m_1,m_2)\in[0,t_1]*[0,t_2]$) (see Figure 6). Then our task is do divide that subrectangle further, which is fully equivalent to the initial task. If a 0 was received, then the message point lies in the $L$-shaped region of Figure 6, and further resolution is necessary. Hagelbargers strategy resolves the remaining uncertainty about the message point upon receiving a 0 by immediately reusing the same $t$-thresholds on the second transmission, now sending $x_j=1$ if $m_j\in[t_j,1]$, otherwise sending $x_j=0$, where $j=1,2$. Figure 7 shows the subdivisions according to Hagelbarger. It is clear that after the second transmission we are back in the old situation, i.e. we have to determine the message point in a rectangular region. (The ambiguity upon receiving 00 is resolved, noting that each receiver knows his own message.)
Obviously we can be in two different situations: Either we must determine a message point in a rectangular region (this is called the $i$-situation) or we must determine the message point in a $L$-shaped region (which is called the $o$-situation). The statistical situation can be described by a Markov-chain where each channel use corresponds to a transition from one state to another. The transition probabilities are:

from the $i$-situation to the $i$-situation $t_1t_2$,
from the $i$-situation to the $o$-situation $1-t_1t_2$,
from the $o$-situation to the $i$-situation 1.

Therefore the Markov-chain is the one of Figure 8, where $q_i$ and $q_o$ are the stationary probabilities of the $i$-situation, respectively the $o$-situation.

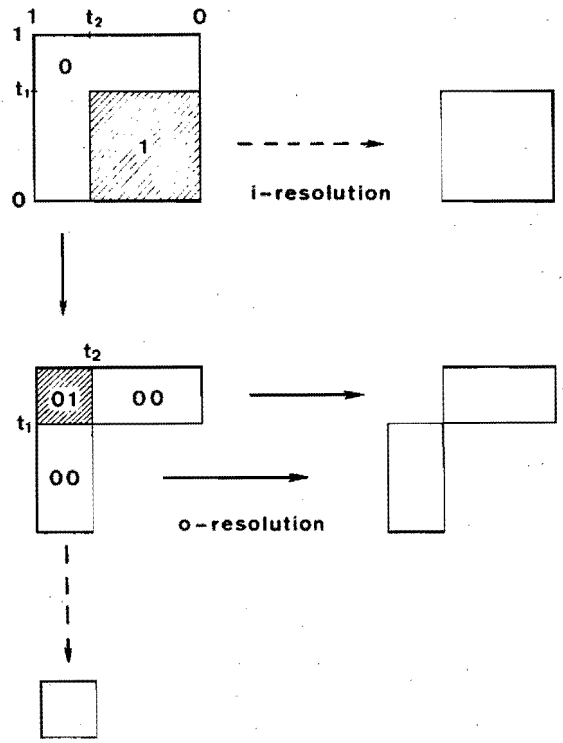



FIGURE 7. Subdivision according to Hagelbarger.

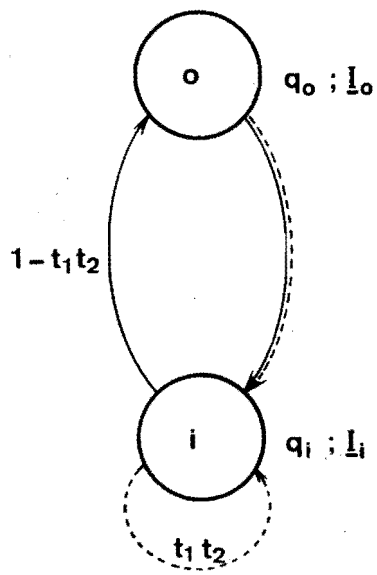



FIGURE 8. Markov-chain according to Hagelbarger strategy.

In the $o$-situation, the amount of information that passes through the channel is:

$$I_o := (I(m_1;Y_2|m_2,o-\text{sit.}),\ I(m_2;Y_1|m_1,o-\text{sit.})) = (\bar{t}_2h(\bar{t}_1),\bar{t}_1h(\bar{t}_2))/(1-t_1t_2),$$

In the $i$-situation, the amount of information that passes through the channel is:

$$I_i := (I(m_1;Y_2|m_2,i-\text{sit.}),I(m_2;Y_1|m_1,i-\text{sit.})) = (t_2h(t_1),t_1h(t_2)).$$

(Where $\bar{t}$ stands for $1-t$).
On the average, the total amount of information passing through the channel is:

$$I = q_oI_o + q_iI_i.$$

Maximizing the sum of the vector components of $I$, we get an average total amount of information passing through the channel equal to 1.1860951 for $(t_1,t_2)=(0.62587,0.62587)$. Therefore the information rate of this Hagelbarger type of strategy is 1.1860951..., (Cf. Shannon [1].).
By using a smarter coding strategy, Schalkwijk obtains as the best total information rate: 1.23828... (Cf. [3].). Later on, by refining his method, he even obtained 1.26112... as total information rate. (Cf. [4].)
The link between strategies and codes was made by Tolhuizen at the Sixth Symposium on Information Theory in the Benelux. (Cf. [5]). He showed that the total information rate obtained by Schalkwijks strategy is achievable as sumrate using block codes with vanishing probability of error. Moreover, the total information rate is achievable as sumrate using variable length codes with probability of error equal to zero.
In the coming sections, we shall generalize part of these results to the case of what is called a three-way channel (which will be defined in the next section).

## 1.2. Three-way channels

In this section, we shall give the definition of a three-way channel. Furthermore, we shall generalize some of the concepts from the previous section.
A typical three-way channel is shown in Figure 1.

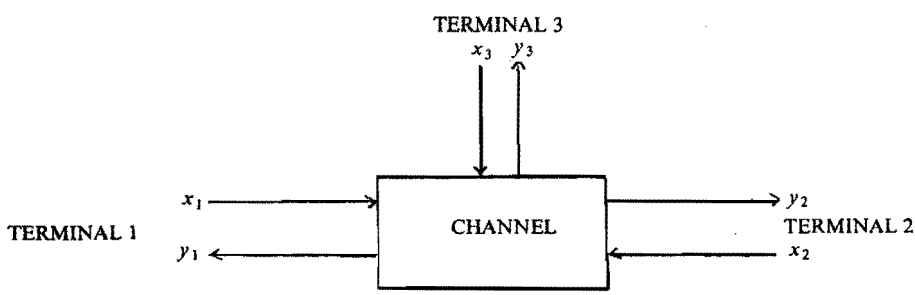



FIGURE 1. A typical three-way channel.

$x_i$ is the input letter at terminal $i(i=1,2,3)$, while $y_i$ is the output letter at

terminal $i(i=1,2,3)$. They come from alphabets $X_i$ and $Y_i$ respectively $(i=1,2,3)$.

Every channel use consists of feeding an input letter at all terminals simultaneously and obtaining an output letter at all terminals. In general, the outputs are related statistically to the inputs and maybe to previous inputs and outputs (if the channel has memory). Again the problem is to communicate as effectively as possible in all directions.

We assume that every user is interested in all the messages of the other two users. Basically, there are now 6 information flows namely from user $i$ to user $j$ for $i \neq j$, $i, j \in \{1,2,3\}$. These six flows can be replaced by three flows, containing for each user $i$ the information which flows from the other two users $j$, $k$ to $i$. The situation of Figure 1 is also described by E.C. van der Meulen [7]. Similar to E.C. van der Meulen [8] the present situation could be called a "restricted 3-way channel".

The case we wish to study is when the channel is discrete and memoryless. This means that the $X_i$ and $Y_i$ are finite sets $(i=1,2,3)$, and the outputs are only dependent on the actual inputs and not on previous inputs or outputs. A block code triple of length $n$ for such a channel with $M_i$ messages at terminal $i(i=1,2,3)$, consists of three sets of functions:

$$f_0(m_1),\ f_1(m_1,y_{11}),\ldots,\ f_{n-1}(m_1,y_{11},\ldots,y_{1\,n-1});$$
$$g_0(m_2),\ g_1(m_2,y_{21}),\ldots,\ g_{n-1}(m_2,y_{21},\ldots,y_{2\,n-1});$$
$$h_0(m_3),\ h_1(m_3,y_{31}),\ldots,\ h_{n-1}(m_3,y_{31},\ldots,y_{3\,n-1}).$$

The functions $f,g$ and $h$ take values in respectively $X_1,X_2,X_3$. $m_i$ takes values in $\{1,\ldots,M_i\}(i=1,2,3)$. The $y_{ij}$ take values in $Y_i$ $(i=1,2,3; j=1,\ldots,n-1)$. The functions $f_j$ prescribe how the input symbol $x_{1j}$ at terminal 1 will be chosen dependent on $m_1$ and $y_{11},\ldots,y_{1j-1}$, and the same for $g$ and $h$ at terminal 2 and 3. A decoding system for a block code triple of length $n$ consists of a triple of functions $\phi(m_1,y_{11},\ldots,y_{1n})$, $\psi(m_2,y_{21},\ldots,y_{2n})$ and $\chi(m_3,y_{31},\ldots,y_{3n})$ which take values in respectively $\{1,\ldots,M_2\}*\{1,\ldots,M_3\}$, $\{1,\ldots,M_1\}*\{1,\ldots,M_3\}$ and $\{1,\ldots,M_1\}*\{1,\ldots,M_2\}$. The decoding functions $\phi,\psi$ and $\chi$ give a rule to decide on the messages that were sent by the other terminals considering the message and the $n$ channel outputs at the own terminal. In the sequel, we will assume that all messages are equiprobable (at all terminals) having probabilities $1/M_i$ (at terminal $i$) $(i=1,2,3)$. The signalling rates for such a block code triple are defined to be:

$$R_k^{ij} := (\log M_i M_j)/n \quad (\{i,j,k\} = \{1,2,3\}).$$

The $R_k^{ij}$ can be seen as the amount of information that passes through the channel from the terminals $i,j$ to the terminal $k$. From a code triple and a decoding system together with the conditional probabilities which define the channel (i.e. the probabilities of the output symbols given the input symbols), one could compute the probability of incorrect decoding. Averaging over all messages, we get the error probabilities $P_{e1},P_{e2}$ and $P_{e3}$ at terminals 1,2 and 3 respectively.

The capacity region of the channel is defined to be the set of points $(R_1,R_2,R_3)$ such that for every $\epsilon$ there exists a block code triple and a decoding system having signalling rates $R_3^{12},R_1^{23},R_2^{13}$ such that $|R_k-R_k^{ij}|<\epsilon$ $(\{i,j,k\}=\{1,2,3\})$ and such that $P_{e1},P_{e2}$ and $P_{e3}$ are less than $\epsilon$. A variable length code triple consists of:

The encoding functions $f_{ir}(m_1,y_{11},\ldots,y_{ir})$ taking values from $X_i(i=1,2,3;r=0,\ldots,n-1)$.

The decoding functions $\phi_{kr}(m_k,y_{k1},\ldots,y_{kr})$ which take values from $\{1,\ldots,M_i\}*\{1,\ldots,M_j\}\cup\{\text{not}\}$ $(\{i,j,k\}=\{1,2,3\};r=0,\ldots,n)$. The value "not" is taken when it is impossible to decode without error at the given stage. (It is not allowed to make decoding errors since otherwise the algorithm does not work.)

A procedure similar to the one given in Section 1, Figure 2.

Now with each message triple $(m_1,m_2,m_3)$, there corresponds a number of transmissions (channel uses) $N(m_1,m_2,m_3)$. The total number of transmissions for a variable length coding procedure is $N:=\Sigma_{(m_1,m_2,m_3)}N(m_1,m_2,m_3)$. The average length is $n=N/(M_1,M_2,M_3)$. The signalling rates for a variable length code are defined to be

$$R_k^{ij}:=(\log M_iM_j)/n \quad (\{i,j,k\}=\{1,2,3\}).$$

## 1.3. Inner and outer bounds for the capacity region of a three-way channel

We are now going to prove the analogue of some results from [1].

THEOREM 1.

*Define* $I_k^{ij}:=I(X_i,X_j;Y_k|X_k)$,

$$I_{kj}^{i}:=I(X_i;Y_k|X_k|X_j,X_k) \qquad (\{i,j,k\}=\{1,2,3\}).$$

$$S:=\{(R_1,R_2,R_3)|(R_i+R_j-R_k)/2<I_{ij}^{k},R_i<I_i^{kj};\{i,j,k\}=\{1,2,3\};$$

*the probability distribution on the input symbols is product distribution*$\}$

$$T:=\{(I_1^{23},I_2^{13},I_3^{12})|\ \textit{the probability distribution on the inputs is arbitrary}\}$$

*Then the convex hull of S can serve as an inner bound and the convex hull of T can serve as an outer bound for the capacity region of the channel.*

PROOF: The method we follow to prove the inner bound is based on random coding techniques similar to those used by Shannon [1] and by El Gamal and Cover [6]. Consider a sequence of $n$ uses of the channel. The inputs are $\mathbf{x}_1=(x_{11},\ldots,x_{1n})$, $\mathbf{x}_2=(x_{21},\ldots,x_{2n})$ and $\mathbf{x}_3=(x_{31},\ldots,x_{3n})$ and the outputs are $\mathbf{y}_1=(y_{11},\ldots,y_{1n}),\mathbf{y}_2=(y_{21},\ldots,y_{2n})$ and $\mathbf{y}_3=(y_{31},\ldots,y_{3n})$, where the $x_{ij}\in X_i$ and the $y_{ij}\in Y_i(i=1,2,3;j=1,\ldots,n)$. The conditional probabilities for the blocks are given by:

$$P(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3|\mathbf{x}_2,\mathbf{x}_3)=\prod_k P(y_{1k},y_{2k},y_{3k}|x_{1k},x_{2k},x_{3k}),$$

taking into consideration that the channel is memoryless, and hence that successive operations are independent. We assume that there is also a probability measure for the blocks $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3$ given by the product measure of the probability measure on the $x_i (i=1,2,3)$:

$$P(\mathbf{x}_i) = \prod_k P(x_{ik}) \quad (i=1,2,3).$$

It then follows that the other probabilities are also the products of these for the individual letters.
The (unaveraged) mutual information between say $\mathbf{x}_i,\mathbf{x}_j$, and $\mathbf{y}_k$ conditional to $\mathbf{x}_k$ may be written as:

$$\begin{aligned} I(\mathbf{x}_i,\mathbf{x}_j\ ;\ \mathbf{y}_k|\mathbf{x}_k) &= \log\,(P(\mathbf{x}_i,\mathbf{x}_j,\mathbf{y}_k|\mathbf{x}_k)/P(\mathbf{x}_i,\mathbf{x}_j|\mathbf{x}_k)P(\mathbf{y}_k|\mathbf{x}_k)) = \\ &= \log\prod_r (P(x_{ir},\, x_{jr},\, y_{kr}|x_{kr})/(P(x_{ir},x_{jr}|x_{kr})P(y_{kr}|x_{kr}))) = \\ &= \sum_r I(x_{ir},x_{jr};\, y_{kr}|x_{kr}), \quad (\{i,j,k\} = \{1,2,3\}). \end{aligned}$$

Similarly for $I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_j,\mathbf{x}_k)$ we get:

$$I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_j,\mathbf{x}_k) = \sum_r I(x_{ir};y_{kr}|x_{jr},x_{kr}) \quad (\{i,j,k\} = \{1,2,3\}).$$

Thus the mutual information is as usual the sum of the individual mutual informations. Now the quantities $I(\mathbf{x}_i,\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k)$ and $I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_j,\mathbf{x}_k)$ take on different values with probabilities given by $P(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k,\mathbf{y}_k)$ $(\{i,j,k\}=\{1,2,3\})$.
The distribution functions for $I(\mathbf{x}_i,\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k)$ and $I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_j,\mathbf{x}_k)$ will be denoted by $\rho_k^{ij}(z)$ and $\rho_{kj}^{i}(z)$ respectively $(\{i,j,k\}=\{1,2,3\})$,

$$\rho_k^{ij}(z) := P(I(\mathbf{x}_i,\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k) \leqslant z) \text{ and}$$

$$\rho_{kj}^{i}(z) := P(I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_k,\mathbf{x}_j) \leqslant z) \quad (\{i,j,k\} = \{1,2,3\}).$$

Since each of the random variables $I(\mathbf{x}_i,\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k)$ and $I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_k,\mathbf{x}_j)$ is the sum of $n$ independent random variables, each with the same distribution, we can apply the central limit theorems and the law of large numbers. The mean of the distributions $\rho_k^{ij}$ and $\rho_{kj}^{i}$ will be $nI_k^{ij}$ and $nI_{kj}^{i}$ respectively $(\{i,j,k\}=\{1,2,3\})$, and the variances are $n$ times the corresponding variances for single letters. As $n$ tends to infinity $\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon))$ and $\rho_k^{ij}(n(I_k^{ij}-\epsilon))$ tend to 0 for any fixed positive $\epsilon$. In fact this approach is exponential in $n$:

$$\rho_k^{ij}(n(I_k^{ij}-\epsilon)) \leqslant \exp(-A(\epsilon)n) \text{ for some } A(\epsilon)>0,$$

$$\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon)) \leqslant \exp(-B(\epsilon)n) \text{ for some } B(\epsilon)>0 \text{ (Cf. Shannon [1]).}$$

After these preliminary definitions and results, we now wish to prove the existence of codes with error probabilities bounded by expressions involving the $\rho_k^{ij}$ and $\rho_{kj}^{i}$. We will construct an ensemble of code triples and establish bounds on the error probabilities $P_{e1},P_{e2}$ and $P_{e3}$ averaged over the ensemble. From this result it follows that there exists a particular code triple in the ensemble with related bounds on its error probabilities. The random ensemble

of code triples for a three-way channel with $M_i$ words at terminal $i(i=1,2,3)$, is constructed as follows. The $M_i$ integers $\{1,...,M_i\}$ are mapped in all possible ways into the set of input words $X_i^n$ of length $n(i=1,2,3)$.
If there were $a_i$ possible input letters at terminal $i$, there will be $a_i^n$ input words of length $n$ and $a_i^{nM_i}$ mappings at terminal $i(i=1,2,3)$. Such a mapping serves as an encoding functions. We consider all $a_1^{nM_1}a_2^{nM_2}a_3^{nM_3}$ code triples obtained in this way. Each code triple is given a probability equal to the occurrence of that triple if the three mappings were done independently, and an integer is mapped into a word with the assigned probability of that word. Thus a code triple is given a probability equal to the product of the probabilities associated with the input words that the integers are mapped into (for all three codes). This set of code triples with these associated probabilities we call the random ensemble of code triples based on the assigned probabilities $P(x_i)$ $(i=1,2,3)$. Any particular code triple of the ensemble could be used to transmit information, if we agreed on a method of decoding. The method of decoding will consist of three functions $\phi(\mathbf{x}_1,\mathbf{y}_1),\psi(\mathbf{x}_2,\mathbf{y}_2)$ and $\chi(\mathbf{x}_3,\mathbf{y}_3)$. Here $\mathbf{x}_i$ varies over the input words of length $n$ at terminal $i$, and $\mathbf{y}_i$ varies over the possible received blocks of length $n(i=1,2,3)$. The functions $\phi,\psi$ and $\chi$ take values from $\{1,...,M_2\}*\{1,...,M_3\},\{1,...,M_1\}*\{1,...,M_3\}$ and $\{1,...,M_1\}*\{1,...,M_2\}$ respectively, and represent the decoded message pair for a received $\mathbf{y}_i$ if $\mathbf{x}_i$ was transmitted $(i=1,2,3)$.
It should be noted that the decoding functions $\phi,\psi$ and $\chi$ need not be the same for all code triples in the ensemble. We also point out that the encoding functions for our random ensemble are more specialized than in the general case. Indeed the sequence of input letters $\mathbf{x}_i$ for a given message $m_i$ does not depend on the received letters at terminal $i(i=1,2,3)$. In any particular code of the ensemble, there is a strict mapping from messages to input sequences. Given an ensemble of code triples as described above and decoding functions, one could compute for each particular code triple three error probabilities: $P_{ei}$ (the probability of a decoding error at terminal $i$), $i=1,2,3$. We assume that the different messages occur with equal probability at each terminal. By the average error probability for the ensemble of code triples, we mean the average $\mathbb{E}P_{ei}(i=1,2,3)$ where each probability of error is weighted according to the probability associated to the code triple. We wish to describe a particular method of decoding, that is a choice for $\phi,\psi$ and $\chi$, and then place upper bounds on these average error probabilities.

LEMMA 1. *Suppose the probability assignments* $P(x_i)$ $(i=1,2,3)$ *in a discrete memoryless channel produce information distributions* $\rho_k^{ij}(z)$ *and* $\rho_{kj}^{i}(z)$. *Let* $R_1,R_2$ *and* $R_3$ *be such that* $(R_k+R_j-R_i)/2<I_{kj}^{i}$ *and* $R_k<I_k^{ij}$ $(\{i,j,k\}=\{1,2,3\})$. *Let* $\epsilon>0$.
*The random ensemble of code triples with* $M_i=2^{n(R_k+R_j-R_i)/2}$ *messages at terminal i has (with appropriate decoding functions) average error probabilities bounded as follows:*

$$\mathbb{E}P_{ek}\leqslant 2^{-n(I_{ki}^{j}-(R_i+R_k-R_j)/2-\epsilon)}+2^{-n(I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon)}+$$

$$+2^{-n(I_k^{ij}-R_k-\epsilon)}+2^{-n(I_k^{ij}-R_k-\epsilon+I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon+I_{ki}^{j}-(R_i+R_k-R_j/2-\epsilon}+$$

$$+\rho_k^{ij}(n(I_k^{ij}-\epsilon))+\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon))+\rho_{ki}^{j}(n(I_{ki}^{j}-\epsilon))$$

*Furthermore there exists at least one code triple in the ensemble for which the individual error probabilities are bounded by three times these expressions* ($\{i,j,k\}=\{1,2,3\}$).

This lemma is a generalization of Theorem 1 in [1] (compare also with Section V of [6]).

PROOF OF LEMMA: The statistical events involved are the following:

1) The choice of the messages $m_1,m_2$ and $m_3$.
2) The choice of the code triple in the ensemble of code triples.
3) The statistics of the channel, i.e. the conditional probabilities $P(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3|\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)$.

The ensemble error probabilities we are calculating are averages over all these statistical events.

We first define decoding systems for the various codes in the ensemble. For each pair $(\mathbf{x}_k,\mathbf{y}_k)$ and for every $\epsilon>0$ define a corresponding set of words:

$$S_k^{\epsilon}(\mathbf{x}_k,\mathbf{y}_k) := \{(\mathbf{x}_i,\mathbf{x}_j)\in X_i^n * X_j^n |\ I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_k,\mathbf{x}_j)\geq n(I_{kj}^{i}-\epsilon),$$

$$I(\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k,\mathbf{x}_i)\geq n(I_{ik}^{j}-\epsilon),$$

$$I(\mathbf{x}_i,\mathbf{x}_k;\mathbf{y}_k|\mathbf{x}_k)\geq n(I_k^{ij}-\epsilon)\}.$$

We will use the sets $S_k^{\epsilon}$ ($k=1,2,3$) to define the decoding procedure and to aid in overbounding the error probabilities. The decoding process will be as follows. In any particular code triple in the random ensemble, suppose message $m_k$ is sent and this is mapped into input word $\mathbf{x}_k$. Suppose that $\mathbf{y}_k$ is received at terminal $k$. Consider the set $S_k^{\epsilon}(\mathbf{x}_k,\mathbf{y}_k)$.

If there is no message pair $(m_i,m_j)$ mapped into $S_k^{\epsilon}(\mathbf{x}_k,\mathbf{y}_k)$ for the code triple in question, then $\mathbf{y}_k$ is decoded by convention as message pair (1,1). If there is exactly one message pair mapped into $S_k^{\epsilon}(\mathbf{x}_k,\mathbf{y}_k)$, decode $\mathbf{y}_k$ as this message pair .

If there is more than one message pair mapped into $S_k^{\epsilon}(\mathbf{x}_k,\mathbf{y}_k)$, then decode as the lexicographically least such message pair.

The error probabilities that we are estimating would normally be calculated in the following way: For each code triple, calculate the error probabilities for all message triples $m_1,m_2,m_3$ and from their averages get the error probabilities for the code triple. Then these error probabilities are averaged over the ensemble of code triples, using the appropriate probabilities. We change the order of averaging (summation).

Fix $\tilde{m}_1,\tilde{m}_2$ and $\tilde{m}_3$ and the $\tilde{\mathbf{x}}_1,\tilde{\mathbf{x}}_2$ and $\tilde{\mathbf{x}}_3$ to which they are mapped, and the received words $\tilde{\mathbf{y}}_1,\tilde{\mathbf{y}}_2$ and $\tilde{\mathbf{y}}_3$. There is still, in the statistical picture, the range of possible code triples, that is mappings of the other $M_k-1$ messages for the code at terminal $k$ ($k=1,2,3$).

The error probabilities will now be bounded by the sum of:

1) the average probability (over the subset of code triples) of the message pairs other than $(\tilde{m}_i,\tilde{m}_j)$ being mapped into $S_k^\epsilon$ and
2) the average probability that $(\tilde{m}_i,\tilde{m}_j)$ is mapped outside $S_k^\epsilon$.

We will now show that averaged over the subset of code triples, the probability of the message pairs $(m_i,m_j)\neq(\tilde{m}_i,\tilde{m}_j)$ being mapped into $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$ does not exceed:

$$2^{-n(I_k^{ij}-R_k-\epsilon)}+2^{-n(I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon)}+2^{-n(I_{ki}^{j}-(R_i+R_k-R_j)/2-\epsilon)}+$$
$$2^{-n(I_k^{ij}-R_k-\epsilon+I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon+I_{ki}^{j}-(R_k+R_i-R_j)/2-\epsilon)}(\{i,j,k\}=\{1,2,3\}).$$

Indeed the probability that all other message pairs will be mapped outside $S_k^\epsilon\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$ is equal to:

$$P((m_i,m_j) \text{ is mapped outside } S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)\ ,\ m_i\neq\tilde{m}_i\ ,\ m_j\neq\tilde{m}_j)\ *$$
$$P((m_i,\tilde{m}_j) \text{ is mapped outside } S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)\ ,\ m_i\neq\tilde{m}_i\ )*$$
$$P((\tilde{m}_i,m_j) \text{ is mapped outside } S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)\ ,\ m_j\neq\tilde{m}_j).$$

First note that if $(\mathbf{x}_i,\mathbf{x}_j)$ belongs to the set $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$, then by definition:

$$I((\mathbf{x}_i,\mathbf{x}_j);\ \tilde{\mathbf{y}}_k|\tilde{\mathbf{x}}_k)\geqslant n(I_k^{ij}-\epsilon).$$

So:

$$\log(P(\mathbf{x}_i,\mathbf{x}_j,\tilde{\mathbf{y}}_k|\tilde{\mathbf{x}}_k)/(P(\mathbf{x}_i,\mathbf{x}_j|\tilde{\mathbf{x}}_k)P(\tilde{\mathbf{y}}_k|\tilde{\mathbf{x}}_k)))\geqslant n(I_k^{ij}-\epsilon),$$

and hence using the fact that $\mathbf{x}_i$ and $\mathbf{x}_j$ are statistically independent of $\tilde{\mathbf{x}}_k$:

$$\log(P(\mathbf{x}_i,\mathbf{x}_j,\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)/(P(\mathbf{x}_i,\mathbf{x}_j)P(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)))\geqslant n(I_k^{ij}-\epsilon).$$

Therefore:

$$P(\mathbf{x}_i,\mathbf{x}_j|\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)\geqslant P(\mathbf{x}_i,\mathbf{x}_j)*2^{n(I_k^{ij}-\epsilon)}.$$

Now summing both sides over the pairs $(\mathbf{x}_i,\mathbf{x}_j)$ belonging to $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$, we obtain:

$$1\geqslant\sum_{(\mathbf{x}_i,\mathbf{x}_j)\in S_k^\epsilon}P(\mathbf{x}_i,\mathbf{x}_j|\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)\geqslant 2^{n(I_k^{ij}-\epsilon)}\sum_{(\mathbf{x}_i,\mathbf{x}_j)\in S_k^\epsilon}P(\mathbf{x}_i,\mathbf{x}_j)$$

The left inequality follows from the fact that a sum of disjoint probabilities cannot exceed one. The sum on the right we may denote by $P(S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k))$ and we find:

$$P(S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k))\leqslant 2^{-n(I_k^{ij}-\epsilon)}.$$

So the total probability associated with any set $S_k^\epsilon(\tilde{x}_k,\tilde{y}_k)$ is bounded by an expression involving $n,I_k^{ij}$ and $\epsilon$ but independent of the particular $(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$. The messages satisfying $m_i\neq\tilde{m}_i,m_j\neq\tilde{m}_j$ and $m_k\neq\tilde{m}_k$ were mapped independently into the input words using the probabilities $P(x_r)$ $(r=1,2,3)$. The probability

of a particular message pair $m_i \neq \tilde{m}_i$ and $m_j \neq \tilde{m}_j$ being mapped into $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$ in the ensemble of code triples is therefore just: $P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$. The probability of being in the complementary set is $1 - P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$.
Therefore:
$P((m_i, m_j)$ is mapped outside $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k), m_i \neq \tilde{m}_i, m_j \neq \tilde{m}_j) =$

$$(1 - P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)))^{M_i M_j - M_i - M_j + 1} \geq$$

$$1 - (M_i M_j - M_i - M_j + 1) P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)) \geq 1 - M_i M_j P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)) \geq$$

$$1 - M_i M_j 2^{-n(I_k'' - \epsilon)} = 1 - 2^{-n(I_k'' - R_k - \epsilon)}.$$

Here we used the inequality $(1-x)^p \geq 1 - px$, the fact that $P(S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)) \leq 2^{-n(I_k'' - \epsilon)}$ and the fact that $M_i M_j = 2^{nR_k}$ by the definition of $M_i$ and $M_j$.
Secondly note that if $(\mathbf{x}_i, \tilde{\mathbf{x}}_j)$ belongs to the set $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$, then by definition: $(I(\mathbf{x}_i; \tilde{\mathbf{y}}_k | \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j) \geq n(I_{kj}' - \epsilon)$, so:

$$\log(P(\mathbf{x}_i, \tilde{\mathbf{y}}_k | \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j) / (P(\mathbf{x}_i | \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j) P(\tilde{\mathbf{y}}_k | \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j))) \geq n(I_{kj}' - \epsilon),$$

and hence using the fact that $x_i$ is statistically independent of $\tilde{\mathbf{x}}_k$ and $\tilde{\mathbf{x}}_j$:

$$\log(P(\mathbf{x}_i, \tilde{\mathbf{y}}_k, \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j) / P(\mathbf{x}_i) P(\tilde{\mathbf{y}}_k, \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j))) \geq n(I_{kj}' - \epsilon).$$

Therefore:

$$P(\mathbf{x}_i | \tilde{\mathbf{y}}_k, \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_j) \geq P(\mathbf{x}_i) 2^{n(I_{kj}' - \epsilon)}.$$

Summing both sides over the pairs $(\mathbf{x}_i, \tilde{\mathbf{x}}_j)$ belonging to $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$, we obtain:

$$1 \geq \sum_{(\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)} P(\mathbf{x}_i | \tilde{\mathbf{y}}_k, \tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_j) \geq 2^{n(I_{kj}' - \epsilon)} \sum_{(\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)} P(\mathbf{x}_i).$$

The left inequality follows from the fact that a sum of disjoint probabilities cannot exceed 1. The sum on the right hand side we may denote by $P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$ and we find: $P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)) \leq 2^{-n(I_{kj}' - \epsilon)}$. The probability that a message pair $(m_i, \tilde{m}_j)$ $(m_i \neq \tilde{m}_i)$ is mapped into $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$, (in the ensemble of code triples) is just $P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$. The probability of being in the complementary set is $1 - P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$. So:
$P((m_i, \tilde{m}_j)$ is mapped outside $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k), m_i \neq \tilde{m}_i) =$

$$= (1 - P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon((\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)))^{M_i - 1} \geq 1 - (M_i - 1) P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)) \geq$$

$$\geq 1 - M_i 2^{-n(I_{kj}' - \epsilon)} \geq 1 - 2^{-n(I_{kj}' - (R_j + R_k - R_i)/2 - \epsilon)}.$$

Here we used the inequality $(1-x)^p \geq 1 - px$, the estimate on $P((\mathbf{x}_i, \tilde{\mathbf{x}}_j) \in S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k))$ and the definition of $M_i$. Thirdly in the same way as above, we obtain the inequality: $P((\tilde{m}_i, m_j)$ is mapped outside $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k), m_j \neq \tilde{m}_j) \geq 1 - 2^{-n(I_{kj}' - (R_k + R_i - R_j)/2 - \epsilon)}$. Therefore the total probability that the messages other than $(\tilde{m}_i, \tilde{m}_j)$ will be mapped outside $S_k^\epsilon(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k)$ is larger than or equal to:

$$(1 - 2^{-n(I_k'' - R_k - \epsilon)})(1 - 2^{-n(I_{kj}' - (R_k + R_j - R_i)/2 - \epsilon)})^\star$$

$$\cdot(1-2^{-n(I'_{ki}-(R_i+R_k-R_j)/2-\epsilon)})\geqslant$$
$$1-2^{-n(I''_k-R_k-\epsilon)}-2^{-n(I'_{kj}-(R_k+R_j-R_i)/2-\epsilon)}-2^{-n(I'_{ki}-(R_k+R_i-R_j)/2-\epsilon)}+$$
$$-2^{-n(I''_k-R_k-\epsilon+I'_{kj}-(R_k+R_j-R_i)/2-\epsilon+I'_{ki}-(R_k+R_i-R_j)/2-\epsilon)}.$$

We have established that in the subset of cases under consideration (i.e. $\tilde{m}_1,\tilde{m}_2$ and $\tilde{m}_3$ are mapped into $\tilde{\mathbf{x}}_1,\tilde{\mathbf{x}}_2$ and $\tilde{\mathbf{x}}_3$ and received as $\tilde{\mathbf{y}}_1,\tilde{\mathbf{y}}_2$ and $\tilde{\mathbf{y}}_3$), with probability at least the above expression, there will be no other message pairs mapped into $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$ than $(\tilde{m}_i,\tilde{m}_j)$. These bounds are independent of the particular $\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k$ (as noted before). This proves the bound on the average total probability of the message pairs $(m_i,m_j)\neq(\tilde{m}_i,\tilde{m}_j)$ being mapped into $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$. We now bound the probability of the actual message pair $(\tilde{m}_i,\tilde{m}_j)$ being within the subset $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$.
From the definition of $\rho_k^{ij}(z)$:

$$\rho_k^{ij}(n(I_k^{ij}-\epsilon)) = P(I(\mathbf{x}_i,\mathbf{x}_j;\mathbf{y}_k|\mathbf{x}_k)\leqslant n(I_k^{ij}-\epsilon),$$

and similarly:

$$\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon)) = P(I(\mathbf{x}_i;\mathbf{y}_k|\mathbf{x}_k,\mathbf{x}_j)\leqslant n(I_{kj}^{i}-\epsilon).$$

In the ensemble of code triples a message $\tilde{m}_k$, say is mapped into words $\tilde{\mathbf{x}}_k$ with probabilities just equal to $P(\tilde{\mathbf{x}}_k)$. Consequently, the probability in the full ensemble of code triples, message choices and channel statistics, that the actual message pair is mapped outside $S_k^\epsilon(\tilde{\mathbf{x}}_k,\tilde{\mathbf{y}}_k)$ is less than or equal to:

$$\rho_k^{ij}(n(I_k^{ij}-\epsilon))+\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon))+\rho_{ki}^{j}(n(I_{ki}^{j}-\epsilon)).$$

Now the probability of a decoding error can be bounded by the sum of the above probabilities.
This proves the first part of the lemma. To prove the second part of the lemma, we observe that we can apply the same lemma as in [1].

LEMMA 2. *Suppose we have a set of objects $B_1,B_2,\ldots,B_n$ with associated probabilities $P_1,P_2,\ldots,P_n$ and a number of numerically valued properties (functions) of the objects $f_1,\ldots,f_d$, that are nonnegative, $f_i(B_j)\geqslant 0$, and suppose that we know the average $A_i$ of these properties over the objects, $\Sigma_j P_j f_i(B_j)=A_i$ $i=1,\ldots,d$.*
*Then there exists an object $B_p$ for which $f_i(B_p)\leqslant dA_i$ $i=1,\ldots,d$. More generally, given any set of $K_i>0$ satisfying $\Sigma_{i=1}^{d}1/K_i\leqslant 1$, there exists an object $B_p$ with $f_i(B_p)\leqslant K_iA_i$ $i=1,\ldots,d$.*

For the proof of this lemma we refer to Shannon [1]. The rest of Lemma 1 is now proved by using Lemma 2 with as objects code triples and as properties the error probabilities $P_{ek}$. These are nonnegative and their averages are bounded as shown before. It follows from Lemma 2 that there exists a code triple for which simultaneously $P_{ek}$ is less than three times the bound on the average $\mathbb{E}P_{ek}$.

This concludes the proof of Lemma 1.

We now prove Theorem 1. Let $(R_1,R_2,R_3)$ be in $S$. There exists some product distribution giving rise to distribution functions $\rho_k^{ij}$ and $\rho_{kj}^{i}$ and mean mutual informations $I_k^{ij},I_{kj}^{i}$ ($\{i,j,k\}=\{1,2,3\}$), and such that

$$\tfrac{1}{2}(R_k+R_j-R_i)\leqslant I_{kj}^{i}-2\epsilon, R_k\leqslant I_k^{ij}-2\epsilon, \text{ for some } \epsilon>0.$$

Now by Lemma 1 we can find, for $n$ arbitrary, at least one code triple for which the individual error probabilities are bounded by:

$$\begin{aligned}P_{ek}\leqslant 3\,(&2^{-n(I_k^{ij}-R_k-\epsilon)}+2^{-n(I_{ki}^{j}-(R_k+R_i-R_j)/2-\epsilon)}+2^{-n(I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon)}\\&+2^{-n(I_k^{ij}-R_k-\epsilon+I_{kj}^{i}-(R_k+R_j-R_i)/2-\epsilon+I_{ki}^{j}-(R_k+R_i-R_j)/2-\epsilon)}\\&+\rho_k^{ij}(n(I_k^{ij}-\epsilon))+\rho_{kj}^{i}((I_{kj}^{i}-\epsilon))+\rho_{ki}^{j}(n(I_{ki}^{j}-\epsilon)))\leqslant\\\leqslant 3(&2^{-n\epsilon}+2^{-n\epsilon}+2^{-n\epsilon}+2^{-3n\epsilon})+\rho_k^{ij}(n(I_k^{ij}-\epsilon))+\rho_{kj}^{i}(n(I_{kj}^{i}-\epsilon))+\\&+\rho_{ki}^{j}(n(I_{ki}^{j}-\epsilon))).\end{aligned}$$

If $n$ tends to infinity, all terms go to 0.

Analogous to the method described in Section 8 of [1], we can prove that in fact we could take the convex hull of $S$ as an inner bound region. This proves the first part of Theorem 1.

The next thing to do is to prove that the convex hull of $T$ is an outer bound region for the capacity region. That will finish the proof of Theorem 1. Suppose we have a code triple starting at time zero with messages $m_1,m_2$ and, $m_3$ at the three terminals. After $n$ transmissions let $\mathbf{y}_1,\mathbf{y}_2$ and $\mathbf{y}_3$ be the received blocks at the three terminals. Let $x_1,x_2,x_3,y_1,y_2$ and $y_3$ be the next transmitted and received letters. Consider the change in "equivocation" of message pairs at the three terminals due to the next received letter. At terminal $k$ for example this change is (making some obvious reductions cf. Shannon [1]):

$$\begin{aligned}\Delta &= H(m_i,m_j|m_k,\mathbf{y}_k)-H(m_i,m_j|m_k,\mathbf{y}_k,y_k)=\\&\mathbb{E}(\log(P(m_k,\mathbf{y}_k)/P(m_i,m_j,m_k,\mathbf{y}_k)))-\mathbb{E}(\log(P(m_k,\mathbf{y}_k,y_k)/P(m_i,m_j,m_k,\mathbf{y}_k,y_k)=\\&\mathbb{E}(\log((P(y_k|m_i,m_j,m_k,\mathbf{y}_k)/P(y_k|x_k))*(P(y_k|x_k)/P(y_k|\mathbf{y}_k,m_k)))).\end{aligned}$$

Now

$$H(y_k|m_i,m_j,m_k,\mathbf{y}_k)\geqslant H(y_k|m_i,m_j,m_k,\mathbf{y}_i,\mathbf{y}_j,\mathbf{y}_k)=H(y_k|x_i,x_j,x_k),$$

since adding a condition variable cannot increase entropy and since

$$P(y_k|m_i,m_j,m_k,\mathbf{y}_i,\mathbf{y}_j,\mathbf{y}_k)=P(y_k|x_i,x_j,x_k)\quad\{i,j,k\}=\{1,2,3\}.$$

Also $H(y_k|x_k)\geqslant H(y_k|\mathbf{y}_k,m_k)$ since $x_k$ is a function of $\mathbf{y}_k$ and $m_k$ by the encoding function.

Therefore

$$\begin{aligned}\Delta\leqslant\ &\mathbb{E}(\log(P(y_k|x_i,x_j,x_k)/P(y_k|x_k)))+H(y_k|\mathbf{y}_k,m_k)-H(y_k|x_k)\leqslant\\&\mathbb{E}(\log(P(y_k|x_i,x_j,x_k)/P(y_k|x_k)))=\end{aligned}$$

$$\mathbb{E}(\log(P(y_k,x_i,x_j,x_k)P(x_k)/(P(y_k,x_k)P(x_i,x_j,x_k)))) =$$
$$\mathbb{E}(\log(P(x_i,x_j|x_k,y_k)/P(x_i,x_j|x_k)) = I(X_i,X_j;Y_k|X_k) = I_k^{ij}.$$

Thus the vector change in equivocation is included in the convex hull of all triples $(I_1^{23},I_2^{13},I_3^{12})$ when the input distribution is varied. In a code of length $n$, the total change in equivocation from beginning to end of the block cannot exceed the sum of $n$ such vectors and hence is included in $n$ times this convex hull (which is the convex hull of $T$). If a code has signalling rates $R_k^{ij}$, then the initial equivocations are $nR_k^{ij}$. If we assume that the point $n(R_1^{23},R_2^{13},R_3^{12})$ is outside the convex hull of $nT$ with nearest distance $n\epsilon$, we can construct (using the convexity) a plane $P$ passing through the nearest point of $nT$ and perpendicular to the nearest approach segment with $nT$ on one side. It is clear that for any point $n(R_1^*,R_2^*,R_3^*)$ on the $nT$ side of $P$ and in particular for any point of $nT$:

$$\Sigma|\, n(R_k^{ij}-R_k^*)| \geqslant n\epsilon \quad \text{(since the shortest distance is } n\epsilon\text{)},$$

and furthermore one of the $|n(R_k^{ij}-R_k^*)|$ is at least $n\epsilon/\sqrt{3}$. (One of the components of a vector is at least as large as the vector length divided by $\sqrt{3}$). Thus after $n$ uses of the channel, if the signalling rates $R_k^{ij}$ are outside the convex hull of $T$, at least one of the final equivocations is at least $\epsilon/\sqrt{3}$. Thus for signalling rates outside of $T$, the equivocations per second are bounded from below independent of the code length. This implies that the error probability is also bounded from below (Cf. Shannon [1] Section 9). This concludes the proof of Theorem 1.

## 1.4. The binary multiplying channel as a three-way channel: Bounds

In the previous section, we studied three-way channels in general and derived inner and outer bounds for the capacity region. We will now look at a more specific example of a three-way channel. The channel we are interested in is depicted in Figure 1.

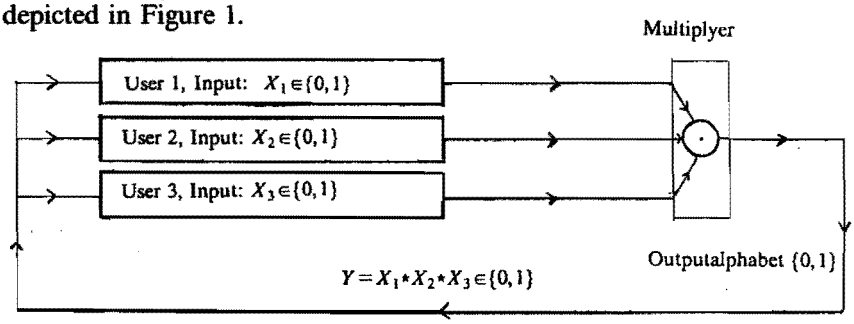



FIGURE 1. The BMC as a three-way channel.

As measure for the information which passes through the channel we take the sumrate:

$$R_{\text{sum}} = R_1^{23} + R_2^{13} + R_3^{12},$$

where as before, $R_k^{ij}$ stand for the signalling rates of a code triple.
Before calculating the bounds for $R_{\text{sum}}$ following from the inner and outer bounds of the previous section, we first give some trivial bounds. It is easy to see that when two users send a 1, the third one can send one bit of information along the channel. However since this information goes into two directions the contribution to $R_{\text{sum}}$ equals 2. The conclusion is that using timesharing, the three senders can communicate with a rate of $R_{\text{sum}}=2$. So we have the following lower bound:

TIMESHARING LOWER BOUND: $R_{\text{sum}} \geqslant 2$.

Of course we have the following trivial observation:

$$R_k^{ij} \leqslant 1.$$

So we have the following upper bound:

TRIVIAL UPPER BOUND: $R_{\text{sum}} \leqslant 3$.

Now define

$I_k^{ij} := I(X_i,X_j;Y|X_k)$, $I_{kj}^{i} := I(X_i;Y|X_k,X_j)$,

$S := \{(R_1,R_2,R_3) | (R_i+R_j-R_k)/2 < I_{ij}^k, R_k < I_k^{ij} \quad (\{i,j,k\} = \{1,2,3\})$

the probability distribution on the inputs is a product distribution$\}$.

$T := \{(I_1^{23},I_2^{13},I_3^{12}) |$ the probabiltiy distribution on the inputs is arbitrary$\}$.

From the previous section, we know that the convex hull of the set $S$ can serve as an inner bound and the convex hull of the set $T$ can serve as an outer bound for the capacity region. This means that we can find a nonconstructive lower bound for $R_{\text{sum}}$ by taking:

$$\max \{(R_1+R_2+R_3) | \tfrac{1}{2}(R_i+R_j-R_k) \leqslant I_{ij}^k,\ R_k \leqslant I_k^{ij} \quad (\{i,j,k\} = \{1,2,3\}),$$

the probability distribution on the inputs is a product distribution$\}$.

The same applies in finding an upper bound for $R_{\text{sum}}$:

$\max \{I_3^{12}+I_2^{13}+I_1^{23} |$ the probability distribution on the inputs is arbitrary$\}$

In the case of independent input probabilities define:

$$p_i := P(X_i=1), \quad i=1,2,3.$$

Then

$$P(X_i=0) = 1-p_i,\ i=1,2,3.$$

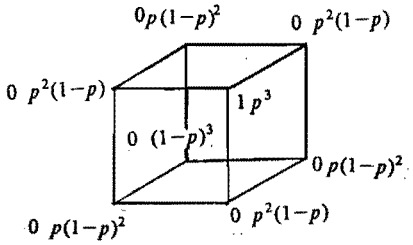
For symmetry reasons, the maximal value of $I_3^{12}+I_2^{13}+I_1^{23}$ is obtained if the $p_i$ are equal:

$$p_1 = p_2 = p_3 =: p.$$

Then

$$I_3^{12} = I_2^{13} = I_1^{23} = ph(p^2) \text{ and } I_{kj}^{i} = p^2 h(p) \ (\{i,j,k\} = \{1,2,3\}).$$

This can be cheked easily from Figure 2.




$P(0,0,0) = (1-p)^3$
$P(1,0,0) = P(0,1,0) = P(0,0,1) = p(1-p)^2$
$P(1,1,0) = P(1,0,1) = P(0,1,1,) = p^2(1-p)$
$P(1,1,1) = p^3$
$I_k^{ij} = ph(p^2),\ I_{kj}^{i} = p^2 h(p).$

FIGURE 2. Statistical situation when using equally distributed independent inputs. Vertices correspond to the various inputs $(x_1, x_2, x_3)$. The numbers involving the $p$'s at the vertices are the probabilities with which they appear. The symbols 0,1 at the vertices are the outputs.

By numerical optimization, we find an absolute maximum for $I_k^{ij} = ph(p^2)$ when $p = 0.79267....$ Therefore by taking $R_k = I_k^{ij} = ph(p^2) = 0.75458...$ and verifying that $(R_k + R_j - R_i)/2 = 0.37729... \leqslant 0.46266... = p^2 h(p) = I_{kj}^{i}$, we find a nonconstructive lower bound for $R_{\text{sum}}$:

$$R_{\text{sum}} \geqslant 2.2637441....$$

If we wish to maximize $I_3^{12} + I_2^{13} + I_1^{23}$ over all input probability distributions, then for symmetry reasons it is clear that the maximum is attained when we take:

$$P(1,0,0) = P(0,1,0) = P(0,0,1)$$
$$P(1,1,0) = P(1,0,1) = P(0,1,1).$$

Define:

$$p := P(0,0,0), q := P(1,0,0) = P(0,1,0) = P(0,0,1), r = P(1,1,0) = P(1,0,1) = P(0,1,1)$$
$$s := P(1,1,1).$$

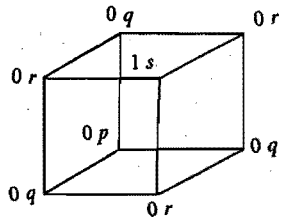



FIGURE 3. Probabilities and output symbols to obtain the maximal value for $I_3^{12}+I_2^{13}+I_1^{23}$.

It follows that $p+3q+3r+s=1$. As can be checked easily from Figure 3, we have in this case:

$$I_3^{12} = I_2^{13} = I_1^{23} = (q+2r+s)h(s/(q+2r+s)).$$

Here again vertices correspond to the various inputs. The letters $p,q$ and $r$ at the vertices correspond to the probabilities with which they appear. The symbols 0,1 are the output symbols.
Now:

$$\max\{(q+2r+s)h(s/(q+2r+s))|p+3q+3r+s=1\} =$$
$$\max\{(q+2r+s)h(s/(q+2r+s))|0\leqslant 3q+3r+s\leqslant 1\}.$$

Differentiating the given function with respect to $s$, we find

$$\log((q+2r+s)/s).$$

Since this is always greater than or equal to 0, the maximum is obtained at the boundary:

$$\max\{(q+2r+s)h(s/(q+2r+s))|0\leqslant 3q+3r+s\leqslant 1\} =$$
$$\max\{(q+2r+s)h(s/(q+2r+s))|3q+3r+s=1\} =$$
$$\max\{(1-2q-r)h((1-3q-3r)/(1-2q-r))|0\leqslant(q+r)\leqslant 1/3\}.$$

Calculating the partial derivative with respect to $q$ of this expression, we find

$$\log((1-3q-3r)^3/((1-2q-r)^2(q+2r))).$$

The partial derivative with respect to $r$ is equal to:

$$\log((1-3q-3r)^3/((1-2q-r)(q+2r)^2)).$$

Therefore the coordinates of the stationary points satisfy the equation

$$q+r = 1/3,$$

which implies that there are no stationary points. Hence the maximum of the function will be taken on the boundary, i.e.: $q=0$ or $r=0$ or $q+r=1/3$. By numerical optimization we find that the maximum is taken at $q=0$ and $r=0.1770088227....$ This gives an absolute upper bound for $R_{\text{sum}}$:

$$R_{\text{sum}}\leqslant 2.434111388... .$$

## 1.5. The binary multiplying channel as a three-way channel: Codes

In this section, we wish to give some explicit variable length codes for the case that we have finitely many messages to be sent over the channel. Let $M_i$ be the cardinality of the message set of user $i$. Let $m_i$ be the message of user $i$. As described before, the users decide to encode a new message when they are sure that all three of them know the messages of the other two (See the definition of variable length code given before).

We will represent the messages as points in a cube. After the first transmission the cube splits into two regions, the first one being the message points where we receive a 0, the second one being the message points where we receive a 1. In general if we do a transmission, the region under consideration splits into two pieces. But sometimes we are lucky and the region splits into three pieces, because for instance the region of messages where we receive a 0 consists of the union of two projection disjoint regions such that every user can decide in which region the message point lies, on account of his own message.

We shall use the following notation:

$$N = \sum_{(m_1,m_2,m_3)} N(m_1,m_2,m_3).$$

$n$ = the average length of the codewords, i.e.:

$$n = N/(M_1M_2M_3).$$

$$R_k^{ij} = (\log(M_iM_j))/n.$$

$$R_{\text{sum}} = R_3^{12} + R_2^{13} + R_1^{23}.$$

The aim is to find the optimal sumrate for a 3*3*3 cube by considering all possible smaller building blocks. However, we shall not give a rigorous proof of the optimality.

In the pictures, the symbols have the following meaning: The vectors on the right represent the symbol to be sent by user 1, user 2 and user 3, $(x_1,x_2,x_3)$. A black node represents a node where a 1 will be received due to the transmission. A white node represents a node where a 0 will be received due to the transmission. The number above the arrow is the contribution to $N$ due to the actual transmission. The depicted situation following the arrow is the new situation depending on the received symbol.

In case $M_1 = 1$, we are in the situation in which the message of one of the users is known and the other two try to communicate their message to each other and to the third one. This situation differs from the situation in Section 1, because the two users 2 and 3 must take into consideration the presence of user 1 (E.g. compare Example 2 of Section 1 to Example 2 of this Section.) The best code possible uses at least a total number of bits equal to the number of bits in a binary Huffman code for $M_2M_3$ equiprobable messages (cf. Section 7). In the examples we show how this number of bits can be reached. In a certain sence these codes can be considered as binary Huffman codes (cf. Gallager [2]).

EXAMPLE 1. $M_1=M_2=1$, $M_3=2$.

First transmission : user 1: 1,
user 2: 1,
user 3: 0 if $m_3=1$
1 if $m_3=2$.

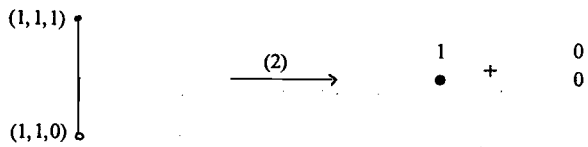



FIGURE 1. Situation before and after the first transmission.

After this first transmission the users are finished, since all message points are uniquely determined by the received symbol and their own message.
In this case: $N=2$; $n=1$; $R_3^{12}=0$; $R_2^{13}=R_1^{23}=1$; $R_{\text{sum}}=2$.

EXAMPLE 2. $M_1=1, M_2=M_3=2$.

First transmission : user 1: 1,
user 2: 1,
user 3: 0 if $m_3=1$
1 if $m_3=2$.

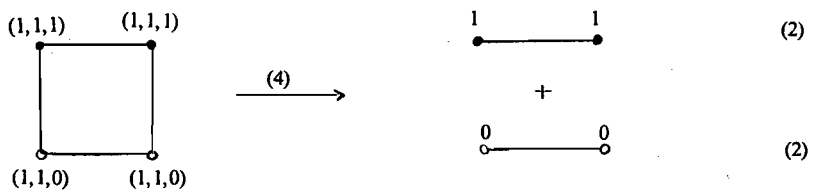



FIGURE 2. Situation before and after the first transmission.

After this first transmission the users are in the situation of Example 1 i.e. $M_1=M_3=1$, $M_2=2$, independent of the received symbol. Therefore they can proceed in the same way as in Example 1. We find: $N=8$; $n=2$; $R_3^{12}=R_2^{13}=0.5$; $R_1^{23}=1$; $R_{\text{sum}}=2$. (In fact this example can be seen as timesharing of two times Example 1).

EXAMPLE 3. $M_1=M_2=M_3=2$.

First transmission : user 1: 1,
user 2: 1,
user 3: 0 if $m_3=1$
1 if $m_3=2$.

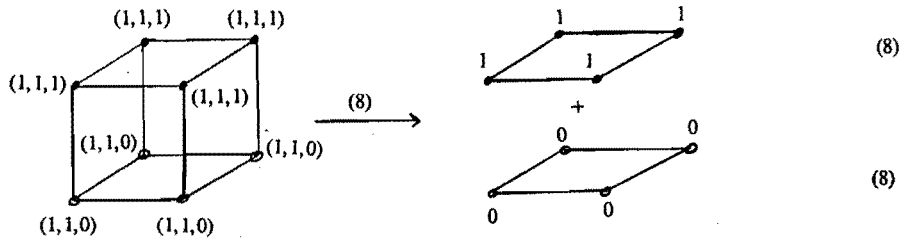



FIGURE 3. Situation before and after the first transmission.

After the first transmission, the users are in the situation of Example 2 i.e. $M_1 = M_2 = 2, M_3 = 1$.
This is independent of the received symbol. The users can therefore proceed in the same way as in Example 2. We find:
$N = 24; n = 3; R_3^{12} = R_2^{13} R_1^{23} = 0.6667$, $R_{\text{sum}} = 2$. In fact this example can be considered as time sharing of two times Example 2.

EXAMPLE 4. $M_1 = M_2 = 1, M_3 = 3$.

First transmission : user 1: 1,
user 2: 1,
user 3: 0 if $m_3 = 1$
1 if $m_3 = 2$ or 3.

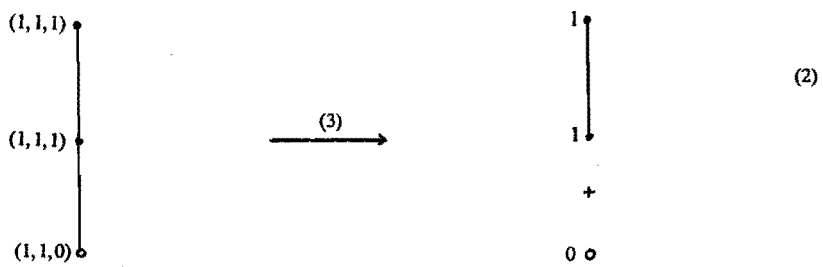



FIGURE 4. Situation before and after the first transmission.

After the first transmission, the users are in the situation of Example 1 upon receiving a 1 and they are finished upon receiving a 0. So upon receiving a 1, they proceed as in Example 1.
We find: $N = 5; n = 5/3; R_3^{12} = 0; R_2^{13} = R_1^{23} = (3\log 3)/5; R_{\text{sum}} = 1.9019$.

EXAMPLE 5. $M_1 = 1; M_2 = 2; M_3 = 3$.

First transmission : user 1 : 1,
user 2 : 0 if $m_2 = 1$,
1 if $m_2 = 2$,
user 3 : 1.

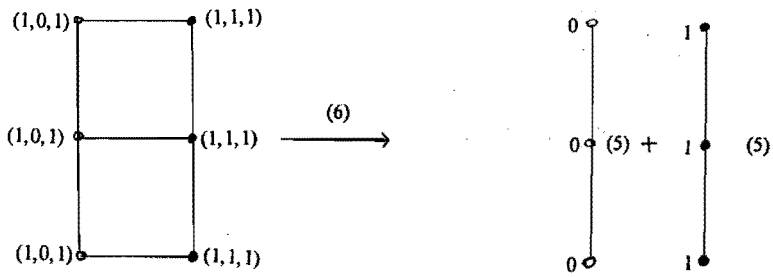



FIGURE 5. Situation before and after the first transmission.

After this first transmission, the users are in the situation of Example 4, independent of the received symbol. Hence they proceed as in Example 4. We find: $N=16$; $n=8/3$, $R_3^{12}=3/8$, $R_2^{13}=(3\log 3)/8$,

$$R_1^{23}=(3\log 6)/8;\ R_{\text{sum}}=1.9387...$$

In fact this example can be considered as time sharing of two times Example 4.

EXAMPLE 6. $M_1=M_2=2, M_3=3$.

First transmission : user 1 : 1,
user 2 : 0 if $m_2=1$,
1 if $m_2=2$,
user 3 : 0 if $m_3=3$
1 if $m_3=1$ or 2.

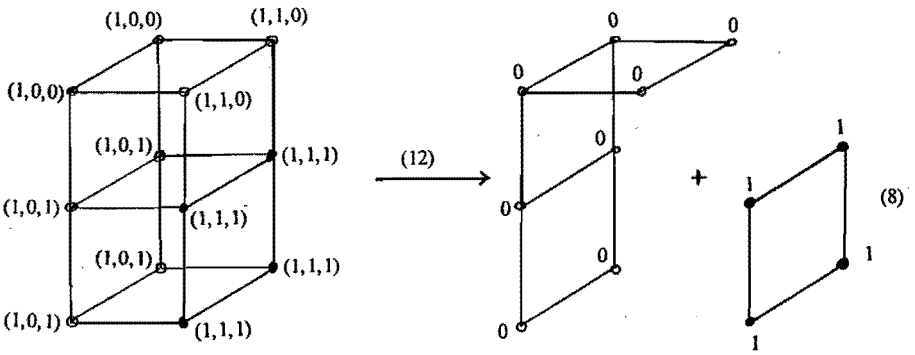



FIGURE 6. Situation before and after the first transmission.

After this first transmission, the users are in the situation of Example 2 if they receive a 1 and then they can proceed as in Example 2. If they receive a 0 they are in a new situation which will be resolved as below:

Second transmission upon receiving a 0:

user 1 : 0 if $m_1 = 2$
1 if $m_1 = 1$,
user 2 : 0 if $m_2 = 2$
1 if $m_2 = 1$,
user 3 : 1.

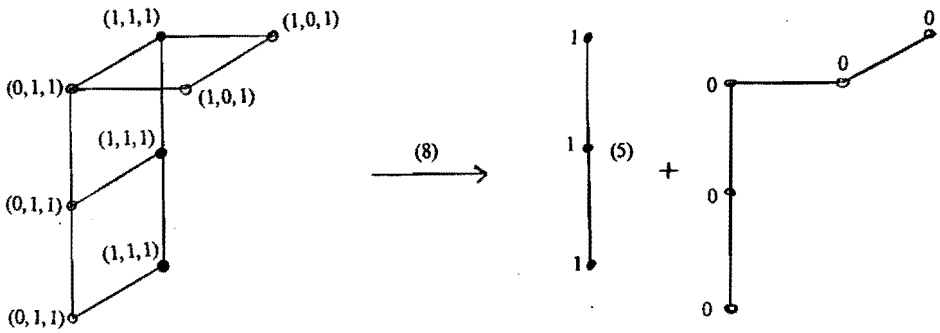



FIGURE 7. Situation before and after the second transmission upon receiving 0.

After this second transmission, the users are in the situation of Example 4 if they receive a 1 and in a new situation if they receive a 0. This new situation will be resolved as follows:

Third transmission upon receiving 00.

user 1 : 0 if $m_1 = 1$
1 if $m_1 = 2$,
user 2 : 1,
user 3 : 0 if $m_3 = 1$ or 2
1 if $m_3 = 3$.

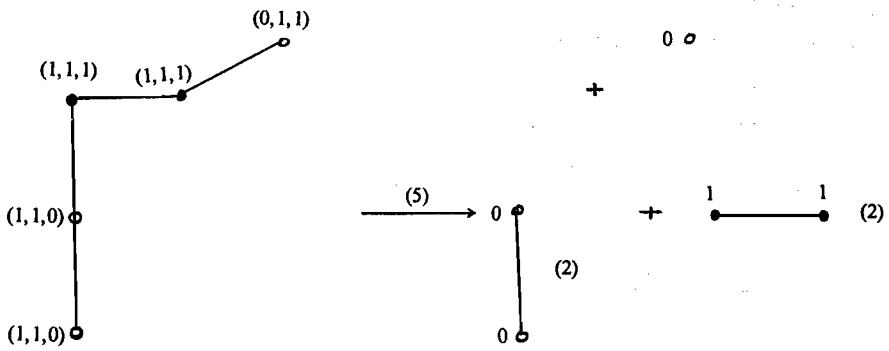



FIGURE 8. Situation before and after the third transmission upon receiving 00.

After this third transmission, the users are in the situation of Example 1 or they are finished. Indeed, if the received symbol is a 1, they are in the situation of Example 1, and if they receive a 0 they know that the set of possible message points is a projection disjoint union of two sets and hence from their own message they can decide which of the two sets the message point has to be in.

We find: $N=42, n=7/2,\ R_3^{12}=4/7, R_2^{13}=R_1^{23}=(2\log 6)/7,$

$R_{\text{sum}}=2.04855\ldots$ .

EXAMPLE 7. $M_1=1,\ M_2=M_3=3.$

First transmission : user 1 : 1,

user 2 : 0 if $m_2 = 1$

1 if $m_2 = 2$ or 3,

user 3 : 0 if $m_3 = 3$

1 if $m_3 = 1$ or 2.

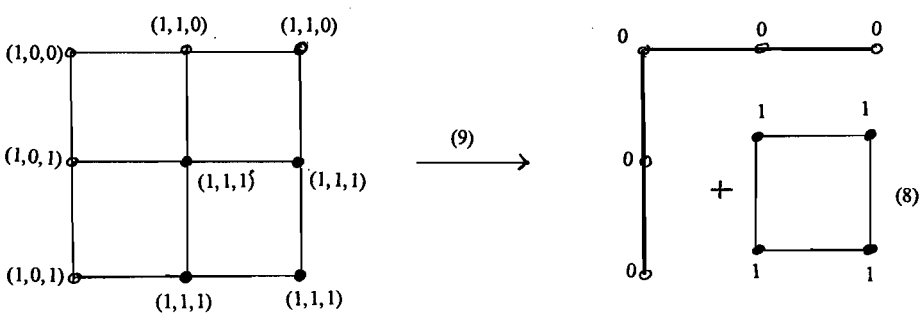



FIGURE 9. Situation before and after the first transmission.

After this first transmission, the users are in the situation of Example 2 if they receive a 1. If they receive a 0, they are in a new situation which will be

resolved as follows:
Second transmission upon receiving a 0:

user 1 : 1,
user 2 : 1 if $m_2 = 1$
0 if $m_2 = 2$ or 3,
user 3 : 1.

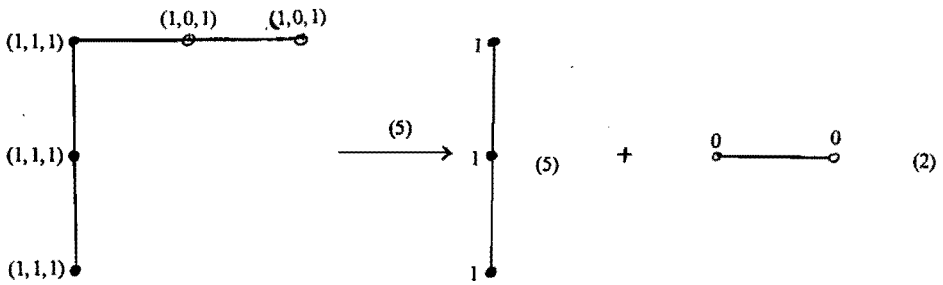



FIGURE 10. Situation before and after the second transmission upon receiving 0.

After this second transmission, the users are in the situation of Example 4 if they receive a 1, and they are in the situation of Example 1 if they receive a 0.

We find: $N=29$, $n=29/9$, $R_3^{12}=R_2^{13}=(9\log 3)/29$,

$R_1^{23}=(18\log 3)/29$, $R_{\text{sum}}=1.9675\ldots$.

EXAMPLE 8. $M_1=M_2=3$, $M_3=2$.

First transmission : user 1 : 0 if $m_1=1$
1 if $m_1=2$ or 3,
user 2 : 0 if $m_2=1$
1 if $m_2=2$ or 3
user 3 : 1.

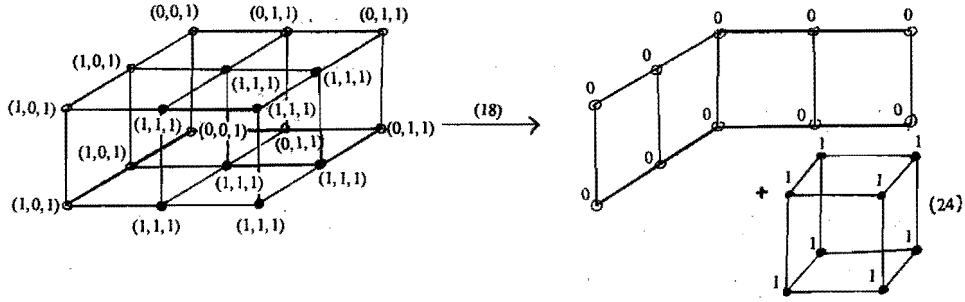



FIGURE 11. Situation before and after the first transmission.

After the first transmission, the users are in the situation of Example 3 if they receive a 1. If they receive a 0, they are in a new situation, which will be resolved as follows:

Second transmission upon receiving a 0:

user 1 : 1,
user 2 : 0 if $m_2 = 2$ or 3
1 if $m_2 = 1$,
user 3 : 0 if $m_3 = 1$
1 if $m_3 = 2$.

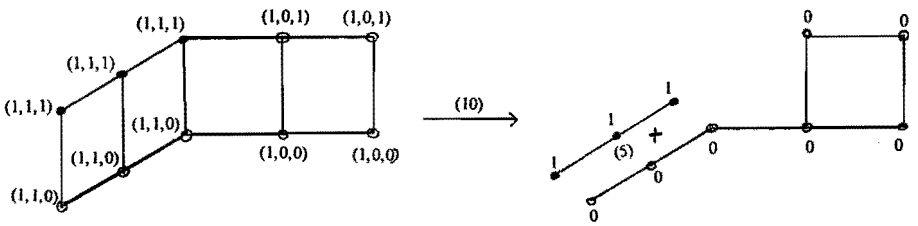



FIGURE 12. Situation before and after the second transmission upon receiving 0.

After this second transmission, the users are in the situation of Example 4 if they receive a 1. If they receive a 0, they are in a new situation which will be resolved as follows:

Third transmission upon receiving 00.

user 1 : 0 if $m_1 = 2$ or 3
1 if $m_1 = 1$,
user 2 : 1,
user 3 : 0 if $m_3 = 2$
1 if $m_3 = 1$.

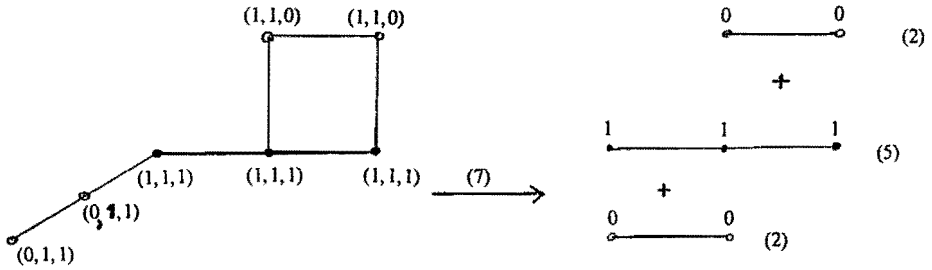



FIGURE 13. Situation before and after the third transmission upon receiving 00.

After this third transmission, the users are in the situation of Example 4 if they receive a 1. If they receive a 0 the set of possible message points is a projection disjoint union of two sets and from their own message the users can decide in which of the two the message point lies. Therefore upon receiving 000, they are in the situation of Example 1.

We find: $N=73$, $n=73/18(=4.056..)$, $R_3^{12}=(36\log 3)/73$,

$$R_2^{13}=R_1^{23}=(18\log 6)/73,\ R_{\text{sum}}=2.0564\ldots.$$

EXAMPLE 9. $M_1=M_2=M_3=3$.

First transmission : user 1 : 0 if $m_1=1$

1 if $m_1=2$ or 3,

user 2 : 0 if $m_2=1$

1 if $m_2=2$ or 3,

user 3 : 1.

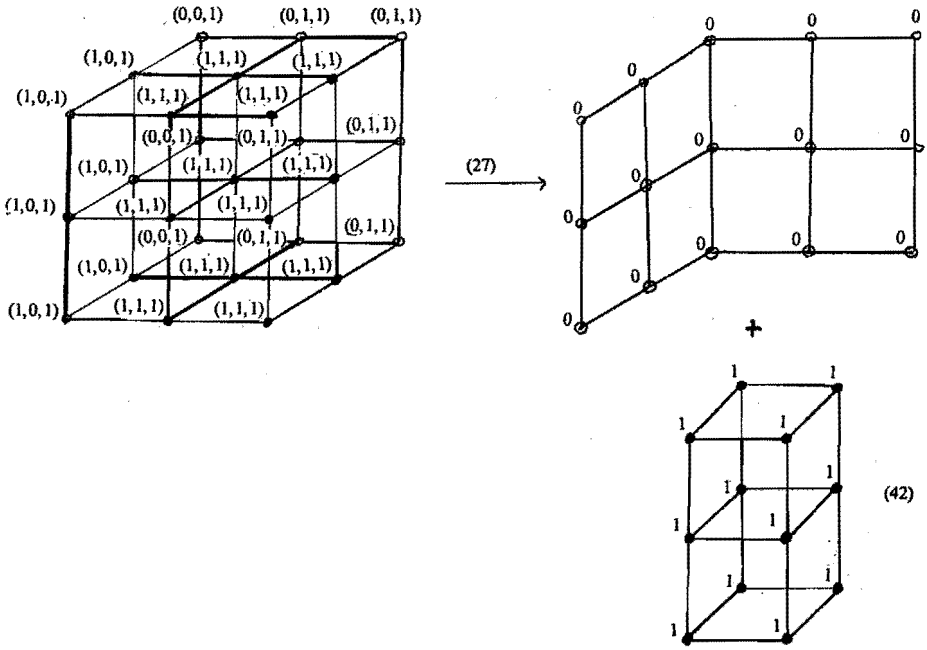



FIGURE 14. Situation before and after the first transmission.

After this first transmission, the users are in the situation of Example 6, if they receive a 1. If they receive a 0, they are in a new situation which will be resolved as follows:

Second transmission upon receiving a 0:

user 1 : 1,

*user* 2 : 0 if $m_2 = 3$

1 if $m_2 = 1$ or 2,

user 3 : 0 if $m_3 = 3$

1 if $m_3 = 1$ or 2.

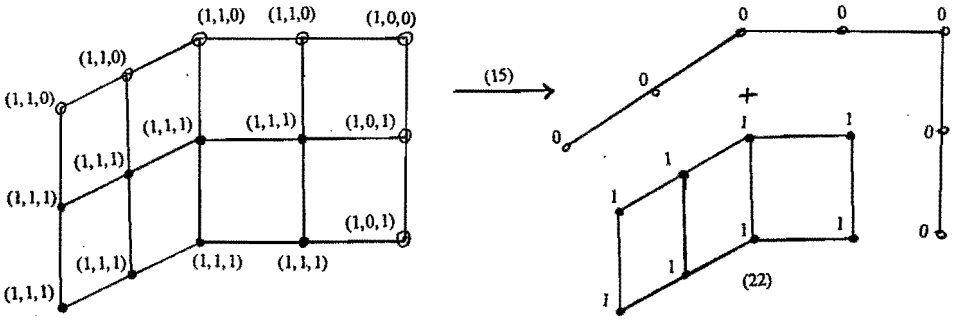



FIGURE 15. Situation before and after the second transmission upon receiving 0.

After this transmission, the users are in a situation similar to the one after the first transmission in Example 6 (which costs 22 bits to resolve), if they receive a 1. If they receive a 0, they are in a new situation which will be resolved as follows:

Third transmission upon receiving 00:

user 1 : 0 if $m_1 = 2$ or 3
1 if $m_1 = 1$,
user 2 : 1,
user3 : 0 if $m_3 = 1$ or 2,
1 if $m_3 = 3$.

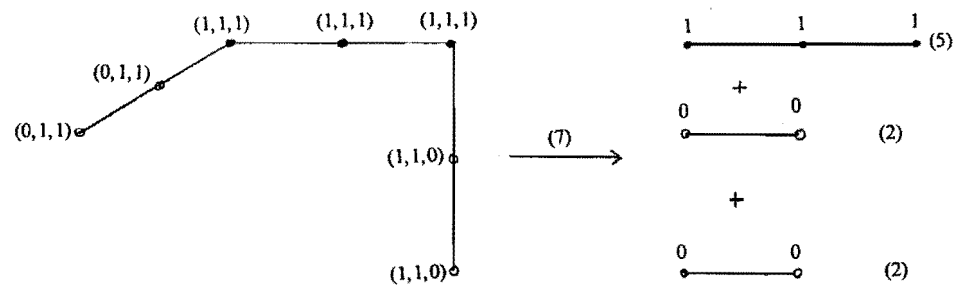



FIGURE 16. Situation before and after the third transmission upon receiving 00.

After this transmission, the users are in the situation of Example 4 if they receive a 1. If they receive a 0, the messagepoint is in a projection disjoint union of two sets and the users can determine in which of the two by inspecting their own message. So in this case, they are in the situation of Example 1.

We find: $N = 122$, $n = 122/27$ $(=4.5185)$, $R_1^{23} = R_2^{13} = R_3^{12} =$
$= (27\log 9)/122$, $R_{\text{sum}} = 2.1046\ldots$.

## 1.6. The binary multiplying channel as a three-way channel: Strategies

Now we are going to generalize the strategy given in Section 1 to three dimensions. In this case, we can represent coding strategies for the BMC as a three-way channel, as strategies for subdiving the unit cube. At the beginning of transmission, the message point $(m_1,m_2,m_3)$ is uniformly distributed over the unit cube.

We now want to divide a unit cube into blocks, such that there are not too much "intermediate" states. After we have described a strategy in which this done, the states can be taken as states of a Markov Chain. Then using the stationary probabilities of the states we can calculate the average information per transmission passing along the channel, using the described strategy.

For the first transmission, sender 3 sends a 1, sender 1 and sender 2 send a 1 if $m_i \in [0,t_i]$, and send a 0 otherwise $(i=1,2)$. The initial thresholds $t_1$ and $t_2$ for $m_1$ and $m_2$ respectively, divide the unit cube into the four subblocks 1,2,3 and 4 of Figure 1.

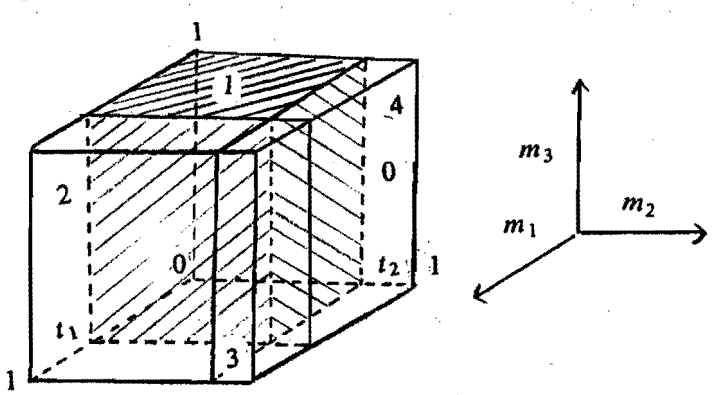



FIGURE 1. Unit cube with initial thresholds $(t_1,t_2)$. State 1.

After receiving a 1, the message point $(m_1,m_2,m_3)$ is in the shaded subblock (i.e. $(m_1,m_2)\in[0,t_1]*[0,t_2]$), see Figure 1). Then our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received, then the message point lies in the corner shaped body consisting of the blocks 2,3 and 4 in Figure 1, and further resolution is necessary. In this strategy, we are going to "resolve" the remaining uncertainty upon receiving a 0 in two steps. In step 1, sender 2 sends a 1, sender 1 sends a 1 if $m_1 \in (t_1,1]$ and a 0 otherwise and sender 3 sends a 1 if $m_3 \in (t_3,1]$ and a 0 otherwise.

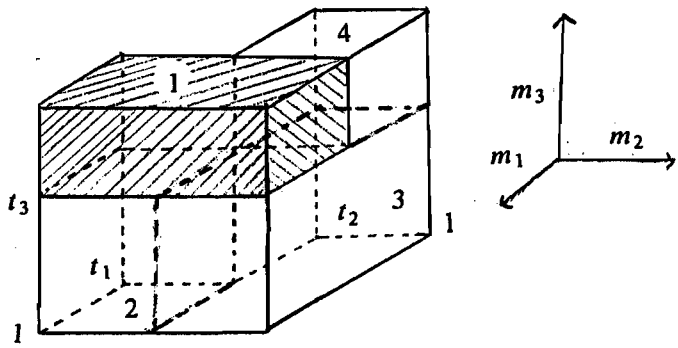



FIGURE 2. Subdivision of the region for the message point upon receiving a 0. Thresholds: $(t_1,t_2,t_3)$. State 2.

After receiving a 1, in this second transmission, the message point $(m_1,m_2,m_3)$ is in the shaded subblock 1 (see Figure 2). Then our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received, then the message point $(m_1,m_2,m_3)$ lies in the double corner shaped body of Figure 2, i.e. the body which is left when the shaded block (1) is removed. Again a further resolution is necessary. This second resolution is step 2. In step 2, sender 1 sends a 1, sender 2 sends a 1 if $m_2 \in (t_2,1]$ and a 0 otherwise and sender 3 sends a 1 if $m_3 \in [0,t_3)$ and a 0 otherwise.

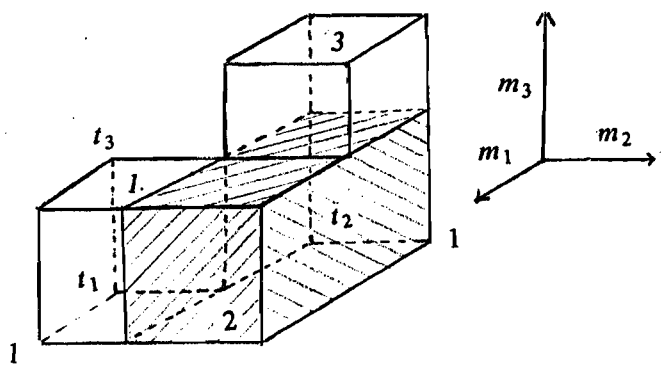



FIGURE 3. Subdivision of the region for the message point upon receiving 00. Thresholds $(t_1,t_2,t_3)$. State 3.

After receiving a 1, the message point $(m_1,m_2,m_3)$ is in the shaded subblock 2. (See Figure 3.) Then our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received, then the message point $(m_1,m_2,m_3)$ lies in one of the subblocks 1 or 3 (see Figure 3). Since these two subblocks are projection disjoint, all three the users know exactly in which of the two subblocks the message point lies, since they know their own message. Again the remaining task is to divide this subblock further, which is fully equivalent to the initial task. It is clear that after the third division, we are

back in the old situation, i.e. we have to determine the message point in a block. Figures 1 to 3 show the subdivisions that are used in our strategy, using thresholds $(t_1,t_2,t_3)$. As before, we can describe this process by a Markov-chain. This Markov-chain is given in Figure 4.

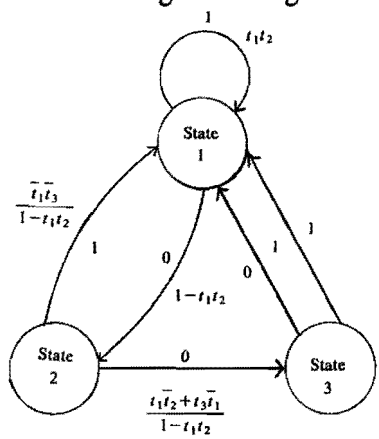



FIGURE 4. Markov-chain in the three dimensional case.

The matrix of transition probabilities for this Markov-chain is:

| | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $t_1t_2$ | $1-t_1t_2$ | 0 |
| 2 | $\bar{t}_1\bar{t}_3/(1-t_1t_2)$ | 0 | $(\bar{t}_2t_1+t_3\bar{t}_1)/(1-t_1t_2)$ |
| 3 | 1 | 0 | 0 |

where $\bar{t}$ is defined as: $\bar{t}:=1-t$.
This gives for the stationary probabilities $q_i$ of state $i$ $(i=1,2,3)$:

$$q_1 = 1/(2+t_1+t_3-2t_1t_2-t_1t_3)$$
$$q_2 = (1-t_1t_2)/(2+t_1+t_3-2t_1t_2-t_1t_3)$$
$$q_3 = (\bar{t}_2t_1+t_3\bar{t}_1)/(2+t_1+t_3-2t_1t_2-t_1t_3)$$

We will now give the quantities $I_k^{ij}$ for the three different states:

*State 1.*

$$I_3^{12}(1) = h(t_1t_2),$$
$$I_2^{13}(1) = t_2(h(t_1),$$
$$I_1^{23}(1) = t_1h(t_2).$$

*State 2.*

$$I_3^{12}(2) = (1-t_3)h((1-t_1)/(1-t_1t_2)),$$
$$I_2^{13}(2) = (t_2(1-t_1)h(t_3) + (1-t_2)h((1-t_1)(1-t_3)))/(1-t_1t_2),$$

$$I_1^{23}(2) = (1-t_1)h(t_3)/(1-t_1t_2).$$

*State 3.*

$$I_3^{12}(3) = t_3(1-t_1t_2)h(t_2(1-t_1)/(1-t_1t_2))/(t_1+t_3-t_1t_3-t_1t_2),$$

$$I_2^{13}(3) = (1-t_2)(t_1+t_3-t_1t_3)h(t_3/(t_1+t_3-t_1t_3))/(t_1+t_3-t_1t_3-t_1t_2),$$

$$I_1^{23}(3) = (t_1(1-t_2)h(t_3) + (1-t_1)t_3h(t_2))/(t_1+t_3-t_1t_3-t_1t_2).$$

The optimal value for:

$$R_{\text{sum}} = \sum_{i=1}^{3} q_i(I_3^{12}(i)+I_2^{13}(i)+I_1^{23}(i)),$$

is reached for $t_1 = 0.6398$, $t_2 = 0.7347$ and $t_3 = 0.3898$.
In this case $R_{\text{sum}} = 2.1822$.
(These results were obtained by numerical optimization). We will now describe a modification of this strategy, which gives a slightly better result. The first step is the same as in the old strategy: sender 3 sends a 1 and senders 1 and 2 send a 1 if $m_i \in [0,t_i]$ and a 0 otherwise ($i = 1,2$). The message point $(m_1,m_2,m_3)$ is uniformly distributed over the unit cube. The initial thresholds $t_1$ and $t_2$ for $m_1$ and $m_2$ respectively, divide the unit cube into four subblocks 1,2,3 and 4. (See Figure 5.)

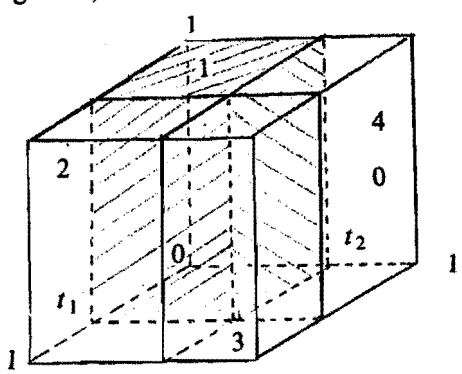



FIGURE 5. Unit cube with initial thresholds $(t_1,t_2)$. State 1.

After receiving a 1, the message point is in the shaded subblock. (i.e. $(m_1,m_2) \in [0,t_1]*[0,t_2]$), see Figure 5). Then our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received then the message point lies in the corner shaped body consisting of the blocks 2,3 and 4 in Figure 5, and further resolution is necessary. In our new strategy we are going to resolve the uncertainty in at most 3 steps. In step 1, sender 2 sends a 1, sender 1 sends a 1 if $m_1 \in (t'_1,1]$, (where $t'_1 < t_1$) and 0 otherwise, and sender 3 sends a 1 if $m_3 \in (t_3,1]$ and a 0 otherwise.

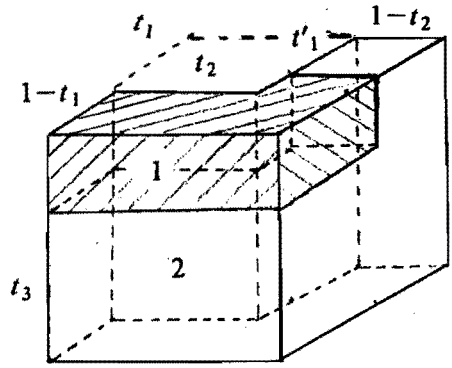



FIGURE 6. Subdivision of the region for the message point upon receiving a 0, thresholds: $(t_1,t'_1,t_2,t_3)$. State 2.

After receiving a 0 for the second time, the message point lies in the double corner shaped body of Figure 6. (Region 2). In this case, one further resolution is necessary. This resolution is shown in Figure 7.

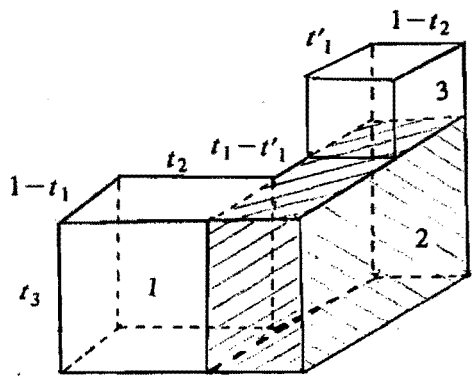



FIGURE 7. Subdivision of the region for the message point upon receiving 00. Thresholds $(t_1,t'_1,t_2,t_3)$. State 3.

Sender 1 sends a 1, sender 2 sends a 1 if $m_2 \in [t_2,1]$, and a 0 otherwise, sender 3 sends a 1 if $m_3 \in [0,t_3]$, and a 0 otherwise, in this case. After receiving a 1, the message point $(m_1,m_2,m_3)$ is in the shaded subblock 2 (see Figure 7). Then our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received, then the message point lies in one of the subblocks 1 or 3. These subblocks being projection disjoint, all three senders know in which subblock the message point lies, since they know their own message. Again the remaining task is to divide this subblock further, which is fully equivalent to the initial task. It is clear that in this case, we are back in the old situation (i.e. we have to determine the message point in a block) after three divisions.

After receiving a 1 in the second transmission (i.e. we have seen 01 now), the

message point lies in the corner shaped body of Figure 6. (Region 1.) The reason that State 4 is different from State 2 is that the proportions of the sizes of the sides are essentially different. We are going to resolve the uncertainty in two steps (as before), the first one being shown in Figure 8.

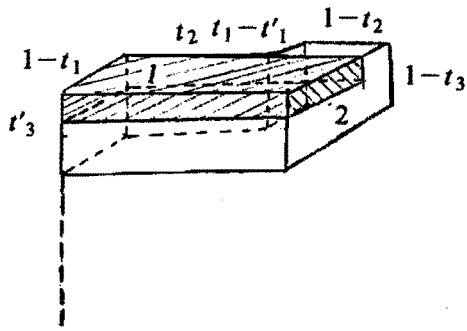



FIGURE 8. Subdivision of the region for the message point upon receiving 01. Thresholds $(t_1,t'_1,t_2,t_3,t'_3)$. State 4.

Sender 2 sends a 1, Sender 1 sends a 1 if $m_1 \in [1-t_1,1]$ and a 0 otherwise, Sender 3 sends 1 if $m_3 \in [t'_3,1]$, and a 0 otherwise, where $t'_3 > t_3$. After receiving a 1, the message point $(m_1,m_2,m_3)$ is in the shaded subblock 1, see Figure 8. Again our task is to divide this subblock further, which is fully equivalent with the initial task. After receiving a 0, the message point is in the double corner shaped body (region 2) of Figure 8, and some more resolution is necessary. This resolution is shown in Figure 9.

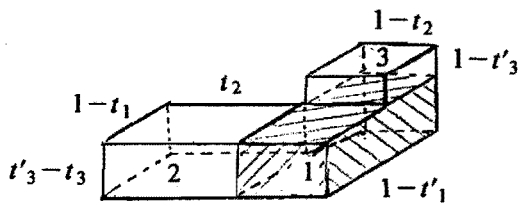



FIGURE 9. Subdivision of the region for the message point upon receiving 010. Thresholds $(t_1,t'_1,t_2,t_3,t'_3)$. State 5.

Sender 1 sends a 1, Sender 2 sends a 1 if $m_2 \in [t_2,1]$ and a 0 otherwise, Sender 3 sends a 1 if $m_3 \in [t_3,t'_3]$ and a 0 otherwise. After receiving a 1 in this case, the message point is in the shaded subblock 1, see Figure 9, and our task is to divide that subblock further, which is fully equivalent to the initial task. If a 0 is received, then the message point is in one of the two subblocks 2 or 3, see Figure 9. These two subblocks being projection disjoint, all users can decide in which subblock the message point is, by inspecting their own message. Again the remaining task is reduced to the initial one, i.e. finding the message point in a block. It is clear that we always come back to the initial state 1. Figures 5 to 9 show which subdivisions are used in this new strategy, using thresholds

$(t_1,t'_1,t_2,t_3,t'_3)$. The Markov-chain corresponding to these subdivisions is given in Figure 10.

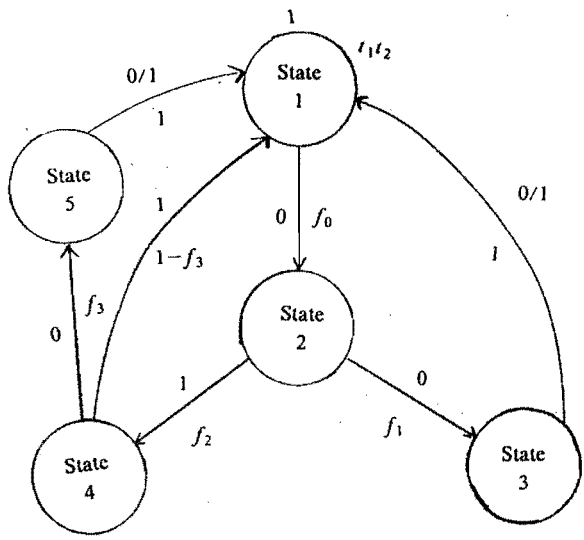



FIGURE 10. Markov-chain for the new strategy.

Later on, we will specify the transition probabilities $f_0,f_1,f_2$ and $f_3$. First we shall now calculate some quantities. $V_i$ stands for the contents of the region under consideration in State $i$, $i=1,2,3,4,5$. $S_i$ is the surface of the bottom of the region under consideration in State $i$, $i=1,2,3,4,5$. $I_k^{ij}(l)$ is the amount of information passing through the channel in State $l$, from users $i$ and $j$ to user $k$. We shall calculate these numbers. The reader can go through the calculations himself.

*State 1.*

Contents of the region: $V_1=1$.

$$I_3^{12}(1) = h(t_1t_2),$$
$$I_2^{13}(1) = t_2h(t_1),$$
$$I_1^{23}(1) = t_1h(t_2).$$

*State 2.*

Contents of the region: $V_2=1-t_1t_2$.

$$I_3^{12}(2) = (1-t_3)h(t'_1(1-t_2)/V_2).$$
$$I_2^{13}(2) = ((1-t_1)t_2h(t_3)+(1-t_2)h((1-t_3)(1-t'_1)))/V_2,$$
$$I_1^{23}(2) = ((1-t'_1-t_1t_2+t'_1t_2)h(t_3))/V_2.$$

*State 3.*

Contents of the region: $V_3 = t_3(1-t_1t_2)+t'_1(1-t_2)(1-t_3)$.

$$I_3^{12}(3) = t_3(1-t_1t_2)h((1-t_2)/(1-t_1t_2))/V_3,$$
$$I_2^{13}(3) = (1-t_2)(t'_1+t_3-t'_1t_3)h(t_3/(t_3+t'_1-t'_1t_3))/V_3,$$
$$I_1^{23}(3) = (t'_1(1-t_2)h(t_3) + (1-t_1)t_3h(t_2))/V_3.$$

*State 4.*

Surface of the bottom: $S_4 = 1-t'_1-t_1t_2+t'_1t_2$.
Contents of the region: $V_4 = (1-t_3)S_4$.

$$I_3^{12}(4) = (1-t'_3)h((1-t_1)/S_4)/(1-t_3)$$
$$I_2^{13}(4) = (t_2(1-t_1)h((1-t'_3)/(1-t_3)))/S_4 +$$
$$+ ((1-t'_1)(1-t_2)h((1-t_1)(1-t'_3)/((1-t'_1)(1-t_3)))/S_4,$$
$$I_1^{23}(4) = (1-t_1)h(1-t'_3)/(1-t_3))/S_4.$$

*State 5.*

Contents of the region: $V_5 = (t'_3-t_3)S_4+(1-t'_3)(1-t_2)(t_1-t'_1)$.

$$I_3^{12}(5) = (t'_3-t_3)S_4h((1-t_1)t_2/S_4)/V_5,$$
$$I_2^{13}(5) = (1-t_2)((1-t'_1)(1-t_3)-(1-t_1)(1-t_3))\star$$
$$\star h((1-t_3)(t_1-t'_1)/((1-t'_1)(1-t_3)-(1-t_1)(1-t'_3)))/V_5,$$
$$I_1^{23}(5) = ((t'_3-t_3)(1-t_1)h(t_2)/V_5 +$$
$$+ (t_1-t'_1)(1-t_2)(1-t_3)h((1-t'_3)/(1-t_3))/V_5.$$

The transition probabilities $f_0, f_1, f_2$ and $f_3$ are given by:

$$f_0 = V_2;\ f_1 = V_3/V_2;\ f_2 = V_4/V_2;\ f_3 = V_5/V_3.$$

For the stationary probabilities $q_i$ of state $i (i = 1,2,3,4,5)$, we find the equations:

$$q_2 = f_0q_1 \quad (1),$$
$$q_3 = f_1q_2 = f_0f_1q_1 \quad (2),$$
$$q_4 = f_2fq_2 = f_0f_2q_1 \quad (3),$$
$$q_5 = f_3q_4 = f_0f_2f_3q_1 \quad (4),$$
$$q_1 + q_2 + q_3 + q_4 + q_5 = 1 \quad (5),$$

which have the following solution:

$$q_1 = 1/(1+f_0 + f_0f_1 + f_0f_2 + f_0f_2f_3),$$

and $q_2, q_3, q_4$ and $q_5$ are given by (1), (2), (3) and (4). The optimal value for

$$R_{\text{sum}} = \sum_{i=1}^{5} q_i(I_3^{12}(i)+I_2^{13}(i)+I_1^{23}(i)),$$

is reached for:

$$t_1 = 0.6902... \; ; \; t_2 = 0.6783... \; ; \; t_3 = 0.3700... \; ;$$

$$t'_1 = 0.4452... \; ; \; t'_3 = 0.6387... \; .$$

In this case $R_{\text{sum}} = 2.2044...$. (this result again was obtained by numerical optimization). By arguments similar to those used by Tolhuizen [5], it probably is possible to show that in this case the calculated rates can be reached by block codes with vanishing probability of error, as well as by variable length codes with zero error probability. However we shall not go into details.

1.7. THE CASE IN WHICH TWO SENDERS KNOW THE MESSAGE OF THE THIRD ONE

In this case, i.e. when two senders know the message of the third one, this third one does not want to pass information along the channel, so he keeps sending a 1. Actually the situation is equivalent to the case where we use the BMC as a multiple access channel with two senders and one receiver and in which there is complete feedback. For general results on multiple access channels with feed-back we refer to van der Meulen [8]. More details can be found in van der Meulen [9]. The situation is like in Figure 1.

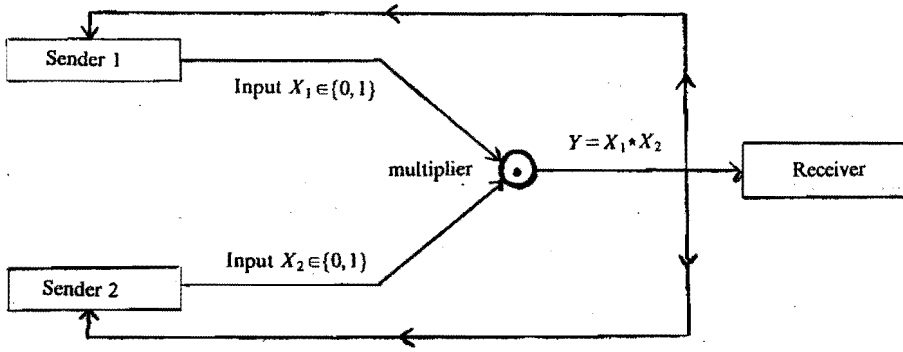



FIGURE 1. An equivalent communication situation.

Again the cardinality of the message set for sender $i$ is $M_i, (i = 1,2)$. The best code possible uses a total number of bits equal to the number of bits in a binary Huffman code for $M_1 M_2$ equiprobable messages. Less is not possible since this number equals the source information. For more details about such Huffman codes we refer to Gallager [2], Exercise 3.15, page 515. The question is whether this number of bits is feasible, and I conjecture that the answer is yes. In this respect, I want to make the following remarks:

Per transmission, the region in which the message point has to be, is divided into two smaller regions: the one where a 0 shall be received and the one where a 1 shall be received. This division should be such that these two

regions are more or less of the same size, i.e. the number of possible message points in each region should be in between the same two powers of 2. It is easy to see that such a division is always possible if the following statement is true:

CONJECTURE 1. *Let $S$ be a positive integer $2^{a+1}<S\leqslant 2^{a+2}$ for some $a\in\mathbb{N}$. Let $M$ be a $(0,1)$-matrix having exactly $S$ ones. Then there exists a submatrix of $M$ containing $S_1$ ones, such that $2^a\leqslant S_1\leqslant 2^{a+1}$ and $2^a\leqslant S-S_1\leqslant 2^{a+1}$.*

Indeed, if we use the messages of user 1 and 2 as indices of rows respectively colums of a matrix, having $(i,j)^{\text{th}}$ entry equal to 1 if a message point is a region which is under consideration and a 0 otherwise, then conjecture 1 asserts that every arbitrary region can be divided into two regions of more or less the same size, if user 1 sends a 1 if his message is the index of one of the rows of the submatrix and a 0 otherwise, and user 2 sends a 1 if his message is one of the column of the submatrix and a 0 otherwise.

Unfortunately this conjecture still needs a proof. This statement is equivalent to the following one:

CONJECTURE 2. *Let $M$ be a $(0,1)$-matrix having $2^{a+1}$ ones. Then there exists a submatrix of $M$ having exactly $2^a$ ones.*

PROOF OF THE EQUIVALENCE OF CONJECTURES 1 AND 2

LEMMA 1. *Let $M$ be a counterexample against Conjecture 1, having a minimal number of ones $S$, such that $2^{a+2}\geqslant S>2^{a+1}$. Then $S\geqslant 2^{a+1}+2^a+1$.*

PROOF: Delete a one in $M$. The resulting matrix has $S-1$ ones and by the minimality of the counterexample, it follows that the new matrix is not a counterexample. So we can find a submatrix for this new matrix, satisfying the conditions of Conjecture 1. The region where the deleted one is in, contains $2^{a+1}$ ones, otherwise the same submatrix would give a legal splitting and the counterexample would not be counterexample. Therefore we can conclude: $S-1-2^{a+1}\geqslant 2^a$. ($S-1-2^{a+1}$ being the number of ones in the other region i.e. not containing the deleted one).

LEMMA 2. *If there exists a counterexample for Conjecture 1, then there exists a counterexample containing $2^a$ ones for some $a\in\mathbb{N}$.*

PROOF: Suppose there exists a counterexample for Conjecture 1. Take the one with the minimal number of ones, say $S$ ones. From the previous lemma, we know $2^{a+2}\geqslant S\geqslant 2^{a+1}+2^a+1$ for some $a\in\mathbb{N}$.
Now it follows that $S=2^{a+1}+2^a+t$, for some $t$. From a counterexample satisfying $S=2^{a+1}+2^a+t<2^{a+2}$ ($S$ is the number of ones in the matrix), $t\geqslant 1$, we can make a counterexample satisfying $S=2^{a+1}+2^a+t+1$ ones by

changing an arbitrary zero into a one. Indeed, suppose the new matrix is not a counterexample. $S'=S+1$ is the new number of ones. We find a submatrix satisfying the conditions of Conjecture 1 for the new matrix. $S'_1=$ the number of ones in the region where we changed something. If this submatrix is used for the original situation, it does not satisfy the conditions of Conjecture 1. Hence $S'_1-1<2^a$ so $S'_1\leqslant 2^a$. But also $S'_1\geqslant 2^a$. So $S'_1=2^a$. Then $S+1-S'_1\leqslant 2^{a+1}$ so $S\leqslant 2^{a+1}+2^a-1$.
But $S\geqslant 2^{a+1}+2^a+1$, a contradiction. Therefore the new matrix has to be a counterexample. By subsequently enlarging the number of ones in a counterexample, we reach a counterexample having $2^a$ ones, for some $a\in\mathbb{N}$. We have now proven that Conjecture 2 implies Conjecture 1. To prove the other implication assume Conjecture 1. Let $M$ be a $(0,1)$-matrix having $2^{a+1}$ ones. From Conjecture 1, it follows that there exists a submatrix of $M$ containing $S_1$ ones with $2^{a-1}\leqslant S_1\leqslant 2^a$, and $2^{a-1}\leqslant 2^{a+1}-S_1\leqslant 2^a$.

So $S_1\leqslant 2^a$ and $S_1\geqslant 2^a$. Therefore $S_1=2^a$. So Conjecture 2 follows from Conjecture 1.

This proves the equivalence of both statements.

## References


1. C.E. Shannon (1961). *Two-way communication channels,* proc. 4th Berkeley Symp. Math. Statist. and Prob., Vol 1, pp. 611-644. Reprinted in Key-papers in the development of Information Theory, (D. Slepian, ed.), New York, *IEEE Press,* (1974), pp. 339-372.
2. R.G. Gallager (1968). *Information Theory and Reliable communication,* John Wiley and Sons, Inc.
3. J.P.M. Schalkwijk (1982). *The binary multiplying channel - A coding scheme that operates beyond Shannon's inner bound region.* IEEE Trans. Inform. Theory, Vol. IT-28, pp. 107-110.
4. J.P.M. Schalkwijk (1984). *On an Extension of an achievable rate region for the Binary Multiplying Channel.* IEEE Trans. Inform. Theory, Vol. IT-29, pp. 445-448.
5. L.M.G.M. Tolhuizen. *Discrete coding for the BMC, based on Schalkswijk's strategy.* Proceedings of the sixth Symposium on Information Theory in the Benelux, A.J. Vinck, ed. Enschede: Werkgemeenschap Informatie en Communicatie theorie.
6. A. El Gamal, T.M. Cover (1980). *Multiple User Information Theory,* Proceedings of IEEE, Vol. 68, No. 12, pp. 1467-1483.
7. E.C. van der Meulen (1971). *Three-terminal communication channels.* Advances in Applied Probability, 3(1) pp. 120-154.
8. E.C. van der Meulen (1977). *A survey of multi-way channels in information theory: 1961-1976,* IEEE Trans. Inform. Th. vol IT-23(1), pp. 1-37.
9 E.C. van der Meulen (1988). *Capacity theorems for multiple-acces channels with feed-back.* To appear: Proc. Int. Symp. on Coding and Information Theory, Campinas, Brazilië.

# Chapter 2

# Error-correcting Codes for the Binary-adder Channel

## 2.0. Introduction

In 1976 Kasami and Lin ([1]-[4]) studied the coding procedures for a multiple access channel known as the binary adder channel. First the study was restricted to channels without noise and the aim was to find so-called uniquely decodable codes with rates as high as possible. Later the presence of noise was taken into consideration and the concept of $\delta$-decodability was introduced. Kasami et al. [1]-[4] give existence results for "good" $\delta$-decodable codes. In [6], Van Tilborg gave upper bounds for codes for this noisy two access binary adder channel. In the coming chapter we shall give some explicit construction methods for the noisy two access binary adder channel. Furthermore, we give tables of the constructed codes. In our construction, a major role is played by ternary codes having high distance $d$ in the $L$-metric (which will be defined in a coming section). Therefore in a separate section, (Section 6), we shall study this kind of codes.

## 2.1. The binary adder channel as a multiple access channel

In the coming sections we wish to study the Binary Adder Channel with noise. The communication situation that we are interested in is given by Figure 1.

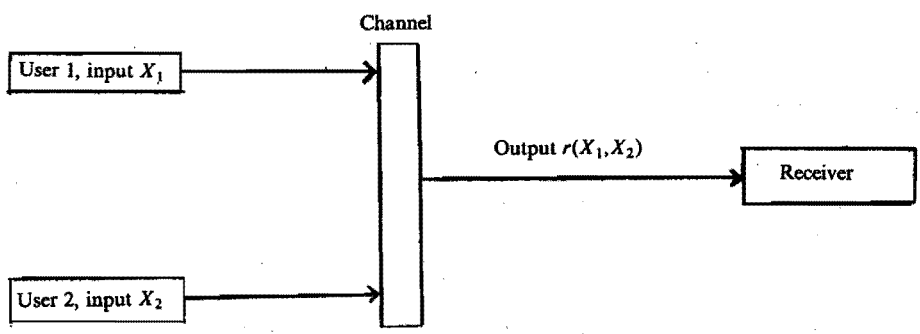



FIGURE 1. A multiple access communication situation.

For the rest of this chapter we assume that there is bit and block synchronization. In the case we wish to study, the channel output is simply the sum of the two bits $X_1$ and $X_2$, i.e. $r = X_1 + X_2$, that is to say when there is no noise. This channel model is depicted in Figure 2a, and is called the noiseless BAC (Cf. [4]). In the presence of noise, $r$ can be any integer from 0,1,2. This channel model is called noisy BAC, and is depicted in Figure 2b.

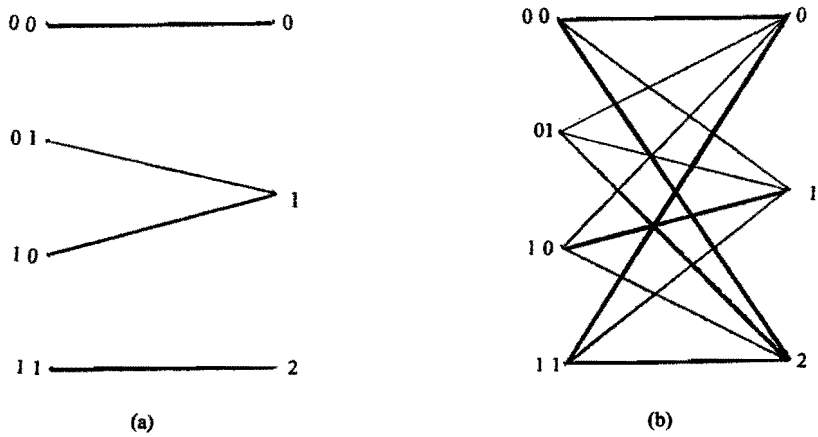



FIGURE 2. (a) Noiseless BAC, possible transitions
(b) Noisy BAC, possible transitions.

A block code pair of length $n$ for this channel with $M_1$ messages at terminal 1 and $M_2$ messages at terminal 2 consists of two encoding functions:

$$f_1 : \{1,...,M_1\} \longrightarrow \{0,1\}^n \text{ and}$$
$$f_2 : \{1,...,M_2\} \longrightarrow \{0,1\}^n .$$

The functions $f_1$ and $f_2$ prescribe how the messages are encoded into sequences of $n$ input symbols. A decoding function is a function $\phi$.

$$\phi : \{0,1,2\}^n \longrightarrow \{1,...,M_1\} \times \{1,...,M_2\}.$$

The decoding function gives a rule for the receiver to decide on the messages that were sent by terminal 1 and terminal 2.
the signalling rates for such a block code pair are:

$$R_1 := (\log M_1)/n \text{ and } R_2 := (\log M_2)/n.$$

As measure for the total amount of information passing through the channel we take $R_{\text{sum}} = R_1 + R_2$.

We define $C_i := \{f_i(m_i) | m_i \in \{1,...,M_i\}\}, \quad (i = 1,2)$.

In the noiseless case, the received output sequences come from the set

$$D = \{f_1(m_1) + f_2(m_2) | (m_1,m_2) \in \{1,...,M_1\} \times \{1,...,M_2\}\} =$$
$$\{c_1 + c_2 | c_1 \in C_1, c_2 \in C_2\}.$$

The coding problem for this case consists of constructing code pairs $(C_1,C_2)$ such that for all pairs $(c_1,c_2) \in C_1 \star C_2$, the real sums $c_1+c_2$ are different elements of $\{0,1,2\}^n$. Then the decoding function can be chosen such that the receiver finds the messages that were sent by users 1 and 2 without making any decoding errors. Such a code pair $(C_1,C_2)$ is called uniquely decodable.

In the presence of noise, we supply the set $\{0,1,2\}^n$ with a metric $d$ (depending on the nature of the noise). A code pair $(C_1,C_2)$ is said to be $d$-decodable if for all $(u,v) \neq (u',v') \in C_1 \star C_2$:

$$d(u+v,u'+v') \geq d.$$

In this case the decoding function $\phi$ is defined as follows. If $y \in \{0,1,2\}^n$, search for the word in $C_1+C_2$ which is at the smallest distance (in the given metric) to $y$. This word is related to a unique pair $(m_1,m_2) \in \{1,\ldots,M_1\} \star \{1,\ldots,M_2\}$. Then $\phi(y)$ is defined to be that unique pair.

In Section 2, we will describe two noise models which give rise to different metrics. Section 3 is concerned with the known results on $d$-decodable codes in the $L$-metric (which will be defined in Section 2). Section 4 gives the known results for the Hamming metric. In Section 5, a new construction method is given using a kind of concatenation. Section 6 contains upper and lower bounds for ternary codes in the $L$-metric. The numerical results that can be obtained from these bounds can be found in the appendices. The lower bounds from Section 6 are used to obtain lower bounds for $d$-decodable code pairs, applying the construction of Section 5. Section 7 shows these results. Furthermore it contains some other constructions. The numerical results can be found in the appendices to Section 7.

### 2.2. Heuristic discussion of two metrics corresponding to noise models

We will now discuss noise models leading to different metrics. In Figure 1 the two situations are depicted. In both cases, we are going to define a metric on the symbols. Then the distance between two words is defined as the sum of the distances between the symbols in the words. We say that one bit error occurs if 00 goes to 1, 11 goes to 1 or 01 or 10 go to 0 or 2, and that two bit errors occur if 00 goes to 2 or 11 goes to 0.

In the model of Figure 1a, we see that the probability of having one bit error is $2p(1-p)$ and the probability of having two bit errors is $p^2$ where $p$ is the probability to generate a 1, for the separate noise generators. We require that error patterns that have the same distance to 0 are about equiprobable. One can easily check that a good choice for the metric in this case is $d(0,1)=d(1,2)=1$; $d(0,2)=2$; $d(u,u)=0$.

This gives rise to the so-called $L$-metric (Cf. [1] and [2].):

$$d_L(u,v) = \sum_{i=1}^{D} |u_i - v_i| .$$

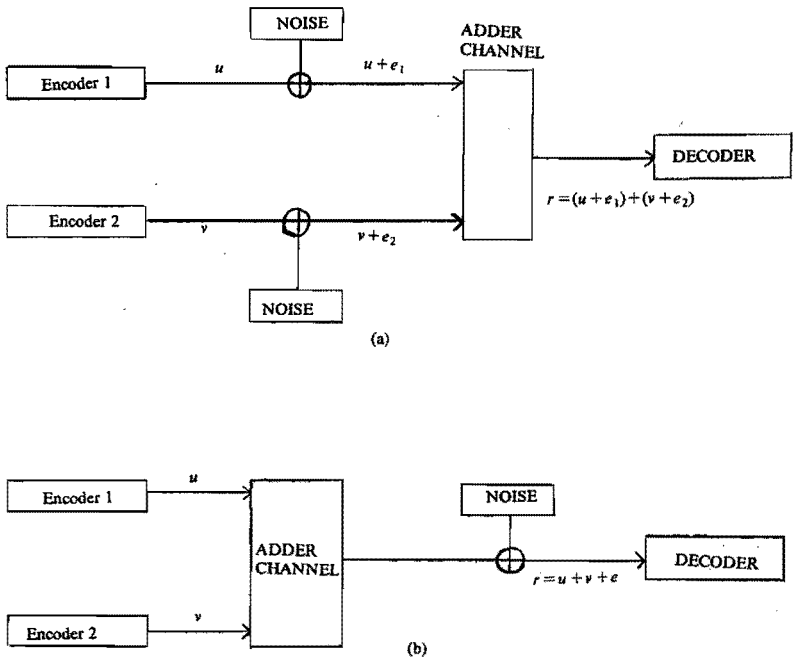



FIGURE 1. Two noise models:
(a) Noise at the input side of the channel.
(b) Noise at the output side of the channel.

In the model of Figure 1b we suppose that all errors are equiprobable. Then there is no point in taking a different measure for double bit errors. So we take $d(0,1)=d(1,2)=d(0,2)=1$, and $d(u,u)=0$. The resulting metric is the ordinary Hamming metric on ternary words:

$$d_H(u,v) = |\{i|u_i \neq v_i\}|.$$

(Cf. [9]).

## 2.3. THE KNOWN RESULTS ON D-DECODABLE CODES IN THE L-METRIC

### UPPER BOUNDS

In [6], Van Tilborg proved the following bound for $d$-decodable code pairs $C_1, C_2$ of length $n$:

THEOREM.

$$|C_1|\,|C_2| \leqslant 2 \sum_{k=0}^{\lfloor n/2 \rfloor - e} \begin{bmatrix} n \\ k \end{bmatrix} \cdot 2^k / W(n,e), \quad \text{where}$$

$$W(n,e) = 2\{ \begin{bmatrix} n/2 \\ e \end{bmatrix} + \begin{bmatrix} n/2 \\ e+1 \end{bmatrix} \}/(2n/(e+1)) \quad \text{if } d \text{ is even,}$$

$$W(n,e) = 2\left(\begin{bmatrix}\lceil n/2\rceil\\ e\end{bmatrix} + \begin{bmatrix}\lfloor n/2\rfloor+1\\ e+1\end{bmatrix} + \begin{bmatrix}\lfloor n/2\rfloor\\ e+1\end{bmatrix}\right)/\lfloor(2n/(e+1))\rfloor$$

if $d$ is odd,

*and* $e=\lfloor(d-1)/2\rfloor$.

This leads to the following upper bound for the sumrate $R^i_{\text{sum}}=R^i_1+R^i_2$ for a series of code pairs $(C_{1i},C_{2i})$ with $d_i/n_i\to\delta$, when $n_i\to\infty$:

(*where as before* $R^i_1=(\log|C_{1i}|)/n_i$; $R^i_2=(\log|C_{2i}|)/n_i, R_{\text{sum}}=\limsup_{i\to\infty} R^i_{\text{sum}}$)

$$R_{\text{sum}}\leqslant 3/2 + (1/2)\delta\log\delta/2-((1+\delta)/2)\log(1+\delta)=$$
$$= 1/2-(\delta/2) + h((1-\delta)/2 - h(\delta)/2\,.$$

From [1], $R_{\text{sum}}$ can also be bounded above by twice the best known upper bound for $[n,d]$-codes (since both $C_1$ and $C_2$ are binary $[n,d]$-codes). If one compares the above asymptotic upper bound with twice the bound in [7], then this latter bound is better for $\delta>0.16$. These two upper bounds are the best known up to now.

LOWER BOUNDS (Constructions and existence results).
In [1], the following code pairs were constructed using Reed-Muller codes:

1) $|C_1|=32;|C_2|=1177$; $d=4$, $n=16$, $d/n=0.25$; $R_{\text{sum}}=0.95005....$
2) $|C_1|=64;|C_2|=65309809$; $d=4$; $n=32$; $d/n=0.125$; $R_{\text{sum}}=0.68201....$
3) $|C_1|=64;|C_2|=58033$; $d=8$; $n=32$; $d/n=0.25$; $R_{\text{sum}}=0.68201....$
4) A class of 4-decodable code pairs:

$$|C_1|=2^{2^{m-1}-1};\ |C_2|=(1/2) + 2^{-(m-1)}\{3^{2^{m-1}}+2(2^{m-1}-1)3^{2^{m-2}}+1\};$$
$$n=2^m,\ d=4.$$

REMARK: Taking $m=5$ in the Construction 4, one gets a code pair having a much better sumrate and with the same $d/n$ as the code under 2) ($R_{\text{sum}}=1.0739$). In Section 7, we will give a much easier description of the class found under 4. In [2], there is given an asymptotic existence result for $d$-decodable code pairs. The arguments used to obtain this lower bound are basically the same as those in the proof of the Gilbert-Varshamov bound. Unfortunately Kasami et al. have made a mistake in the calculation, and the results in their paper [2] should be weakened somewhat. In Section 7 we will give a simpler description of the codes constructed here, and we will do the calculations correctly. Kasami et al. [3] contains a generalization of the construction in [2], however, this paper contains the same mistake. We will give the adjusted results in Section 7. In Kasami et al. [4], a graph theoretic approach is used instead of a Gilbert-Varshamov argument. The result obtained by this method is much stronger (asymptotically) than the other results. However, from the viewpoint of constructivity this method is far too complicated and unsuitable.

Khachatrian [8] gives a construction method using Kasami and Lin's coset approach (i.e. choose a linear code $C_1$, split $\mathbf{F}_2^n$ into cosets of $C_1$ and look how many words of each coset can included in $C_2$, such that $(C_1, C_2)$ is a $d$-decodable code pair). He constructs code pairs having the following parameters ($N =$ length, $d =$ distance):

| $N$ | $d$ | $d/N$ | $R_{\text{sum}}$ |
|---|---|---|---|
| 77 | 6 | 0.0779 | 0.8317 |
| 79 | 8 | 0.1013 | 0.7114 |
| 87 | 6 | 0.689 | 0.8701 |
| 89 | 8 | 0.0899 | 0.7782 |
| 445 | 10 | 0.0225 | 0.9472 |
| 447 | 12 | 0.0268 | 0.9162 |
| 459 | 10 | 0.0218 | 0.9512 |
| 461 | 12 | 0.0260 | 0.9211 |

TABLE 1. Code parameters for Khachatrians construction.

In Section 7, we will construct much better codes, i.e. shorter codes having the same distance and a higher sumrate. For uniquely decodable codes the results of Coebergh van den Braak et al. [17] are improved in one of the Theorems of Van Pul [15].

## 2.4. THE KNOWN RESULTS ON $d$-DECODABLE CODE PAIRS IN THE HAMMING METRIC

Since code pairs in the Hamming metric give rise to code pairs in the $L$-metric ($d_L(u,v) \geqslant d_H(u,v)$), we can take the same upper bound as in Section 3. In [9] Weldon suggests the following concatenation technique.

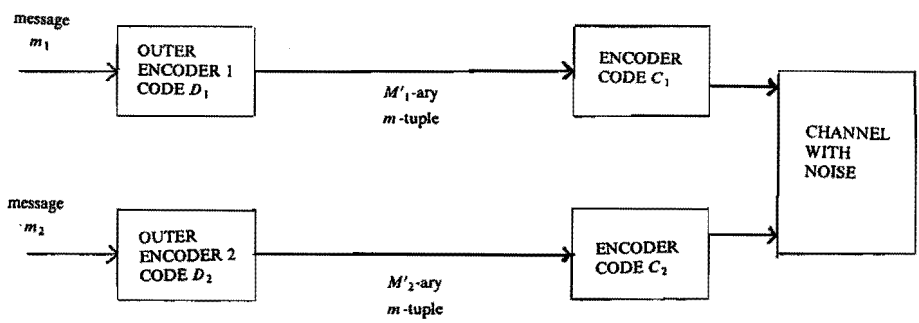



FIGURE 1. Encoding scheme using concatenated codes.

In Figure 1, we have an inner code pair $(C_1, C_2)$ which is $d_i$-decodable; $|C_1| = M'_1$, $|C_2| = M'_2$. The minimum distance of $C_t$ is $d_t^i$ $(t = 1,2)$.
$D_t$ is an $M_t$-ary code of length $n_o$ having Hamming distance $d_t^o, (t = 1,2)$. The codes $D_t$ are used as outer codes to encode the messages $m_t$ from the sources

($t=1,2$). W.l.o.g. we can assume that $d_1^o \geq d_2^o$. Then we have the following theorem. The code pair constructed by concatenating two outer codes $(D_1,D_2)$ with an inner code pair $(C_1,C_2)$ is called $(E_1,E_2)$.

THEOREM. *The minimum Hamming distance of such a code pair* $(E_1,E_2)$ *is bounded below by* $d=\min(d_i d_1^o, d_2^i d_2^o)$.

PROOF: (Weldon [9]). Let $e_1,e_2,f_1$ and $f_2$ be in $E_1,E_2,E_1,E_2$ respectively. We have to prove:
If $(e_1,e_2)\neq(f_1,f_2)$ then $d_H(e_1+e_2,f_1+f_2)\geq d$. We prove this by considering three cases: (i) The case that $e_2=f_2$, (ii) The case that $e_1=f_1$ and (iii) The case that $e_1\neq f_1$ and $e_2\neq f_2$.

(i) If $e_2=f_2$: $d_H(e_1+e_2,f_1+f_2) = d_H(e_1,f_1)\geq d_1^i d_1^o$.
(ii) If $e_1=f_1$: $d_H(e_1+e_2,f_1+f_2)=d_H(e_2,f_2)\geq d_2^i d_2^o$.
(iii) If $e_1\neq f_1$ and $e_2\neq f_2$: since $\hat{e}_1$ and $f_1$ differ in at least $d_1^o$ places (where $\hat{e}_1$ and $f_1$ are the codewords in $D_1$ corresponding to $e_1$ and $f_1$) we have that:

$$d_H(e_i+e_2,f_1+f_2)\geq d_i d_1^o.$$

Since $d_1^i \geq d_i$ we have that $d_i d_1^o \leq d_1 d^i d_1^o$. Hence in all three cases we find $d_H(e_1+c_2,f_1+f_2)\geq d$. Q.E.D.

EXAMPLE. Take $C_1=\{(0,0);(0,1);(1,0)\}$, $C_2=\{(0,0);(1,1)\}$. Then $(C_1,C_2)$ is a uniquely decodable code pair with distance $d_i=1$.
$|C_1|=3;|C_2|=2$. The minimum distance of $C_i$ is $d_1^i=1$. The minimum distance of $C_2$ is $d_2^i=2$. Let $D_2$ be a binary code of length $n$ and Hamming distance $d_2^o=d/2$. Let $D_1$ be a ternary code of length $n$ and Hamming distance $d_1^o=d$. The codes $D_t$ are used as outer codes to encode the messages $m_t$ from the sources ($t=1,2$). We have that $d_1^o=d\geq d/2=d_2^o$.
Then the code pair $(E_1,E_2)$ constructed by concatenation has length $2n$ and minimum distance: $\min(d_1 d_2^o, d_2^1 d_1^o)=\min(d,2d/2)=d$.
Furthermore $|E_1|=|D_1|=:M_1,|E_2|=|D_2|=:M_2$.
Choose an infinite sequence of binary codes $D_2^k$ of type $(n_k,M_2^k,d_k/2)$ satisfying $R_2^k = (\log M_2^k)/n_k \gtrsim 1-h(d_k/(2n_k))(n_k\to\infty)$. (Existence of such a sequence is guaranteed by a Gilbert-Varshamov bound cf. [12], [10]).
Furthermore choose an infinite sequence of ternary codes $D_1^k$ of type $(n_k,M_1^k,d_k)$ satisfying $R_1^k=\log M_1^k)/n_k \gtrsim \log 3-h(d_k/n_k)-d_k/n_k(n_k\to\infty)$. (Existence of such a sequence is guaranteed by a Gilbert-Varshamov argument cf. [10]).
Now by applying the concatenation construction we get an infinite sequence of code pairs $(E_1^k,E_2^k)$ of length $2n_k$, minimum distance $d_k$ and having sumrate:

$$R_{\text{sum}} = (\log M_1^k)/2n_k + (\log M_2^k)/2n_k = (R_1^k + R_2^k)/2 \geq$$
$$\geq (1-h(d_k/2n_k))/2 + (\log 3)/2 - h(d_k/n_k)/2 - d_k/2n_k \quad (n_k\to\infty).$$

If $d_k/n_k$ tends to $\delta$ for $n_k \to \infty$, we have found a sequence of code pairs with sumrates satisfying:

$$R_{\text{sum}} \geqslant 1/2\log 3 + 1/2 - \delta/2 - h(2\delta)/2 - h(\delta)/2$$

and the ratio distance to length tends to $\delta$ as the length tends to infinity. $(0 \leqslant \delta \leqslant 1/4)$.

## 2.5. Using the concatenation idea for constructing code pairs in the $L$-metric

Using the ideas of the previous section and especially the last example, we can give some constructions of good code pairs in the $L$-metric. The following new theorem is basic to these constructions.

THEOREM. *Let $\bar{C}$ be a binary $(n,d/2)$-code in the Hamming metric. Let $\bar{D}$ be a ternary $(n,d)$-code in the L-metric. If $C$ is the code obtained from $\bar{C}$ by replacing 0 by 00 and 1 by 11, $D$ is the code obtained from $\bar{D}$ by replacing 0 by 01, 1 by 00 and 2 by 10, then $(C,D)$ is a d-decodable code pair of length $2n$ in the L-metric.*

PROOF: Let $u,u',v,v'$ be in $C,C,D,D$ respectively. We must prove: $d_L(u+v,u'+v') \geqslant d$. We do this by considering the two possibilities (i) $v = v'$ and (ii) $v \neq v'$.

(i) If $v = v'$ then $d_L(u+v,u'+v') = d_L(u,u') \geqslant 2d/2 = d$.

(ii) If $v \neq v'$ then define $\bar{v}$ and $\bar{v'}$ to be the corresponding words in $\bar{D}$.
If $\bar{v}_i = \bar{v'}_i$ position $i$ gives no contribution to $d_L(\bar{v},\bar{v'})$.
If $\bar{v}_i \in \{0,2\}, \bar{v'}_i = 1$, position $i$ contributes 1 to $d_L(\bar{v},\bar{v'})$. Now:

$$01+00 = 01;\ 01+11=12;\ 10+00=10;\ 10+11=21;$$

$$00+00 = 00;\ 00+11=11.$$

So the contribution of position $i$ to $d_L(u+v,u'+v')$ is larger than or equal to 1.
If $\bar{v}_i = 0$; $\bar{v'}_i = 2$, position $i$ contributes 2 to $d_L(\bar{v},\bar{v'})$. Now:

$$01+00=01;\ 01+11=12;\ 10+00=10;\ 10+11=21.$$

So the contribution of position $i$ to $d_L(u+v,u'+v')$ is 2. Therefore we can conclude:

$$d_L(u+v,u'+v') \geqslant d_L(\bar{v},\bar{v'}) = d.$$

Hence we may conclude that in both cases we have $d_L(u+v,u'+v') \geqslant d$.

Q.E.D.

We are now left with two coding problems:
(i) construct good binary codes of length $n$ and Hamming distance $d/2$.
(ii) construct good ternary codes of length $n$ and distance $d$ in the $L$-metric.
The first problem is treated in many text books (e.g. [10,12] etc.), whereas the

second problem is new and therefore not yet discussed in the literature. We will treat this subject in the next section.

The results for $d$-decodable code pairs obtained by this construction and the results of the next section together with the results on the best known codes in [10] and [12], can be found in Section 7.

### 2.6. Bounds and constructions for ternary codes in the $L$-metric

This section is devoted to ternary codes in the $L$-metric. We will give (asymptotic) lower and upper bounds for these codes. Also we will discuss a construction method. In the appendices we will give some tables containing the best codes known having length less than 20 and distance less than 16. Also in the appendices the numerical results for the asymptotic bounds can be found. Finally the asymptotic bounds are plotted. In analogy to $A(n,d)$ for binary codes (Cf. [12], Chapter 17), we define: $A_L(n,d)$ to be the maximum number of words in a ternary code of length $n$ and with $L$-distance $\geqslant d$.

Trivial observations:
*(1)* $A_L(n, 1) = 3^n$,
*(2)* $A_L(n, 2) = (3^n+1)/2$.

Proof: The first statement follows from the fact that any two different words have $L$-distance $\geqslant 1$. To prove the second statement define:

$$n = 1 : S_n = S_1 := \{\{0,1\}\},$$

$$n>1 : S_n := \{\{0u,0v\}, \{1u,1v\}, \{2u,2v\} | \{u,v\} \in S_{n-1}\} \cup$$

$$\cup \{\{02...2,12...2\}\}.$$

The following properties of $S_n$ follow by induction:

$$\forall_{\{u,v\} \in S_n} [d_L(u,v) = 1], \tag{1}$$

$$\bigcup_{\{u,v\} \in S_n} \{u,v\} = \{0,1,2\}^n - \{(2,...,2)\}, \tag{2}$$

$$|\{u,v\} \cap \{u',v'\}| = 0 \text{ or } 2 \text{ for all } \{u,v\}, \{u',v'\} \in S_n, \tag{3}$$

$$|S_n| = (3^n-1)/2 \tag{4}$$

From (1), (2), (3) and (4) it follows that a code in $\{0,1,2\}^n$ having distance 2 can have at most one word from each pair in $S_n$. Hence $A_L(n, 2) \leqslant (3^n-1)/2+1=(3^n+1)/2$.

We can reach this number of words by taking all words $c$ with $\Sigma c_i \equiv 0 mod 2$. There are $(3^n+1)/2$ such words and their $L$-distance is even. Q.E.D.

We will now introduce a notation which will be used frequently.

Notation:
Let $c \in \{0,1,2\}^n$, $c=(c_1,...,c_n)$. Then:

$$\bar{c} := (c_1 \bmod 2, \ldots, c_n \bmod 2),$$

$$\hat{c} := (c_{i_1}, \ldots, c_{i_t})$$

where the $i_j$ are the coordinate positions where $c$ does not have a 1, (i.e. $c$ has a 0 or a 2), $i_1 < i_2 < \ldots < i_t$ ($\hat{c}$ is possibly empty).

LOWER BOUNDS

We will now derive a lower bound on $A_L(n,d)$. Define:

$$V_w := \{c \in \{0,1,2\}^n \mid wt(\bar{c}) = w\}.$$

For $c \in V_w$:

$$B_w(c,d) := \{x \in V_w \mid d_L(x,c) \leqslant d\}.$$

We have the following theorem.

THEOREM 1.

*$|B_w(c,d)|$ is independent of the centre $c \in V_w$.*

$$|B_w(c,d)| = \sum_{0 \leqslant i+j \leqslant d/2} 2^i \cdot \begin{bmatrix} w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w-i \\ j \end{bmatrix}.$$

PROOF: If $c \in V_w$, w.l.o.g. the $w$ ones are in the first $w$ places (reordering coordinates). Again without loss of generality we can take the last $n-w$ places to have zeroes. Indeed, adding the 0/2-pattern of our word to all other words and reducing coordinates mod4 where furthermore 3 becomes 1, the distance between words is not changed. This shows the first statement. Now

$$c = (1 \ldots 10 \ldots 0) = (1^w 0^{n-w}).$$

The number of words $e$ having $w$ ones, $i$ zeroes or twos among the first $w$ places and $j$ twos among the last $n-w$ places, is equal to:

$$2^i \begin{bmatrix} w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w-i \\ j \end{bmatrix}.$$

For these words $e$ we have: $d_L(c,e) = 2(i+j)$, and look like $e = (1^{w-i}0/2^i|1^i 2^j 0^{n-w-i-j})$ were the front $w$ and back $n-w$ places may be permuted.

Therefore $|B_w(c,d)| = \sum_{0 \leqslant i+j \leqslant d/2} 2^i \cdot \begin{bmatrix} w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w \\ i \end{bmatrix} \cdot \begin{bmatrix} n-w-i \\ j \end{bmatrix}$. Q.E.D.

Theorem 1 leads in a natural way to the following lower bound:

THEOREM 2. *(Analogue of the Gilbert bound)*

$$A_L(n,d) \geqslant \max_{0 \leqslant w \leqslant n} (|V_w|)/(|B_w(c,d-1)|), \quad \textit{where } c \in V_w.$$

PROOF: Let $C$ be a code consisting of words in $V_w$, having minimal $L$-distance $d$, and a maximal number of words. If $|C|\cdot|B_w(c,d-1)|<|V_w|$, where $c\in V_w$, then we can find a word in $V_w$ having distance $\geqslant d$ to all words in $C$. But since $C$ has the maximum number of words of words in $V_w$, the last statement is false so $|C|\cdot|B_w(c,d-1)|\geqslant|V_w|$, from which the theorem follows. Q.E.D.

In Appendix 1 to this section, one can find a table showing this bound for small values of $n$ and $d$. Of course, as with the Gilbert bound, the strength of this bound is in its asymptotics. We will now calculate the asymptotic behaviour, assuming $d/n<1/2$ and $w/n<1/2$. Define:

$$R_L(\delta) := \limsup_{n\to\infty}(\log A_L(n,\delta n))/n$$

Now writing $x$ for $i/n$,$y$ for $j/n$,$z$ for $w/n$ and $t$ for $d/n$($z\leqslant 1/2$ and $t\leqslant 1/2$) and using the elementary estimates for binomial coefficients (see [12], Chapter 11) we find:

$$R_L(t)\geqslant 1-z+h(z)-\max_{0\leqslant x+y<t/2}(x+zh(x/z)+(1-z)h(x/(1-z)) +$$
$$+ (1-z-x)h(y/(1-z-x))) + o(1) \text{ (for all } z\leqslant 1/2,\ n\to\infty).$$

Define:

$$f(x,y) := x+zh(x/z) + (1-z)h(x/(1-z)) + (1-z-x)h(y/(1-z-x)).$$

Since $d/n<1/2$ and $w/n<1/2$ we have that $t+z<1$. Therefore: $t/2-x<(1-z-x)/2$. Hence maximizing $f(x,y)$ over $y$ it is easily seen that this maximum is taken for $y=t/2-x$, since $h(u)$ is monotonically increasing for $0\leqslant u\leqslant 1/2$. This maximum is equal to:

$$f(x) = x+zh(x/z)+(1-z)h(x/(1-z))+1(1-z-x)h((t/2-x)/(1-z-x)).$$

By standard analysis, this function has a maximum in:

$$x=z+t/2-\sqrt{(z^2+t^2/4)} = (zt)/(z+t/2+\sqrt{z^2+t^2/4}) =: x_0$$

and we find:

$$R_L(t)\geqslant 1-z-h(z)-f(x_0)+o(1)\ \ (n\to\infty)\ \text{ for } \textit{all}\ z<1/2.$$

Taking $z=1/3$ we find the following asymptotic result:

THEOREM 3. *(asymptotic analogue of the Gilbert bound)*

$$R_L(\delta)\geqslant \log 3-x_0-h(3x_0)/3-2h(3x_0/2)/3 +$$
$$- (2/3-x_0)h((\delta/2-x_0)/(2/3-x_0))\quad (n\to\infty).$$

*where* $x_0=1/3+\delta/2-\sqrt{\delta^2/4+1/9}$.

(The reason that we take $z=1/3$ is that if $t$ is small compared to $z$, we have that $x_0$ is approximately $t$, and therefore $R$ is approximately $1-z-h(z)$ (with

a small correction term $f(t)$). This takes a maximum at $z=1/3$. If $t$ is about equal to $z$, we also find that the maximum is near $z=1/3$). In Appendix 8 we will give some numerical results using the bound of Theorem 3. Also the bound will be plotted in Appendix 9. We shall now give another lower bound based on the following Theorem of Turán [16].

THEOREM 4. (Turán) *Let $G=(V,E)$ be a graph, $V$ the set of points, $E$ the set of edges. If $y$ and $y'$ are the sizes of respectively the largest clique and the largest coclique then:*

$$y \geqslant (|V|^2)/(|V|^2 - 2|E|), \; y' \geqslant (|V|^2)/(|V| + 2|E|)$$

For a proof of this theorem we refer to Appendix 10. Now define the following graph on $\{0,1,2\}^n$

$$x \sim y \text{ iff } d_L(x,y) \leqslant d-1.$$

If we have a coclique in this graph, then this is a ternary code having $L$-distance $d$. Since $V=\{0,1,2\}^n$, we have $|V|=3^n$. Calculating the number of edges we find:

$$|E| = |\{\{x,y\} : 0<d_L(x,y)\leqslant d-1\}| =$$

$$1/2|\{(x,y) : 0<d_L(x,y)\leqslant d-1\}| =$$

$$1/2\sum_{k=0}^{n}|\{(x,y) : 0<d_L(x,y)\leqslant d-1 \, , \, wt(\bar{x}) = k\}| =$$

$$1/2\sum_{k=0}^{n}\binom{n}{k}\cdot 2^{n-k}\cdot|\{y:0<d_L(x_0,y)\leqslant d-1, \; x_0=(1^k0^{n-k})\}| =$$

$$1/2\sum_{k=0}^{n}\binom{n}{k}\cdot 2^{n-k}\cdot\sum_{l=0}^{n}|\{y:0<d_L(x_0,y)\leqslant d-1, x_0=(1^k0^{n-k}), \; wt(y)=l\}|$$

$$= 1/2\sum_{k=0}^{n}\binom{n}{k}\cdot 2^{n-k}\cdot\{[\sum_{l=0}^{n}\sum_{C}2^{k-i}\binom{k}{i}\cdot\binom{n-k}{l-1}\cdot\binom{n-k-l+i}{j}]-1\},$$

where $C$ stands for the conditions:

$$C : 0\leqslant k+l-2i+2j\leqslant d-1$$

$$0\leqslant i\leqslant l$$

$$0\leqslant i\leqslant k$$

$$0\leqslant j\leqslant n-k-l+i.$$

The last equality follows from the following figure:

$$x_0 = \underset{k}{1.....1}\underset{n-k}{0.....0}$$

$$y = \underset{k-i}{*.....*}\underset{i}{1.....1}\underset{l-i}{1.....1}\underset{j}{2.....2}\underset{n-k-l+i-j}{0.....0} \qquad * \in \{0,2\}$$

FIGURE 1. Two words having $L$-distance $d_L(x_0,y)=k+l-2i+2j$.

From the above formula we find:

$$|V| + 2|E| = \sum_{k=0}^{n} \sum_{l=0}^{n} \sum_{C} 2^{n-i} \begin{bmatrix} n \\ k \end{bmatrix} \cdot \begin{bmatrix} k \\ i \end{bmatrix} \cdot \begin{bmatrix} n-k \\ l-i \end{bmatrix} \cdot \begin{bmatrix} n-k-l+i \\ j \end{bmatrix} =: E(n,d).$$

Applying Turáns Theorem we obtain:

THEOREM 5. *(Turán Lower bound)*

$$A_L(n,d) \geqslant (3^{2n})/E(n,d).$$

In Appendix 2 to this section, we will give a table showing this bound for small values of $n$ and $d$. Of course the strength of this bound also lies in its asymptotic behaviour. From Theorem 5 it follows that:

$$\log A_L(n, \delta n))/n \geqslant (\log 3^{2n} - \log E(n,\delta n))/n = 2\log 3 - (\log E(n,\delta n))/n$$

In order to find an estimate for the right-hand side we have to find an upper bound for $S(\delta n) := (\log E(n,\delta n))/n$. It is easy to see that for $n \to \infty$:

$$S(\delta n) \leqslant (\log( \max_{\substack{0 \leqslant k \leqslant n \\ 0 \leqslant l \leqslant n \\ 0 \leqslant k+l-2i+2j \leqslant \delta n - 1}} ( \begin{bmatrix} n \\ k \end{bmatrix} \cdot \begin{bmatrix} k \\ i \end{bmatrix} \cdot \begin{bmatrix} n-k \\ l-i \end{bmatrix} \cdot \begin{bmatrix} n-k-l+i \\ j \end{bmatrix} \cdot 2^{n-i} ))) / n + o(1)$$

By elementary reasoning it can be seen that the maximum is taken when $l = k$, so for $n \to \infty$:

$$S(\delta n) \leqslant (\log( \max_{\substack{0 \leqslant k \leqslant n \\ 0 \leqslant 2(k-i+j) \leqslant \delta n - 1}} ( \begin{bmatrix} n \\ k \end{bmatrix} \cdot \begin{bmatrix} k \\ i \end{bmatrix} \cdot \begin{bmatrix} n-k \\ k-i \end{bmatrix} \cdot \begin{bmatrix} n-2k+i \\ j \end{bmatrix} \cdot 2^{n-i} ))) / n + o(1)$$

Now putting $x = i/n, y = j/n, \lambda = k/n$, defining:

$$f(x,y,\lambda) := h(\lambda) + \lambda h(x/\lambda) + (1-\lambda) h((\lambda - x)/(1-\lambda)) + (1 - 2\lambda + x) h(y/(1-2\lambda+x)) + 1 - x,$$

and using the standard estimates for binomial coefficients we get:

$$S(\delta n) \leqslant \max_{\substack{0 \leqslant x \leqslant \lambda \\ 0 \leqslant \lambda - x + y \leqslant \delta/2}} f(x,y,\lambda) + o(1) \qquad (n \to \infty).$$

Under the assumption that $\delta/2 + x - \lambda \leqslant 1/2 - \lambda + x/2$ (i.e. $x \leqslant 1 - \delta$), it follows from the fact that $h(z)$ is monotonically increasing for $0 < z < 1/2$, that the maximum is attained for $y = \delta/2 + x - \lambda$. Hence:

$$S(\delta n) \leqslant \max_{\substack{0 \leqslant x \leqslant \lambda \\ 0 \leqslant \lambda \leqslant 1 \\ 0 \leqslant x \leqslant 1-\delta \\ 0 \leqslant \delta/2 - \lambda + x}} f_1(x,\lambda) + o(1) \qquad (n \to \infty),$$

where :

$$f_1(x,\lambda) := f(x, \delta/2 - \lambda + x, \lambda)$$

Doing some calculus, we find:

$$f_1(x,\lambda) = -x\log x - 2(\lambda-x)\log(\lambda-x) - (\delta/2-(\lambda-x))\log(\delta/2-(\lambda-x)) +$$
$$-(1-\lambda-\delta/2)\log(1-\lambda-\delta/2)+1-x.$$

Substituting $t$ for $\lambda-x$ we find:

$$S(\delta n) \leqslant \max_{\substack{0\leqslant t\leqslant \delta/2 \\ 0\leqslant x\leqslant 1-\delta}} (-x\log x - 2t\log t - (\delta/2-t)\log(\delta/2-t) +$$
$$-(1-x-\delta/2-t)\log(1-x-\delta/2-t) + 1-x).$$

Differentiating with respect to $t$ we find:

$$\log((1-x-\delta/2-t)(\delta/2-t)/t^2).$$

Therefore the maximum is attained for $t=\delta(1-x-\delta/2)/(2(1-x))$. Hence:

$$S(\delta n) \leqslant M(\delta) := \max_{0\leqslant x\leqslant 1-\delta} (-x\log x +$$
$$-\delta(1-\delta/(2(1-x))\log(\delta(1-\delta/(2(1-x)))/2 +$$
$$-\delta^2/(4(1-x))\log(\delta^2/(4(1-x))) +$$
$$-(1-x-\delta+\delta^2/(4(1-x)))\log(1-x-\delta+\delta^2/(4(1-x))) +$$
$$+1-x).$$

Therefore we have the following theorem.

THEOREM 6. *(Asymptotic Turán lower bound)*

$$R_L(\delta) \geqslant 2\log 3 - M(\delta) + o(1) \quad (n\to\infty).$$

Appendix 8 contains the numerical values for the asymptotic bounds of Theorems 3 and 6. It appears that for small values of $n$ and $d$ the Turán bound is better than the generalized Gilbert bound. However, asymptotically they are very close. With the aid of an extension trick, we can sharpen the bounds of Theorems 2 and 5 somewhat in the case that $d$ is even.

Namely: Let $C$ be a $(n,d)$ ternary code in the $L$-metric, where $d$ is odd. If we extend $C$ in the following way: put a 1 in front of a word $c$ if $wt(\bar{c})$ is odd and a 0 otherwise, then we find a code of length $n+1$, distance $\geqslant d+1$ and the same number of words. (Extending in this way makes the distance of the resulting code even.)

Next we want to describe a construction method for ternary codes in the $L$-metric which ameliorates the bounds of Theorem 2 and 5 for small values of $n,d$. In Appendix 3 we will give for certain values of $n$ and $d$ the results we get using this construction. Moreover we will describe an asymptotic bound which can be obtained applying the construction. However, asymptotically the

bound loses its strength i.e. the asymptotic bounds of Theorem 3 and 6 are stronger than Theorem 8.

CONSTRUCTION. Let $D'$ be a binary code of length $n$ and Hamming distance $d$. Let $w_1,...,w_t$ be the nonzero weights occurring in $D'$. Let $D_{w_j}$ be binary codes of length $n-w_j$ and Hamming distance $\geq d/2 (j=1,...,t)$. Define:

$$D := \{c \in \{0,1,2\}^n | \bar{c} \in D';\ 1/2 \cdot \hat{c} \in D_{w_j} \text{ if } wt(\bar{c}) = w_j\}.$$

THEOREM 7. *D is a ternary code having L-distance d.*

PROOF: Let $c_1,c_2 \in D, c_1 \neq c_2$. If $\bar{c}_1 \neq \bar{c}_2$ then $d_L(c_1,c_2) \geq d_H(\bar{c}_1,\bar{c}_2) \geq d$. If $\bar{c}_1 = \bar{c}_2$ then $\hat{c}_1 \neq \hat{c}_2$ and $d_L(c_1,c_2) = 2d_H(1/2 \star \hat{c}_1, 1/2 \star \hat{c}_2) \geq 2d/2 = d$. Q.E.D.

For examples and numerical results see Appendix 3.
The construction described above gives rise to an asymptotic bound (using asymptotic lower bounds on $A(n,d)$ and $A(n,d,w)$ cf. [12]).
Let $C$ be a code with weight distribution $A_i, A_0 = 1, A_1 = 0, ..., A_{d-1} = 0, A_d > 0$. Applying the above construction, we get a ternary code in $L$-metric having cardinality:

$$|C| = \sum_{i=0}^{n} A_i A(n-i,d/2).$$

So $A_L(n,d) \geq \Sigma_{i=0}^{n} A_i A(n-i,d/2)$.
In particular $A_L(n,d) \geq A(n,d,w) A(n-w,d/2)$ for all $w$ between 0 and $n-d/2$. Now:

$$\limsup_{\substack{n\to\infty \\ d/n\to\delta \\ w/n\to\omega}} (\log A(n,d,w))/n \geq \limsup_{\substack{n\to\infty \\ d/n\to\delta \\ w/n\to\omega}} (\log 1/2^n \begin{bmatrix} n \\ w \end{bmatrix} \cdot A(n,d))/n \geq$$

$$h(\omega) + 1 - h(\delta) - 1 = h(\omega) - h(\delta)$$

(where the first inequality follows from [12], p. 558 Theorem 33), and:

$$\limsup_{\substack{n\to\infty \\ d/n\to\delta \\ w/n\to\omega}} (\log A(n-w,d/2))/n \geq (1-\omega)(1-h(\delta/(2(1-\omega)))).$$

Conclusion:

THEOREM 8. *(Asymptotic lower bound using concatenation techniques)*

$$R_L(\delta) \geq 1 - h(\delta) + h(\omega) - \omega - (1-\omega) h(\delta/(2(1-\omega))),\ (0 < \omega \leq 1 - \delta/2),$$

$$R_L(\delta) \geq 1 - h(\delta/2),\ (\omega = 0).$$

The numerical results of this bound can be found in Appendix 8. (of course we maximize over $\omega$).

We will now discuss some upper bounds for codes in the $L$-metric. The first one is an analogue of the sphere packing bound for binary codes. The second one is a slight improvement of the first. The last one uses some ideas of H.C.A. van Tilborg [6], and leads to a kind of linear programming bound. Let $V(x,r)$ denote the cardinality of a sphere with center $x$ and radius $r$:

$$V(x,r) = |\{y | d_L(x,y) < r\}|.$$

Let $C$ be a ternary code having length $n$, $L$-distance $d$. Spheres of radius $d/2$ around codewords are disjoint, and therefore the following inequality is obvious:

$$\sum_{x \in C} V(x,d/2) \leqslant 3^n.$$

Now

$$V(x,d/2) \geqslant V(0,d/2) = \sum_{0 \leqslant i+2j \leqslant d/2} \binom{n}{i} \binom{n-i}{j}.$$

Hence we have the following two Theorems:

UPPERBOUNDS

THEOREM 9. *(sphere packing bound)*

$$A_L(n,d) \leqslant 3^n / \left( \sum_{0 \leqslant i+2j < d/2} \binom{n}{i} \cdot \binom{n-i}{j} \right).$$

THEOREM 10. *(sphere packing bound asymptotic version).*

$$R_L(\delta) \leqslant \log 3 - h(x_0) - (1-x_0) h((\delta/4 - x_0/2)/(1-x_0)),$$

*where* $x_0 = -1/3 + 1/3\sqrt{(1+3\delta-3\delta^2/4)}$.

PROOF: Theorem 9 follows by the above argument. From Theorem 9, we may conclude (using the standard estimates for binomial coefficients):

$$R_L(\delta) \leqslant \log 3 - \max_{0 \leqslant x+2y \leqslant \delta/2} (h(x) + (1-x) h(y/(1-x))).$$

Since $h(z)$ is monotonically increasing on $(0,1/2)$ and $y \leqslant \delta/4 - x/2$, this maximum is taken when $y = \delta/4 - x/2$. So we have to calculate:

$$\max_{0 \leqslant x \leqslant \delta/2} \{h(x) + (1-x) h((\delta/4 x/2(1-x))\}.$$

Let $f(x) = h(x) + (1-x) h((\delta/4 - x/2)/(1-x)) =$

$$= -x \log x - (\delta/4 - x/2) \log(\delta/4 - x/2) +$$

$$- (1 - \delta/4 - x/2) \log(1 - \delta/4 - x/2).$$

Then $f'(x) = 1/2 \log((\delta/4 - x/2)(1 - \delta/4 - x/2)/x^2)$, which is equal to 0 for $x = x_0 = -1/3 + 1/3\sqrt{(1+3\delta-3\delta^2/4)}$, and it is easy to see that the function

$f$ attains its maximum in $x_0$. Q.E.D.

The numerical values for these two bounds can be found in Appendices 4 and 8.

Improved sphere packing bound:

Let $C$ be a code with $L$-distance $d$.

$$A_w = |\{x \in C | wt(x) = w\}|; \; B_L(x,d) = \{y | d_L(x,y) < d\}.$$

$|B_L(x,d)|$ depends only on $wt(\bar{x})$. Now: $\Sigma_{x \in C} |B_L(x,d/2)| \leqslant 3^n$. (Spheres with radius $d/2$ are disjoint.) Moreover:

$$\sum_{x \in C} |B_L(x,d/2)| = \sum_{w=0}^{n} \sum_{x \in C, wt(\bar{x})=w} |B_L(x,d/2)| =$$

$$= \sum_{w=0}^{n} A_w B_L(n,w,d/2) \leqslant 3^n,$$

where $B_L(n,w,d/2) = |B_L(x,d/2)|$ for some word $x$ of weight: $wt(\bar{x}) = w$. Hence we have the following Theorem:

THEOREM 11. *(Improved sphere packing bound)*

$$A_L(n,d) \leqslant \max \left\{ \sum_{w=0}^{n} A_w \Big| \sum_{w=0}^{n} A_w B_L(n,w,d/2) \leqslant 3^n, A_w \leqslant \binom{n}{w} 2^{n-w} \right\}.$$

The results of this theorem can be found in Appendix 4. There is no asymptotic version of this bound.

The last bound we will discuss is based on ideas of H.C.A. Van Tilborg [5] and [6]. Let $C$ be a ternary code having $L$-distance $d$ and length $n$. Define:

$$W_n = \{0,1,2\}^n; \; W_{n,k} = \{u \in W_n | wt(u) = k\}.$$

$$R_k = |W_{n,k} \cap C|.$$

$$f(k,i) = |\{(u,v) | u \in W_{n,k}, \; v \in C, d_L(u,v) = i\}|$$

$$N(k,i) = \max_{x \in W_{n,k}} |\{v \in C | d_L(v,x) = i\}|.$$

It is easy to see that: $|C| = \Sigma_{k=0}^{n} R_k$.

LEMMA 12.

$$f(k,i) = \sum_{\substack{p,q,r \\ r+q+2p=i}} \binom{k+r-q}{r} \cdot \binom{n-k-r+q}{q} \cdot \binom{n-k-r}{p} \cdot 2^r \cdot R_{k+r-q}.$$

PROOF: We have to count the numer of pairs $(u,v)$ such that (w.l.o.g.) the following situation occurs:

$$v = \underbrace{1 \ldots\ldots 1}_{k+r-q} \; 0/2 \underbrace{\ldots\ldots 0/2}_{n-k-r+q} \quad (r+q+2p=1)$$

$$u = \underbrace{1 \ldots\ldots 1}_{k-q} \; \underbrace{0/2 \ldots 0/2}_{r} \; \underbrace{1 \ldots\ldots 1}_{q} \; \underbrace{2/0 \ldots 2/0}_{p} \; 0/2 \ldots 0/2$$

The number of choices for $v$ is $R_{k+r-q}$. The number of choices for a suitable $u$ is:

$$\binom{k+r-q}{r} \cdot \binom{n-k-r+q}{q} \cdot \binom{n-k-r}{p} \cdot 2^r. \qquad \text{Q.E.D.}$$

In [5] Van Tilborg proved the following two lemmas.

LEMMA 13. *If* $d=2e+2$ *then* $N(k,e+1) \leqslant \lfloor 2n/(e+1) \rfloor$

LEMMA 14.

$$\sum_{i=0}^{e} f(k,i) \leqslant \binom{n}{k} \cdot 2^{n-k} \quad (= \# W_{n,k}), \quad d=2e+1.$$

$$\sum_{i=0}^{e} f(k,i) + f(k,i)/(\lfloor 2n/(e+1) \rfloor) \leqslant \binom{n}{k} \cdot 2^{n-k}, d=2e+2.$$

From the above Lemmas we get the following theorem.

THEOREM 15. *(L.P.-bound)*

$A_L(n,d) \leqslant \max\{\Sigma R_k \mid \textit{Condition}\}$,

*where Condition stands for:*

$$\sum_{i=0}^{e} \sum_{\substack{p,q,r \\ r+q+2p=i}} \binom{k+r-q}{r} \cdot \binom{n-k-r+q}{q} \cdot \binom{n-k-r}{p} \cdot 2^r \cdot R_{k+r-q} \leqslant \binom{n}{k} \cdot 2^{n-k},$$

$R_k \geqslant 0$,

*if* $d=2e+1$:

$$\sum_{i=0}^{e} \sum_{\substack{p,q,r \\ r+q+2p=i}} \binom{k+r-q}{r} \cdot \binom{n-k-r+q}{q} \cdot \binom{n-k-r}{p} \cdot 2^r \cdot R_{k+r-q} +$$

$$\sum_{\substack{p,q,r \\ r+q+2p=e+1}} \left( \binom{k+r-q}{r} \cdot \binom{n-k-r+q}{q} \cdot \binom{n-k-r}{p} \cdot 2^r \cdot R_{k+r-q} \right) / (\lfloor 2n/(e+1) \rfloor)$$

$$\leqslant \binom{n}{k} \cdot 2^{n-k} \quad , \text{ if } \quad d=2e+2.$$

Some numerical results due to this bound can be found in Appendix 5. Appendix 6 gives some ad hoc improvements on the lower bounds. Appendix 7 is a final table containing best known lower and upper bounds for codes up to length 20 and minimum distance $\leqslant 8$. Appendix 9 contains the graphs for the derived asymptotic bounds for the codes in the $L$-metric.

## Appendix 1. Table of numerical values for the lower bound of Theorem 2 (Gilbert bound)

| $n$ | $d$ | 3 | | 4,5 | | 6,7 | | 8,9 | | 10,11 | | 12,13 | | 14,15 | | 16,17 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $w$ | | $w$ | | $w$ | | $w$ | | $w$ | | $w$ | | $w$ | | $w$ |
| 2 | | 2 | 0 | 1 | 0 | - | - | - | - | - | - | - | - | - | - | - | - |
| 3 | | 2 | 0 | 2 | 0 | 1 | 0 | - | - | - | - | - | - | - | - | - | |
| 4 | | 4 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | - | - | - | - | - | - | - | - |
| 5 | | 7 | 1 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | - | - | - | - | - | - |
| 6 | | 12 | 1 | 3 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | - | - | - | - |
| 7 | | 26 | 2 | 5 | 1 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | - | - |
| 8 | | 58 | 2 | 9 | 1 | 3 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 |
| 9 | | 128 | 2 | 15 | 2 | 5 | 1 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| 10 | | 308 | 3 | 29 | 2 | 7 | 1 | 3 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| 11 | | 742 | 3 | 56 | 3 | 11 | 2 | 4 | 1 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| 12 | | 1760 | 3 | 117 | 3 | 19 | 2 | 6 | 1 | 3 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| 13 | | 4465 | 4 | 245 | 3 | 34 | 3 | 9 | 2 | 4 | 1 | 2 | 0 | 2 | 0 | 2 | 0 |
| 14 | | 11264 | 4 | 560 | 4 | 63 | 3 | 14 | 2 | 5 | 1 | 3 | 0 | 2 | 0 | 2 | 0 |
| 15 | | 27956 | 4 | 1188 | 4 | 120 | 3 | 24 | 3 | 8 | 2 | 4 | 1 | 2 | 0 | 2 | 0 |
| 16 | | 73326 | 5 | 2649 | 4 | 239 | 4 | 41 | 3 | 12 | 2 | 5 | 1 | 3 | 0 | 2 | 0 |
| 17 | | 190572 | 5 | 6095 | 5 | 485 | 4 | 73 | 4 | 19 | 3 | 7 | 2 | 4 | 1 | 2 | 0 |
| 18 | | 487424 | 5 | 14319 | 5 | 990 | 5 | 135 | 4 | 30 | 3 | 10 | 2 | 5 | 1 | 3 | 0 |
| 19 | | 1307444 | 6 | 33389 | 5 | 2222 | 5 | 254 | 4 | 50 | 4 | 15 | 3 | 6 | 2 | 4 | 1 |
| 20 | | 3470185 | 6 | 80203 | 6 | 4550 | 5 | 488 | 5 | 88 | 4 | 24 | 3 | 9 | 2 | 5 | 1 |

## Appendix 2. Table of numerical values for the lower bound of Theorem 5 (Turán bound)

| $n$ | $d$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | | 2 | 2 | - | - | - | - | - | - | - | - | - | - |
| 3 | | 3 | 2 | 2 | 2 | - | - | - | - | - | - | - | - |
| 4 | | 5 | 3 | 2 | 2 | 2 | 2 | - | - | - | - | - | - |
| 5 | | 9 | 4 | 2 | 2 | 2 | 2 | 2 | 2 | - | - | - | - |
| 6 | | 19 | 7 | 4 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | - | - |
| 7 | | 42 | 13 | 6 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 8 | | 99 | 27 | 10 | 5 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 9 | | 238 | 57 | 19 | 8 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 2 |
| 10 | | 585 | 127 | 38 | 15 | 7 | 4 | 3 | 2 | 2 | 2 | 2 | 2 |
| 11 | | 1467 | 291 | 79 | 27 | 12 | 6 | 4 | 3 | 2 | 2 | 2 | 2 |
| 12 | | 3734 | 681 | 169 | 54 | 21 | 10 | 6 | 4 | 3 | 2 | 2 | 2 |
| 13 | | 9624 | 1624 | 372 | 109 | 39 | 17 | 9 | 5 | 3 | 3 | 2 | 2 |
| 14 | | 25071 | 3936 | 838 | 227 | 75 | 30 | 14 | 7 | 5 | 3 | 2 | 2 |
| 15 | | 65922 | 9677 | 1924 | 486 | 150 | 55 | 24 | 12 | 7 | 4 | 3 | 2 |
| 16 | | 174750 | 24088 | 4492 | 1064 | 307 | 106 | 42 | 20 | 10 | 6 | 4 | 3 |
| 17 | | 466585 | 60618 | 10646 | 2372 | 644 | 207 | 78 | 34 | 16 | 9 | 6 | 4 |
| 18 | | 1253788 | 154031 | 25563 | 5377 | 1377 | 418 | 147 | 60 | 27 | 14 | 7 | 5 |
| 19 | | 3388518 | 394805 | 62103 | 12373 | 3000 | 861 | 287 | 109 | 47 | 23 | 12 | 7 |
| 20 | | 9205357 | 1019893 | 152466 | 28852 | 6639 | 1807 | 570 | 206 | 84 | 38 | 19 | 11 |

## Appendix 3. Examples and numerical results for the construction method of Theorem 7

Example 1. Let $D'$ be an extended binary Hamming code of length $n=2^{m-1}$, and minimum distance 4. $D'$ has weight distribution function:

$$\sum_{i=0}^{2^{m-1}} A_i x^i = 2^{-m}((1+x)^{2^{m-1}} + (1-x)^{2^{m-1}} + 2(2^{m-1}-1)(1-x^2)^{2^{m-2}}),$$

(see [10]).
For the codes $D_j$, we choose even weight codes having Hamming distance 2 and length $n-w_j$. $|D_j|=2^{n-w_j-1}$, $w_j \neq n$. The code $D$ is a code having length $n=2^{m-1}$ and $L$-distance 4. The cardinality of $D$ is:

$$|D| = 1/2+2^{-(m+1)}(3^{2^{m-1}}+2(2^{m-1}-1)3^{2^{m-2}}+1).$$

Example 2. $d=3, n=2,\ldots,16$.

$n=2$: $D'=\{0\}$,

| wt | 0 |
|---|---|
| # | 1 |

$D_0=\{0,1\}, D=\{(00),(22)\}$,
$|D|=1*2=2$.

$n=3$: $D'=\{0,1\}$,

| wt | 0 | 3 |
|---|---|---|
| # | 1 | 1 |

$D_0=<(110),(101)>$,
$D_3=\phi$.
$D=\{(000),(220),(202),(022),(111)\}$
$|D|=1*|D_0|+1=1*4+1=5$.

$n=4$: $D'=\{0,1\}$,

| wt | 0 | 4 |
|---|---|---|
| # | 1 | 1 |

$D_0=$ even weight code of length 4,
$D_4=\phi$,
$|D|=1*2^3+1=9$.

$n=5$: $D'=<(11100),(10011)>$

| wt | 0 | 3 | 4 |
|---|---|---|---|
| # | 1 | 2 | 1 |

$D_0=$ even weight code of length 5; $|D_0|=2^4=16$,
$D_3=$ even weight code of length 2 ; $|D_3|=2$,
$D_4=\phi$,
$|D|=1*16+2*2+1=21$.

$n=6$: $D'=$ Hamming code

| *wt* | 0 | 3 | 4 |
|---|---|---|---|
| # | 1 | 4 | 3 |

$D_j$ = even weight code of length $6-j, j=0,3,4$,
$|D|=1*32+4*4+3*2=54$.

$n=7$: $D'$ = Hamming code

| *wt* | 0 | 3 | 4 | 7 |
|---|---|---|---|---|
| # | 1 | 7 | 7 | 1 |

$D_j$ = even weight code of length $7-j, j=0,3,4,7$,
$|D|=1*64+7*8+7*4+1*1=149$

$n=8$: $D'$ = the (8,20,3)-code of [12], p. 72.

| *wt* | 0 | 3 | 4 | 8 |
|---|---|---|---|---|
| # | 1 | 8 | 10 | 1 |

$D_j$ = even weight code of length $8-j, j=0,3,4,8$,
$|D|=1*128+8*16+10*8+1=337$.

$n=9$: $D'$ = the Best code [13],

| *wt* | 0 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| # | 1 | 8 | 14 | 8 | 4 | 4 | 1 |

$D_j$ = even weight code of length $9-j, j=0,3,4,5,6,7,8$,
$|D|=825$.

$n=10$: $D'$ = <(1110000000),(1001100000),(1000011000),(1000000110),
(0101010000),(0100000101)>

| *wt* | 0 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| # | 1 | 10 | 16 | 12 | 12 | 10 | 3 |

$D_j$ = even weight code of length $10-j, j=0,3,4,5,6,7,8$,
$|D|=1998$

$n=11$: $D'$ = <(11100000000),(10011000000),(10000110000),(10000001100),
(10000000011),(01010100000),(01000000101)>

| *wt* | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 11 |
|---|---|---|---|---|---|---|---|---|
| # | 1 | 13 | 26 | 24 | 24 | 26 | 13 | 1 |

$D_j$ = even weight code of length $11-j, j=0,3,4,5,6,7,8,11$
$|D|=5765$ .

$n=12$: $D'$ = <(111000000000),(100110000000),(100001100000),
(100000011000), (100000000110),(010101000000),
(010000001010),(000001000101)>,

| *wt* | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 19 | 36 | 46 | 50 | 52 | 35 | 10 | 6 | 1 |

$D_j$ = even weight code of length $12-j, j=0,3,4,5,6,7,8,9,10,11$,
$|D|=17229$.

$n=13$: $D'$ = <(1110000000000),(1001100000000),(1000011000000),
(1000000110000),(1000000001100),(1000000000011),
(0101010000000),(0100000101000),(0101000100010)>,

| wt | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 22 | 55 | 72 | 96 | 96 | 87 | 60 | 16 | 6 | 1 | |

$D_j$ = even weight code of length $13-j, j = 0,3,4,5,6,7,8,9,10,11,12$,
$|D| = 49821$.

$n = 14$: $D' = <(11100000000000),(10011000000000),(10000110000000),(10000001100000),(10000000011000),(10000000000110),(01010100000000),(01000001010000),(01010001000100),(01010101010101)>$,

| wt | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 28 | 77 | 112 | 168 | 212 | 203 | 132 | 56 | 28 | 7 |

$D_j$ = even weight code of length $14-j, j = 0,3,4,5,6,7,8,9,10,11,12$,
$|D| = 149214$.

$n = 15$: $D'$ = Hamming code

| wt | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 35 | 105 | 168 | 280 | 435 | 435 | 280 | 168 | 105 | 35 | 1 |

$D_j$ = even weight code of length $15-j$, $0,3,4,5,6,7,8,9,10,11,12,15$,
$|D| = 449429$.

$n = 16$: $D'$ = Code constructed in [12] p. 76-77.

| wt | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 32 | 108 | 160 | 320 | 320 | 678 | 320 | 320 | 160 | 108 | 32 | 1 |

$D_j$ = even weight code of length $16-j$,
$j = 0,3,4,5,6,7,8,9,10,11,12,13,16$,
$|D| = 915681$.

EXAMPLE 3. $d = 4; n = 2,3,\ldots,17$.

$n = 2,3$: $D' = \{0\}$. $D_0$ = even weight code of length 2, resp. 3, $|D| = 2$ resp. 4,

$n = 4$: $D' = <1>$, $D_j$ = even weight code of length $4-j$, $j = 0,4$. $|D| = 9$.

$n = 5$: $D' = <(11110)>$, $D_j$ = even weight code of length $5-j$ $j = 0,4$. $|D| = 17$.

$n = 6$: $D'$ = Extended length 5, distance 3 code of Example 2.
$D_j$ = even weight code of length $6-j, j = 0,4$.
$|D| = 38$.

$n = 7$: $D'$ = Extended length 6, distance 3 code of Example 2.
$D_j$ = even weight code of length $7-j, j = 0,4$.
$|D| = 92$.

$n = 8$: $D'$ = Extended length 7, distance 3 code of Example 2.
$D_j$ = even weight code of length $8-j, j = 0,4,8$.
$|D| = 241$.

$n = 9$: $D'$ = Extended length 8, distance 3 code of Example 2.

$D_j$ = even weight code of length $9-j, j=0,4,8$.
$|D| = 545$.

$n=10$: $D'$ = the union of the constant weight codes with $w=0,4,8$. and $d=4$. (see [15]).

| *wt* | 0 | 4 | 8 |
|---|---|---|---|
| # | 1 | 30 | 5 |

$D_j$ = even weight code of length $10-j, j=0,4,8$.
$|D| = 1482$.

$n=11$: $D'$ = the union of the constant weight codes with $w=0,4,8$ and distance 4, (see [15]),

| *wt* | 0 | 4 | 8 |
|---|---|---|---|
| # | 1 | 35 | 17 |

$D_j$ = even weight code of length $11-j, j=0,4,8$.
$|D| = 3332$.

$n=12$: $D'$ = the union of the constant weight codes with $w=0,4,8,12$ and distance 4, (see [15]),

| *wt* | 0 | 4 | 8 | 12 |
|---|---|---|---|---|
| # | 1 | 51 | 51 | 1 |

$D_j$ = even weight code of length $12-j, j=0,4,8,12$.
$|D| = 8985$.

$n=13,14,15,16,17$:
$D'$ = Extended length $n-1$, distance 3 code of Example 2.
$D_j$ = even weight code of length $n-j, j$ ranges over the weights in $D'$
$n=13$, $|D|=25777$.
$n=14$, $|D|=75598$.
$n=15$, $|D|=215052$.
$n=16$, $|D|=621281$.
$n=17$, $|D|=1431361$.

EXAMPLE 4. $d=5, n=5,6,\ldots,17$.

$n=5$: $D' = \langle 1 \rangle, D_0$ = the well known [5,2,3]-code. $|D|=5$.

$n=6$: $D' = \langle (111110) \rangle$, $D_0$ = Hamming code of length 6, distance 3, $|D_0|=8$
$|D|=9$.

$n=7$: $D' = \langle (1111100) \rangle$, $D_0$ = Hamming code of length 7, distance 3, $|D_0|=16$
$|D|=17$.

$n=8$: $D' = \langle (11111000),(11000111) \rangle$,

| *wt* | 0 | 5 | 6 |
|---|---|---|---|
| # | 1 | 2 | 1 |

$D_0$ = the (8,20,3) code of [12] p.72.

$D_5 =$ the (3,2,3) code $<1>$.
$D_6 = <(00)>$.
$|D| = 25$.

$n = 9$: $D' =$ a nonlinear Hadamard code consisting of the following words: {(000000000),(111011010),(101110001),(100101110), (010010111),(011101101)}.

| *wt* | 0 | 5 | 6 |
|---|---|---|---|
| # | 1 | 3 | 2 |

$D_0 =$ the (9,40,3)-Best code [13].
$D_5 =$ the (4,2,4)-code $<(1111)>$.
$D_6 =$ the (3,2,3)-code $<(111)>$.
$|D| = 50$.

$n = 10$: $D' =$ a nonlinear Hadamard code consisting of the following words : {(0000000000),(1101000111),(1110110100),(0111011010),(0011101101), (1000011101),(1011100010),(0101110001),(1001011100),(0100101110), (0010010111),(1110001001)}.

| *wt* | 0 | 5 | 6 |
|---|---|---|---|
| # | 1 | 6 | 5 |

$D_0 =$ the (10,72,3)-code of [12] p. 71.
$D_5 =$ a (5,4,3)-code.
$D_6 =$ a (4,2,3)-code.
$|D| = 106$

$n = 11$: $D' =$ a nonlinear (11,24,5) Hadamard code (see [12], p 49).

| *wt* | 0 | 5 | 6 | 11 |
|---|---|---|---|---|
| # | 1 | 11 | 11 | 1 |

$D_0 =$ the (11,144,3)-code of [12], p.71.
$D_5 =$ the (6,8,3) Hamming code.
$D_6 =$ a (5,4,3)-code.
$|D| = 277$.

$n = 12$: $D' =$ the nonlinear (12,32,5)-code of [12], p.75, Problem 16.

| *wt* | 0 | 5 | 6 | 8 | 9 |
|---|---|---|---|---|---|
| # | 1 | 12 | 12 | 3 | 4 |

$D_0 =$ the (12,256,3) linear Hamming code.
$D_5 =$ the (7,16,3) linear Hamming code.
$D_6 =$ the (6,8,3) linear Hamming code.
$D_8 =$ a (4,2,3) linear code.
$D_9 =$ a (3,2,3) linear code.
$|D| = 558$.

$n = 13$: $D' =$ the shortened Nordstrom-Robinson code (13,64,5), ([12], p .75).

| *wt* | 0 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|
| # | 1 | 18 | 24 | 4 | 3 | 10 | 4 |

$D_0 =$ the (13,512,3)-Hamming code.

$D_5 =$ the (8,20,3)-code of [12], p.72.
$D_6 =$ the (7,16,3 )-linear Hamming code.
$D_7 =$ the (6,8,3)-Hamming code.
$D_8 =$ a (5,4,3) linear code.
$D_9 =$ a (4,2,3) linear code.
$D_{10} =$ a (3,2,3) linear code.
$|D| =$ 1328.

$n = 14$: $D' =$ the shortened Nordstrom-Robinson code (14,128,5), ([12], p.75.).

| *wt* | 0 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|
| # | 1 | 28 | 42 | 8 | 7 | 34 | 8 |

$D_0 =$ the (14,1024,3)-Hamming code.
$D_5 =$ the (9,40,3)-Best code.
$D_6 =$ the (8,20,3)-code of ([12],p.72.).
$D_7 =$ the (7,16,3)-Hamming code.
$D_8 =$ a (6,8,3)-Hamming code.
$D_9 =$ the (5,4,3)-Hamming code.
$D_{10} =$ a (4,2,3) linear code.
$|D| =$ 3320.

$n = 15$: $D' =$ the shortened Nordstrom-Robinson code.

| *wt* | 0 | 5 | 6 | 7 | 8 | 9 | 10 | 15 |
|---|---|---|---|---|---|---|---|---|
| # | 1 | 42 | 70 | 15 | 15 | 70 | 42 | 1 |

$D_0 =$ the (515,2048,3)-Hamming code.
$D_5 =$ the (10,72,3)-code of ([12], p.71.).
$D_6 =$ the (9,40,3)-Best code.
$D_7 =$ the (8,20,3)-code of [12], p.71.
$D_8 =$ a (7,16,3)-Hamming code.
$D_9 =$ the (6,8,3)-Hamming code.
$D_{10} =$ a (5,4,3) linear code.
$|D| =$ 9141.

$n = 16$: $D' =$ the shortened Nordstrom-Robinson code with all codewords extended by a 0.
$D_0 =$ the (16,2560,3)-code of [12], p.77.
$D_5 =$ the (11,144,3)-code of [12], p.71.
$D_6 =$ the (10,72,3)-code of [12], p.71.
$D_7 =$ the (9,40,3)-Best code.
$D_8 =$ a (8,20,3)-code of [12], p.71.
$D_9 =$ the (7,16,3)-Hamming code.
$D_{10} =$ a (6,8,3)-Hamming code.
$|D| =$ 16005.

$n = 17$: $D' =$ a [17,9,5]-BCH code.

According to [14] the even weight subcode has weight distribution:

| wt | 0 | 6 | 8 | 10 | 12 |
|---|---|---|---|---|---|
| # | 1 | 68 | 85 | 68 | 34 |

Since 1 is in the code and has odd weight, the weight distribution of this code is:

| wt | 0 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|
| # | 1 | 34 | 68 | 68 | 85 | 85 | 68 | 68 | 34 | 1 |

$D_0$ = the (17,5120,3) code of [12], p.77.
$D_5$ = the (12,256,3) Hamming code.
$D_6$ = the (11,144,3) code of [12], p.71.
$D_7$ = the (10,72,3) code of [12], p.71.
$D_8$ = the (9,40,3) Best code.
$D_9$ = the (8,20,3) code of [12], p.71.
$D_{10}$ = the (7,16,3) Hamming code.
$D_{11}$ = the (6,8,3) Hamming code.
$D_{12}$ = a (5,4,3) linear code.
$|D| = 35381$.

EXAMPLE 5. $d=6$, $n=5,\ldots,18$.

$n=5$: $D' = 0, D_0$ = a (5,4,3) code, $|D|=4$.

$n=6$: $D' = \langle 1 \rangle, D_0$ = the (6,8,3) Hamming code, $|D|=9$.

$n=7$: $D'$ = Extended (6,2,5) Hamming code of Example 4.
$D_0$ = the (7,16,3) Hamming code.
$|D| = 17$.

$n=8$: $D'$ = Extended (7,2,5) code of Example 4.
$D_0$ = the (8,20,3) code of [12], p.71.
$|D| = 21$.

$n=9$: $D'$ = Extended (8,4,5) code of Example 4.
$D_0$ = the (9,40,3) Best code.
$D_6$ = the (3,2,3) repitition code.
$|D| = 46$.

$n=10$: $D'$ = Extended (9,6,5) code of Example 4.
$D_0$ = the (10,72,3)-code of [12], p.71.
$D_6$ = the (4,2,3) code.
$|D| = 82$.

$n=11$: $D'$ = Extended (10,11,5) code of Example 4.
$D_0$ = the (11,144,3)-code of [12], p.71.
$D_6$ = the (5,4,3) code.
$|D| = 188$.

$n=12$: $D'$ = Extended (11,24,5) code of Example 4.
$D_0$ = the (12,256,3) Hamming code.
$D_6$ = the (6,8,3) Hamming code.
$|D| = 433$.

$n=13$: $D'$ = the union of the constant weight codes of length 13, distance 6, and $w = 0, 6, 12$. (See [12] , p. 686.).

| *wt* | 0 | 6 | 12 |
|---|---|---|---|
| # | 1 | 26 | 1 |

$D_0$ = the (13,512,3) Hamming code.
$D_6$ = the (7,16,3) Hamming code.
$|D|$ = 929.

$n = 14$: $D'$ = Extended (13, 64,5) code of Example 4.
$D_0$ = the (14,1024,3) Hamming code.
$D_6$ = the (8,20,3) code of [12]. p.71.
$D_8$ = the (6,8,3) Hamming code.
$D_{10}$ = a (4,2,3) code.
$|D|$ = 1948.

$n = 15$: $D'$ = Extended (14, 128,5) code of Example 4.
$D_0$ = the (15,2048,3) Hamming code.
$D_6$ = the (9,40,3) Best code.
$D_8$ = the (7,16,3) Hamming code.
$D_{10}$ = a (5,4,3) code.
$|D|$ = 5256.

$n = 16$: $D'$ = the Nordstrom Robinson code.
$D_0$ = the (16,2561,3) code of [12], p.77.
$D_6$ = the (10,72,3) code of [12], p.71.
$D_8$ = the (8,20,3) code of [12], p.71.
$D_{10}$ = the (6,8,3) Hamming code.
$|D|$ = 12121.

$n = 17$: $D'$ = Extended (16,256,5) code of Example 4.
$D_0$ = the (17,5120,3) code of [12], p.77.
$D_6$ = the (11,144,3) code of [12], p.71.
$D_8$ = the (9,40,3) Best code.
$D_{10}$ = the (7,16,3) Hamming code.
$|D|$ = 24241.

$n = 18$: $D'$ = Extended (17,256,5) code of Example 4.
$D_0$ = the (18,9728,3) code of [12], p.77.
$D_6$ = the (12,256,3) Hamming code.
$D_8$ = the (10,72,3) code of [12], p.71.
$D_{10}$ = the (8,20,3) code.
$D_{12}$ = the (6,8,3) Hamming code.
$|D|$ = 50733.

EXAMPLE 6. $d = 7, n = 6, \ldots, 18$.
$n = 6$: $D' = \{0\}, D_0$ = a (6,4,4) code, $|D| = 4$.
$n = 7,8,9,10$: $D' = <(11111110\ldots0)>$.
$n = 7$: $D_0$ = a (7,8,4) Hamming code. $|D| = 9$.
$n = 8$: $D_0$ = a (8,16,4) Hamming code. $|D| = 17$.
$n = 9$: $D_0$ = the (9,20,4) code obtained by extending the (8,20,3) code.
$|D| = 21$.
$n = 10$: $D_0$ = the (10,40,4) Best code. $|D| = 41$.

$n = 11,12$: $D' = <(111111100000),(111000011110)>$.

| *wt* | 0 | 7 | 8 |
|---|---|---|---|
| # | 1 | 2 | 1 |

$n = 11$: $D_0 =$ a (11,72,4) code (extending the (10, 72,3) code).
$D_7 =$ the (4,2,4) repitition code.
$|D| = 77$.

$n = 12$: $D_0 =$ the (12,144,4) code (see [12]).
$D_7 =$ a (5,2,4) code.
$D_8 =$ a (4,2,4) code.
$|D| = 150$.

$n = 13$: $D' = <(1111111000000),(1110000111100),(1001100110011)>$,

| *wt* | 0 | 7 | 8 |
|---|---|---|---|
| # | 1 | 4 | 3 |

$D_0 =$ a (13,256,4) code.
$D_7 =$ a (6,4,4) code.
$D_8 =$ a (5,2,4) code.
$|D| = 278$.

$n = 14$: $D' =$ a shortened first order Reed-Muller code.

| *wt* | 0 | 7 | 8 |
|---|---|---|---|
| # | 1 | 8 | 7 |

$D_0 =$ a (14,512,4) code.
$D_7 =$ a (7,8,4) code.
$D_8 =$ a (6,4,4) code.
$|D| = 604$.

$n = 15$: $D'$ a (15,32,7) Reed-Muller code.

| *wt* | 0 | 7 | 8 | 15 |
|---|---|---|---|---|
| # | 1 | 15 | 15 | 1 |

$D_0 =$ a (15,1024,4) code.
$D_7 =$ a (8,16,4) code.
$D_8 =$ a (7,8,4) code.
$|D| = 1385$.

$n = 16$: $D' =$ the previous code extended by a 0. The weight distribution remains the same.
$D_0 =$ a (16,2048,4) code.
$D_7 =$ a (9,20,4) code.
$D_8 =$ a (8,16,4) code.
$|D| = 2589$.

$n = 17$: $D' =$ a (17,64,7) shortened Golay code.

| *wt* | 0 | 7 | 8 | 11 | 12 |
|---|---|---|---|---|---|
| # | 1 | 20 | 25 | 12 | 6 |

$D_0 =$ a (17,2560,4) code.
$D_7 =$ a (10,40,4) code.

$D_8 =$ a (9,20,4) code.
$D_{11} =$ a (6,4,4) code.
$D_{12} =$ a (5,2,4) code.
$|D| = 3920$.

$n = 18$: $D' =$ a (18,128,7) shortened Golay code.

| *wt* | 0 | 7 | 8 | 11 | 12 | 15 |
|---|---|---|---|---|---|---|
| # | 1 | 32 | 46 | 30 | 18 | 1 |

$D_0 =$ a (18,5120,4) code.
$D_7 =$ a (11,72,4) code.
$D_8 =$ a (10,40,4) code.
$D_{11} =$ a (7,8,4) code.
$D_{12} =$ a (6,4,4) code.
$|D| = 9577$.

EXAMPLE 7. $d = 8, n = 7,\ldots,19$.
For all $n \neq 17$, we take $D'$ to be the extensions of the codes in Example 6.
$n = 7$: $D_0 =$ a (7,8,4) code. $|D| = 8$.
$n = 8$: $D_0 =$ a (8,16,4) code. $|D| = 17$.
$n = 9$: $D_0 =$ a (9,20,4) code. $|D| = 21$.
$n = 10$: $D_0 =$ a (10,40,4) code. $|D| = 41$.
$n = 11$ $D_0 =$ a (11,72,4) code. $|D| = 73$.
$n = 12$: $D_0 =$ a (12, 144,4) code. $D_8 =$ a (4,2,4) code. $|D| = 150$.
$n = 13$: $D_0 =$ a (13,256, 4) code. $D_8 =$ a (5,2,4) code. $|D| = 262$.
$n = 14$: $D_0 =$ a (14,512,4) code . $D_8 =$ a (6,4,4) code. $|D| = 540$.
$n = 15$: $D_0 =$ a (15,1024,4) code. $D_8 =$ a (7,8,4) code. $|D| = 1144$ .
$n = 16$: $D_0 =$ a (16,2048,4) code. $D_8 =$ a (8,16,4) code. $|D| = 2529$.
$n = 18$: $D_0 =$ a (18,5120,4) code. $D_8 =$ a (10,40,4) code. $D_{12} =$ a (6,4,4) code $|D| = 6992$.
$n = 19$: $D_0 =$ a (19,9728,4) code. $D_8 =$ a (11,72,4) code. $D_{12} =$ a (7,8,4) code.
$|D| = 15729$.
$n = 17$: $D' =$ the union of the constant weigth codes of length 17, distance 8, and weights $w = 0, 8, 16$.

| *wt* | 0 | 8 | 16 |
|---|---|---|---|
| # | 1 | 34 | 1 |

$D_0 =$ a (17,2560,4) code . $D_8 =$ a (9,20,4) code. $|D| = 3241$.

EXAMPLE 8. $d = 9, n = 8,\ldots,20$.
$n = 8$: $D' = 0, D_0$ is a (8,4,5)-code (see [12]), $|D| = 4$.
$n = 9,\ldots,13$: $D' = <(1111111110\ldots0)>$,
$n = 9$: $D_0 =$ a (9,6,5) code. $|D| = 7$.
$n = 10$: $D_0 =$ a (10,12,5) code. $|D| = 13$.
$n = 10$: $D_0 =$ a (11,24,5) code. $|D| = 25$.
$n = 12$: $D_0 =$ a (12,32,5) code. $|D| = 33$.

$n=13$: $D_0=$ a (13,64,5) code. $|D|=65$.
$n=14,15$: $D'=$ <(1111111110000000),(0000011111111110)>,

| wt | 0 | 9 | 10 |
|---|---|---|---|
| # | 1 | 2 | 1 |

$n=14$: $D_0=$ a (14,128, 5) code , $D_9=$ a (5,2,5) code, $|D|=133$.
$n=15$: $D_0=$ a (15,256,5) code , $D_9=$ a (6,2,5) code, $D_{10}=$ a (5,2,5) code. $|D|=262$.
$n=16$: $D'=$ {(0000000000000000),(1011100011001011),(0110111001100101), (1011011101110010),(0101101110111001),(1100010111011100)}.

| wt | 0 | 9 | 10 |
|---|---|---|---|
| # | 1 | 3 | 2 |

$D_0=$ a (16,256,5 ) code, $D_9=$ a (7,2,5) code, $D_{10}=$ a (6,2,5) code. $|D|=266$.
$n=17$: $D'$ is the code constructed with the Hadamard-Levenstein construction (see [12] Ch.2, section 3.)

| wt | 0 | 9 | 10 |
|---|---|---|---|
| # | 1 | 5 | 4 |

$D_0=$ a (17,512,5) code, $D_9=$ a (8,4,5) code, $D_{10}=$ a (7,2,5) code. $|D|=540$.
$n=18$: $D'$ is a (18,20,9) Hadamard code.

| wt | 0 | 9 | 10 |
|---|---|---|---|
| # | 1 | 10 | 9 |

$D_0=$ a (18,1024,5) code, $D_9=$ a (9,6,5) codes, $D_{10}=$ a (8,4,5) code. $|D|=1120$.
$n=19$: $D'=$ is a (19,40,9) Hadamard code.

| wt | 0 | 9 | 10 | 19 |
|---|---|---|---|---|
| # | 1 | 19 | 19 | 1 |

$D_0=$ a (19,2048,5) code, $D_9=$ a (10,12,5) code, $D_{10}=$ a (9,6,5) code. $|D|=2391$.
$n=20$: $D'=$ the extension of the previous $D'$ by a 0.
$D_0=$ a (20,2560,5) code, $D_9=$ a (11,24,5) code, $D_{10}=$ a (10,12,5) code.
$|D|=3105$.

EXAMPLE 9. $d=10$, $n=9,\ldots,20$.
For all $n$ we take $D'$ to be the extensions of the $(n-1,5)$ codes of Example 8.

$n=9$: $D_0=$ a (9,6,59) code, $|D|=6$.
$n=10$: $D_0=$ a (10,12,5) code, $|D|=13$.
$n=11$: $D_0=$ a (11,24,5) code, $|D|=25$.
$n=12$: $D_0=$ a (12,32,5) code, $|D|=33$.
$n=13$: $D_0=$ a (13,64,5) code, $|D|=65$.

$n=14$: $D_0=$ a (14,128,5) code, $|D|=129$.
$n=15$: $D_0=$ a (15,256,5) code , $D_{10}=$ a (5,2,5) code, $|D|=262$.
$n=16$: $D_0=$ a (16,256,5) code, $D_{10}=$ a (6.2,5) code, $|D|=262$.
$n=17$: $D_0=$ a (17,512,5) code, $D_{10}=$ a (7,2,5) code, $|D|=532$.
$n=18$: $D_0=$ a (18,1024,5) code , $D_{10}=$ a (8,4,5) code, $|D|=1060$.
$n=19$: $D_0=$ a (19,2048,5) code , $D_{10}=$ a (9,6,5) code, $|D|=2162$.
$n=20$: $D_0=$ a (20,2560,5) code, $D_{10}=$ a (10,12,5) code, $|D|=3071$.

For the values $d=11,12,13,14, n=10,\ldots,20$: All codes to be chosen are trivial and we only list the results in the final table.

TABLE OF THE CONSTRUCTED CODES IN $L$-METRIC

| $n$ \ $d$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 2 | - | - | - | - | - | - | - | - | - | - |
| 3 | 5 | 4 | 2 | 2 | - | - | - | - | - | - | - | - |
| 4 | 9 | 9 | 2 | 2 | 2 | 2 | - | - | - | - | - | - |
| 5 | 21 | 17 | 5 | 4 | 2 | 2 | 2 | 2 | - | - | - | - |
| 6 | 54 | 38 | 9 | 9 | 4 | 4 | 2 | 2 | 2 | 2 | - | - |
| 7 | 149 | 92 | 17 | 17 | 9 | 8 | 2 | 2 | 2 | 2 | 2 | 2 |
| 8 | 337 | 241 | 25 | 21 | 17 | 17 | 4 | 4 | 2 | 2 | 2 | 2 |
| 9 | 825 | 545 | 50 | 46 | 21 | 21 | 7 | 6 | 4 | 4 | 2 | 2 |
| 10 | 1998 | 1482 | 106 | 82 | 41 | 41 | 13 | 13 | 6 | 6 | 2 | 2 |
| 11 | 5765 | 3332 | 277 | 188 | 77 | 73 | 25 | 25 | 13 | 12 | 4 | 4 |
| 12 | 17229 | 8985 | 558 | 433 | 150 | 150 | 33 | 33 | 25 | 25 | 4 | 4 |
| 13 | 49821 | 25777 | 1328 | 929 | 278 | 262 | 65 | 65 | 33 | 33 | 9 | 8 |
| 14 | 149241 | 75598 | 3320 | 1948 | 604 | 540 | 133 | 129 | 65 | 65 | 17 | 17 |
| 15 | 449429 | 215052 | 9141 | 5256 | 1385 | 1144 | 262 | 262 | 129 | 129 | 33 | 33 |
| 16 | 915681 | 621281 | 16005 | 12121 | 2589 | 2529 | 266 | 262 | 257 | 257 | 37 | 37 |
| 17 | ... | 1431361 | 35381 | 24241 | 3920 | 3241 | 540 | 532 | 261 | 257 | 65 | 65 |
| 18 | ... | ... | ... | 50733 | 9577 | 6992 | 1120 | 1060 | 518 | 518 | 129 | 129 |
| 19 | ... | ... | ... | ... | ... | 15729 | 2391 | 2162 | 1034 | 1030 | 257 | 257 |
| 20 | ... | ... | ... | ... | ... | ... | 3105 | 3017 | 2070 | 2058 | 517 | 513 |

APPENDIX 4. TABLE OF NUMERICAL VALUES FOR THE UPPER BOUND OF THEOREM 9 (SPHERE PACKING BOUND)

| $n$ \ $d$ | 3,4 | 5,6 | 7,8 | 9,10 | 11,12 | 13,14 | 15,16 |
|---|---|---|---|---|---|---|---|
| 2 | 3 | 1 | - | - | - | - | - |
| 3 | 6 | 2 | 1 | - | - | - | - |
| 4 | 16 | 5 | 2 | 1 | - | - | - |
| 5 | 40 | 11 | 4 | 2 | 1 | - | - |
| 6 | 104 | 26 | 9 | 4 | 2 | - | - |
| 7 | 273 | 60 | 19 | 7 | 4 | 2 | 1 |
| 8 | 729 | 145 | 41 | 15 | 7 | 3 | 2 |
| 9 | 1968 | 357 | 93 | 31 | 13 | 6 | 3 |
| 10 | 5368 | 894 | 213 | 66 | 25 | 11 | 5 |
| 11 | 14762 | 2271 | 501 | 143 | 50 | 21 | 10 |
| 12 | 40880 | 5840 | 1199 | 319 | 104 | 40 | 18 |
| 13 | 113880 | 15184 | 2914 | 725 | 221 | 79 | 33 |
| 14 | 318864 | 39858 | 7181 | 1678 | 480 | 163 | 63 |
| 15 | 896806 | 105506 | 17913 | 3946 | 1065 | 340 | 125 |
| 16 | 2532160 | 281351 | 45169 | 9413 | 2404 | 726 | 253 |
| 17 | 7174453 | 755205 | 114995 | 22739 | 5512 | 1580 | 523 |
| 18 | 20390552 | 2039055 | 295290 | 55560 | 12816 | 3496 | 1100 |
| 19 | 58113073 | 5534578 | 764142 | 137139 | 30177 | 7852 | 2357 |
| 20 | 166037352 | 15094304 | 1991310 | 341640 | 71877 | 17880 | 5131 |

A NUMBER OF ENTRIES IN THIS TABLE CAN BE IMPROVED BY APPLYING THEOREM 11. (IMPROVED SPHERE PACKING BOUND)

| $d$ | $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3,4 | | 6 | 16 | 39 | 99 | 257 | 681 | 1835 | 5006 | 13750 | 37613 | 103997 | 290025 | 814455 | 2300109 |
| 5,6 | | - | - | - | - | - | - | - | - | 2245 | 5651 | 14466 | 37563 | 98717 | 262081 |

APPENDIX 5. TABLE OF NUMERICAL VALUES FOR THE UPPER BOUND OF THEOREM 15 (L.P. BOUND)

| $n$ | $d$ 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 2 | - | - | - | - | - | - | |
| 3 | 5 | 5 | 4 | 2 | - | - | - | - | - |
| 4 | 13 | 12 | 9 | 5 | 4 | 2 | - | - | - |
| 5 | 32 | 27 | 18 | 10 | 9 | 4 | 4 | 2 | - |
| 6 | 82 | 66 | 39 | 22 | 17 | 9 | 7 | 4 | 4 |
| 7 | 213 | 170 | 85 | 48 | 37 | 17 | 14 | 7 | 7 |
| 8 | 572 | 441 | 189 | 103 | 81 | 37 | 28 | 14 | 12 |
| 9 | 1538 | 1166 | 427 | 237 | 179 | 81 | 57 | 28 | 24 |
| 10 | 4176 | 3128 | 974 | 570 | 396 | 179 | 124 | 57 | 46 |
| 11 | 11445 | 8483 | 2301 | 1407 | 898 | 396 | 278 | 124 | 92 |
| 12 | 31618 | 23216 | 5623 | 3475 | 2020 | 898 | 600 | 278 | 191 |
| 13 | 87872 | 64023 | 14135 | 8826 | 4632 | 2020 | 1263 | 600 | 414 |
| 14 | 245510 | 177766 | 36128 | 22716 | 11345 | 4632 | 2937 | 1263 | 937 |
| 15 | 689383 | 496014 | 93678 | 58901 | 27870 | 11345 | 7091 | 2937 | 1976 |
| 16 | 1943533 | 1391292 | 245439 | 151978 | 68129 | 27870 | 16218 | 7091 | 4435 |
| 17 | 5499245 | 3919400 | 649667 | 404991 | 161001 | 68129 | 38261 | 16218 | 9988 |
| 18 | 15610619 | 11083899 | 1733402 | 1099016 | 386395 | 171454 | 89994 | 38261 | 22145 |
| 19 | 44442699 | 31442989 | 4658836 | 2879528 | 943854 | 417663 | 215604 | 89994 | 51328 |
| 20 | 126856499 | 89479129 | 12602189 | 7819634 | 2343686 | 1082561 | 889793 | 215604 | 614784 |

REMARK:
It appeared that the results for even distances were much stronger than the results obtained for odd distances (if $d$ is large). Therefore in making the table, the observation that $A_L(n,d) \leqslant A_L(n+1,d+1)$ if $d$ is odd could be succesfully applied as one can see at several places in the table.

APPENDIX 6. MISCELLANEOUS CONSTRUCTIONS
By a computer search, we obtained the following codes: (numbers are to be replaced by their ternary representation)

$d=3$; $n=4$: 0,5,15,19,26,61,56,63,77,78,40
$n=5$: 0,5,15,19,26,34,38,48,55,71,94,111,131,140,144,159,163,168,179, 180,214,223,227
$n=6$: 0,.....,227 (as above) 279,295,300,316,326,331,345,366,391,404 425,432,437,477,487,492,500,504,521,533,542,555,592,598,621, 642,651,656,658,691,698,708.

The following construction sometimes gives good results: Let $C$ be a $(n,M,d)$ code over $GF(5)$. Replace 0 by 00,1 by 02,2 by 20,3 by 11, and 4 by 22. Then we get an $(2n,M,2d)$ code in the $L$-metric over $\{0,1,2\}$.

EXAMPLE: The (4,25,3) *MDS* code over $GF(5)$ gives a (8,25,6) code in the $L$-

metric over $\{0,1,2\}$. The $(5,125,3)$ *MDS* code over *GF* (5) gives a $(10,125,6)$ code in the $L$-metric over $\{0,1,2\}$. The ternary Hamming codes have parameters: length $n=(q^m-1)/(q-1)$, dimension $n-m$ and minimum distance 3. Therefore they have distance at least 3 in the $L$-metric.

EXAMPLE: The Hamming codes: [13,10,3], [12,9,3], [11,8,3], [10,7,3] (taking $m=3$ and shortening) give better results than the other constructions.

From the ternary Golay code, we can construct the following codes having Hamming distance 5 (6) and hence having $L$-distance at least 5 (6): (all examples give better results than the other constructions): [12,6,6], [11,5,6], [11,6,5], [10,5,5], [9,4,5], [8,3,5]. For $n=16$, we can make a BCH code having Hamming distance 5 by taking as zeroes $'\alpha^0,\alpha^1,\alpha^2$: the cyclotomic cosets are $\{0\},,\{1,3,9,11\}$ and $\{2,6\}$ and hence the dimension of this code is $16-7=9$. This code also has $L$-distance at least 5 and is better than the one obtained by previous constructions. It is easy to prove that the upper bound for $n=3$ and $d=5$ is 2 and for $n=3$ $d=6$ is 2.

## APPENDIX 7. BEST KNOWN UPPER AND LOWER BOUNDS

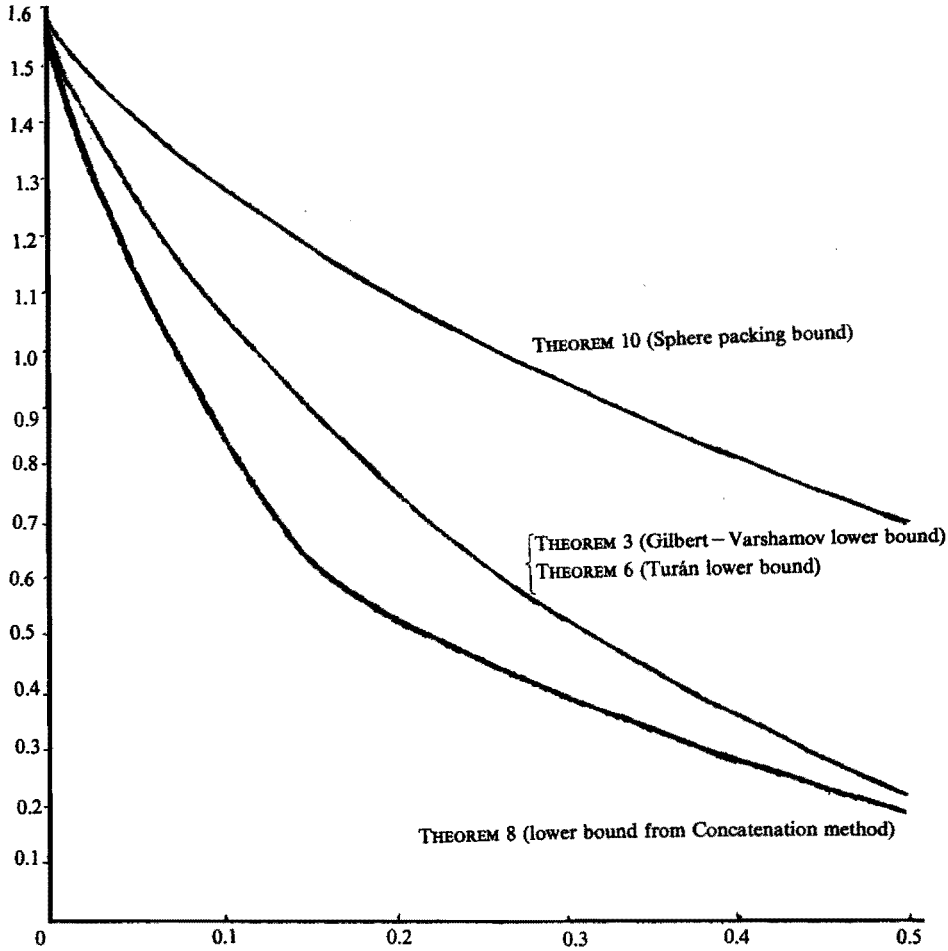
| $n$ | $d$ | 3 | | | 4 | | | 5 | | | 6 | | | 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | | | 5 | | | 4 | | | 2 | | | 2 | | | - | |
| 4 | | 11 | - | 13 | 9 | | - | 12 | 3 | | - | 5 | 2 | | - | 5 |
| 5 | | 23 | | - | 32 | 17 | | - | 27 | 5 | | - | 11 | 4 | | - |
| 6 | | 55 | | - | 82 | 38 | | - | 66 | 9 | | - | 26 | 9 | | - |
| 7 | | 149 | | - | 213 | 92 | | - | 170 | 17 | | - | 60 | 17 | | - |
| 8 | | 337 | | - | 572 | 241 | | - | 441 | 27 | | - | 145 | 25 | | - |
| 9 | | 825 | | - | 1538 | 545 | | - | 1166 | 81 | - | 357 | 46 | - | 237 | 21 |
| 10 | | 2787 | - | 4170 | 1482 | - | 3128 | 243 | - | 894 | 125 | - | 570 | 41 | - | 213 |
| 11 | | 6561 | - | 11445 | 3332 | - | 8483 | 729 | - | 2245 | 188 | - | 1407 | 77 | - | 501 |
| 12 | | 19683 | - | 31618 | 8985 | - | 23216 | 558 | - | 5651 | 729 | - | 3475 | 15 | - | 1199 |
| 13 | | 59049 | - | 87872 | 25777 | - | 64023 | 1328 | - | 14135 | 929 | - | 8826 | 278 | - | 2914 |
| 14 | | 149214 | - | 245510 | 75598 | - | 177766 | 3320 | - | 36128 | 1948 | - | 22716 | 604 | - | 7181 |
| 15 | | 449429 | - | 689383 | 215052 | - | 496014 | 9141 | - | 93678 | 5256 | - | 58901 | 1315 | - | 17913 |
| 16 | | 915681 | - | 1943533 | 621281 | - | 1391292 | 19683 | - | 649667 | 12121 | - | 151978 | 2589 | - | 45169 |
| 17 | | ..... | | | 1431361 | - | 3919400 | 35381 | - | 1733407 | 24241 | - | 404991 | 3920 | - | 114995 |
| 18 | | ..... | | | | ..... | | | ..... | | | ..... | | | ..... | |
| 19 | | ..... | | | | ..... | | | ..... | | | ..... | | | ..... | |
| 20 | | ..... | | | | ..... | | | ..... | | | ..... | | | ..... | |

## APPENDIX 8. ASYMPTOTIC BOUNDS

| | Lower bounds | | | Upper bound |
|---|---|---|---|---|
| $\delta$ | Theorem 3 | Theorem 6 | Theorem 8 | Theorem 10 |
| 0.00 | 1.58496 | 1.58496 | 1.58496 | 1.58496 |
| 0.01 | 1.49997 | 1.49996 | 1.46169 | 1.53951 |
| 0.02 | 1.43503 | 1.43500 | 1.36865 | 1.50402 |
| 0.03 | 1.37770 | 1.37762 | 1.28714 | 1.47228 |
| 0.04 | 1.32532 | 1.32519 | 1.21322 | 1.44295 |
| 0.05 | 1.27664 | 1.27644 | 1.14499 | 1.41542 |
| 0.06 | 1.23092 | 1.23063 | 1.08135 | 1.38930 |
| 0.07 | 1.18766 | 1.18728 | 1.02156 | 1.36437 |
| 0.08 | 1.14653 | 1.14604 | 0.96511 | 1.34044 |
| 0.09 | 1.10725 | 1.10664 | 0.91161 | 1.31739 |
| 0.10 | 1.06963 | 1.06890 | 0.86077 | 1.29511 |
| 0.11 | 1.03351 | 1.03264 | 0.81233 | 1.27353 |
| 0.12 | 0.99875 | 0.99774 | 0.76611 | 1.25258 |
| 0.13 | 0.96524 | 0.96408 | 0.72195 | 1.23221 |
| 0.14 | 0.93290 | 0.93158 | 0.67970 | 1.21237 |
| 0.15 | 0.90163 | 0.90016 | 0.63926 | 1.19302 |
| 0.16 | 0.87138 | 0.86974 | 0.60053 | 1.17413 |
| 0.17 | 0.84207 | 0.84028 | 0.58044* | 1.5568 |
| 0.18 | 0.81367 | 0.81170 | 0.56353* | 1.13762 |
| 0.19 | 0.78612 | 0.78398 | 0.54705* | 1.11995 |
| 0.20 | 0.75937 | 0.75706 | 0.53100* | 1.10263 |
| 0.21 | 0.73339 | 0.73092 | 0.51535* | 1.08566 |
| 0.22 | 0.70816 | 0.70551 | 0.50008* | 1.06901 |
| 0.23 | 0.68362 | 0.68080 | 0.48518* | 1.05267 |
| 0.24 | 0.65976 | 0.65678 | 0.47063* | 1.03663 |
| 0.25 | 0.63655 | 0.63340 | 0.45643* | 1.02087 |
| 0.30 | 0.52946 | 0.52555 | 0.39015* | 0.94591 |
| 0.35 | 0.43567 | 0.43114 | 0.33098* | 0.87670 |
| 0.40 | 0.35346 | 0.34851 | 0.27807* | 0.81232 |
| 0.45 | 0.28158 | 0.27645 | 0.23080* | 0.75219 |
| 0.50 | 0.21909 | 0.21403 | 0.18872* | 0.69584 |

For the entries marked with *, in the bound from Theorem 8, $1-H(\delta/2)$ appeared the largest bound of the two.

## APPENDIX 9. GRAPHS OF THE ASYMPTOTIC BOUNDS




DISCUSSION: In this picture the asymptotic upperbound of Theorem 10 is shown together with the asymptotic lower bounds of Theorem 3,6 and 8. Unfortunately there is a wide gap between upper and lower bounds.

APPENDIX 10. PROOF OF TURÁNS THEOREM

The following proof of Turáns was refound by me and H.A. Wilbrink and follows the proof given in a lecture by A. Schrijver.

THEOREM ( Turán). *Let $G=(V,E)$ be a graph, $V$ the set of points, $E$ the set of edges. If $l$ and $l'$ are the sizes of respectively the largest clique and the largest coclique then:*

$$l \geq \frac{|V|^2}{|V|^2-2|E|}$$

$$l' \geq \frac{|V|}{|V|^2+2|E|}$$

PROOF: The second statement is a direct consequence of the first one to the complementary graph). To prove the first statement, we will introduce a kind of weight function on the edges (depending on a weight function on the vertices). Then we are going to maximize the total weight of the edges by varying the weights on the vertices. It will be shown that without loss of generality the maximum is taken for a weight function which has its nonzero weigths on a clique. Moreover in Lemma 2 we will show that such a weight function (i.e. which maximizes the total weight) can be taken to be constant on this clique. Then it follows that this clique is a largest clique. By taking a trivial weight function we get a bound for the size at such a largest clique. Let $w(x)$ be a function on $V, w:V\rightarrow R_0^+$, $\Sigma_{x\in V} w(x)=1$. With this function on $V$, we associate a function $W$ on $E$ defined by: $W(e):=w(x)w(y)$ where $e=\{x,y\}\in E$. Since the set $\{w|w:V\rightarrow R_0^+, \Sigma_{x\in V} w(x)=1\}$ is compact, it follows that $\Sigma_{e\in E} W(e)$ attains a maximum for some function $w$. Let $w'$ be the function that maximizes this sum and define $S(w'):=\{x\in V|w'(x)\neq 0\}$.

LEMMA 1. *Without loss of generality, we may assume $S(w')$ to be a clique.*

PROOF: Let $x$ and $y$ be non-adjacent. Suppose $w(x)>0, w(y)>0$, w.l.o.g $\Sigma_{z\sim x} w'(z)\geq \Sigma_{z\sim y} w'(z)$. Define

$$w''(x):=w'(x)+w'(y),\ w''(y):=0,\ w''(z);=w'(z), z\neq x,y,$$

then $\Sigma_{e\in E} W''(e)\geq \Sigma_{e\in E} W'(e)$. This last inequality can easily be seen, if one writes out the summations replacing $W(e)$ by $w(x)w(z)$, $e=\{x,z\}$ and using the first inequality. Q.E.D.

LEMMA 2. *If $S(w')$ is a clique then w.l.o.g. $w'(x)=w'(y)$ for $x,y$ for $x,y\in S(w')$.*

PROOF: Let $x,y\in S(w')$ and $w'(x)\neq w'(y)$. Then w.l.o.g. $w'(y)>w'(x)$. Now define $w''(x):=\ w''(y):=(w'(x)+w'(y)))/2$, $w''(z):=\ w'(z), z\neq x,y$. Then $\Sigma_{e\in E} W''(e)\geq \Sigma_{e\in E} W'(e)$ since:

$$a\,b\leq a^2/4 + b^2/4 + a\,b/2. \qquad \text{Q.E.D.}$$

So we may assume $S(w')$ is a clique and $w$ is constant on this clique. But then obviously the maximum is equal to:

$$|(S(w')|(|S(w')|-1))/2\cdot 1/(|S(w')|^2 = 1/2\cdot(1-1/|S(w')|),$$

which is maximal whenever $S(w')$ is a largest clique. Hence the maximum is equal to: $1/2(1-1/l)$.
If we now take $(w(x)=1/|V|,\ (x\in V)$, then it follows that:

$$1/2(1-1/l)\geqslant |E|/|V|^2. \qquad \text{Q.E.D.}$$

2.7. Bounds for code pairs following from Sections 3,5 and 6

We will now give the upper and lower bounds for code pairs using the results of Sections 3,5 and 6. First we will give the asymptotical results, then in the appendices to this section one can find tables containing code pairs of length $\leqslant 20$ and $L$-distance $\leqslant 8$, which are obtained by the construction of Section 5 applied to the ternary codes of Section 6 and the binary codes in [10] and [12]. For the rest of this section, we will fix the following notation: $A(n,d)$ will have its usual meaning, i.e. the number of words in the best binary code of length $n$ and distance $d$. $A_L(n,d)$ will have the meaning we attached to it in the previous section: the number of words in the best ternary code of length $n$ and $L$-distance $d$.
$R(\delta)$ will be defined to be:

$$R(\delta) := \limsup_{\substack{n\to\infty \\ d/n\to\delta}} (\log A(n,d))/n,$$

$R_L(\delta)$ will be defined to be :

$$R_L(\delta) := \limsup_{\substack{n\to\infty \\ d/n\to\delta}} (\log A_L(n,d))/n,$$

For code pairs, we have the additional definitions: $M(n,d)$ will denote the maximal cardinality of the cartesian product of the code pairs of length $n$ and $L$-distance $d$. The rates of a code pair $(C_1,C_2)$ will be defined by:

$$R_i := (\log|C_i|)/n \qquad (i=1,2),$$

$$R_{\text{sum}} := R_1 + R_2 \quad \text{(As before : cf. Section 3.)}.$$

Now

$$R_{\text{sum}}(\delta) = \limsup_{\substack{n\to\infty \\ d/n\to\delta}} (\log M(n,d))/n$$

Asymptotic upper bounds for $R_{\text{sum}}(\delta)$:

1) From Section 3 we recall Van Tilborgs upper bound (cf. [6]):

$$R_{\text{sum}}(\delta) \leqslant 1/2 - \delta/2 + H(1/2\star(1-\delta)) - H(\delta)/2.$$

It is obvious that $R_{\text{sum}}(\delta)$ is less than or equal to twice the McEliece-

Rodemich-Rumsey-Welch upper bound (Cf. [7]) for binary codes.

2) If $(C_1,C_2)$ is a $d$-decodable code pair, then $C:=\{u+v|u\in C_1,v\in C_2\}$ is a ternary code of length $n$ and $L$-distance $d$. Therefore from Theorem 10, Sphere packing bound:

$$R_{\text{sum}}(\delta)\leqslant \log 3-h(x_0)-(1-x_0)h((\delta/4-x_0/2)/(1-x_0)),$$

where

$$x_0=-1/3+1/3\sqrt{(1+3\delta-3\delta^2/4)}.$$

The numerical results of these bounds can be found in Appendix 1.

Asymptotic lower bounds for $R_{\text{sum}}(\delta)$.

From the construction of Section 4, we get the following observation:

$$M(n,d)\geqslant A(n/2,d/2)A_L(n/2,d).$$

Now:

$$\log M(n,d)/n\geqslant \log A(n/2,d/2)/n+\log A_L(n/2,d)/n.$$

Therefore (taking the limsup's in the right order), we find:

$$R_{\text{sum}}\geqslant 1/2R(\delta)+1/2R_L(2\delta).$$

Recalling from [12] Ch.17 that: $R(\delta)\geqslant 1-H(\delta)$, and combining this with the results of Section 6, we find the following asymptotic bounds:

3) From Section 6, Theorem 3:

$$R_{\text{sum}}(\delta)\geqslant$$
$$1/2+1/2\log 3-1/2h(\delta)-x_0/2-1/6h(3x_0)+$$
$$-1/3h(3x_0/2)-(1/3-x_0/2)h((\delta-x_0)/(2/3-x_0)),$$

where

$$x_0=1/3+\delta-\sqrt{\delta^2+1/9}.$$

4) From Section 6, Theorem 6:
$R_{\text{sum}}(\delta)\geqslant 1/2+\log 3-1/2h(\delta)-1/2M(2\delta)$, where $M(\delta)$ is as in Section 6.

5) From Section 6, Theorem 8.
$R_{\text{sum}}(\delta)\geqslant 1-h(\delta)-h(2\delta)+h(\omega)-(1-\omega)\,h(\delta/(1-\omega))-\omega, 0<\omega\leqslant 1-\delta$,
$R_{\text{sum}}(\delta)\geqslant 1-h(\delta)-h(2\delta)\qquad(\omega=0)$.

The numerical results for these bounds can be found in Appendix 1 to this section. The bound 3) is the corrected version of the bound by Kasami [2]. The bound 4) is the corrected version of the bound found in [3], by Kasami et al. The above bounds are plotted in Appendix 2.

In Section 3 we promised to give a simple construction of a class of 4-decodable code pairs having the same parameters as those in Section 3,4.

Take $C$ = the even weight from Example 1 in Appendix 3 to Section 5.
$D$ = the code from Example 1 in Appendix 3 to Section 5.

Applying the construction of Section 5, we get code pairs having the right parameters.

In Section 3 we promised to give some results which are better than the results found by Kachatrian. In order to do this, we apply the construction method of Section 5 to the Binary Golay code having weight distribution:

| *wt* | 0 | 8 | 12 | 16 | 24 |
|---|---|---|---|---|---|
| # | 1 | 759 | 2576 | 759 | 1 |

From [12] we get by construction: $A(24,4) \geq 2^{19} = 524288$, $A(16,4) \geq 2048$, $A(12,4) \geq 1024$, $A(8,4) \geq 16$. Hence we can construct a ternary code with length 24, $L$-distance 8 and having 4728689 words. Combining this with the (24,524288,4) code and applying the construction method of Section 4, we get a code pair having length 48, sumrate 0.8577 and $d = 8$. This is much better than the parameters given in Section 3 since the rate is higher, the length smaller and the distance 8.

The code pairs constructed by the method in Section 5 all have even length. Applying the following shortening techniques will give codes of odd length:

S1. If $(C,D)$ is a $d$-decodable code pair, define:

$$C_i = \{c \in C | c_1 = i\} \qquad (i = 0,1),$$

$$D_i = \{d \in D | d_1 = i\} \qquad (i = 0,1).$$

The pairs $(\hat{C}_i, \hat{D}_j)$ $i,j \in \{0,1\}$ are $d$-decodable , where :

$\hat{C}_i$ (resp.$\hat{D}_j$)$= \{c = (c_2,\ldots,c_n) | (i\ c)$ (resp.$(j\ c)) \in C_i$ (resp.$D_j)\}$.

S2. Notation the same as under S1. The pairs $(\hat{C}_i, \hat{D})(i \in \{0,1\}$ and $(\hat{C}, \hat{D}_i), i \in \{0,1\}$, are $(d-1)$-decodable.

$$(\hat{C} \text{ (resp. } \hat{D}) = \{c = (c_2,\ldots,c_n) | (0c) \text{ or } (1c) \in C \text{ (resp.} D)\}).$$

S3. Notation the same as under S2. The pair $(\hat{C}, \hat{D})$ is $(d-2)$-decodable. Additional to this we can use the following technique (Subcode Construction).

SC. Let $(C_1, C_2)$ be a (2e-1)-decodable code pair, and assume that all weights of

$C_1$ have the same parity.

Define: $C_2^e := \{c \in C_2 | wt(c)$ is even$\}$, and analogous for $C_2^o$

Now $(C_1, C_2^e)$ and $(C_1, C_2^o)$ are 2e-decodable code pairs.

Furthermore, we have the following extension trick:

E. Let $(C,D)$ be a (2e-1)-decodable code pair. Let $\hat{C}$ be the code obtained by extending $C$ such that all weights in $C$ are even. Let $\overline{D}$ be the code obtained from $D$ by extending all words such that all weights in $D$ are even. Then $(\hat{C}, \overline{D})$ is a 2e-decodable code pair.

Proofs of the statements made under S1,S2,S3,SC, and E.: To prove S1,S2,S3, let $B = C + D$ be the ternary code $\{c + d | c \in C, d \in D\}$. In case S1, the shortened codes $\hat{B}$ can be obtained from $B$ by taking the words having the same first

symbol (or a subcode thereof) and deleting this first symbol, hence the minimum distance remains the same. In case S2 the shortened codes $\hat{B}$ can be obtained from $B$ by taking subsets of the words beginning with $i$ or $i+1$ and deleting symbol. Hence the minimum distance goes down by at most 1.

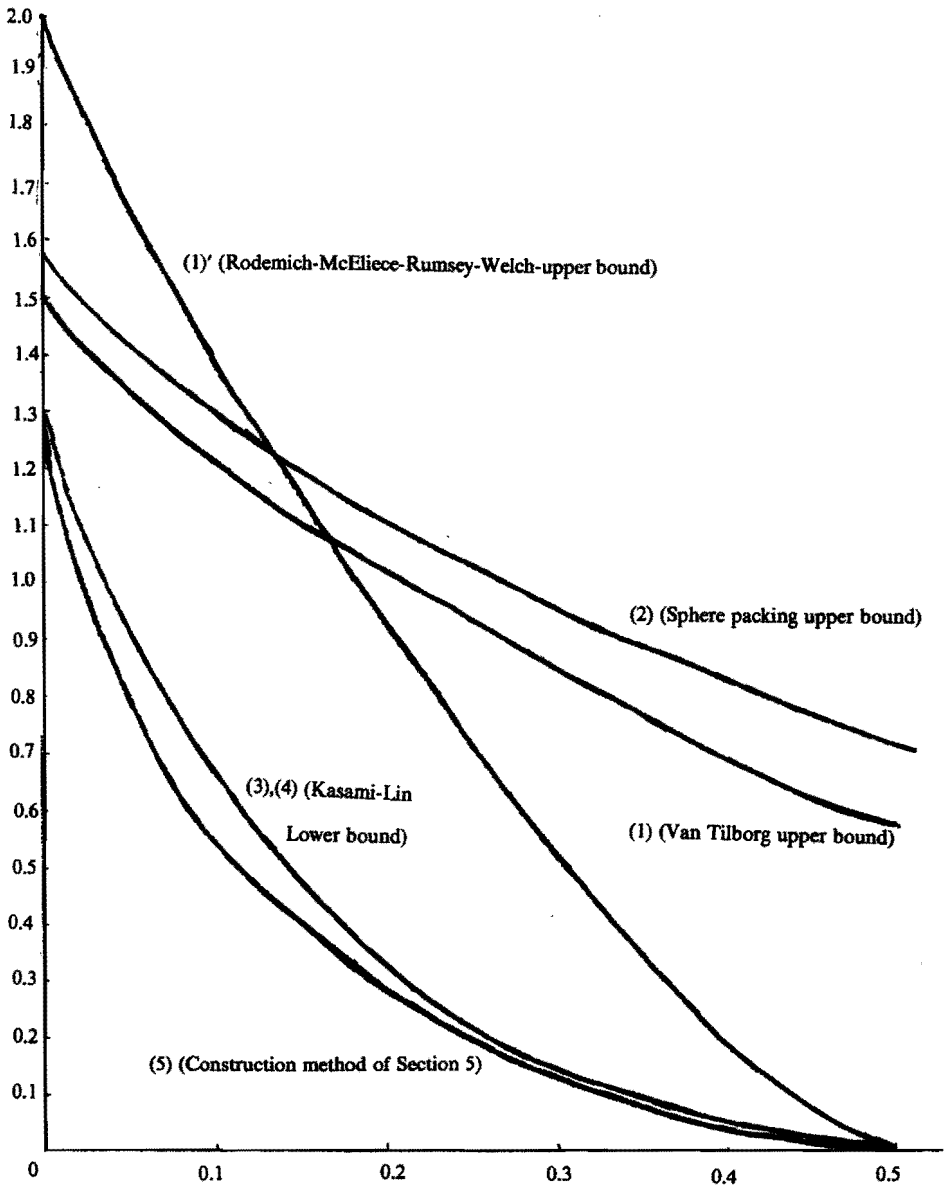
In case S3 the shortened code $\hat{B}$ can be obtained from $B$ by deleting the first symbol of the words. The minimum distance goes down by at most 2. To prove the statements under SC and E it is enough to remark that the numer of ones of the code words of the code $B$ are all even and hence the minimum distance is even. The results of all kinds of constructions can be found in Appendix 3.

## APPENDIX 1. ASYMPTOTIC BOUNDS FOR $R_{\text{sum}}(\delta)$

| | Upper bounds | | | Lower bounds | |
|---|---|---|---|---|---|
| $\delta$ | (1) | (1) | (2) | (3),(4) | (5) |
| 0.00 | 1.5000 | 2.000 | 1.58496 | 1.2925 | 1.2925 |
| 0.01 | 1.4545 | ----- | 1.53951 | 1.1771 | 1.1439 |
| 0.02 | 1.4190 | 1.836 | 1.50402 | 1.0918 | 1.0358 |
| 0.03 | 1.3872 | ----- | 1.47228 | 1.0182 | 0.9068 |
| 0.04 | 1.3577 | 1.708 | 1.44295 | 0.9520 | 0.8613 |
| 0.05 | 1.3300 | ----- | 1.41542 | 0.8916 | 0.7871 |
| 0.06 | 1.3037 | 1.594 | 1.38930 | 0.8355 | 0.7192 |
| 0.07 | 1.2785 | ----- | 1.36437 | 0.7834 | 0.6568 |
| 0.08 | 1.2543 | 1.488 | 1.34044 | 0.7345 | 0.5991 |
| 0.09 | 1.2309 | ----- | 1.31739 | 0.6885 | 0.5634 |
| 0.10 | 1.2083 | 1.386 | 1.29511 | 0.6461 | 0.5310 |
| 0.11 | 1.1863 | ----- | 1.27353 | 0.6040 | 0.5000 |
| 0.12 | 1.1649 | 1.280 | 1.25258 | 0.5651 | 0.4706 |
| 0.13 | 1.1441 | ----- | 1.23221 | ..... | 0.4424 |
| 0.14 | 1.1237 | 1.180 | 1.21237 | ..... | 0.4156 |
| 0.15 | 1.1038 | ----- | 1.19302 | 0.4597 | 0.3900 |
| 0.16 | 1.0834 | 1.110 | 1.17413 | ...., | 0.3656 |
| 0.17 | 1.0652 | ----- | 1.15568 | ..... | 0.3422 |
| 0.18 | 1.0465 | 1.010 | 1.13762 | ..... | 0.3198 |
| 0.19 | 1.0281 | ----- | 1.11995 | ..... | 0.2984 |
| 0.20 | 1.0010 | 0.922 | 1.10263 | 0.3157 | 0.2780 |
| 0.21 | 0.9922 | ----- | 1.08566 | ..... | 0.2584 |
| 0.22 | 0.9747 | 0.836 | 1.06901 | ..... | 0.2398 |
| 0.23 | 0.9575 | ----- | 1.05267 | ..... | 0.2218 |
| 0.24 | 0.94505 | 0,750 | 1.03663 | ..... | 0.2048 |
| 0.25 | 0.9238 | ----- | 1.02087 | 0.2038 | 0.1886 |
| 0.30 | 0.8434 | 0.500 | 0.94591 | ..... | 0.1186 |
| 0.35 | 0.7677 | 0.350 | 0.87670 | ..... | 0.0658 |
| 0.40 | 0.6958 | 0.162 | 0.81232 | ..... | 0.0290 |
| 0.45 | 0.6271 | 0.070 | 0.75218 | ..... | 0.0072 |
| 0.50 | 0.5613 | 0.000 | 0.69584 | ..... | ..... |

The values in the first column are obtained by calculating the bound, 1) the second column is obtained by using the values in McEliece et al. [7], the third column is obtained by using Appendix 8 of Section 6, the fourth column can be obtained by using Appendix 8 and the results for binary codes, the bounds from (3) and (4) are almost the same. The bound (5) can be calculated easily. The numbers (1), (2), (3), (4) and (5) correspond to the numbers in the text of Section 7.

## APPENDIX 2. GRAPH OF THE ASYMPTOTIC BOUNDS




DISCUSSION: In this picture the asymptotic upper bounds (1), (1)' and (2) are shown together with the lower bounds (3), (4) and (5). Like in Appendix 9 to Section 6 also in this case there is a big gap between upper and lower bounds.

## APPENDIX 3. CODE PAIRS HAVING LENGTHS UP TO 20 AND DISTANCES UP TO 8

| | $d=3$ | | | | $d=4$ | | | | $d=5$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $\|C\|$ | $\|D\|$ | $R_{sum}$ | Con | $\|C\|$ | $\|D\|$ | $R_{sum}$ | Con | $\|C\|$ | $\|D\|$ | $R_{sum}$ | |
| 3 | 2 | 1 | 0.3333 | S1 | - | - | - | - | - | - | - | - |
| 4 | 2 | 2 | 0.5000 | * | 2 | 2 | 0.5000 | * | - | - | - | - |
| 5 | 2 | 4 | 0.6000 | S2 | 2 | 2 | 0.4000 | E,SC | 1 | 2 | 0.2000 | S2 |
| 6 | 4 | 5 | 0.7203 | * | 4 | 4 | 0.6667 | * | 2 | 2 | 0.333 | * |
| 7 | 8 | 5 | 0.7605 | S2 | 4 | 5 | 0.6174 | E,SC | 2 | 2 | 0.2857 | - |
| 8 | 8 | 11 | 0.8074 | * | 8 | 9 | 0.7712 | * | 2 | 3 | 0.3231 | * |
| 9 | 16 | 9 | 0.7967 | S2 | 8 | 11 | 0.7177 | E | 2 | 4 | 0.3333 | S2 |
| 10 | 16 | 23 | 0.8523 | * | 16 | 17 | 0.8087 | * | 4 | 5 | 0.4322 | * |
| 11 | 32 | 21 | 0.8538 | S2 | 16 | 23 | 0.7748 | E | 8 | 5 | 0.4838 | S2 |
| 12 | 32 | 55 | 0.8984 | * | 32 | 38 | 0.8540 | * | 8 | 9 | 0.5142 | * |
| 13 | 64 | 54 | 0.9042 | S2 | 32 | 55 | 0.8293 | E | 8 | 17 | 0.5452 | S2 |
| 14 | 64 | 149 | 0.9442 | * | 64 | 92 | 0.8945 | * | 16 | 17 | 0.5777 | * |
| 15 | 128 | 149 | 0.9479 | S2 | 64 | 149 | 0.8812 | E,SC | 10 | 25 | 0.5310 | S2 |
| 16 | 128 | 337 | 0.9623 | * | 128 | 241 | 0.9321 | * | 20 | 27 | 0.5628 | * |
| 17 | 256 | 345 | 0.9665 | S2 | 128 | 345 | 0.9077 | SC | 20 | 54 | 0.5928 | S1 |
| 18 | 256 | 825 | 0.9827 | * | 256 | 545 | 0.9495 | * | 40 | 81 | 0.6479 | * |
| 19 | 512 | 772 | 0.9785 | S2 | 256 | 825 | 0.9310 | E | 36 | 162 | 0.6584 | S1 |
| 20 | 512 | 2187 | 1.0047 | * | 512 | 1482 | 0.9766 | * | 72 | 243 | 0.7047 | * |

| | $d=6$ | | | | $d=7$ | | | | $d=8$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $\|C\|$ | $\|D\|$ | $R_{sum}$ | Con | $\|C\|$ | $\|D\|$ | $R_{sum}$ | Con | $\|C\|$ | $\|D\|$ | $R_{sum}$ | Con |
| 6 | 2 | 2 | 0.3333 | * | - | - | - | - | - | - | - | - |
| 7 | 2 | 2 | 0.2857 | E | 1 | 2 | 0.1492 | S2 | - | - | - | - |
| 8 | 2 | 2 | 0.2500 | * | 2 | 2 | 0.2500 | * | 2 | 2 | 0.2500 | * |
| 9 | 4 | 3 | 0.2872 | E | 2 | 2 | 0.2222 | S2 | 2 | 2 | 0.2222 | E |
| 10 | 4 | 4 | 0.4000 | * | 2 | 2 | 0.2000 | * | 2 | 2 | 0.2000 | * |
| 11 | 8 | 5 | 0.3929 | E | 2 | 4 | 0.2727 | S2 | 2 | 2 | 0.1818 | E |
| 12 | 8 | 9 | 0.5142 | * | 4 | 4 | 0.3333 | * | 4 | 4 | 0.3333 | * |
| 13 | 16 | 9 | 0.4746 | E | 4 | 8 | 0.3846 | S2 | 4 | 4 | 0.3077 | E |
| 14 | 16 | 17 | 0.5777 | * | 8 | 9 | 0.4407 | * | 8 | 8 | 0.4286 | * |
| 15 | 20 | 17 | 0.5391 | E | 8 | 17 | 0.4725 | S2 | 8 | 9 | 0.4113 | E |
| 16 | 20 | 25 | 0.5603 | * | 16 | 17 | 0.5055 | * | 16 | 17 | 0.5055 | * |
| 17 | 20 | 27 | 0.5339 | E | 20 | 11 | 0.4577 | S2 | 16 | 17 | 0.4757 | E |
| 18 | 40 | 46 | 0.6025 | * | 20 | 21 | 0.4841 | * | 20 | 21 | 0.4841 | * |
| 19 | 40 | 81 | 0.6167 | E | 40 | 21 | 0.5113 | S2 | 20 | 21 | 0.4586 | E |
| 20 | 72 | 125 | 0.6567 | * | 40 | 41 | 0.5340 | * | 40 | 41 | 0.5340 | * |

Key to the tables, $|C|,|D|$ are the cardinalities of the codes of the code pair constructed by the construction under the column Con. $R_{sum}$ is the sumrate. If the Con-entry is * then the code is constructed using the construction of Section 5 with the best known binary and ternary codes (which can be found in the tables in [12] and in Appendix 7 to Section 6). If the Con-entry is equal to S1,S2,S,E or SC then the code pair is constructed by applying construction S1,S2,S,E respectively SC as described in Section 7 to an appropriate code pair obtained from construction *.

## 2.8. CONCLUSIONS

The results discussed in this chapter show that asymptotically the results for the binary adder channel of Kasami et al. [1] to [4], are the best known lower bounds, and that the upper bound of Van Tilborg [6] is the best one known. However for small $n$ and $d$ the construction method of Section 2.5 gives better results than those of Kasami [1] to [4] and Khachatrian [8].

REFERENCES


1. T. KASAMI, S. LIN (1976). *Coding for a multiple access channel* IEEE Trans. on Inf. Theory 22, pp. 129-137.
2. T. KASAMI, S. LIN (1978). *Bounds on the achievable rates of block coding for a memoryless multiple access channel.* IEEE Trans. on Inf. Theory 24, pp. 187-197.
3. T. KASAMI, S. LIN, S. YAMAMURA (1980). *Existence of good δ-decodable codes for the two user multiple access adder channel.* IBM J. Res. Development Vol. 24 no. pp. 486-495.
4. T. KASAMI, S. LIN, S. YAMAMURA, U.K. WEI (1983). *Graph theoretic approaches to the code construction for the two user multiple access binary adder channel.* IEEE Trans. on Inf. Theory 29, pp. 114-130.
5. H.C.A. VAN TILBORG (1978). *An upper bound for codes for the noisy two access binary adder channel.* IEEE Trans. on Inf. Theory 24, pp. 112-116.
6. H.C.A. VAN TILBORG (1986). *An upper bound for codes for the noisy two access binary adder channel.* IEEE Trans. on Inf.6 Theory 32, pp. 436-440.
7. R.J. MCELIECE, E.R. RODEMICH, H. RUMSEY JR., L.R. WELCH (1977). *New Upper Bounds on the rate of a code via the Delsarte-MacWilliams inequalities.* IEEE Trans. on Inf. Theory 23, pp. 157-166.
8. G.H. KHACHATRIAN (1984). *Codes for the 2-user adder channel.* Probl. Control and Info. Th. 13, pp. 275-279.
9. E.J. WELDON JR. (1978). *Coding for a multiple access channel.* Info. and Control 36, pp. 256-274.
10. J.H. VAN LINT (1982). *Introduction to coding theory.* Springer Verlag, NY.
11. W.W. PETERSON (1966). *Error correcting codes.* Cambridge, Mass; MIT-Press, pp. 69.
12. F.J. MACWILLIAMS, N.J.A. SLOANE (1977). *The theory of Error-Correcting Codes.* North-Holland publishing company, Amsterdam.
13. M. BEST (1980). *Binary codes with a minimum distance of 4.* IEEE Trans. on Inf. Theory 25, pp. 738-743.
14. A. DÜR (1986). *The weight distribution of double error correcting Goppa codes,* Proceedings of A.A.E.C.C.-4, Karlsruhe, Springer LN in Computer Science.
15. C.L.M. VAN PUL (1986). *Some distance problems in coding theory,* Ph. D. Thesis, Eindhoven University of Technology.
16. P. TURÁN (1941). *Egy Grafelmeleto Szelsoertekfeladatrol.* Mat.es Fizikai Lopok, Vol. 48, pp. 436-452, (See also *On the theory of graphs.* Colloq. Math, Vol. 3, pp. 19-30, (1954) or J. Denes, *Latin squares and codes.* Proc. Int. Symp. Information Theory, Cachan, France, July (1977).
17. P.B.A.M. COEBERGH VAN DEN BRAAK and H.C.A. VAN TILBORG (1985). *A Family of Good Uniquely Decodable Code Pains for the Two-Acces Binary Adder Channel.* IEEE Trans, on Info. Theory, Vol. IT-31, pp. 3-9.

# Chapter 3

# Codes Constructed from Algebraic Geometry

## 3.0. Introduction

In this chapter we will give some results on codes constructed from Curves. First in Section 1, we will give some basic facts from Algebraic Geometry. Then in Section 2, we will give the general theory for codes from curves. Also in this section, we will describe a decoding algorithm for these codes, that was recently discovered by Justesen et al. [6], and generalized by Vlădut and Skorobogatov [7]. In Section 3, the codes on Hermitian curves are studied. Section 4 contains some examples of codes over GF (4). For the proofs and details about the concepts and theorems from Algebraic Geometry, the reader is referred to the books of Fulton [1] and Hartshorne [2].

## 3.1. Basic facts from Algebraic Geometry

*Algebraic sets*

Let $K$ be an algebraically closed field, $A^n$ the $n$-dimensional affine space over $k$ (i.e. the set of $n$-tuples of elements of $k$). The elements of $A^n$ will be called points. $A^1$ is the affine line, $A^2$ the affine plane etc.. For $F \in k[X_1, X_2, ..., X_n]$ (The polynomial ring in $n$ variables over $k$), a point $P=(a_1,...,a_n) \in A^n$ is a zero of $F$ if $F(a_1,...,a_n)=0$ (write $F(P)=0$). The set of zeroes of $F$ is called the hypersurface defined by $F$, and is denoted by $V(F)$. A hypersurface in $A^2$ is called an affine plane curve. More generally if $S$ is a set of polynomials in $k[X_1,...,X_n]$, we define $V(S):\{P \in A^n | F(P)=0, \forall F \in S\} = \cap_{F \in S} V(F)$. A subset $X$ of $A^n$ is called an affine algebraic set if $X=V(S)$ for $S \subseteq k[X_1,...,X_n]$. Instead of taking $S$, we can look at the ideal I generated by $S$. Obviously $V(I)=V(S)$. Therefore every affine algebraic set can be written as $V(I)$ with I some ideal in $k[X_1,...,X_n]$. To a set $X \subseteq A^n$, we can associate an ideal in $k[X_1,...,X_n]: I(X)=\{F(P)=0, \forall P \in X\}$. It is called the ideal of $X$.

An algebraic set $V \subseteq A^n$ is called reducible if it is the union of two smaller algebraic sets $V_1, V_2 \neq V, V=V_1 \cup V_2$. Otherwise $V$ is called irreducible. An algebraic set $V$ is irreducible if and only if $I(V)$ is a prime ideal. For every algebraic set $V$, there is a unique decomposition of $V$ into irreducible components (unique up to order of the components). $V=V_1 \cup V_2 \cup ..... \cup V_m$, and $V_i \not\subset V_j i \neq j$. If $F \in k[X_1,...,X_n]$ and $F=F^{n_1}....F_r^{n_r}$ is the decomposition of $F$ into

irreducible factors then $V(F)=V(F_1)\cup\ldots\cup V(F_r)$ is the decomposition of $V(F)$ into irreducible components. Furthermore there is a 1-1 correspondence between irreducible polynomials and irreducible hypersurfaces.

*Affine varieties, coordinate rings and function fields*

An irreducible algebraic set is called an affine variety. For a variety $V\subseteq A^n$, let $I(V)$ be the corresponding prime ideal in $k[X_1,\ldots,X_n]/I(V)$. Then $k[X_1,\ldots,X_n]/I(V)$ is a domain. Now $\Gamma(V)=k[X_1,\ldots,X_n/I(V)$ is called the coordinate ring of $V$. We can form the quotient field of $\Gamma(V)$ and denote it by $k(V)$. It is called the field of rational funtions on $V$. The elements of $k(V)$ are called rational functions. For $P\in V$, a rational function $f$ is said to be defined at $P$ if there exist $a$ and $b$ in $\Gamma(V)$ such that $f=a/b$ and $b(P)\neq 0$. In this case the value of $f$ at $P$ is defined to be $a(P)/b(P)$. The set of points where a rational function is not defined is called the pole set of $f$. It is an algebraic subset of $V$. For $P\in V$, by $O_p(V)$ we mean the ring of rational functions which are defined at $P$. It is called the local ring of $V$ at $P$ and it is a subring of $k(V)$ containing $\Gamma(V)$.

By $M_P(V)$ is meant the unique maximal ideal of $O_P(V)$. $M_P(V)=\{f\in O_P(V)|f(P)=0\}$. It is the kernel of the evaluation homomorphism from $O_P(V)$ to $k$ defined by $f\mapsto f(P)$. Therefore $O_P(V)/M_P(V)\simeq k$. The elements $f\in O_P(V)$ such that $f(P)=0$ are called the units. The elements of $M_P(V)$ are called nonunits.

*Projective varieties*

Let $k$ be a field. Projective $n$-space over $k$ written $P^n$ is defined to be the set of all lines passing through the origin $(0,\ldots,0)$ in $A^{n+1}$. A point $(x_1,\ldots,x_{n+1})\neq(0,\ldots,0)$ in $A^{n+1}$ uniquely determines the line $\{\lambda(x_1,\ldots,x_{n+1})|\lambda\in k\}$. If we define the equivalence relation $\sim$ to be: $(x_1,\ldots,x_{n+1})\sim(y_1,\ldots,y_{n+1})$ if and only if $\exists_{\lambda\in k}(x_1,\ldots,x_{n+1})=\lambda(y_1,\ldots,y_{n+1})$, then $P^n$ can be seen as the set of equivalence classes of points in $A^{n+1}\setminus\{(0,\ldots,0)\}$. The elements of $P^n$ will be called projective points. If a projective point is determined by some $(x_1,\ldots,x_{n+1})\in A^{n+1}$, we call this a set of homogeneous coordinates for $P$. A projective point $P=(a_1,\ldots,a_{n+1})$ is called a zero of a polynomial $F\in k[X_1,\ldots,X_{n+1}]$ if $F(a_1,\ldots,a_{n+1})=0$ for every choice of homogeneous coordinates for $P$. Again we write $F(P)=0$. If $S$ is a set of polynomials in $k[X_1,\ldots,X_{n+1}]$, then $V(S)$ is defined as $V(S)=\{P\in P^n|F(P)=0,\forall F\in S\}$. If $I$ is the ideal generated by $S$, then $V(I)=V(S)$. If $I=(F^{(1)},\ldots,F^{(r)})$, where the $F^{(i)}=\Sigma F_j^{(i)}$ and the $F_j^{(i)}$ are forms of degree $j$, then $V(I)=V(\{F_j^{(i)}\})$. Such a set is called a projective algebraic set. An ideal is called homogeneous if for every $F=\Sigma F_i\in I$, $F_i$ a form of degree $i$, it is true that $F_i\in I$. For any set $X\subseteq P^n$, $I(X)$ is a homogeneous ideal. An ideal $I\subset k[X_1,\ldots,X_{n+1}]$ is homogeneous if and only if it is generated by a finite set of forms. An algebraic set $V\subset P^n$ is irreducible if it is not the union of two smaller algebraic sets. Again $V$ is irreducible if and only if $I(V)$ is prime. An irreducible algbraic set in $P^n$ is called a projective variety. A projective algebraic set can be written uniquely (up to the order of the components) as a union of projective varieties, its

irreducible components. If $V$ is a projective variety in $P$,$I(V)$ is a prime ideal; so $\Gamma_h(V)=k[X_1,...,X_{n+1}]/I(V)$ is a domain. It is the homogeneous coordinate ring of $V$. More generally, let $I$ be any homogeneous ideal in $k[X_1,...,X_{n+1}]$ and $\Gamma=k[X_1,...,X_{n+1}]/I$, then an element of $\Gamma$ is called a form of degree $d$ if there is a form $F$ degree $d$ in $k[X_1,...,X_{n+1}]$ whose residue mod $I$ is $f$. Let $k_h(V)$ be the quotient field of $\Gamma_h(V)$. It is called the homogeneous function field of $V$. The elements of $k_h(V)$ cannot be viewed as functions, but if $f\in k_h(V)$ is the quotient to two forms of the same degree $d$, say $h$ and $g$,$f=h/g$, then $:h(\lambda x)/g(\lambda x)=\lambda^d h(x)/\lambda^d g(x)=h(x)/g(x)$. So in this case, $f$ can be viewed as a function on the projective variety since the value of $f$ is independent of the choice of the homogeneous coordinates. In this case, the function field $k(V)$ is a subfield of $k_h(V)$ and is defined by: $\{f\in k_h(V)|f$ is quotient of two forms of the same degree $\}$. The elements of $k(V)$ are called rational functions on $V$. If $P\in V,f\in k(V)$, we say that $f$ is defined at $P$ if there exist forms $g,h$ of the same degree such that $f=g/h$ and $h(P)\neq 0$. The ring $0_P(V)=\{f\in k(V)|f$ is defined at $P\}$ is a subring of $k(V)$. It is a local ring with maximal ideal $M_P(V)=\{f|f=h/g,g(P)\neq 0,h(P)=0\}$.

*Varieties, morphisms, rational maps and function fields*

Let $X$ be a projective space. The Zariski toplogy has as open sets the sets $X-U$ is an algebraic subset of $X$. If $V$ is an irreducible algebraic set, every $U$ for which open subset $U$ of $V$ will be called a variety, and we can give it the induced topology. We can define $k(U)=k(V)$ to be the field of rational functions on $U$ and for $P\in U$ we define $O_P(U)=O_P(V)$, the local ring of $U$ at $P$. Every open subset of $U$ is also a variety and therefore we call it an open subvariety of $U$. If $V$ is a variety, $U\subset V,U\neq\varnothing,U$ is open in $V$, then we define $\Gamma(U,O_V):=\cap_{P\in U}O_P(V)$. $\Gamma(U,O_V)$ is a subring of $k(V)$. It is the subring of rational functions defined at each $P\in U$. If $V_1$ and $V_2$ are varieties, $\phi:V_1\to V_2$ a function, then $\phi$ is called a morphism if it has the following two properties:

1) $\phi$ is continuous with respect to the Zariski topology
2) For every open set $U\subset V_2$, if $f\in\Gamma(U,O_{V_2})$ then $\tilde{\phi}(f)=f\circ\phi\in\Gamma(\phi^{\leftarrow}(U),O_{V_1})$, (i.e. $f\circ\phi$ is again a rational function.).

Let $K$ be a finitely generated field extension of $k$. The transcendence degree of $K$ over $k$ is defined to be the smallest integer $n$, such that for some $X_1,...,X_n$, $K$ is algbraic over $k(X_1,...,X_n)$. $K$ is called an algebraic function field in $n$ variables over $k$. If $K$ is an algebraic function field in one variable over $k$ and $x\in K,x\notin k$, then $K$ is an algebraic over $k(x)$. (Assuming $k$ to be algebraically closed). If $V$ is a variety, then $k(V)$ is a finitely generated extension of $k$ and $dim(V)$ is defined to be the transcendence degree of $k(V)$. A variety of dimension 1 is called a curve. If $V_1$ and $V_2$ are varieties $f_i:U_i\to V_2$ two morphisms from subvarieties $U_i$ of $V_1$ to $V_2,i=1,2$, then $f_1$ and $f_2$ are called equivalent if the restrictions of $f_1$ and $f_2$ to $U_1\cap U_2$ are the same. An equivalence class of morphisms is called a rational map from $V_1$ to $V_2$. It is said to be

birational if there are open sets $U \subseteq V_1$, and $U' \subseteq V_2$ and an isomorphism $f: U \to U'$ that represents the equivalence class defining the rational map. $V_1$ and $V_2$ are called birationally equivalent in this case. Two varieties are birationally equivalent if and only if their function fields are isomorphic. Every curve is birationally equivalent to an irreducible plane curve. Therefore it is useful to study plane curves.

*Discrete valuation rings*

In studying plane curves, it is pleasant to have at hand the concept of discrete valuation ring. If $R$ is a domain, that is not a field, then $R$ is Noetherian and local and the maximal ideal is a principal ideal if and only if there is an irreducible element $t \in R$ such that every nonzero $z \in R$ is of the form $z = ut^n$, where $u$ is a unit in $R$ and $n$ is a nonnegative integer. A ring having this property is called a discrete valuation ring. The element $t$ is called a uniformizing parameter for $R$. Any other uniformizing parameter is of the form $ut$, where $u$ is a unit in $R$. In the quotient field $K$ of $R$, every element $z$ can be written uniquely as $z = ut^n$, where $u$ is a unit and $n$ is an integer. The exponent $n$ is called the order of $z$ and is denoted $ord(z)$. We define $ord(0) = \infty$.
$R = \{z \in K | ord(z) \geqslant 0\}$, $M = \{z \in K | ord(z) > 0\}$. $M$ is the maximal ideal of $R$. The function $ord$ is an order function on $K$, i.e. it has the properties:

1) $ord$ is a function from $K$ onto $\mathbb{Z} \cup \{\infty\}$,
2) $ord(a) = \infty$ if and only if $a = 0$,
3) $ord(ab) = ord(a) + ord(b)$,
4) $ord(a+b) \geqslant \min(ord(a), ord(b))$.

*Plane curves*

We are now going to discuss affine plane curves. For projective plane curves the theory is more or less the same because if we dehomogenize the projective plane curve in a suitable way, we get an affine plane curve and the local properties of corresponding points are the same for both curves since they depend only on the local ring.

On $k[X, Y]$ we define the following equivalence relation: $F \sim G$ if and only if $F = \lambda G$ for some $\lambda \in k, \lambda \neq 0$. An affine plane curve is defined to be an equivalence class under this relation. The degree of the curve is the degree of a defining polynomial. If $F = \Pi F_i^{e_i}$ where the $F_i$ are the irreducible factors of $F$, we say that the $F_i$ are the components of $F$ and that they appear with multiplicities $e_i$. $F_i$ is a simple component if $e_i = 1$, and multiple otherwise. If $F$ is irreducible then $V(F)$ is a variety in $A^2$. A point $P = (a,b) \in F$ is called a simple point of $F$ if $F_x(P) \neq 0$ or $F_y(P) \neq 0$ where $F_x$ and $F_y$ are the derivatives of $F$ with respect to $X$, respectively $Y$. The line: $F_x(P)(X-a) + F_y(P)(Y-b) = 0$ is called a tangent line to the curve at $P$. A point that is not simple is called multiple or singular. A nonsingular curve is a curve with simple points only. If $F$ is a curve, $P = (0,0)$, we can write $F = F_m + F_{m+1} + \cdots + F_n$, where $F_i$ is a form of degree $i$ in $k[X, Y]$. $F_m \neq 0$. The multiplicity of $F$ at $P$ is defined to be $m_P(F) := m$. Now $P \in F$ if and only if $m_P(F) > 0$. Furthermore $P$ is a simple

point if and only if $m_P(F)=1$. If $m_P(F)=2$, then $P$ is called a double point, if $m_P(F)=3$ a triple point etc.. The form $F_m$ can be written $F_m=\Pi L_i^{r_i}$ where the $L_i$ are distinct lines (assuming $k$ to be algebraically closed). The $L_i$ are called the tangent lines to $F$ at $P(=(0,0))$. $r_i$ is the multiplicity of the tangent $L_i$. $L_i$ is a simple, respectively double tangent if $r_i=1$, respectively 2.

If $F$ has $m_P(F)$ distinct simple tangents at $P$, then $P$ is an ordinary multiple point of $F$. The above definitions can be extended to a point $P=(a,b)\neq(0,0)$ by using translations (First translate the affine plane such that $P$ is at the origin and then apply the given definitions.). Now $P$ is a simple point of $F$ if and only if $O_P(F)$ is a discrete valuation ring. If $L$ is a line through $P$ not tangent to $F$ at $P$, then the image $l$ of $L$ in $O_P(F)$ is a uniformizing parameter for $O_P(F)$. If $P$ is a point on $F$ and $F$ is irreducible, then $m_P(F)=dim(M_P(F)^n/M_P(F)^{n+1})$ for $n$ sufficiently large. So $m_P(F)$ only depends on $O_P(F)$. If $F$ and $G$ are plane curves, then $I(P,F\cap G):=dim_k(O_P(A^2)/(F,G))$. This number $I(P,F\cap G)$ is called the intersection number of $F$ and $G$ at $P$. It is equal to $\infty$ if $F$ and $G$ do not intersect properly, i.e. have a common component through $P$. If $F$ and $G$ do not intersect at $P, I(P,F\cap G)=0$ and otherwise if $F$ and $G$ intersect properly at $P$, then $I(P,F\cap G)$ is a nonnegative integer satisfying $I(P,F\cap G)\geq m_P(F)m_P(G)$. If $F$ and $G$ have no tangent in common at $P$, then $I(P,F\cap G)=m_P(F)m_P(G)$ and the converse also holds. For projective plane curves we have the Theorem of Bezout: If $F$ and $G$ are projective plane curves of degree $m$ and $n$ respectively, $F$ and $G$ have no common component, then $\Sigma_P\, I(P,F\cap G)=mn$.

*The Riemann-Roch Theorem, Differentials and Residues*

Let $C$ be an irreducible projective curve, $X$ its nonsingular model, $K=k(C)=k(X)$ its function field. The point $P\in X$ will be identified with the places of $K$ and $ord_P$ denotes the order function. A divisor on $X$ is a formal sum $D=\Sigma_{P\in X}n_PP$ where the $n_P\in\mathbb{Z}$ and $\{P|n_P\neq 0\}$ is a finite set. The degree of a divisor is the sum of the coefficients: $deg(D)=\Sigma_{P\in X}n_P$. $D$ is called effective if $n_P\geq 0$ for all $P$. Furthermore $\Sigma n_PP>\Sigma m_PP$ if $n_P\geq m_P$ for all $P$. For $f\in K$, we can define the divisor of $f$ to be $div(f)=\Sigma_{P\in X}ord_P(f)P$. The degree of such a divisor is 0, since a rational function has the same number of zeroes as poles. Two divisors $D$ and $D'$ are called linearly equivalent if there exists a function $f\in K$ such that $D=D'+div(f)$. It can easily be seen that this in fact is an equivalence relation. Now with a divisor, we can associate the vector space $L(D)=\{f\in K|div(f)+D>0\}$. Now Riemann's Theorem states the existence of a constant $g$ such that $dim(L(D))\geq deg(D)+1-g$ for all divisors $D$. The smallest such $g$ is called the genus of the curve. For the definition of differentials, we refer to Fulton [1] p.204. We recall that $\Omega_k(K)$ denotes the space of differentials of $K$ over $k$. It is a 1-dimensional vector space over $K$. If $\omega\in\Omega_k(K),\omega\neq 0, P\in X$ a place and $t$ an uniformizing parameter in $O_P(X)$ and $\omega=fdt, f\in K$, then $ord_P(\omega):=ord_P(f)$. Therefore to every differential $\omega$, we can associate a divisor: $div(\omega)=\Sigma_{P\in X}ord_P(\omega)P=W$. $W$ is called a canonical divisor and its degree is $2g-2$, where $g$ is the genus of the curve. Now the Riemann-Roch Theorem states that for every canonical divisor $W$ on $X$ and

for every divisor $D$: $dim(L(D))=deg(D)+1-g+dim(L(W-D))$. The space of differentials $\Omega(D)$ is defined as: $\Omega(D)=\{\omega\in\Omega_k(K)|\, div(\omega)>D\}$. Another formulation of the Riemann-Roch Theorem is: $dim(L(D))=deg(D)+1-g+dim(\Omega(D))$. We know that $\Omega(D)\cong L(W-D)$. For the definition of residue of a differential at a place ($res_P(\omega)$), we refer to Lang [11]. We recall the most important theorem concerning residues: $\Sigma_{P\in X} res_P(\omega)=0$.

*Rational points on a curve*

A point on a curve is called rational over $GF(q)$ if it can be represented by vector that has all coordinates in $GF(q)$ where $q$ is a power of $char(k)$, $k$ the functionfield of the curve. If $C$ is a curve of genus $g$ and $N$ is the number of rational points over $GF(q)$, then the Theorem of Hasse-Weil states that $|N-q-1|\leq 2g\sqrt{q}$. (cf. Serre [12].).

## 3.2. Codes from Algebraic Geometry

Let $C$ be an irreducible smooth curve. Let $P_1,...,P_n$ be rational points over $GF(q)$ on this curve. Define $D:=\Sigma_{i=1}^{n} P_i$. Choose a rational divisor $G$ disjoint from $D$ (i.e. $G$ is invariant under the Galois automorphism $\sigma$ fixing $GF(q)$). Consider the vector space $L(G)$ of rational functions $f$ satisfying $div(f)>-G$. Define the map $\psi: L(G)\to GF(q)^n$ by: $f \xrightarrow{\phi} (f(P_1),...,f(P_n))$. Furthermore consider the vector space $\Omega(D-G)$ of differentials $\omega$ satisfying $div(\omega)>G-D$. Define the map $\psi:\Omega(D-G)\to GF(q)^n$ by:

$$\omega \xrightarrow{\psi} (Res_{P_1}(\omega),...,Res_{P_n}(\omega))$$

Then $\phi$ and $\psi$ are injective:

$Ker(\phi)=\{f\in L(G)|f(P_i)=0 \text{ for all } i\}=L(G-D)=\{0\}$, assuming $n>deg(G)$,

$Ker(\psi)=\{\omega\in L(D-G)|Res_{P_i}(\omega)=0 \text{ for all } i\}=\Omega(-G)=\{0\}$, if $deg(G)\geq 0$.

Let $C:=\mathrm{Im}(\phi)$ and $C':=\mathrm{Im}(\psi)$, and assume $deg(G)\geq 2g-2$ where $g$ is the genus of the curve.

THEOREM 1. *$C$ and $C'$ are linear codes having parameters:*

*length: $n$,*

*dimension: $deg(G)=g+1$ resp. $n-deg(G)+g-1$,*

*distance: $\geq n-deg(G)$ resp. $\geq deg(G)-2g+2$.*

*Moreover $C$ and $C'$ are dual codes. (provided $2g-2<deg(G)\leq n$.).*

PROOF: Since $deg(G)>2g-2$, we can apply the Riemann-Roch Theorem to find the dimension of $L(G)$:$dim(L(G))=deg(G)-g+1$ and from the injectivety of $\phi$ this is the dimension of $C$. To obtain the minimum distance of $C$, remark that the number of zeroes of a rational function is equal to the number of poles, and hence the functions in $L(G)$ can have at most $deg(G)$ zeroes among the points $P_i$ (since $D$ and $G$ are disjoint). Therefore the minimum weight of the code $C$ (and hence the minimum distance of the code $C$)

$\geqslant n-deg(G)$. The dimension of $C'$ follows from the fact that $\Omega(D-G)\cong L(W+D-G)$, where $W$ is a canonical divisor ($deg(W)=2g-2$). From the injectivety of the map $\psi$ and the Riemann-Roch Theorem, we obtain: $dim(C')=dim(\Omega(D-G))=dim(L(W+D-G))=n-deg(G)+g-1$. The minimum distance of $C'$ can be found by observing that if $\omega\in\Omega(D-G)$ and $Res_{P_{ij}}(\omega)=0$ for $j=1,\ldots,t, \omega\neq 0$, then: $\omega\in\Omega(D-\Sigma_{j=1}^{t}P_{i_j}-G)$ and hence

$$n-t-deg(G)+2g-2\geqslant 0 \quad (deg(W+D-\Sigma_{j=1}^{t}P_{i_j}-G)\geqslant 0).$$

So $t\leqslant n-deg(G)+2g-2$, and therefore $wt(\psi(\omega))\geqslant deg(G)-2g+2$. To prove that $C$ and $C'$ are dual codes, note that from the Residue Theorem, and the fact that for $f\in L(G)$ and $\omega\in\Omega(D-G)$ the differential $f\omega$ has its poles among the points $P_i$, we can conclude:

$$0=\sum_{i=1}^{n} Res_{P_i}(f\omega) = \sum_{i=1}^{n} f(P_i)\, Res_{P_i}(\omega) = (\phi(f), \psi(\omega)).$$

Q.E.D.

The numbers $\delta_1=n-deg(G)$ and $\delta_2=deg(G)-2g+2$ are called the designed distances of the codes. We want now to investigate under which conditions the codes are (weakly) selfdual. We already remarked that $L(W+D-G)\cong\Omega(D-G)$, the isomorphism given by $f\to f\omega$, where $W=div(\omega)$, $\omega$ a differential. If we can find a differential $\omega$ such that $Res_{P_i}(\omega)=1$ for all $i=1,\ldots,n$, then $W+D-G$ and $D$ are disjoint since $\omega$ has poles of order 1 at the points $P_i$, so the code $C$ obtained from $L(W+D-G)$ and the code $C'$ obtained from $\Omega(D-G)$ are equal:

$$\begin{aligned}
\phi(L(W+D-G) &= \{(f(P_1),\ldots,f(P_n)|f\in L(W+D-G)\} = \\
&= (f(P_1)Res_{P_1}(\omega),\ldots,f(P_n)Res_{P_n}(\omega)|f\in L(W+D-G)\} = \\
&= \{(Res_{P_1}(f\omega),\ldots,Res_{P_n}(f\omega))|f\in L(W+D-G)\} = \\
&= \{(Res_{P_1}(\omega'),\ldots,Res_{P_n}(\omega'))|\omega'\in\Omega(D-G)\} = \\
&= \psi(\Omega(D-G)).
\end{aligned}$$

Suppose $G$ is effective, $G\geqslant 0$. If we now have that $W+D-G\supseteq G$, then for the codes $C_1$ obtained from $L(G)$ and $C_2$ obtained from $\Omega(D-G)$ we have $C_1\subseteq C_2$, $C_1^{\perp}=C_2$ and hence $C_1\subseteq C_1^{\perp}$, meaning $C_1$ is weakly selfdual. Furthermore $C_1=C_1^{\perp}$ if $W+D-G=G$, or equivalently if $div(\omega)=2G-D$. Stichtenoth and Lachaud [5] prove that under certain circumstances these conditions are necessary to have (weakly) self-duality.
We have now obtained the following theorem

THEOREM 2. *If $G$ is effective and there exists a differential $\omega$ satisfying $div(\omega)\geqslant 2G-D$, $Res_{P_i}(\omega)=1$ for all $i=1,\ldots,n$, then the code $C$ constructed from $L(G)$ is weakly self-dual, and $C$ is self-dual if $div(\omega)=2G-D$.*

Next we want to make some remarks on the automorphism group of the code. Let $G$ be the automorphism group of the curve, and let the curve be embedded in some projective space $PG(n,q)$. Then $G$ can be seen to be a subgroup of $PGL(n+1,q)$. The elements of $G$ also act on rational functions (since they act on coordinates). Therefore the elements of $G$ act on codewords. In order to have that codewords go into codewords under the action of $\sigma \in G$, it is enough to require that the divisor $G$ is stabilized by $\sigma$, because then the rational functions in $L(G)$ go into rational functions in $L(G)$. Conclusion:

THEOREM 3. *The stabilizer of the divisor $G$ in the automorphism group of the curve is a subgroup of the automorphism group of the code.*

REMARK: The automorphism group of the curve is usually larger but not always (see Section 3, Example 1).

In the rest of this section, we will describe a decoding algorithm for AG-codes which was in essence found by Justesen et al. [6]. In our description we will follow the paper by Skorobogatov and Vlădut [7].
The idea is as follows:

1. Find an error location set, i.e. a (small) set of coordinate positions $L=\{i_1,...,i_t\}$, such that all errors occur in positions in $L$.
2. Solve the equations $\Sigma_{ij\in L}\alpha_{ij}h_{ij}=s$, where $H=[h_1,...,h_n]$ is a parity check matrix, $s$ is the syndrome: $s=Hr$ ($r$ is the received word), and thus find the error pattern.

In order to ensure that the above method works, we make the restriction that $t$ is not too big. Then the equations will have a unique solution, which must be the error pattern.
We will now give a detailed description of the algorithm. The code which we are going to decode is $\mathrm{Im}(\Omega(G-D))$, and from Theorem 1, we know that $\mathrm{Im}(L(G))$ is the dual of this code. Choose a basis $f_1,...,f_r$ for $L(G)$. Then

$$H = \begin{bmatrix} f_1(P_1) & \cdots & f_1(P_n) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ f_r(P_1) & \cdots & f_r(P_n) \end{bmatrix}$$

is a parity check matrix for $\mathrm{Im}(\Omega(G-D))$. For $u=(u_1,...,u_n)\in GF(q)^n$, $f$ a rational function define:
$s(u,f):=\Sigma_{i=1}^{n}u_i f(P_i)$.
Now the syndrome $s$ of $u$ is: $s=(s(u,f_1),...,s(u,f_r))^T=Hu^T$.

LEMMA 4.
$u\in\Omega(G-D)\Leftrightarrow s=0\Leftrightarrow s(u,f_i)=0,\ i=1,...,r\Leftrightarrow s(u,f)=0$ *for all* $f\in L(G)$.
Let $c$ be a codeword; $e$ an error pattern, $wt(e)=t$, and $u$ is the received word:

$u=c+e$. Suppose $Q_1,...,Q_t$ correspond to the locations where an error occurs.

Choose a divisor $F$. Let $g_1,...,g_l$ be a basis for $L(F)$ and let $h_1,...,h_k$ be a basis for $L(G-F)$. Define $s_{ij}(u):=s(u,g_ih_j)$, $i=1,...l$, $j=1,...,k$.

LEMMA 5.

$$s_{ij}(u) = s_{ij}(c+e) = s_{ij}(c) + s_{ij}(e) = s_{ij}(e).$$

PROOF: Since $c$ is a codeword and $g_ih_j \in L(G)$, the statement follows from Lemma 4.
Consider the system of equations:

$$\sum_{i=1}^{l} s_{ij}(u)X_i = 0 \qquad j=1,...,k. \tag{$\star$}$$

THEOREM 6. *If $t<dim(L(F))$ then ($\star$) has a nontrivial solution.*

PROOF: $dim(L(F-\Sigma_{i=1}^{t}Q_i)) \geq dim(L(F))-t>0$, (the first inequality follows from the fact that requiring that the $Q_i$ are zeroes imposes at most $t$ linear conditions.). Since $dim(L(F-\Sigma_{i=1}^{t}Q_i))>0$, we can find a nonzero $g$ in this space. This $g$ is an element of $L(F)$ and satisfies $g(Q_i)=0, i=1,...,t$. since $g\in L(F)$ we can write $g=y_1g_1+\cdots+y_lg_l$, with $y_i\in GF(q)$, $i=1,...,l$. Now:

$$\sum_{i=1}^{l} s_{ij}(u)y_i = \sum_{i=1}^{l} s_{ij}(e)y_i = \sum_{i=1}^{l}\sum_{m=1}^{n} e_m g_i(P_m)h_j(P_m)y_i = \sum_{m=1}^{n} e_m g(P_m)h_j(P_m) = 0,$$

since $g(P_m)=0$ if $P_m\in\{Q_1,...,Q_t\}$ and $e_m=0$ if $P_m\notin\{Q_1,...,Q_t\}$. Hence the $y_i$ are a solution to the system of equations ($\star$) and since $g\neq 0$ this solution is non-trivial. Q.E.D.

THEOREM 7. *If $deg(G-F)>t+2g-2$, then for every non-trivial $y_1,...,y_l$ of ($\star$) the rational function $g=\Sigma_{i=1}^{l}y_ig_i$ has the property that $g(Q_1)=\cdots=g(Q_t)=0$.*

PROOF: Consider the sequence:

$$0 \qquad L(G-F-\sum_{i=1}^{t}Q_i) \qquad L(G-F) \xrightarrow{\pi} GF(q)^t.$$

The last map is the map $\pi: f \qquad (f(Q_1),...,f(Q_t))$.
The map $\pi$ is surjective:

$$Ker(\pi)=\{f\in L(G-F) | f(Q_1)=\cdots=f(Q_t)=0\}=L(G-F-\sum_{i=1}^{t}Q_i).$$

Now $dim(L(G-F))=deg(G-F)+1-g$ (applying the Riemann-Roch Theorem),

$dim(L(G-F-\Sigma_{i=1}^{t}Q_i))=deg(G-F)-t+1-g$ (applying the Riemann-Roch Theorem), (The conditions in the Riemann-Roch Theorem are satisfied

since $deg(G-F)>t+2g-2$).
Hence:

$$dim(\mathrm{Im}(\pi))=dim(L(G-F))-dim(Ker(\pi))=$$

$$dimL(G-F)-dim(L(G-F-\sum_{i=1}^{t}Q_i))=t.$$

Choose a basis $h'_1,...,h'_k$ for $L(G-F)$ such that $h'_j(Q_j)=1$ and $h'_j(Q_m)=0$ for $m=j$. (Such a basis exists since $\pi$ is surjective.). Write $h'_j=\sum_{r=1}^{k}\gamma_r h_r$.

Now suppose $(y_1,...,y_l)$ is a solution to (*) and define $g=\Sigma_{i=1}^{l}y_i g_i$. Then: for $j=1,...,k: 0=\Sigma_{i=1}^{l}s(e,g_i h_j)y_i$, and hence for $j=1,...,k$: $0=\Sigma_{i=1}^{l}s(e,g_i h'_j)y_i$.
Therefore

$$0=\sum_{i=1}^{l}\sum_{m=1}^{t}e_m g_i(Q_m)h'_j(Q_m)y_i=\sum_{i=1}^{l}e_j g_i(Q_j)y_i = e_j g(Q_j)\,,\quad j=1,...,t.$$

Thus $e_j\neq 0\Rightarrow g(Q_j)=0, j=1,...,t$. Q.E.D.
From Theorem 7, it follows that if $u$ is the received vector and $y_1,...,y_l$ is a non-trivial solution to (*), then the errors are located among the zeroes of $g=\Sigma_{i=1}^{l}y_i g_i$. Let $g\in L(F)$ such that $g(Q_i)=0, i=1,...,t$. Let $R_1,...,R_P$ be all the zeroes of $g$ among $P_1,...,P_n$. Consider the equations:

$$\sum_{i=l}^{P}f_j(R_i)Z_i = s(u,f_j),\quad j=1,...,r. \qquad (**)$$

LEMMA 8. *The error vector e corresponds to a solution of this system of equations.*

LEMMA 9. *The system of equations (**) has a unique solution, under the assumption that* $(G-F)>t+2g-2$.

PROOF: Suppose we have two solutions to (**) and $e_1$ and $e_2$ are the error patterns corresponding to these solutions. Then $wt(e_1-e_2)\leqslant p$; $e_1-e_2\in\mathrm{Im}(\Omega(G-D))$, since $s(e_1-e_2,f)=0$ for $f\in L(G)$ (cf. Lemma 4). Now $p=\#$ zeroes of $g$ among the points $P_i\leqslant deg(F)<deg(G)-t-2g+2\leqslant$
$\leqslant$ minimum distance of $\mathrm{Im}(\Omega(G-D))$, (Cf. Theorem 1.).

CONCLUSION: $e_1-e_2=0$, hence $e_1=e_2$. Q.E.D.

We have made the following restrictions on the number of errors $t$:
1) In Theorem 6: $t<dim(L(F))$,
2) In Theorem 7 and Lemma 9: $t+2g-2<deg(G-F)$. For the rest, $F$ might be chosen arbitrarily. In order to maximize the number of errors to be corrected by this procedure, we have to choose $F$ such that $\min(deg(F)-g+1, deg(G)-deg(F)-2g+2)$ is as high as possible. This is the case when $deg(F)=\lceil(deg(G)-g+1)/2\rceil$. The number of errors that

can be corrected by this procedure is then:

$$\lfloor (deg(G)-3g+1)/2 \rfloor = \lfloor (\delta_2 - g + 1)/2 \rfloor,$$

where $\delta_2$ is the designed distance of the code $\mathrm{Im}(\Omega(G-D))$.

The previous results are stated in the following theorem.

THEOREM 10. *If $F$ is a divisor of degree $deg(F) = \lceil (deg(G)-g+1)/2 \rceil$, $g_1,...,g_l$ is a basis for $L(F)$, $h_1,...,h_k$ is a basis for $L(G-F)$, then the following procedure is a decoding procedure for the code $\mathrm{Im}(\Omega(G-D))$, which corrects $\lfloor (\delta - g - 1/2 \rfloor$ errors, where $\delta = deg(G) - 2g + 2$ is the designed distance.*

PROCEDURE:

1) Calculate the syndromes: $s_{ij}(u) = s(u, g_i h_j)$; $s(u, f_s)$, $s = 1,...,r$, $i = 1,...,l$, $j = 1,...,k$. The basis $\{f_s\}$ for $L(G)$ can be chosen such that it contains the $\{g_i h_j\}$ as far as they are independent; this diminishes the amount of calculations.
2) Find a non-trivial solution to

$$\sum_{i=1}^{l} s_{ij}(u) X_i = 0 \quad (j = 1,...,k) \qquad (\star)$$

(Existence is guaranteed by Theorem 6 if the number of errors is less than $deg(F)$).
3) If $y_1,...,y_l$ is the non-trivial solution to $(\star)$, then calculate the zeroes $Q_1,...,Q_P$ of the rational function $g = \Sigma_{i=1}^{l} y_i g_i$.
From Theorem 7, it follows that the errors are located among these zeroes).
4) Solve the system of equations:

$$\sum_{i=1}^{P} f_j(Q_j) Z_i = s(u, f_j) \ , \ j = 1,...,r. \qquad (\star\star)$$

(From Lemma 9, there is a unique solution to $(\star\star)$ which corresponds to the error pattern).

REMARKS:

1. For $g = 0$, the number of errors which can be decoded is $\lfloor (\delta - 1)/2 \rfloor$, and from the Singleton bound, the real minimum distance of the code is equal to the designed distance (i.e. $deg(G)+2$). Therefore, the decoding algorithm reaches complete decoding. It appears that this decoding algorithm is equivalent to the one given by Peterson [8].
2. The algorithm can be generalized (extended) in several ways to give better results. One of these ways is by processing the algorithm several times with different $F$'s. This idea is explored in detail by Pellikaan [9].
The second way is to make the estimates more precise:
$dim(L(F)) = deg(F) - g + 1 + dim(\Omega(F))$. The contribution of $dim(\Omega(F))$

will only be positive if $deg(F)<2g-2$. A more detailed discussion can be found in Skorobogatov and Vlădut [7].
A third way will be given below.

Let $F_1$ and $F_2$ be divisors. Let $g_1,...,g_l$ be a basis for $L(F_1)$,$h_1,...,h_k$ be a basis for $L(F_2)$. Consider the space $L$ spanned by $g_ih_ji=1,...,l;j=1,...,k$. It is a subspace of $L(F_1+F_2)$. The space $L$ can be mapped to $GF(q)^n$ by using the same mapping $\phi$ as in the beginning of this section. This mapping is injective. Let $C(L)$ be the code obtained in this way and let $D(L)$ be its dual. Lemma 4 can be generalized to:

LEMMA 11.

$$u\in D(L)\Leftrightarrow s(u,g_ih_j) = 0;\ \ i=1,...,l;\ j=1,...,k.$$

Let $c\in D(L)$ be a codeword. $e$ and error pattern, $wt(e)=t$ and $u$ the received word: $u=c+e$. Suppose $Q_1,...,Q_t$ correspond to the locations where the errors occur. Then Lemma 5 and Theorem 6 hold again. If we replace $G-F$ by $F_2$ in Theorem 7, then this new theorem can be proven in the same way. If we replace the second system of equations (**) by:

$$\sum_{s=1}^{P} g_ih_j(R_s)Z_s = s_{ij}(u) \quad i=1,...,l;\ j=1,...,k, \qquad (***)$$

then Lemma 8 holds for this new system of equations. Furthermore generalizing Lemma 9 we get:

LEMMA 12. *Under the assumptions*
1. $deg(F_2)>t+2g-2$
2. *min dist* $(D(L))>deg(F_1)$,
*the system of equations (***) has a unique solution.*

So, we have the following conditions on $t$:

1) $t<deg(F_1)-g+1=dim(L(F_1))$,
2) $t<deg(F_2)-2g+2$
And a condition on $F_1$:$deg(F_1)<$min *dist* $(D(L))$.

### 3.3. CODES FROM HERMITIAN CURVES

We now will apply the theory from Section 2 to the class of Hermitian curves over $GF(q)$, $q$ a power of 2. The Hermitian curves are plane curves satisfying the Hasse-Weil bound with equality, i.e. the number of rational points is maximal. We consider the projective plane $PG(2,q)$, where $q=r^2$, $r$ and $q$ powers of 2. The Hermitian curve is given by the equation:

$$H:X^{r+1}+Y^{r+1}+Z^{r+1} = 0.$$

The genus of the curve is $g=r(r-1)/2$. There are $r^3+1$ rational points on it (cf. Hirschfeld [10]). In what follows, $y=Y/X$ and $z=Z/X$. We can classify

the points on the curve:

1) Uniformizing parameter $t=y$; Points $P_{y,i}=(1,0,\alpha^{(r-1)i})i=0,\ldots,r$;
2) Uniformizing parameter $t=1/y$; Points $Q=(0,1,1)$
   $P_{1/y,i}=(0,1,\alpha^{(r-1)i})i=1,\ldots,r$;
3) Uniformizing parameter $t=z$; Points $P_{z,i}=(1,\alpha^{(r-1)i},0)i=0,\ldots,r$;
4) Uniformizing parameter $t=\beta y+z$; Points $P_{t,1,i,j}=(1,\alpha^{(r-1)i+i_o},\alpha^{(r-1)j+j_o})$,
   $i_o=1,\ldots,r-2; i=0,\ldots,r; j=0,\ldots,r; j_o$ and $\beta$ are uniquely determined by the equations: $1+\alpha^{i_o(r+1)}=\alpha^{j_o(r+1)}$,
   $\beta=\alpha^{(r-1)(j-i)+(j_o-i_o)}$.

PROOF: The fact that the equation $1+\alpha^{i_o(r+1)}=\alpha^{j_o(r+1)}$ uniquely determines $j_o$ follows form the fact that $1+\alpha^{i_o(r+1)}$ is an $(r-1)^{th}$ root of unity as can easily be verified. It is also easy to see that the given points are indeed points on $H$ and to verify that in each case, $t$ is indeed a uniformizing parameter corresponding to those points. (Cf. Fulton [1], p.70). Furthermore the list contains $(r-2)(r+1)^2+3r+3=r^3+1$ points so it is complete. Q.E.D.

Now we take $G=mQ, D=\Sigma_{i=1}^{r^3}P_i$ where the $P_i$ are the other rational points. According to Theorem 1 of Section 2, the codes from $L(mQ)$ and $\Omega(\Sigma P_i-mQ)$ are dual codes having parameters:
length: $r^3$,
distance: $m-g+1$ resp. $n-m+g-1$.
distance: $n-m$ resp. $m-2g+2$.

We will now construct a basis for the space $L(mQ), 2g-2<m<r^3$. We need the followoing facts:

$$ord_Q^H(X)=1,\ ord_Q^H(Y+Z)=r+1.$$

PROOF: Since $X$ is not a tangent line at $Q$: $ord_Q^H(X)=1$. Now

$$ord_Q^H(Y+Z)=ord_Q^H((Y^{r+1}+Z^{r+1})/(Y^r+Y^{r-1}Z+\cdots+YZ^{r-1}+Z^r))=$$

$$=ord_Q^H(Y^{r+1}+Z^{r+1})=ord_Q^H(X^{r+1})=r+1. \qquad \text{Q.E.D.}$$

Now the following rational functions are linearly independent in $L(mQ)$. $(m>2g-2)$:

$$f_{ij}=X^iY^j/(Y+Z)^{i+j} \quad ,0\leq(i+j)(r+1)-i\leq m, 0\leq i\leq r,\ j\geq 0.$$

PROOF: $ord_Q^H(f_{ij})=i-(i+j)(r+1)$ for $i,j$ in the given ranges. Furthermore $ord_Q^H(f_{ij}\geq 0, P\neq Q, P\in H$. Since the $f_{ij}$ have different orders at $Q$ for different values of $i,j$ in the range $0\leq i\leq r, j\geq 0$, they are linearly independent in $L(mQ)$.

Defining $N(m):=|\{(i,j)|0\leq(i+j)(r+1)-i\leq m,\ 0\leq i\leq r,\ 0\leq j\}|$,

it is easy to see that:

$$N(m) = \begin{cases} a(a+1)/2+a-b+1 & \text{if } a<r,\ b\leqslant a \\ a(a+1)/2 & \text{if } a<r,\ b>a, \\ (r+1)a-r(r+1)/2+r-b+1 & \text{if } a\geqslant r \end{cases}$$

where $a,b$ are determined by $m$ and $r$ as follows:

$$m = a(r+1)-b,\ \ 0\leqslant b\leqslant r.$$

Recalling that $g=r(r-1)/2$ we find (using the above formula), $N(m)=m-g+1$, for $m\geqslant r^2-r-1$. since the dimension of $L(mQ)=m-g+1$ by Riemann-Roch Theorem, and since the $f_{ij}$ are independent in $L(mQ)$ and there are $N(m)=m-g+1$ of them if $m\geqslant r^2-r-1=2g-1$, these functions form a basis for the space $L(mQ)$. This basis allows us to construct a generator matrix and parity check matrix) for the codes involved, since we already classified the rational points on the curve. The next thing we are giving to do is to apply Theorem 2 of the previous section to our codes. In order to do this define:

$$\omega := ((Y+Z)^{r^2-r-1}X(Y^{r+1}+Z^{r+1})/(YZ(Y^{r^2-1}+Z^{r^2-1})))\ (dy+dz).$$

Then:

1) $Res_{P_i}(\omega) = 1$ for all $P_i, i=1,\ldots,r^3$,
2) $ord_Q^H(\omega) = r^3+r^2-r-2$,
3) $div(\omega) = (r^3+r^2-r-2)Q-\Sigma P_i$

PROOF:

1) i) $P = (1,0,\alpha^{(r-1)i})$ (for some $i=0,\ldots,r$). The corresponding uniformizing parameter is $t=y$, and we have:

$$dy+dz=((Y+Z)^r/Z^r)dt.$$

Therefore:

$$Res_P(\omega) =$$

$$Res_P((((Y+Z)^{r^2-r-1}X(Y^{r+1}+Z^{r+1})(Y+Z)^r)/(YZ(Y^{r^2-1}+Z^{r^2-1})Z^r))dt)$$

$$= Res_P((((Y+Z)^{r^2-1}(Y^{r+1}+Z^{r+1}))/(Z^{r+1}(Y^{r^2-1}+Z^{r^2-1})))1/t dt).$$

$$= ((\alpha^{(r-1)i})^{r^2-1}\alpha^{(r^2-1)i})/(\alpha^{(r^2-1)i}(\alpha^{i(r-1)})^{r^2-1}) = 1.$$

ii) $P=(0,1,\alpha^{(r-1)i})$ (for some $i=1,\ldots,r$). The corresponding uniformizing parameter is $t=1/y$ and we have:

$$dy+dz = ((Y+Z)^r)/Z^r)(Y^2/X^2)dt.$$

Therefore:

$$Res_P(\omega)=$$

$$Res_P(((Y+Z)^{r^2-r-1}X(Y^{r+1}+Z^{r+1})\ (Y+Z)^r)Y^2/(YZ(Y^{r^2-1}+Z^{r^2-1})Z^rX^2))dt)$$

$= Res_P((((Y+Z)^{r^2-1}/(Z^{r+1}(Y^{(r+1)(r-2)}+Y^{(r+1)(r-3)}Z+\cdots+Z^{(r+1)(r-2)})dt/t)$

$= (1+\alpha^{(r-1)i})^{r^2-1}/\alpha^{i(r^2-1)}(1+\alpha^{i(r^2-1)}+\cdots+\alpha^{i(r^2-1)(r-2)}) = 1.$

iii) $P=(1,\alpha^{(r-1)i},0)$ (for some $i=0,\ldots,r$).
In the same way as under i) we find $Res_P(\omega)=1$.

iv) $P=(1,\alpha^{(r-1)i+i_o},\alpha^{(r-1)j+j_o})$, (for some $i_o=1,\ldots,r-2$; $i=0,\ldots,r$; $j=0,\ldots,r$; $j_o$ and $\beta$ are uniquely determined as before). The corresponding uniformizing parameter is $t=\beta y+z$ and we have that: $dy+dz=((y^r+z^r)/(\beta z^r+y^r))dt$. Define $y_P=\alpha^{(r-1)i+i_o}$ and $z_P=\alpha^{(r-1)j+j_o}$ and note that $\beta=z_P/y_P$.
Now:

$Res_P(\omega) = Res_P((y+z)^{r^2-r-1}(y^{r+1}+z^{r+1})/(yz(y^{r^2-1}+z^{r^2-1})\,(dy+dz)) =$

$Res_P((y+z)^{r^2-1}(y^{r+1}+z^{r+1})(\beta y+z)/(yz^{r+1}(\beta+(y/z)^r)\,(y^{r^2-1}+z^{r^2-1}))dt/t) =$

$Res_P((y+z)^{r^2-1}(y^{r+1}+z^{r+1})(\beta+(z/y))/(z^{r+1}(\beta+(y/z)^r)y^{r^2-1}(\beta^{r^2-1}+(z/y)^{r^2-1})$

$*dt/t)$ (Since $\beta^{r^2-1}=1$).

$= (y_P+z_P)^{r^2-1}(y_P^{r+1}+z_P^{r+1})/(z_P^{r+1}(\beta+(1/\beta)^r)y_P^{r^2-1}(\beta^{r^2-1}+\ldots+(z_P/y_P)^{r^2-2})$

$= 1/(z_P^{r+1}(1+(1/\beta)^{r+1}\beta\cdot\beta^{r^2-2}) = 1/(y_P^{r+1}+z_P^{r+1}) = 1.$

2)

$ord_Q^H(\omega) = ord_Q^H((Y+Z)^{r^2-r-1})X(Y^{r+1}+Z^{r+1})/(YZ(Y^{r^2-1}+Z^{r^2-1})(dy+dz)) =$

$ord_Q^H((Y+Z)^{r^2-r-1}X(Y^{r+1}+Z^{r+1})/(Y^{r^2-1}+Z^{r^2-1})) + ord_Q^H(dy+dz) =$

$= (r^2-r-1)(r+1) + 1 + r^2-r-2 = r^3+r^2-r-2.$

Since $ord_Q^H(dy+dz)=ord_Q^H(((z+y)^r/z^r)dy)=ord_Q^H(((z+y)^r/z^r)*1/t^2dt)$ (where $t$ is the uniformizing parameter at $Q$) $=r(r+1)-2=r^2+r-2$.

3) Follows from 1) and 2).

We can now conclude, using the above facts and Theorem 2 of the previous section:

THEOREM 1. *The code constructed from $L(mQ)$ is weakly self-dual, if $m<(r^3+r^2-r-2)/2$. The code constructed from $L(mQ)$ is self-dual if $m=(r^3+r^2-r-2)/2$.*

To gain some information about the automorhism group of the code recall that the automorhism group of the curve is $PGU(3,r^2)$, which has order $(r^3+1)r^3(r^2-1)$ (see Hirschfeld [10] p. 147). Therefore the stabilizer of the point $Q$ has order $r^3(r^2-1)$. Applying Theorem 3 of the previous section we get:

THEOREM 2. *The stabilizer of $Q=(0,1,1)$ in $PGU(3,r^2)$ is a subgroup of the automorphism group of the code obtained from $L(mQ)$. It has order $r^3(r^2-1)$.*

*Decoding*
Applying the decoding algorithm described in the previous section becomes very simple in this case: choose $F=\lfloor (m-g+1)/2 \rfloor Q$.
For $g_1,\ldots,g_l$ we can take the functions as decribed before, $h_1,\ldots,h_k$ we can take the functions as described before. (Because both $L(F)$ and $L(G-F)$ are of the form $L(aQ)$). Then apply the algorithm. We now wish to look at the third way of extending the algorithm: suppose that we want to decode $t$ errors. Choose $F_1=F_2=(t+1)Q$. A basis for $L(F_1)$: the first $t+1$ $f_{ij}$'s. $L(F_2)$ has the same basis.
Therefore (if $t>1$), $<L(F_1)L(F_2)>$ has a basis: the first $2(t+1)$ $f_{ij}$'s. Hence: $<L(F_1)L(F_2)>=L(2(t+1)Q = L(F_1+F_2)$.
The minimum distance of the dual code is $2(t+1)-2g+2 \geqslant t+1=deg(F_1)$. So the decoding method works (but gives nothing new).

EXAMPLE 1. Let $q=4=r^2$. Let $\omega$ be a primitive element of $GF(4)$. The Hermitian curve $H$ with equation $X^3+Y^3+Z^3=0$ has no multiple points and hence genus 1. The number of rational points of the curve is 9. We can describe them by: $Q=(0,1,1)$; $P_1=(1,0,\bar{\omega})$; $P_2=(1,0,\omega)$; $P_3=(1,0,1)$; $P_4=(1,\bar{\omega},0)$; $P_5=(1,\omega,0)$; $P_6=(1,1,0)$; $P_7=(0,\bar{\omega},1)$ and $P_8=(0,\omega,1)$. It is easy to see that the permuations (162435)(78) and (138467)(25) stabilizer $Q$ and are in the unitary group. They correspond to the linear mappings

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \omega \\ 0 & \omega & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 1 \\ \bar{\omega} & \omega & 1 \\ \bar{\omega} & 1 & \omega \end{bmatrix} \quad \text{respectively.}$$

Define: $f_1=1, f_2=X/(Y+Z)$, $f_3=Y/(Y+Z)$, $f_4=X^2/(Y+Z)^2$, $f_5=XY/(Y+Z)^2$ $f_6=Y^2/(Y+Z)^2$, $f_7=X^2Y/(Y+Z)^3$, $f_8=XY^2/(Y+Z)^3$.
Then $ord_Q^H(f_i)=-i$, $i\neq 1$; $ord_Q^H(f_1)=0$.
So the bases for the spaces $L(mQ)$ can be given by $\{f_i\}_{i=1}^{m}, 1\leqslant m\leqslant 8$. The following table contains information about weight distribution, generator matrices and automorphism groups of the codes $C$ corresponding to $L(mQ)$, $1\leqslant m\leqslant 4$. The weight distributions in Table 1 are found by inspection. $C_1$ has automorphism group $S_8$. $C_2$ has minimum distance 6. Its automorphism group is of order 192. This can be seen as follows. The stabilizer of the codeword $(01\bar{\omega}01\bar{\omega}\omega\omega)$ in the automorphism group of the code is: $<(14),(25),(36),(78)>$. This stabilizer is isomorphic to $C_2*C_2*C_2*C_2$ and hence has order 16. From the complete weight distribution of this code it follows that the orbit of the given word can have at most 12 elements. Since $<(162435)(78), (138467)(25)>$ is a subgroup of the automorphism group of the code, we find again by inspection, that all codewords having the same weight structure are in the orbit.

TABLE I. Some facts about Codes $C_m, 1 \leqslant m \leqslant 4$

| | | Complete Weight Distribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Space | Generator Matrix | 0 | 1 | $\omega$ | $\bar{\omega}$ | Number | Weight | Number | Order of Automorphism Group |
| $L(Q)$ | 11111111 | 8 | 0 | 0 | 0 | 1 | 0 | 1 | 8!=40320 |
| | | 0 | 8 | 0 | 0 | 1 | | | |
| | | 0 | 0 | 8 | 0 | 1 | 8 | 3 | |
| | | 0 | 0 | 0 | 8 | 1 | | | |
| $L(2Q)$ | 11111111 | 8 | 0 | 0 | 0 | 1 | 0 | 1 | |
| | $\omega\bar{\omega}1\omega\bar{\omega}100$ | 0 | 8 | 0 | 0 | 1 | | | |
| | | 0 | 0 | 8 | 0 | 1 | 8 | 3 | 192 |
| | | 0 | 0 | 0 | 8 | 1 | | | |
| | | 2 | 2 | 2 | 2 | 12 | 6 | 12 | |
| $L(3Q)$ | 11111111 | 8 | 0 | 0 | 0 | 1 | 0 | 1 | |
| | $\omega\bar{\omega}1\omega\bar{\omega}100$ | 0 | 8 | 0 | 0 | 1 | | | |
| | $000111\omega\bar{\omega}$ | 0 | 0 | 8 | 0 | 1 | 8 | 3 | |
| | | 0 | 0 | 0 | 8 | 1 | | | |
| | | 1 | 1 | 3 | 3 | 8 | | | |
| | | 1 | 3 | 1 | 3 | 8 | 7 | 24 | 24 |
| | | 1 | 3 | 3 | 1 | 8 | | | |
| | | 2 | 2 | 2 | 2 | 12 | 6 | 12 | |
| | | 3 | 1 | 3 | 1 | 8 | | | |
| | | 3 | 3 | 1 | 1 | 8 | 5 | 24 | |
| | | 3 | 1 | 1 | 3 | 8 | | | |
| $L(4Q)$ | 11111111 | 8 | 0 | 0 | 0 | 1 | 0 | 1 | |
| | $\omega\bar{\omega}1\omega\bar{\omega}100$ | 0 | 8 | 0 | 0 | 1 | | | |
| | $000111\omega\bar{\omega}$ | 0 | 0 | 8 | 0 | 1 | | | |
| | $\omega\bar{\omega}1\omega\bar{\omega}100$ | 0 | 0 | 0 | 8 | 1 | | | |
| | | 0 | 4 | 4 | 0 | 6 | 8 | 21 | |
| | | 0 | 4 | 0 | 4 | 6 | | | |
| | | 0 | 0 | 4 | 4 | 6 | | | |
| | | 1 | 3 | 3 | 1 | 32 | | | |
| | | 1 | 3 | 1 | 3 | 32 | 7 | 96 | 192 |
| | | 1 | 1 | 3 | 3 | 32 | | | |
| | | 2 | 2 | 2 | 2 | 24 | 6 | 24 | |
| | | 3 | 3 | 1 | 1 | 32 | | | |
| | | 3 | 1 | 3 | 1 | 32 | 5 | 96 | |
| | | 3 | 1 | 1 | 3 | 32 | | | |
| | | 4 | 4 | 0 | 0 | 6 | | | |
| | | 4 | 0 | 4 | 0 | 6 | 4 | 18 | |
| | | 4 | 0 | 0 | 4 | 6 | | | |

We conclude that the order of the automorphism group is $16 \times 12 = 192$. $C_3$ has minimum distance five. Its automorphism group is of order 24. This can be seen as follows. The stabilizer of $(\omega\bar{\omega}1\,\omega\bar{\omega}1\,0\,0)$ contains just two elements: (1) and (14) (25) (36) (78), and the orbit of this word has 12 elements (both facts by inspection). In this case the automorphism group equals the stabilizer of $Q$ in the unitary group. $C_4$ has minimum distance four. Its automorphism group is of order 192. This can be seen as follows. The stabilizer of $(\omega\ \bar{\omega}\ 1\ \omega\ \bar{\omega}\ 1\ 0\ 0)$ is equal to $<(14)(25), (14)(36), (14)(78)>$ (by inspection). Therefore it is isomorphic to $C_2 * C_2 * C_2$ and has order 8. Again by inspection all 24 codewords having the same weight structure are in the orbit of this word. We conclude that the order of the automorphism group is $8 * 24 = 192$.

DECODING

The correct three errors, choose $F_1=F_2=4Q$. $L(F_1+F_2)=L(8Q)$. The dual code is 0 which is not particularly interesting to decode. To correct two errors, choose $F_1=F_2=3Q$, $L(F_1+F_2)=L(6Q)$. The dual code is in this case $C_2$.

A basis for $C_2$ is: $f_1$ $\begin{bmatrix}1&1&1&1&1&1&1&1\\ \omega&\bar{\omega}&1&\omega&\bar{\omega}&1&0&0\end{bmatrix}$ $f_2$

parity check matrix is:

$$\begin{matrix}f_1\\f_2\\f_3\\f_4\\f_5\\f_6\end{matrix}\begin{bmatrix}1&1&1&1&1&1&1&1\\ \omega&\bar{\omega}&1&\omega&\bar{\omega}&1&0&0\\ 0&0&0&1&1&1&\omega&\bar{\omega}\\ \bar{\omega}&\omega&1&\bar{\omega}&\omega&1&0&0\\ 0&0&0&\omega&\bar{\omega}&1&0&0\\ 0&0&0&1&1&1&\bar{\omega}&\omega\end{bmatrix}$$

Storage of this matrix costs $2*6*8=96$ memory bits. Decoding proceeds as follows:

1) Calculate $s_1,\ldots,s_6$. This costs 14 multiplications and 27 additions.
2) find a non-trivial solution of:

$$\begin{bmatrix}s_1&s_2&s_3\\ s_2&s_4&s_5\\ s_3&s_5&s_6\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0.$$

   a) Calculate

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}s_2s_5+s_3s_4\\ s_1s_5+s_2s_3\\ s_1s_4+s_2s_2\end{bmatrix}.$$

   Cost: 6 multiplications and 3 additions.
   If $(y_1,y_2,y_3)\neq 0$ then test whether $s_3y_1+s_5y_2+s_6y_3=0$.
   If 0 then go to step 3 else alarm.

   b) If $(y_1,y_2,y_3)=0$ calculate $y_1=s_4s_6+s_5s_5$, $y_2=s_2s_6+s_3s_5$, $y_3=0$.
   Cost: 4 multiplications and 2 additions. If $(y_1,y_2,y_3)=0$ then c) else step 3.

   c) $(y_1,y_2,y_3)=(s_2,s_1,0)$. go to step 3.
   Total cost of step 2: ca. 10 multiplications and 5 additions.

3) Calculate the zeroes of $g=y_1f_1+y_2f_2+y_3f_3$.

$$(y_1,y_2,y_3)\begin{bmatrix}1&1&1&1&1&1&1&1\\ \omega&\bar{\omega}&1&\omega&\bar{\omega}&1&0&0\\ 0&0&0&1&1&1&\omega&\bar{\omega}\end{bmatrix}$$

4) Solve the equations (**) corresponding to the error locations found under step 3.
Cost: ca. 13 multiplications and 12 additions.

The total cost of this decoding method is: 96 memory bits, 43 multiplications and 55 additions. Decoding method 2 for this code: calculate all words of the code and search among $u+c$ the word with the lowest weight. Storage of all words costs $2*16*8=256$ memory bits. Calculation of $u+c$ for every codeword costs $8*16=128$ additions. Therefore the cost of this method is worse than the total cost of the above method. Decoding method 3 for this code:
Precompute a parity check matrix of the form $(I_6 H)$.
The storage of this matrix costs: $2*12=24$ memory bits. Find an automorphism $\pi_1$ permuting the coordinates in such a way that the last two coordinates move into the six front places. Find a second automorphism which moves the coordinates in such a way that the last two coordinates and the positions $\pi_1^{\leftarrow}(7)$ and $\pi_1^{\leftarrow}(8)$ are among the first six places. Choose for example $\pi_1=(138467)(25)$ and $\pi_2=(186)(347)$. Compute the syndrome. Cost: 12 multiplications and 12 additions. If the weight of the syndrome $\leqslant 2$ then the error is known; otherwise permute coordinates according to $\pi_1$ and compute syndromes again. The cost is again 12 multiplications and 12 additions. Again if the weight of the syndrome $\leqslant 2$ then the error is known; otherwise there are too many errors. The average number of computations is: $(15*12+9*24+4*36)/28=19.28$. multiplications and the same amount of additions. So this method is to be preferred above the other two methods in this case. To correct one error choose $F_1$ and $F_2$ both equal to $2Q$. $L(F_1+F_2)=L(4Q)$. A basis for $<L(F_1)L(F_2)>$ is: $f_1, f_2, f_4$. A parity check matrix for the dual code corresponding to this space is:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \omega & \bar{\omega} & 1 & \omega & \bar{\omega} & 1 & 0 & 0 \\ \bar{\omega} & \omega & 1 & \bar{\omega} & \omega & 1 & 0 & 0 \end{bmatrix}.$$

Now for instance the word (10010000) is in this code so the minimum distance of this code is less than or equal to 2 and hence less than or equal to $deg(F_2)$. Therefore the last part of Section 2 cannot be applied. Let us consider $L(4Q)$. It has basis $\{f_1, f_2, f_3, f_4\}$. Its minimum distance is 4. A basis for $L(2Q)=\{f_1, f_2\}$. We can apply Theorem 10 of Section 2 to decode the code: choose $F=2Q$. A useful parity check matrix for the code is:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \omega & \bar{\omega} & 1 & \omega & \bar{\omega} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \omega & \bar{\omega} \\ \bar{\omega} & \omega & 1 & \bar{\omega} & \omega & 1 & 0 & 0 \end{bmatrix} \begin{matrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{matrix}$$

Storage of this parity check matrix costs $2\times 8\times 4=64$ memory bits. The decoding algorithm of Section 2 proceeds as follows:

1) calculate the syndromes $s_1, s_2, s_3, s_4$. This costs: 10 multiplications and 21 additions.
2) Find a non-trivial solution of:

$$\begin{bmatrix} s_1 & s_2 \\ s_2 & s_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

Calculate: $s_1 s_4 + s_2^2$ (Cost: 2 multiplication and 1 addition),
If the result is 0, then $x_1 = s_2$ and $x_2 = s_1$.
If the result is not 0, then there are too many errors.
3) Calculate the zeroes of $g = x_1 f_1 + x_2 f_2$ by computing:

$$(x_1 x_2) \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \omega & \bar{\omega} & 1 & \omega & \bar{\omega} & 1 & 0 & 0 \end{bmatrix}.$$

(Cost: 4 multiplications, 8 additions and 32 bits of memory.)
4) Solve the corresponding equations (**). Cost: 4 multiplications and 3 additions.

Total cost of this algorithm is: 96 bits of memory, 20 multiplications and 33 additions.
Decoding method 2: Store the generator matrices (and at the same time parity check matrices).

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & \bar{\omega} & \bar{\omega} \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & \omega & \bar{\omega} \end{bmatrix} \quad H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ \bar{\omega} & 1 & 1 & \omega & 0 & 0 & 1 & 0 \\ \omega & 1 & 1 & \bar{\omega} & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Cost: $2*4*8 = 64$ bits of memory. Calculate there syndromes $s = Gu$ and $t = Hu$.
Cost: 8 multiplications and 28 additions. If $wt(s) \leqslant 1$, then the error is known. If $wt(s) > 1$ and $wt(t) \leqslant 1$, then the error is known. Otherwise there are too many errors.
Total cost: 64 bits of memory, 8 multiplications and 28 additions. So again the last algorithm has my preferences. However, in larger examples the first algorithm will do better probably.

### 3.4. Other examples over $GF(4)$

In this section we wish to study some examples of codes over $GF(4)$. A curve of degree 1 is non-singular, has genus 0 and 5 rational points. The codes constructed from such a curve are Reed-Solomon codes. A non-singular curve of degree 2 over $GF(4)$ also has genus 0. so this gives nothing new.

A non-singular curve of degree 3 over $GF(4)$ has genus $g = 1/2*2*1 = 1$. The Hasse-Weil bound gives in this case that there are less than or equal to 9 rational on the curve. This number of rational points is realized by the Hermitian curve and this example was studied extensively in the previous section. We

are now going to look at non-singular plane curves of degree 4. A non-singular plane curve of degree 4 has genus $g=3$. The Hasse-Weil bound gives that it contains less than or equal to 17 rational points over $GF(4)$. We shall now prove that a non-singular plane curve of degree 4 contains less than or equal to 14 points. Let $F$ be a non-singular plane curve of degree 4. Since $F$ is irreducible it contains no lines. Therefore every line contains a nonzero of $F$ and so the nonzeros of $F$ form a blocking set. (A blocking set is a set of points such that every line contains at least one point of the set.). A blocking set not containing a line has $\geqslant 7$ points. Of course this is a well known fact (cf. [13]), but since the proof is not too long, we shall give it. Take $P_1, P_2$ in the blocking set. Let $l$ be the line through these two points. Then $l$ contains a point $Q$ not in the blocking set. Otherwise $l$ would be contained in the blocking set. Through $Q$ there are 4 other lines and each of these lines contains a point of the blocking set. So the number of points in the blocking set is $\geqslant 6$. ($P_1$ and $P_2$ and at least 4 other points). Suppose there exists a blocking set not containing a line and having 6 points exactly. Then every line hits the blocking set in at most 2 points. (If there is a line having 3 or 4 points of the blocking set choose a point on this line not in the blocking set, there are 4 other lines through this point and each of them has to contain a point of the blocking set, so the blocking set has $\geqslant 7$ points.). Let $P_1$ and $P_2$ be two points in the blocking set and $l$ the line through these points. Let $Q_1$ and $Q_2$ be 2 other points in the blocking set and $m$ be the line through these 2 points. The lines $l$ and $m$ intersect in a point $Q$ not in the blocking set since every line contains at most 2 points of the blocking set. There are 3 lines going trough $Q$ and different from $l$ and $m$. On each of these 3 lines we must find a point of the blocking set making a total of at least 7 points in the blocking set, which contradicts the assumption. So we may conclude: a blocking set in $PG(2,4)$ not containing a line has $\geqslant 7$ points.

Therefore it follows that a non-singular plane curve in $PG(2,4)$ of degree 4 and genus 3 has $\leqslant 14$ rational points. The following curve realizes 14. It is the complement of a Baer-subplane.

$$F(X,Y,Z) = X^4+Y^4+Z^4+X^2Y^2+Y^2Z^2+X^2Z^2+X^2YZ+XY^2Z+XYZ^2.$$

The rational points are:

$$Q=(0,1,\omega);P_1=(0,1,\bar{\omega});P_2=(1,0,\omega);P_3=(1,0,\bar{\omega});P_4=(1,1,\omega);P_5=(1,1,\bar{\omega});$$

$$P_6=(1,\omega,0);P_7=(1,\omega,1);P_8=(1,\omega,\omega);P_9=(1,\omega,\bar{\omega});P_{10}=(1,\bar{\omega},0);P_{11}=(1,\bar{\omega},1)$$

$P_{12}=(1,\bar{\omega},\omega);P_{13}=(1,\bar{\omega},\bar{\omega})$. (i.e. all points having at least one component not in $GF(2)$).

In the plane there are 21 lines. They intersect the curve according to the following table:

| LINE | | INTERSECTION DIVISOR | TANGENT |
|---|---|---|---|
| $L_1$ | $X$ | $2Q+2P_1$ | $Y$ |
| $L_2$ | $X+Y$ | $2P_4+2P_5$ | $Y$ |
| $L_3$ | $X+\omega Y$ | $P_{10}+P_{11}+P_{12}+P_{13}$ | $N$ |
| $L_4$ | $X+\bar{\omega}Y$ | $P_6+P_7+P_8+P_9$ | $N$ |
| $L_5$ | $X+Z$ | $2P_7+2P_{11}$ | $Y$ |
| $L_6$ | $X+Y+Z$ | $2P_9+2P_{12}$ | $Y$ |
| $L_7$ | $X+\omega Y+Z$ | $P_5+P_8+P_{10}+Q$ | $N$ |
| $L_8$ | $X+\bar{\omega}Y+Z$ | $P_1+P_4+P_6+P_{13}$ | $N$ |
| $L_9$ | $X+\omega Z$ | $P_3++P_5+P_{13}+P_9$ | $N$ |
| $L_{10}$ | $X+Y+\omega Z$ | $P_1+P_3+P_8+P_{11}$ | $N$ |
| $L_{11}$ | $X+\omega Y+\omega Z$ | $P_3+P_7+P_{10}+P_{11}$ | $N$ |
| $L_{12}$ | $X+\bar{\omega}Y+\omega Z$ | $Q+P_3+P_6+P_{12}$ | $N$ |
| $L_{13}$ | $X+\bar{\omega}Z$ | $P_2+P_4+P_8+P_{12}$ | $N$ |
| $L_{14}$ | $X+Y+\bar{\omega}Z$ | $Q+P_2+P_7+P_{13}$ | $N$ |
| $L_{15}$ | $X+\omega Y+\bar{\omega}Z$ | $P_1+P_2+P_9+P_{10}$ | $N$ |
| $L_{16}$ | $X+\bar{\omega}Y+\bar{\omega}Z$ | $P_2+P_5+P_6+P_{11}$ | $N$ |
| $L_{17}$ | $Y$ | $2P_2+2P_3$ | $Y$ |
| $L_{18}$ | $Z$ | $2P_6+2P_{10}$ | $Y$ |
| $L_{19}$ | $Y+\omega Z$ | $P_1+P_5+P_7+P_{12}$ | $N$ |
| $L_{20}$ | $Y+Z$ | $2P_8+2P_{13}$ | $Y$ |
| $L_{21}$ | $Y+\bar{\omega}Z$ | $Q+P_4+P_9+P_{11}$ | $N$ |

TABLE 1. Intersection divisors of the lines with the curve.

The tangent lines are the lines of the Baer-subplane.
From this table it can be seen that the following functions have the described divisors:

| | function $f$ | $div(f)$ |
|---|---|---|
| $f_0$ | 1 | 0 |
| $f_1$ | $(L_8L_{10}L_{15}L_{19})/(L_1^2L_{11}L_{16})$ | $-4Q+P_{13}+P_8+P_9+P_{12}$ |
| $f_2$ | $(L_8^2L_{10}^2L_{15}L_{19}L_{14})/(L_1^3L_4L_{11}L_{13}L_{16})$ | $-5Q+P_3+P_{11}+3P_{13}$ |
| $f_3$ | $(L_8^2L_{10}^2L_{15}L_{19})/(L_1^3L_3L_4L_{16})$ | $-6Q+2P_3+2P_4+P_8+P_{13}$ |
| $f_4$ | $(L_8^2L_{10}^2L_{15}^2L_{19}^2)/(L_1^4L_2L_5L_6L_{17})$ | $-7Q+P_3+3P_6+2P_{10}+P_{12}$ |

TABLE 2. Some important functions and their divisors.

Define $f_5=f_1^2, f_6=f_1f_2, f_7=f_1f_3, f_8=f_1f_4, f_9=f_1^3$. Then $ord_Q(f_i)=-(i+3)$ $(i\neq 0)$. Therefore $L(mQ)$ has as a basis $\{f_0,\ldots,f_{m-3}\}$. $(m=3,4,\ldots,12)$. In order to calculate the values of the $f_j$ at the points $P_1,\ldots,P_{13}$ we remark that for a point $P_i$ we can choose a nontangent $L_{i1}$ passing through the point and a line $L_{i2}$ not through $P_i$. Then $u_i=L_{i1}/L_{i2}$ is a uniformizing parameter for $P_i$.

Now since $f_j$ is of the form $(N_1....N_t)/(M_1....M_t)$ where the $N_1,...,N_t$ $M_1,...,M_t$ are polynomials of degree 1, we can write:

$$f_j = ((N_1/L_{i2}).....(N_t/L_{i2}))/((M_1/L_{i2}).....(M_t/L_{i2})).$$

The value of $f_j$ at $P_i$ is uniquely determined by the first terms in the power series of $f_j$ in $u_i$, which can be easily calculated once the first terms of the powerseries of $L/L_{i2}$ are known for all relevant lines $L$.

After tedious calculations we found for the values of the $f_i$ at the $P_j$:

| Point | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ | $P_8$ | $P_9$ | $P_{10}$ | $P_{11}$ | $P_{12}$ | $P_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Function | | | | | | | | | | | | | |
| $f_0$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $f_1$ | $\bar{\omega}$ | $\omega$ | $\omega$ | $\omega$ | $\omega$ | $\bar{\omega}$ | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| $f_2$ | $\bar{\omega}$ | $\omega$ | 0 | 1 | $\bar{\omega}$ | $\omega$ | 1 | $\bar{\omega}$ | $\omega$ | $\bar{\omega}$ | 0 | 1 | 0 |
| $f_3$ | $\bar{\omega}$ | $\bar{\omega}$ | 0 | 0 | $\bar{\omega}$ | $\bar{\omega}$ | $\omega$ | 0 | $\omega$ | $\omega$ | $\bar{\omega}$ | $\omega$ | 0 |
| $f_4$ | $\bar{\omega}$ | $\omega$ | 0 | $\omega$ | 1 | 0 | $\omega$ | $\omega$ | $\bar{\omega}$ | 0 | 1 | 0 | $\omega$ |
| $f_5$ | $\omega$ | $\bar{\omega}$ | $\bar{\omega}$ | $\bar{\omega}$ | $\bar{\omega}$ | $\omega$ | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| $f_6$ | $\omega$ | $\bar{\omega}$ | 0 | $\omega$ | 1 | 1 | 1 | 0 | 0 | $\bar{\omega}$ | 0 | 0 | 0 |
| $f_7$ | $\omega$ | 1 | 0 | 0 | 1 | $\omega$ | $\omega$ | 0 | 0 | $\omega$ | $\bar{\omega}$ | 0 | 0 |
| $f_8$ | $\omega$ | $\bar{\omega}$ | 0 | $\bar{\omega}$ | $\omega$ | 0 | $\omega$ | 0 | 0 | 0 | 1 | 0 | 0 |
| $f_9$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |

TABLE 3. The basis functions and their values at the points.

This table enables us to write down generator matrices for the codes corresponding to $L(mQ)$. Since the length of the codes is odd and the all one word is in the code is odd and the all one word is in the code, the code is not weakly self-dual. From Theorem 1 of Section 2 we get the following parameters:

| Space | dimension | distance | dual dimension | dual distance |
|---|---|---|---|---|
| $L(mQ)$ | $m-2$ | $\geqslant 13-m$ | $15-m$ | $\geqslant m-4$ |

The code obtained from $L(5Q)$ has generator matrix equivalent to:

$$1\ 1\ \bar{\omega}\ 0\ 0\ 0\ 1\ \omega\ \bar{\omega}\ \bar{\omega}\ 0\ 0\ 1$$

$$\bar{\omega}\ \omega\ 0\ 1\ \bar{\omega}\ \omega\ 1\ \bar{\omega}\ \omega\ \bar{\omega}\ 0\ 1\ 0$$

$$\bar{\omega}\ \omega\ \omega\ \omega\ \omega\ \bar{\omega}\ 1\ 0\ 0\ 1\ 1\ 0\ 0$$

And therefore the words are multiplies of the following:

$$
\begin{array}{cccccccccccc}
1 & 1 & \bar{\omega} & \omega & 0 & 0 & 1 & \omega & \bar{\omega} & \bar{\omega} & 0 & 0 & 1 \\
\omega & \bar{\omega} & \bar{\omega} & \bar{\omega} & \bar{\omega} & \omega & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & \omega & \bar{\omega} & 0 & 1 & \bar{\omega} & \bar{\omega} & \bar{\omega} & 0 & \omega & 0 & \omega & 1 \\
\bar{\omega} & 0 & \bar{\omega} & 1 & \omega & 1 & \omega & 0 & \omega & 1 & 0 & \bar{\omega} & 1 \\
\omega & \bar{\omega} & 1 & 0 & \omega & \bar{\omega} & 0 & \omega & \bar{\omega} & \omega & 1 & 0 & 1 \\
0 & \omega & 0 & 1 & \bar{\omega} & 1 & \bar{\omega} & \omega & \bar{\omega} & 1 & \omega & 0 & 1 \\
\bar{\omega} & 0 & \bar{\omega} & \bar{\omega} & 1 & \omega & \omega & \omega & \bar{\omega} & 0 & \bar{\omega} & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\bar{\omega} & 0 & 0 & 0 & 0 & \bar{\omega} & \omega & 1 & 1 & \omega & \omega & 1 & 1 \\
0 & \omega & \omega & \omega & \omega & 0 & \bar{\omega} & 1 & 1 & \bar{\omega} & \bar{\omega} & 1 & 1 \\
\bar{\omega} & 0 & 1 & \omega & \bar{\omega} & 0 & \omega & \bar{\omega} & 0 & \bar{\omega} & 1 & \omega & 1 \\
1 & 1 & 0 & \bar{\omega} & \omega & \omega & 1 & \bar{\omega} & 0 & 0 & \omega & \omega & 1 \\
\omega & \bar{\omega} & \omega & 1 & 0 & 1 & 0 & \bar{\omega} & 0 & 1 & \bar{\omega} & \omega & 1 \\
0 & \omega & 1 & \bar{\omega} & 0 & \omega & \bar{\omega} & 0 & \omega & 0 & 1 & \bar{\omega} & 1 \\
\omega & \bar{\omega} & 0 & \omega & 1 & 0 & 0 & 0 & \omega & \bar{\omega} & \omega & \bar{\omega} & 1 \\
1 & 1 & \omega & 0 & \bar{\omega} & \bar{\omega} & 1 & 0 & \omega & \omega & \bar{\omega} & \bar{\omega} & 1 \\
\bar{\omega} & \omega & 0 & 1 & \bar{\omega} & \omega & 1 & \bar{\omega} & \omega & \bar{\omega} & 0 & 1 & 0 \\
0 & 0 & \omega & \bar{\omega} & 1 & 1 & 0 & \bar{\omega} & \omega & \omega & 1 & 1 & 0 \\
\omega & 1 & \bar{\omega} & \omega & 0 & \bar{\omega} & \bar{\omega} & \bar{\omega} & \omega & 1 & \omega & 1 & 0 \\
1 & \bar{\omega} & 1 & 0 & \omega & 0 & \omega & \bar{\omega} & \omega & 0 & \bar{\omega} & 1 & 0 \\
\bar{\omega} & \omega & \omega & \omega & \omega & \bar{\omega} & 1 & 0 & 0 & 1 & 1 & 0 & 0
\end{array}
$$

Therefore we may conclude that the real minimum distance of this code is 9, which is bigger than the estimate we made by using the Riemann-Roch Theorem. The code obtained from $L(12Q)$ has distance 1, indeed the function

$$f = (L_8^3 L_{10}^3 L_{15}^3 L_{19}^3)/(L_1^6 L_3 L_4 L_9 L_{11} L_{13} L_{16}) \in L(12Q),$$

since

$$div(f) = -12Q + \sum_{i=1}^{13} P_i.$$

Moreover this shows $wt(f)=1$. Therefore the minimum distance being bigger for $L(5Q)$ is not a property of the curve.

This example showed three things:

1) the Hasse-Weil bound cannot always be reached by taking non-singular plane curves;
2) the minimum distance is sometimes bigger than the designed distance and this does not depend on specific properties of the curve;
3) there are a lot of difficulties if one wants to construct the codes from the curves explicitely, particularly in finding a basis for the space $L(G)$, and evaluating the functions of the basis at the points.

REFERENCES

1. W. FULTON (1969). *Algebraic curves,* Mathematics Lecture Notes series, The Benjamin/Cummings publishing company, Inc..
2. R.D. HARSTHORNE *Algebraic Geometry,* Graduate texts in Mathematics Vol. 52 Springer Verlag 1977.
3. J.H. VAN LINT, T.A. SPRINGER (1987). *Generalized Reed-Solomon codes from algebraic geometry,* IEEE Trans. on Information Theory Vol. IT-33 pp. 305-309.
4. H.J. TIERSMA (1987). *Remarks on codes from Hermitian Curves,* IEEE Trans. on Information Theory, Vol. IT-33 pp. 605-609.
5. H. STICHTENOTH, Y. DRIENCOURT *A criterion for self duality of geometric codes,* to appear.
6. J. JUSTESEN, K.J. LARSEN, H.E. JENSEN, A. HAVEMOSE, T. HØHOLDT (1988). *Construction and decoding fo algebraic geometric codes,* Mat. Rep. 1988-10, Danmarks Tekniske Hojskole.
7. A.N. SKOROBOGATOV, S.G. VLĂDUT (1988). *On the decoding of algebraic-geometric codes,* preprint Inst. for Problems of Information and Transmission.
8. H.H. PETERSON, E.J. WELDON JR. (1979). *Error correcting codes,* Second edition, MIT Press.
9. R. PELLIKAAN (1988). *On a decoding algorithm for codes on maximal curves,* to appear.
10. J.H.P. HIRSCHFELD *Projective geometries over finite fields,* Oxford university Press 1979.
11. S. LANG (1982). *Introduction to algebraic and abelian functions,* Graduate texts in mathematics 89. Springer Verlag.
12. J.P. SERRE (1983). *Sur le nombre des points rationelles d'une courbe algebrique sur un corps fini,* C.R. Acad. Sc. Paris, t. 296 (7 Mars 1983), Serie I-pp. 397-402.
13. A. BLOKHUIS, A.E. BROUWER (1986). *Blocking sets in desarguesian projective planes,* Bull. London. Math. Soc. 18 pp. 132-134.

# Samenvatting

In het proefschrift komen grenzen aan en constructies van codes aan de orde. Het eerste hoofdstuk behandelt grenzen aan en constructies van codes voor een zogenaamd three-way channel. De gebruikte techniek voor de afleiding van de grenzen staat bekend onder de naam 'random coding argument' en blijkt ook in dit geval toepasbaar te zijn. De constructie van codes in dit geval is tamelijk ad hoc. Verder is in dit geval gekeken naar de toepasbaarheid van strategieën analoog aan Schalkwijk's idee voor het two-way channel. Het tweede hoofdstuk behandelt grenzen aan en constructies van codes voor het binary adder (multiple access)channel met ruis. De gebruikte technieken voor het afleiden van de ondergrenzen zijn bekend onder de naam 'Gilbert-Varshamov argument'. De constructie van de codes maakt gebruik van een concatenatie techniek. Verder worden in dit hoofdstuk codes geconstrueerd over een ternair alphabet met als afstand niet de gebruikelijke Hamming afstand maar de zogenaamde Manhattan metriek. Deze codes worden gebruikt als bouwblok in de toepasssing van de concatenatietechniek. In het derde hoofdstuk wordt gekeken naar codes die geconstrueerd kunnen worden met behulp van technieken uit de algebraïsche meetkunde. In het bijzonder worden een aantal van dergelijke codes min of meer expliciet gemaakt door een beschrijving van generatormatrices en parity check matrices. Tevens wordt enige aandacht besteed aan decodeeralgoritmen voor dit soort codes.

# Curriculum vitae

De schrijver van dit proefschrift werd op 3 october 1956 te Eindhoven geboren. In 1975 behaalde hij zijn Atheneum-B diploma aan het St. Joris College te Eindhoven. Daarna studeerde hij Wiskunde aan de Technische Hogeschool te Eindhoven. In Augustus 1982 behaalde hij het diploma wiskundig ingenieur. Na vervulling van zijn dienstplicht was hij gedurende 4 jaar (1983-1987) als Z.W.O. medewerker verbonden aan de Technische Hogeschool van Eindhoven. Daarna volgde een half jaar als toegevoegd docent aan de Technische Universiteit van Eindhoven.

# Stellingen

behorende bij het Proefschrift
Constructing Codes

H.J. Tiersma

1. De gegeneraliseerde tweede orde Reed-Muller codes bevatten subcodes die in zekere zin maximaal zijn en waarvan de gewichtsverdeling op relatief eenvoudige wijze bepaald kan worden.
   H.J. Tiersma, *On subcodes of generalized second order Reed-Muller codes,* SIAM J. Alg. Disc. Meth. Vol. 6, No. 4 October 1985.

2. Een aantal van de resultaten uit J. Körner and V.K. Wei: *Odd and even Hamming spheres also have minimum boundary,* Disc. Math. 51, (1984) 147-165, kunnen op een eenvoudige manier bewezen worden door toepassing van standaardtechnieken.
   H.J. Tiersma, *A note on Hamming spheres,* Disc. Math. 54 (1985) 225-228.

3. De constructies en grenzen voor radar arrays die voorkomen in J.P. Robinson: *Golomb rectangles,* IEEE Trans. Inform. Theory Vol. IT-31, No. 6, November 1985, kunnen aanzienlijk verbeterd worden:
   Robinson: $2.416 \leqslant GR(n)/n \leqslant 3$,
   Tiersma en Blokhuis: $2.5 \leqslant GR(n) \leqslant 2.9629$
   A. Blokhuis and H.J. Tiersma, *Bounds for the size of radar arrays,* IEEE Trans. Inform. Theory, Vol. IT-34, No. 1, January, 1988.

4. Laat $\Gamma(L,g)$ de extended Goppa code zijn, waar $L=\{\alpha_1,...,\alpha_n\} = GF(q^m)$, $q$ een priemmacht en $m$ een natuurlijk getal. Van de uitspraak '$\Gamma(L,g)$ is cyclisch dan en slechts dan als $g(z)=(z-\beta_1)^a(z-\beta_2)^a$ waar $\beta_1$ en $\beta_2$ geconjugeerd zijn in $GF(q^{2m})/GF(q^m)$' bestaan momenteel minstens drie gepubliceerde en twee ongepubliceerde foutieve bewijzen, en geen correcte bewijzen.
   (i) Feng Gui-Liang. *The sufficient and necessary condition for extending Goppa with $L=GF(q^m)$ to cyclic codes.* Santa Monica.
   (ii) K.K. Tzeng and C.Y. Yu. *Characterization Theorems for extending Goppa codes to cyclic codes.* IEEE Trans. Inform. Theory, Vol. IT-25 maart 1979.
   (iii) A.L. Vishnevetskii. *Cyclicity of extended Goppa codes over $GF(q)$.* Problemy Peredachi Informatsii. Vol. 18, No. 3, July-Sept. 1982.
   (iv) J.A. Thiong-Ly. *Symmetries of cyclic extended Goppa codes over $GF(q)$.* Proceedings of the 4th AECCC conference. Springer Lecture Notes.

5. Laat $m$ een positief geheel getal zijn en definieer $N=q^{m-1}$, $n=q^m+1$, waar $q$ een priemmacht is. Laat $\psi$ een primitief element zijn van $GF(q^{2m})$ en defineer $\theta=\psi^N$, $\eta=\psi^n$. Laat $C$ de irreducibele cyclische code zijn over $GF(q)$ met lengte $n$, dimensie $2m$ en nonzeroes $\theta,\theta^q,\ldots,\theta^{q^{2m-1}}$. Laat $A(z)$ het gewichtsverdelingspolynoom zijn van deze code. Laat $D$ de cyclische code zijn over $GF(q)$ met lengte $N$ en dimensie $2m$ en nonzeroes $\eta,\eta^{\pm q},\ldots,\eta^{\pm q^{m-1}}$ en laat $B(z)$ het gewichtsverdelingspolynoom zijn van deze code. De polynomen $a(z)=(A(z)-1)/n$ en $b(z)=(B(z)-1)/N$ hebben gehele coefficienten en staan op de volgende manier met elkaar in verband:

$$b(z)=a(z^{-1})z^{2(q-1)q^{m-1}}+2z^{(q-1)q^{m-1}}$$

J.C.C.M. Remijn and H.J. Tiersma. *A duality Theorem for the weight distribution of some cyclic codes,* IEEE Trans. Inform. Theory, Vol. IT-34, No. 5, September 1988.

6. Laat $X_1,X_2,X_3,Y_1,Y_2,Y_3$ de input en output symbolen van een three-way channel zijn. De bij het kanaal horende mutuele informaties zijn:
$I_k^{ij}:=I(X_i,X_j;Y_k|X_k)$ en $I_{ij}^k:=I(X_i;Y_k|X_j,X_k)$.
Laat $S=\{(R_1,R_2,R_3)|(R_i+R_j-R_k)/2<I_{ij}^k,R_k<I_k^{ij}\ \{i,j,k\}=\{1,2,3\}$,
de kansverdeling op de inputtriples is het produkt van de kansen op de afzonderlijke input symbolen$\}$,
$T=\{(I_1^{23},I_2^{13},I_3^{12})|$ De kansverdeling op de inputtriples is willekeurig$\}$.
Dan ligt het capaciteitsgebied $G$ van het kanaal tussen het convex omhulsel van $S$ en dat van $T$: $co(S)\subseteq G\subseteq co(T)$.
(Proefschrift Hoofdstuk 1, paragraaf 3.)

7. De door Kasami en Lin afgeleide grenzen voor codes voor het binaire optelkanaal zijn niet allemaal correct.
(Proefschrift Hoofdstuk 2, paragraaf 3.)

8. De codes die geconstrueerd worden in Hoofdstuk 3, paragraaf 3 van dit proefschrift zijn zwak zelf duaal, en indien de dimensie groot genoeg is zelf duaal.

9. Het is historisch gezien onjuist om de sopraanstem en de altstem aan te duiden met de verzamelnaam vrouwenstemmen.

10. Het spel dat in Hofstadters boek Metamagical Themes op blz. 70 en volgende beschreven wordt en de naam Nomic draagt, ontaardt in dobbelen, wanneer het gespeeld wordt door spelers die allen uitblinken in een ander specialisme en het spel proberen te winnen.

11. Het benoemen van vrouwen in managementsfuncties werkt rolbevestigend.