Lecture 3

Bounds on Codes

3.1. Introduction

One of the fundamental problems in coding theory is to determine, for a given q, the possible set of values for triples (n, k, d) for which there exists an $[n, k, d]_q$-code. More precisely, we define the function

$$A_q(n, d) := \max\{k \mid \exists\ [n, k, d]_q\text{-code}\}.$$

Exact values for $A_q(n, d)$ are known only for small values of n and d. A good source for looking up results like this would be the server maintained by Andries Brouwer's which is available at

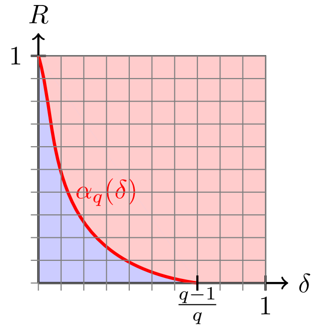
```
http://www.win.tue.nl/˜aeb/voorlincod.html
```

Because of the intractability of exact values for this function, one might want to look at asymptotic assertion. Typically, this is done by fixing for example the ratio d/n, and looking at the upper limit of $A_q(n, d)/n$ as n goes to infinity. We define

$$\alpha_q(\delta) := \limsup_{n\to\infty} \frac{A_q(n, \lfloor \delta n \rfloor)}{n}.$$

The asymptotic theory is concerned with the determination of this function.

Despite its easy looks, we do not as of yet know the values of this function on the interval $[0, (q-1)/q]$, save for two points: the point $\delta = 0$, and the point $\delta = (q-1)/q$, where the value is 1, and 0, respectively. Nevertheless, we know provable upper and lower bound for this function, and some coding theorists even believe that they know the shape of $\alpha_2(\delta)$. This lecture is devoted to some of the things that we can prove about $\alpha_q(x)$. The situation of the function $\alpha_q(x)$ is as depicted here.



The pink region (northeast of $\alpha_q(\delta)$) is not achievable, but the blue region (southwest of $\alpha_q(\delta)$) is achievable.

Lecture and scribes by Amin Shokrollahi

3.2. First Bounds

Theorem 3.1 (Singleton Bound)**.** *For an $[n,k,d]_q$-code we have $k+d \le n+1$. Codes for which equality holds are called Maximum Distance Separable (MDS) codes.*

Proof. Consider the vector space $V = \mathbb{F}_q^{d-1} \times 0^{n-d+1}$ of dimension $d-1$. Since the code C is of minimum distance d, we have $V \cap C = \{0\}$, so that $\dim V + \dim C \le n$, i.e., $k+d-1 \le n$. □

Example 3.2. 1. The parity code is defined as $P = \langle(1,1,\dots,1)\rangle^{\perp}$. The minimum distance of the code is 2, and its dimension is $n-1$. It is MDS.

2. Consider the dual of the parity code. Its distance is n and its dimension is 1. It is MDS.

The Singleton bound is of somewhat limited use, since it does not take into account the dependence on q. To get to more serious bounds, we look at the asymptotic behavior of the set of triples (n,k,d) for which there exists an $[n,k,d]_q$-code. The Singleton bound states that $\alpha_q(\delta) \le 1-\delta$. A different bound is the following.

Theorem 3.3 (Plotkin bound)**.** *We have*

$$\alpha_q(\delta) \begin{cases} \le -\delta\frac{q}{q-1}+1 & \textit{if } \delta \le \frac{q-1}{q}, \\ = 0 & \textit{if } \delta \ge \frac{q-1}{q}. \end{cases}$$

Proof. Let $\theta := (q-1)/q$. We first show that $\alpha_q(\delta) = 0$ for $\delta > \theta$. Let C be an $[n,k,d]_q$-code, where $d \ge \theta n$. We will show that

$$q^k(q^k-1)d \le \sum_{x,y\in C, x\ne y} d(x,y) \le n\theta q^{2k}, \tag{3.1}$$

where $d(x,y)$ is the Hamming distance between x and y, so that

$$k \le \log_q\left(\frac{d}{d-n\theta}\right),$$

and hence $\limsup_{n\to\infty} k/n = 0$.

The inequality $\sum_{x,y\in C,x\ne y} d(x,y) \ge q^k(q^k-1)d$ is clear. The other inequality is obtained by writing down all the q^k codewords of C into a $q^k \times n$-matrix. Fix a column, and let n_j denote the number of times the element j of the alphabet occurs in that column. The contribution of the column to the sum is $\sum_{j=1}^q n_j(q^k-n_j)$, so that

$$\begin{aligned} \sum_{x,y\in C,x\ne y} d(x,y) &= n\sum_{j=1}^{q} n_j(q^k-n_j) \\ &= n\left(q^{2k} - \sum_j n_j^2\right) \qquad (\textstyle\sum_j n_j = q^k) \\ &\le n\left(q^{2k} - \frac{q^{2k}}{q}\right) \qquad \text{(Cauchy-Schwarz)} \\ &= n\theta q^{2k}. \end{aligned}$$

The proof of the other assertion is left as an exercise. □

The *Hamming bound* has a simple interpretation. Suppose that we have an $[n,k,d]_q$-code, and consider the Hamming balls of radius $(d-1)/2$ around the codewords (Hamming ball of radius $(d-1)/2$ around x = set of all words of distance $\le (d-1)/2$ from x). Since the words of the code are distance d apart, these balls are mutually disjoint. If V denotes their volume, we have $q^kV \le q^n$, and hence $k/n \le 1 - \log_q(V)/n$. Let $H_q(x) := x\log_q(q-1) - x\log_q(x) - (1-x)\log_q(1-x)$ denote the q-ary entropy function. Then we have the following result the proof of which we leave as an exercise.

Proposition 3.4. *Let V be a Hamming ball of radius rn around the word $0 \in \mathbb{F}_q^n$. Then $|V| = q^{nH_q(r)+o(n)}$.*

From this proposition we obtain the following bound.

Theorem 3.5 (Hamming bound)**.** $\alpha_q(\delta) \le 1 - H_q(\delta/2)$.

3.3. The Linear Programming Bound

The Krawtchouk polynomials are defined as

$$K_k(x) := \sum_{j=0}^{k} (-1)^j \binom{x}{j}\binom{n-x}{k-j}(q-1)^{k-j}.$$

The following theorem gives the linear programming bound for linear codes.

Theorem 3.6. *Let $q, n, d \in \mathbb{N}$, $q \geq 2$. Then we have*

$$\begin{aligned} A_q(n,d) \quad &\leq \quad \max\left\{ \left\lfloor \log_q\left(\sum_{i=0}^{n} A_i\right)\right\rfloor \;\middle|\; A_0 = 1, A_1 = \cdots = A_{d-1} = 0, A_d, \ldots, A_n \geq 0, \right. \\ &\qquad\qquad \left. \forall k = 0, \ldots, n-1\colon \quad \sum_{i=0}^{n} A_i K_k(i) \geq 0 \right\} \end{aligned}$$

The proof of this theorem proceeds in several steps. Let $Q = \mathbb{Z}/q\mathbb{Z}$. In a first step, we will show the following lemma.

Lemma 3.7. *Let ω be a primitive qth root of unity in $\mathbb{C}$, and let $x \in Q^n$ be a fixed word of weight i. Then we have*

$$\sum_{\substack{y \in Q^n \\ \mathrm{wgt}(y)=k}} \omega^{\langle x,y\rangle} = K_k(i),$$

for $k = 0, 1, \ldots, n$.

Proof. The proof is very simple, but needs a little bit of notation. Let $Q(a,b)$ denote the set of all vectors of weight b in Q^a, and let $\mathrm{Pow}_b(a)$ denote the set of all subsets of size b of $\{1, \ldots, a\}$. It is immediately clear that $|Q(a,b)| = \binom{a}{b}(q-1)^b$.

The assertion to be proved translates to

$$\sum_{y \in Q(n,k)} \omega^{\langle x,y\rangle} = K_k(i),$$

for $k = 0, \ldots, n$. Fix k. Without loss of generality, we can assume that $x = (x_1, \ldots, x_i, 0, \ldots, 0)$ with $x_1, \ldots, x_i \neq 0$. Note that $Q(n,k) = \sqcup_{j=0}^{k} Q(i,j) \times Q(n-i, k-j)$, where $\sqcup$ denotes disjoint union. Thi decomposition can be seen immediately by considering the first i and the last $n-i$ positions of a vector in $Q(n,k)$. As a result:

$$\sum_{y \in Q(n,k)} \omega^{\langle x,y\rangle} = \sum_{j=0}^{k} |Q(n-i, k-j)| \sum_{z \in Q(i,j)} \omega^{\langle x,z\rangle} = \sum_{j=0}^{k} \binom{n-i}{k-j}(q-1)^{k-j} \overbrace{\sum_{z \in Q(i,j)} \omega^{\langle x,z\rangle}}^{=:T},$$

so that it remains to show that $T = (-1)^j\binom{i}{j}$.

To see this, look at the decomposition $Q(i,j) = \sqcup_{D \in \mathrm{Pow}_j(i)} (\mathbb{F}_q^\times)^j$, which is obtained by fixing first the set of positions of a potential vector of weight j in $\mathbb{F}_q^i$, and then letting all elements of $\mathbb{F}_q^\times$ run independently over those positions. This shows us that

$$\begin{aligned} \sum_{j=0}^{k} \sum_{z \in Q(i,j)} \omega^{\langle x,z\rangle} \quad &= \quad \sum_{\substack{D \in \mathrm{Pow}_j(i) \\ D = \{d_1, \ldots, d_j\}}} \sum_{z_1, \ldots, z_j \in \mathbb{F}_q^\times} \omega^{x_{d_1} z_1 + \cdots + x_{d_j} z_j} \\ &= \quad \sum_{\substack{D \in \mathrm{Pow}_j(i) \\ D = \{d_1, \ldots, d_j\}}} \prod_{\ell=1}^{j} \sum_{z \in \mathbb{F}_q^\times} \omega^{x_{d_\ell} z} \\ &= \quad |\mathrm{Pow}_j(i)|(-1)^j \\ &= \quad \binom{i}{j}(-1)^j, \end{aligned}$$

since $\sum_{t \in \mathbb{F}_q^\times} \omega^t = -1$. This proves the assertion. □

The second step of the proof of Theorem 3.6 is the following lemma which immediately implies the theorem.

Lemma 3.8. *Suppose that the code $C \subseteq Q^n$ has weight distribution $(A_0, A_1, \ldots, A_n)$. Then we have*

$$\sum_{i=0}^{n} A_i K_k(i) \geq 0,$$

for $k = 0, 1, \ldots, n$.

Proof.

$$\begin{aligned}
|C| \sum_{i=0}^{n} A_i K_k(i) &= \sum_{i=0}^{n} \sum_{\substack{(x,z)\in C^2 \\ d(x,z)=i}} \sum_{\substack{y\in Q^n \\ \mathrm{wgt}(y)=k}} \omega^{\langle x-z, y\rangle} \\
&= \sum_{\substack{y\in Q^n \\ \mathrm{wgt}(y)=k}} \sum_{i=0}^{n} \sum_{\substack{(x,z)\in C^2 \\ d(x,z)=i}} \omega^{\langle x-z, y\rangle} \\
&= \sum_{\substack{y\in Q^n \\ \mathrm{wgt}(y)=k}} \sum_{(x,z)\in C^2} \omega^{\langle x,y\rangle} \overline{\omega}^{\langle z,y\rangle} \\
&= \sum_{\substack{y\in Q^n \\ \mathrm{wgt}(y)=k}} \left(\sum_{x\in C} \omega^{\langle x,y\rangle} \right) \overline{\left(\sum_{z\in C} \omega^{\langle z,y\rangle} \right)} \\
&= \sum_{\substack{y\in Q^n \\ \mathrm{wgt}(y)=k}} \left| \sum_{x\in C} \omega^{\langle x,y\rangle} \right|^2 \\
&\geq 0,
\end{aligned}$$

where $\overline{a}$ is the complex conjugate of a. □

Example 3.9. We prove the optimality of the $[8, 5, 3]_4$-code that we encountered in the last lecture. The weight distribution of this code involves only the nonnegative parameters $A_0, A_3, A_4, A_5, A_6, A_7, A_8$. The inequalities for these parameters are

$$\begin{aligned}
A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 &\geq 0 \\
24\,A_0 + 12\,A_3 + 8\,A_4 + 4\,A_5 - 4\,A_7 - 8\,A_8 &\geq 0 \\
252\,A_0 + 48\,A_3 + 12\,A_4 - 8\,A_5 - 12\,A_6 + 28\,A_8 &\geq 0 \\
1512\,A_0 + 44\,A_3 - 40\,A_4 - 28\,A_5 + 16\,A_6 + 28\,A_7 - 56\,A_8 &\geq 0 \\
5670\,A_0 - 150\,A_3 - 74\,A_4 + 50\,A_5 + 30\,A_6 - 70\,A_7 + 70\,A_8 &\geq 0 \\
13608\,A_0 - 252\,A_3 + 120\,A_4 + 44\,A_5 - 96\,A_6 + 84\,A_7 - 56\,A_8 &\geq 0 \\
20412\,A_0 + 216\,A_3 + 108\,A_4 - 144\,A_5 + 100\,A_6 - 56\,A_7 + 28\,A_8 &\geq 0 \\
17496\,A_0 + 324\,A_3 - 216\,A_4 + 108\,A_5 - 48\,A_6 + 20\,A_7 - 8\,A_8 &\geq 0 \\
6561A_0 - 243A_3 + 81A_4 - 27A_5 + 9A_6 - 3A_7 + A_8 &\geq 0
\end{aligned}$$

The linear program which maximizes $A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8$ subject to the constraints above, and subject to the non-negativity of the parameters gives a solution

$$A_0 = 1, A_3 = 72, A_4 = 210, A_5 = 432, A_6 = 792, A_7 = \frac{4152}{7}, A_8 = \frac{1683}{7}.$$

The sum of these values is $16384/7$, so that $A_4(8, 3) \leq \lfloor \log_4(16384/7) \rfloor = 5$, which shows the optimality of the code.

We mention without proof the following upper bound on α_2 due to McEliece, Rodemich, Ramsey, and Welch (called the MRRW-bound). Its proof is based on the linear programming approach.

Theorem 3.10 (MRRW-bound)**.** *We have*

$$\alpha_2(\delta) \leq H_2\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right).$$

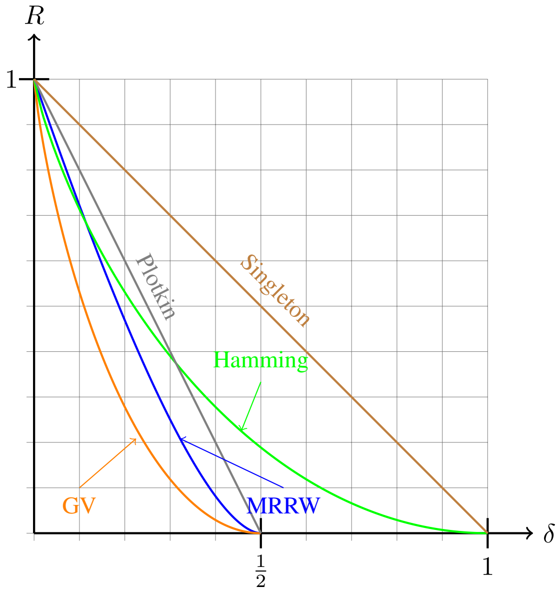


Figure 3.1: Upper and lower bounds for $\alpha_2(x)$.

3.4. The Gilbert-Varshamov Bound

The bounds we have discussed so far are all upper bounds on α_q, i.e., they lead to asymptotic non-existence theorems. The Gilbert-Varshamov bound below is a lower bound. We will prove it for nonlinear codes, and will later show that there are also “more explicit” linear codes that achieve this bound. However, to make matters more precise

Theorem 3.11 (Gilbert-Varshamov bound)**.** *For $\delta \in [0,(q-1)/q]$ we have $\alpha_q(\delta) \geq 1 - H_q(\delta)$.*

Proof. Fix $\delta \in [0,(q-1)/q]$, and $\varepsilon > 0$, and set $R = 1 - H_q(\delta) - \varepsilon$. Let H be a random $(1-R)n \times n$-matrix over $\mathbb{F}_q$, i.e., every entry of H is independently and uniformly distributed over $\mathbb{F}_q$. Let C be the right-kernel of H, i.e., H is a check matrix for C. It suffices to prove that

$$\Pr[d(C) < \delta n] \leq q^{-\varepsilon n + o(n)},$$

where $d(C)$ is the minimum distance of C. To this end, we use the union-bound to obtain

$$\Pr[d(C) < \delta n] \leq \sum_{\substack{0 \neq x \in \mathbb{F}_q^n \\ \mathrm{wgt}(x) < \delta n}} \Pr[x \in C].$$

Since H is random, any nonzero $x \in \mathbb{F}_q^n$ has probability $1/q^{n-Rn}$ to be in C, hence

$$\Pr[d(C) < \delta n] \leq |V(\delta)| q^{-(1-R)n},$$

where $V(\delta)$ is the set of all words in $\mathbb{F}_q^n$ of weight $\leq \delta n$. From Proposition 3.4 we thus obtain

$$\Pr[d(C) < \delta n] \leq q^{(H_q(\delta)-(1-R)+o(1))n} = q^{-n\varepsilon + o(n)},$$

which completes the proof. □

Some of the upper bounds of the previous sections, and the lower GV-bound are depicted in Figure 3.1. The above theorem not only shows that the GV-bound is asymptotically achievable, it also shows that random codes achieve it. We now show a simpler construction of linear codes achieving the GV bound over $\mathbb{F}_2$. We first construct 2^n pairwise non-intersecting n-dimensional subspaces of $\mathbb{F}_2^{2n}$. Consider the field $\mathbb{F}_{2^n}$, and fix a basis of this field as an $\mathbb{F}_2$-vector space. Elements of this field can be represented as n-dimensional vectors over $\mathbb{F}_2$. Let $r(a)$ denote this binary representation of $a \in \mathbb{F}_{2^n}$. For $\alpha \in \mathbb{F}_{2^n}$, let $C_\alpha := \{(r(x), r(\alpha x)) \mid x \in \mathbb{F}_{2^n}\}$. This vector space is obviously n-dimensional (because $\mathbb{F}_{2^n}$ injects into it as a vector space). Moreover, suppose that $0 \neq (a,b) \in C_\alpha \cap C_\beta$. If $a = r(x)$, this means that $b = r(\alpha x) = r(\beta x)$, i.e., $\alpha x = \beta x$, which shows that $\alpha = \beta$, since $x \neq 0$ (otherwise $a = b = 0$).

Theorem 3.12. *Suppose that* $1/2 > \varepsilon > 0$*. If* $h(\delta) = 1/2 - \varepsilon$*, then for large enough* n *an overwhelming fraction of the* C_α*'s have a minimum distance* $\geq 2\delta n$*.*

Proof. The number of words of relative weight $\leq \delta$ in $\mathbb{F}_2^{2n}$ is $2^{2nh(\delta)+o(n)} = 2^{n(1-2\varepsilon)+o(n)}$, and each such word can belong to at most one C_α. So the fraction of C_α's that have a low-weight word is $2^{-2n\varepsilon+o(n)}$, which is exponentially small. □

Using shortening the bound can be made to work for any rate (not just 1/2).

Despite this, we do not know of a single sequence of explicitly describable binary codes which meets the GV-bound! This is a challenging open problem. We will provide later in this class sequences of non-binary codes which surpass the GV-bound. For binary codes it is conjectured that the GV-bound is sharp, though I am not aware of deeper reasons for this conjecture. It is very much an open problem to prove or disprove this conjecture.

On another note, even though we know that there are sequences of binary codes attaining this bound, we do not know of a single explicit such sequence. This is another exciting open problem. Both these problems seem to be very very difficult (not a good idea to bank your PhD on them). An interesting question that may be possible to answer is the following:

Question 3.13. *Let* $\delta \in (0, 1/2)$*. Exhibit an explicit family of sets* S_i *of binary codes of length* n_i *such that* $|S_i| = 2^{o(n_i)}$*,* $n_i \to \infty$*, and such that for all* i *there is a code in* S_i *of rate at least* $1 - h(\delta) - o(1)$ *and minimum distance* $\geq \lfloor \delta n_i \rfloor$*.*

Of course, even if the $o(n_i)$ term above is $O(\log(n_i))$, one does not necessarily have a polynomial time algorithm (polynomial in n_i) to compute the correct codes, unless there is a way of telling what the minimum distance of the code in question is without actually going through all the codewords. Nevertheless, it would be interesting to exhibit such families of codes.

3.5. Final Remarks

The upper bounds derived in this lecture are also valid if we take nonlinear codes, though the proofs are slightly different than the ones we gave here. The lower bound given by the Gilbert-Varshamov curve is obviously also valid for nonlinear codes, and it is perhaps interesting to know that it is also conjectured to be sharp for nonlinear binary codes.