6.897 Algorithmic Introduction to Coding Theory September 19, 2001

Lecture On BCH Codes

Lecturer: Madhu Sudan *Scribe: Madhu Sudan*

In Lecture 4, we spoke about BCH codes in vague terms. Here we give a self-contained account of the codes with full proofs of its parameters. Note that these may not be the most general possible BCH codes, but it shows the specific version that is interesting.

To remind the reader where we are going, recall our definition of a q-ary BCH code. In these notes, we will focus only on the case where the block length of the code equals q^m for some integer m.[1]

Definition 1 *For prime power q, integer m and integer d, the* BCH code $\mathrm{BCH}_{q,m,d}$ *is obtained as follows: Let $n = q^m$ and let $\mathbb{F}_{q^m}$ be an extension of $\mathbb{F}_q$ and let C' be the (extended) $[n, n-(d-1), d]_{q^m}$ Reed-Solomon code obtained by evaluating polynomials of degree at most $n-d$ over $\mathbb{F}_{q^m}$ at all the points of $\mathbb{F}_{q^m}$. Then the code* $\mathrm{BCH}_{q,m,d}$ *is the $\mathbb{F}_q$-subfield subcode of C'. (In other words,* $\mathrm{BCH}_{q,m,d} = C' \cap \mathbb{F}_q^n$*.)*

It is evident from the definition that the codes have distance *at least* d. (The code could have a larger distance, so the literature in coding theory refers to d as the *designed distance* of the code.) Our goal is to lower bound the dimension of this code as a function of n and d (and m and q). The reader may not realize immediately why this code even has positive dimension for small d. We will explain why this is so — and in doing so, we will lower bound the dimension by $n - m(d-1)$. The rest of the writeup will focus on a *mild* improvement by adding roughly $\frac{1}{q}m(d-1)$ to the dimension. The interest in this mild improvement is that in the binary case, this brings us close to the Hamming bound (for small d) as opposed to the Gilbert-Varshamov bound. So these shows that the BCH codes *beat* the Gilbert-Varshamov bound (as do every other family of codes we know).

Let us start with a simple argument showing this code has dimension at least $n - m(d-1)$. First recall that every function from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$ is a polynomial over $\mathbb{F}_{q^m}$ of degree at most $n-1$. Thus the space of polynomials from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$ is a $\mathbb{F}_q$-linear space of dimension exactly n. We wish to know what is the dimension of the subspace of polynomials of degree at most $n-d$. But now note that the restriction that a polynomial $f(x) = \sum_{i=0}^{n-1} f_i x^i$ has degree at most $n-d$ is equivalent to saying that the coefficients f_i must equal zero, for $i \in \{n-(d-1), \ldots, n-1\}$. Each such condition is a single linear constraint over $\mathbb{F}_{q^m}$, but this translates into a block of m linear constraints over $\mathbb{F}_q$. Since we have $d-1$ such blocks of constraints, the restriction that the functions have degree at most $n-d$ places at most $m(d-1)$ linear constraints. Thus the resulting space has dimension at least $n - m(d-1)$. The argument above is informal, but easy to formalize (and in fact some of the ingredients are even included later in these notes). Formalizing the argument leads to the following bound.

Proposition 2 (Weak dimension bound) *The code* $\mathrm{BCH}_{q,m,d}$ *has dimension at least $q^m - m(d-1)$.*

To improve the bound we will argue that not all the constraints above are independent. In particular we will show that the condition that $f_i = 0$ for $i = (q^m - 1) - j$ implies that coefficient $f_{i'} = 0$ for $i' = (q^m - 1) - qj$ under the restriction that f maps $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$. This allows us to recover some of the lost dimensions, and gives us the required bound. To do so we first need to understand the role that $\mathbb{F}_q$ plays within $\mathbb{F}_{q^m}$.

1 $\mathbb{F}_q$ versus $\mathbb{F}_{q^m}$

We have seen how we can go from the field $\mathbb{F}_q$ to its extension $\mathbb{F}_{q^m}$. (See Lecture Notes of Algebra - Lecture 2.5). In Section 6 there, we described a canonical representation of $\mathbb{F}_{q^m}$ as an m-dimensional vector space over $\mathbb{F}_q$, which contains the original field $\mathbb{F}_q$ as a one-dimensional subspace.

[1] Strictly speaking, the classical BCH codes fix $n = q^m - 1$, but I see no reason to do so. If you want the classical codes, just puncture the codes we obtain and you'll get a code matching the classical parameters!

Now we invert the question. Suppose someone gives us the field $\mathbb{F}_{q^m}$ and asks for the elements corresponding to the base field $\mathbb{F}_q$. Can we obtain this subfield without referring to our construction? What is the structure of a subfield $\mathbb{F}_q$ in the bigger field $\mathbb{F}_{q^m}$? Are these well-defined objects inside a field, or does a field $\mathbb{F}_{q^m}$ contain many copies of $\mathbb{F}_q$ floating around inside it? These are questions that no self-respecting algebraist would ask — so we will. (A reader comfortable with finite fields can safely skip this section.)

Recall that every element $\alpha \in \mathbb{F}_q$ satisfies the property $\alpha^q = \alpha$. We can look for elements in $\mathbb{F}_{q^m}$ satisfying this property. Turns out there are exactly q of these and these form the subfield $\mathbb{F}_q$ in $\mathbb{F}_{q^m}$. We will prove this below, but first let us see why it is reasonable that the roots of $x^q - x = 0$ should form a field. It is evident that if $\alpha^q = \alpha$ and $\beta^q = \beta$, then $(\alpha\beta)^q = \alpha\beta$ and so the roots of $x^q - x$ are closed under multiplication. But why is $(\alpha+\beta)^q = \alpha+\beta$? We prove this in a proposition that will be several times in this lecture.

Proposition 3 *Let $\mathbb{F}_q$ be a field of characteristic p. In other words, p is a prime and $q = p^r$. Then for every non-negative ℓ, and $\alpha, \beta \in F_q$, it is the case that $(\alpha+\beta)^{p^\ell} = \alpha^{p^\ell} + \beta^{p^\ell}$.*

Proof We start with an even more elementary fact. For every $\gamma \in \mathbb{F}_q$, $p \cdot \gamma = 0$ (where $p \cdot \gamma$ is just shorthand for $\gamma + \cdots + \gamma$, with p copies of γ in there). The simplest way to see this fact is that addition in $\mathbb{F}_q$ is just vector addition in $\mathbb{F}_p^r$ and the sum of p copies of any vector in $\mathbb{F}_p^r$ is zero.

Now we move to proving the proposition. Note that it is trivial for $\ell = 0$. So, we prove the propostion for the case $\ell = 1$. In this case $(\alpha+\beta)^p = \sum_{i=0}^{p} \binom{p}{i} \alpha^{p-i}\beta^i$. Now note that p divides $\binom{p}{i}$ for every $i \in [p-1]$. Since $p\gamma = 0$ for every $\gamma \in \mathbb{F}_q$, we have $\binom{p}{i}\alpha^{p-i}\beta^i = 0$ for every $i \in [p-1]$ and we are only left with the cases $i \in \{0, p\}$ which gives: $(\alpha+\beta)^p = \alpha^p + \beta^p$.

The proposition now follows by induction on ℓ. For $\ell \geq 2$, we have

$$\begin{aligned}
(\alpha+\beta)^{p^\ell} &= ((\alpha+\beta)^p)^{p^{\ell-1}} \\
&= (\alpha^p + \beta^p)^{p^{\ell-1}} \\
&= (\alpha^p)^{p^{\ell-1}} + (\beta^p)^{p^{\ell-1}} \\
&= \alpha^{p^\ell} + \beta^{p^\ell}.
\end{aligned}$$

■

Now we argue formally that the roots of $x^q - x$ do form the subfield $\mathbb{F}_q$ of $\mathbb{F}_{q^m}$.

Proposition 4 *There are exactly q elements $\alpha \in \mathbb{F}_{q^m}$ satisfying $\alpha^q = \alpha$ and these form the subfield $\mathbb{F}_q$ of $\mathbb{F}_{q^m}$.*

Proof We've already argued that the roots are closed under addition and multiplication. Thus the roots of $x^q - x$ form a field. It suffices to argue that the cardinality of this field is q.

It is easy to see there are at most q elements α in $\mathbb{F}_{q^m}$ satisfying $\alpha^q - \alpha = 0$, since these are roots of the degree q polynomial $x^q - x$. To see that this polynomial has q roots, note that every element of $\mathbb{F}_{q^m}$ is a root of $x^{q^m} - x$. Furthermore we have $x^{q^m} - x = (x^q - x)\cdot h(x)$, where $h(x) = x^{q^m - q} + x^{q^m - 2q} + \cdots + x^q + 1$ is a polynomial of degree $q^m - q$. Thus every element of $\mathbb{F}_{q^m}$ is a either a root of the polynomial $x^q - x$ or of $h(x)$. Since at most $q^m - q$ elements can be roots of $h(x)$, we have that $h(x)$ has exactly $q^m - q$ roots, and $x^q - x$ has exactly q roots. ■

We conclude that the roots of $x^q - x$ in $\mathbb{F}_{q^m}$ define the subfield $\mathbb{F}_q$ of $\mathbb{F}_{q^m}$ uniquely. Thus the operation of restricting a code in $\mathbb{F}_{q^m}^n$ to $\mathbb{F}_q^n$ is not arbitrary (such as say, as restricting the code to S^n for some arbitrary subset S of $\mathbb{F}_{q^m}$). Next we extend our understanding of subfields to functions mapping $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$.

2 Functions from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$

We first extend Propositions 3 and 4 to the case of functions.

Proposition 5 *For functions $f, g : \mathbb{F}_{q^m} \to \mathbb{F}_{q^m}$, the following hold:*

1. $(f+g)^q = f^q + g^q$.
2. $f^q = f$ *if and only if the range of f is contained in $\mathbb{F}_q$.*

We omit the proof since it is straightforward given Propositions 3 and 4. We now move to the proposition that is the crux of the improved dimension analysis.

Proposition 6 *Suppose $f(x) = \sum_{i=0}^{n-1} f_i x^i$ is a polynomial over $\mathbb{F}_{q^m}$ mapping $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$. Suppose $j, \ell \in \{1, \ldots, n-2\}$ are such that $\ell = qj (\mathrm{mod}\ (n-1))$. Then $f_\ell = 0$ if $f_j = 0$. Furthermore $f_{n-1} \in \mathbb{F}_q$.*

Proof Since f maps $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$, we must have $f^q = f$ (Proposition 5, Part 2). Using Proposition 5, Part 1, we get that $f(x)^q = \sum_{i=0}^{n-1} (f_i)^q x^{iq}$. Since this expansion has terms of degree larger than $n-1$, we reduce terms above using the identity $x^n = x$ to get that f^q takes the form $\sum_{i=0}^{n-2} (f_i)^q x^{iq (\mathrm{mod}\ (n-1))} + (f_{n-1})^q x^{n-1}$. (The coefficient f_{n-1} needs to be treated separately since we don't have $x^{n-1} = 1$, but only $x^n = x$. Nevertheless each term can be verified to be correct.) Now using the fact that the functions f^q and f are identical over $\mathbb{F}_{q^m}$, we get that $f_0^q = f_0$, $f_{n-1}^q = f_{n-1}$ and $f_{qj (\mathrm{mod}\ (n-1))} = f_j^q$. The proposition follows as a special case. ■

Before going on to the final claim giving the better bound on the dimension of BCH codes, let us just ensure that the constraints we are looking at are all *linear*. We already insisted that the collection of functions mapping $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$ is linear, but in what representation? The natural representation of such a function $f : \mathbb{F}_{q^m} \to \mathbb{F}_q$ is $\mathbf{f} = \langle f(\alpha) \rangle_{\alpha \in \mathbb{F}_{q^m}}$. But, in such a case, are constraints of the form $f_i = 0$ linear in this representation, (where f_i is the coefficient of x^i in the polynomial representation of f)? The answer is affirmative, and the proposition below asserts this.

Proposition 7 *Let $f = \sum_{i=0}^{n-1} f_i x_i$ be a polynomial mapping $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$. Then the constraint $f_i = 0$ is a conjunction of at most m linear constraints on the the vector $\mathbf{f} = \langle f(\alpha) \rangle_{\alpha \in \mathbb{F}_{q^m}}$. The condition $f_{n-1} = 0$ is one linear constraint on the vector $\mathbf{f}$.*

Proof The first part is easy if we recall the right relationships between $\mathbb{F}_q$ and $\mathbb{F}_{q^m}$, and in particular that elements of $\mathbb{F}_{q^m}$ form a m-dimensional vector space over $\mathbb{F}_q$. Viewing $\mathbf{f}$ as a vector from $\mathbb{F}_{q^m}^n$ and recalling that interpolation is linear, we obtain that f_i is a $\mathbb{F}_{q^m}$ is a linear form in $\mathbf{f}$. But since $\mathbb{F}_q^m$ can be embedded in $\mathbb{F}_q^m$ preserving linearity, we get that f_i is given my m $\mathbb{F}_q$-linear forms in $\mathbf{f}$. The condition $f_i = 0$ just forces all these linear forms to be zero.

For the furthermore part, recall by Proposition 6 that f_{n-1} is already an element of $\mathbb{F}_q$ which is a one-dimensional subspace of $\mathbb{F}_q^m \cong \mathbb{F}_{q^m}$. Thus the $f_{n-1} = 0$ forces one additional linear constraint on this space, and thus on the vector $\mathbf{f}$. ■

3 Bounding the dimension of BCH codes

We are now ready to prove a stronger bound on the dimension of BCH codes.

Lemma 8 *The dimension of the code $\mathrm{BCH}_{q,m,d}$ is at least $q^m - 1 - m \left\lceil \frac{(d-2)(q-1)}{q} \right\rceil$.*

Proof The proof is straightforward given all our work so far. The basic idea is to consider the space of all functions from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$, which forms a $\mathbb{F}_q$-vector space of dimension n. Viewing these functions as polynomials from $\mathbb{F}_{q^m}[x]$, we then restrict them to have zero as the coefficients of x^i for $i \in \{n-(d-1), \ldots, n-1\}$.

The condition that the coefficient of x^{n-1} is zero imposes one linear constraint and reduces the dimension of the space to $n-1$. The remaining conditions, corresponding to coefficients of x^i for $i \in \{n-(d-1), \ldots, n-2\}$, lead to at most m conditions each. However, we don't need to impose all such conditions. In particular we can skip every qth condition (starting at $n-2$ and going down) since these are exponents of the form $\ell = (n-1) - qj$, where j is a positive integer. Since $\ell = q(n-1-j) (\bmod (n-1))$, by Proposition 6, the condition that the coefficient of x^ℓ equals zero is implied by the condition that the coefficient of x^{n-1-j} equals zero. Thus the constraints corresponding to the coefficients of x^i for $i \in \{n-(d-1), \ldots, n-2\}$, lead to at most $m \left\lceil \frac{(d-2)(q-1)}{q} \right\rceil$ linear constraints. Thus the dimension of the remaining space is at least $n - 1 - m \left\lceil \frac{(d-2)(q-1)}{q} \right\rceil$ as claimed. ■

We conclude with the following theorem summarizing the properties of the BCH codes.

Theorem 9 ([1, 3, 2]) *For prime power q, integers m and d, the code $\mathrm{BCH}_{q,m,d}$ is an $\left[n, n-1-m\lceil \frac{(d-2)(q-1)}{q} \rceil, d\right]_q$ code, for $n = q^m$.*

Of particular interest is the case $q = 2$ and even d. In this case, the messy ceilings and floor disappear and we the following nice corollary:

Corollary 10 *For every integer m and t, the code $\mathrm{BCH}_{2,m,2t}$ is a $[n, n-1-(t-1)\log n, 2t]$-code, for $n = 2^m$.*

As pointed out earlier, this is particularly nice, since it is very close to the Hamming bound, for constant d, and in particular matches the Hamming code of distance 4 (and by puncturing matches parameters of the Hamming code of distance 3 as well).

Bibliographic notes

Binary BCH codes were discovered independently by Hocquenghem [3] and Bose and Ray-Chaudhuri [1]. The extension to the general q-ary case is due to Gorenstein and Zierler [2].

The proofs in these notes are original. This proof was inspired by a simple self-contained description of BCH codes by Rüdiger Urbanke [4]. (The only short and self-contained description I could find!) These description here is different in that it does not go to the Fourier transform (something I am allergic to). An earlier version of these notes had a more elaborate proof. It was also buggy! Thanks to Amnon Ta-Shma for pointing out the bug. Fixing the bug resulted in the final (simpler) proof.

References

[1] R. C. Bose and D. K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3:68–79, 1960.

[2] Daniel Gorenstein and Neal Zierler. A class of error-correcting codes in p^m symbols. *Journal of the Society for Industrial and Applied Mathematics*, 9:207–214, June 1961.

[3] A. Hocquenghem. Codes correcteurs d'erreurs. *Chiffres (Paris)*, 2:147–156, 1959.

[4] Rüdiger Urbanke. *Modern Coding Theory – SS2001.* EPFL, DSC-LTHC, Available from `http://lthcwww.epfl.ch/content.php?title=coding2001`, May 15 2001.