Elements of Coding Theory

Error-detecting and -correcting codes

Radu Trîmbiţaş

January 2013

Outline

Contents

1 Hamming's Theory

Hamming's Problem

- Hamming studied magnetic storage devices. He wanted to build (out of magnetic tapes) a reliable storage medium where data was stored in blocks of size 63 (this is a nice number, we will see why later). When you try to read information from this device, bits may be corrupted, i.e. flipped (from 0 to 1, or 1 to 0). Let us consider the case that at most 1 bit in every block of 63 bits may be corrupted. How can we store the information so that all is not lost? We must design an encoding of the message to a codeword with enough redundancy so that we can recover the original sequence from the received word by decoding. (In Hamming's problem about storage we still say "'received word"')
- Naive solution – repetition code – to store each bit three times, so that any one bit that is erroneously ipped can be detected and corrected by majority decoding on its block of three.

1.1 Definitions

Basic Notions

The Hamming encoding tries to do better with the following matrix:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Given a sequence of bits, we chop it into 4 bit chunks. Let b be the vector representing one such chunk, then we encode b as the 7 bit bG, where all arithmetic is performed mod 2. Clearly the efficiency is a much better $\frac{4}{7}$, though we still need to show that this code can correct one bit errors.

Claim 1. *$\forall\, b_1 \neq b_2$, b_1G and b_2G differ in ≥ 3 coordinates.*

First we present some definitions. We denote by Σ the alphabet, and the ambient space Σ^n represents the set of n letter words over the alphabet Σ.

Definition 2. The *Hamming distance* $\Delta(x, y)$ between $x, y \in \Sigma^n$ is the number of coordinates i where $x_i \neq y_i$.

- The Hamming distance is a metric since it is easy to verify that:

$$\begin{aligned} \Delta(x, y) &= \Delta(y, x) \\ \Delta(x, z) &\leq \Delta(x, y) + \Delta(y, z) \\ \Delta(x, y) &= 0 \Leftrightarrow x = y \end{aligned}$$

- In our case, consider the space of all possible encoded words $\{0, 1\}^7$.
- If we can prove our claim, then this means that in this space and under the Hamming metric, each code word bG will have no other code word within a radius of 2 around it.
- In fact, any point at Hamming distance 1 from a code word is guaranteed to be closer to that code word than any other, and thus we can correct one bit errors.

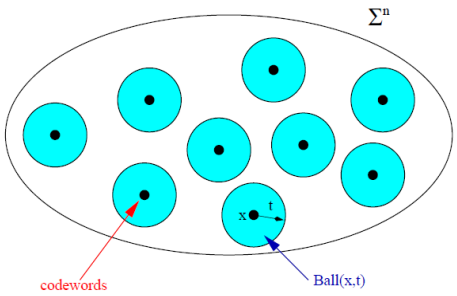
Definition 3. An *Error Correcting Code* is a set of code words $C \subseteq \Sigma^n$. The *minimum distance* of C, written $\Delta(C)$, is the minimum Hamming distance between pairs of different code words in C.

Definition 4. An Error Correcting Code is said to be *e error detecting* if it can tell that an error occurred when there were $\leq e$ errors, and at least one error occurred. It is said to be *t error correcting* if it can tell where the errors are when there were $\leq e$ errors, and at least one error occurred.

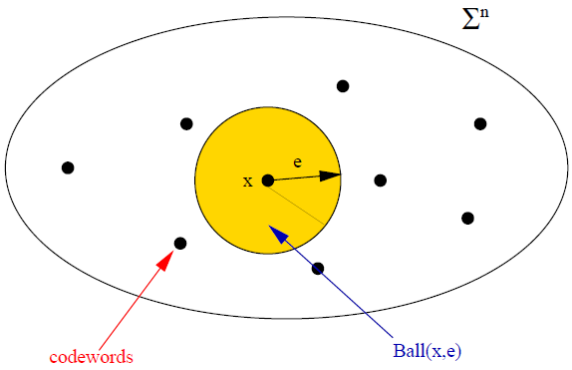
To formalize these notions we define the ball of radius t centered at x:

$$\text{Ball}(x,t) = \{y \in \Sigma^n | \Delta(x,y) \leq t\}$$

Definition 5. Formally: A code C is *t-error correcting* if $\forall x \neq y \in C$, $\text{Ball}(x,t) \cap \text{Ball}(y,t) = \emptyset$.



Definition 6. Formally: A code is *e-error detecting* if $\forall x \in C$, $\text{Ball}(x,e) \cap C = \{x\}$.



Definition 7. The *Hamming weight* $wt(v)$ of a vector v is the number of non-zero coordinates of v.

Proposition 8. $\Delta(C) = 2t+1 \Leftrightarrow$ *the code C is* $2t$ *error detecting and* t *error correcting.*

Proof. : $\Delta(C) = 2t+1 \Rightarrow$ 2t error detecting since the word would simply not be a code word. $\Delta(C) = 2t+1 \Rightarrow$ t error correcting since the word would be closer to the original code word than any other code word. We omit the reverse implications for now, though we note that the case for t error correcting is easy. □

- We now present some key code parameters:
  - $q = |\Sigma|$

- $n =$ block length of code(encoded length)
- $k =$ message length(pre-encoded length) = $\log_q |C|$
- $d = \Delta(C)$

- Usually, we fix three of the above parameters and try to optimize the fourth.
- Clearly, larger k and d values, and smaller n values are desirable. It also turns out that smaller q values are desirable.
- We may also try to maximize the rate of the code(efficiency ratio) $\frac{k}{n}$ and the relative distance $\frac{d}{n}$. We denote an error correcting code with these parameters as a $(n, k, d)_q$ code.

Thus, proving our claim boils down to showing that $\{bG|b \in \{0,1\}^4\}$ is a $(7, 4, 3)_2$ code.

Proof. : Assume that $\Delta(b_1G, b_2G) < 3$ for $b_1 \neq b_2 \Rightarrow \Delta((b_1 - b_2)G, 0) < 3 \Rightarrow \exists$ non-zero $c \in \{0,1\}^4$ s.t. $wt(cG) < 3$. Consider the matrix

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^T$$

It can be shown that 1 that $\{bG|b\} = \{y|yH = 0\}$. Hence, it suffices to prove that: if a non-zero $y \in \{0,1\}^7$ s.t. $yH = 0 \Rightarrow wt(y) \geq 3$, since this would contradict that $wt(cG) < 3$ for some non-zero c.

Assume $wt(y) = 2 \Rightarrow 0 = yH = h_i + h_j$, where $h_1, h_2, ..., h_7$ are the rows of the matrix H. But by the construction of H, this is impossible. Assume $wt(y) = 1 \Rightarrow$ some row h_i has all zeros. Again, impossible by construction. Thus $wt(y)$ is at least 3 and we are done. □

- From the properties of the matrix used above, we see that we may generalize the H matrix for codes of block length $n = 2^\ell - 1$ and minimum distance ≥ 3 simply by forming the $2^\ell - 1$ by ℓ matrix where the rows are the binary representations of all integers between 1 and $2^l - 1$. The message length k that this generalized H corresponds to is left as an exercise.
- Error correcting codes find application in a variety of different fields in mathematics and computer science.
  - In Algorithms, they can be viewed as interesting data structures.
  - In Complexity, they are used for pseudo-randomness, hardness amplification, and probabilistically checkable proofs.

– In Cryptography, they are applied to implement secret sharing schemes and proposal cryptosystems.
– Finally, they also arise in combinatorics and recreational mathematics.

1.2 Hamming Bound

Hamming Bound

Lemma 9. *For $\Sigma = \{0,1\}$ and $x \in \Sigma^n$,*

$$|\text{Ball}(x,t)| = \sum_{i=0}^{n} \binom{n}{i} \equiv Vol(n,t)$$

Proof. A vector $y \in \text{Ball}(x,t)$ if the number of coordinates of y that differ from x is at most t. Since $\Sigma = \{0,1\}$, they can only differ in one way, namely by being the opposite bit (0 if the bit of x is 1 and 1 if the bit of x is 0). Thus, to count the number of ways to be in the ball, we choose i of the n coordinates for i from 0 to t. We define $Vol(n,t)$ to be the number of points in the ball $\text{Ball}(x,t)$. □

Theorem 10. *For $\Sigma = \{0,1\}$, if C is t-error correcting, then*

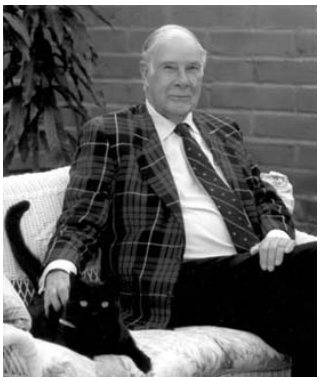
$$|C| \leq \frac{2^n}{Vol(n,t)} = \frac{2^n}{\sum_{i=0}^{n} \binom{n}{i}}.$$

Proof. If the code C is t-error correcting, then $\forall x \neq y \in C$, $Ball(x,t) \cap Ball(y,t) = \emptyset$, namely, the balls do not intersect. Thus, $|C|Vol(n,t) \leq Vol(AmbientSpace)$. Note that $Vol(AmbientSpace) = 2^n$, so dividing gives the result. □

- To decode a Hamming code (of minimum distance 3), we note that multiplying the received codeword by the parity check matrix H associated to the code will give 0 if no error occurred, while if 1 error occurred it will give the binary representation of the index of the bit where this error occurred.
- This is true because suppose we receive the codeword c with an error e_i (where e_i is the $0/1$ vector which is 1 only in the i-th coordinate). Then

  $$(c + e_i)H = cH + e_iH = e_iH.$$

  Now note that e_iH is simply the i-th row of H which by construction is the binary representation of i.
- Click here for examples with Haming codes `html/hamex.html`



Richard Wesley Hamming (1915-1998) American mathematician and computer scientist. Contributions to Information Theory, Coding Theory and Numerical Analysis. Involved in Manhattan project. Fellow of the ACM.

2 Linear Codes

2.1 Basics

Linear Codes

Definition 11. If the alphabet Σ is a finite field, then we say that a code C is *linear* if it is a linear subspace of Σ^n. That is, $x, y \in C$ implies $x + y \in C$ and $x \in C, a \in \Sigma$ implies $ax \in C$.

- Notationally, we represent linear codes with square brackets: $[n, k, d]_q$. All the codes we will see in this class are linear.
- Linear codes are interesting for many reasons. For example, the encoding function is simple, just matrix multiplication. It is also easy to detect errors: since the code is linear there is a parity check matrix H such that $C = \{y : Hy = 0\}$, and therefore we can detect errors again by simple matrix multiplication.
- Another interesting feature of a linear code is its *dual*. Let C be a linear code generated by the matrix G. C has a parity check matrix H. We define the dual code of C as the code generated by H^T, the transpose of the matrix H. It can be shown that the dual of the dual of C is C itself.
- For example, consider the Hamming code with block length $n = 2^\ell - 1$. Its dual is the code generated by a $\ell \times n$ matrix whose columns are all the non zero binary strings of length ℓ. It is easy to see that the encoding of a message b is

  $$< b, x >_{x \in \{0,1\}^\ell - 0},$$

  where $< ., . >$ denotes inner product modulo 2. In other words, the encoding of b is the parity check of all the non-empty subsets of the bits of b.

- It can be shown that if $b \neq 0$ then $< b, x >= 1$ for at least $2^{\ell-1}$ of the x's in $\{0,1\}^\ell - 0$. This implies that the dual code is a $[2^\ell - 1, \ell, 2^{\ell-1}]_2$ code. It can be shown that this is the best possible, in the sense that there is no $(2^\ell - 1, \ell + \epsilon, 2^{\ell-1})_2$ code. This code is called the *simplex code* or the *Hadamard code*. The second name comes from the french mathematician Jacques Hadamard who studied the $n \times n$ matrices M such that $MM^T = nI$, where I is the $n \times n$ identity matrix. It can be shown that if we form a matrix with the codewords from the previous code, we replace 0 with -1, and we pad the last bit with 0, then we obtain a Hadamard matrix.

2.2 Singleton bounds

Singleton bounds

So far we have discussed Shannon's theory, Hamming's metric and Hamming codes, and Hadamard codes. We are looking for aymptotically good codes to correct some constant fraction of errors while still trasmitting the information through the noisy channel at a positive rate. In our usual notation of $[n, k, d]_q$-codes, Hamming's construction gives a $[n, n - \log_2 n, 3]_2$-code while Hadamard codes were $[n, \log_2 n, \frac{n}{2}]_2$-codes. But we are looking for codes where $\frac{k}{n}$ and $\frac{d}{n}$ have a lower bound independent of n and q is not growing to ∞. In this scenario, we discuss the following simple impossibility result:

Theorem 12. *For a code $C : \Sigma^k \to \Sigma^n$ with minimum distance d, $n \geq k + d - 1$.*

Proof. In other words, we want to prove that $d \leq n - (k-1)$. Just project all the codewords on the first $(k-1)$ coordinates. Since there are q^k different codewords, by pigeon-hole principle at least two of them should agree on these $(k-1)$ coordinates. But these then disagree on at most the remaining $n - (k-1)$ coordinates. And hence the minimum distance of the code C, $d \leq n - (k-1)$. □

2.3 Reed-Solomon Codes

Reed-Solomon Codes

Definition 13. Let $\Sigma = \mathbb{F}_q$ a finite field and $\alpha_1, \ldots, \alpha_n$ be distinct elements of $\mathbb{F}_q$. Given n, k and $\mathbb{F}_q$ such that $k \leq n \leq q$, we define the encoding function for *Reed-Solomon codes* as: $C : \Sigma^k \to \Sigma^n$ which on message $m = (m_0, m_1, \ldots, m_{(k-1)})$ consider the polynomial $p(X) = \sum_{i=0}^{(k-1)} m_i X^i$ and $C(m) = \langle p(\alpha_1), p(\alpha_2), \ldots, p(\alpha_n) \rangle$.

Theorem 14. *Reed-Solomon code matches the singleton bound. i.e. it's a $[n, k, n - (k-1)]_q$-code.*

Proof. The proof is based only on the simple fact that a non-zero polynomial of degree l over a field can have at most l zeroes.

For Reed-Solomon code, two codewords (with corresponding polynomials p_1 and p_2) agree at i-th coordinate iff $(p_1 - p_2)(\alpha_i) = 0$. But $(p_1 - p_2)$, from

the above fact, can have at most $(k-1)$ zeros which means that the minimum distance $d \geq n-(k-1)$. $\square$

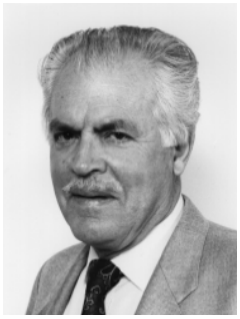
- Reed-Solomon codes are linear. This can be easily verified by the fact that the polynomials of degree $\leq (k-1)$ form a vector space (i.e. if p_1, p_2 are polynomials of degree $\leq (k-1)$ then similarly are βp_1 and $p_1 + p_2$). Since the polynomials $p(X) \equiv 1, p(X) = X, \ldots, p(X) = X^{(k-1)}$ form the basis for this vector space, we can also find a generator matrix for Reed-Solomon codes.

$$G = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \vdots & \vdots & \vdots & \vdots \\ {\alpha_1}^{(k-1)} & {\alpha_2}^{(k-1)} & \ldots & {\alpha_n}^{(k-1)} \end{bmatrix}$$

- One can also prove the theorem about the minimum distance of Reed-Solomon codes by using the fact that any k columns of G are linearly independent (because α_i's are distinct and thus the Vandermonde matrix formed by the k columns is non-singular).

- Using Reed-Solomon codes with $k = \frac{n}{2}$ we can get a $[n, \frac{n}{2}, \frac{(n+2)}{2}]_q$-code which means that we can correct much more erros than before. Typically Reed-Solomon codes are used for storage of information in CD's because they are robust against bursty errors that come in contiguous manner, unlike the random error model studied by Shannon. Also if some information is erased in the corrupted encoding, we can still retrieve the original message by interpolating the polynomial on the remaining values we get.

- A way to visualize Reed-Solomon code on binary alphabet is to consider $q = n = 2^m$. This gives $C : \Sigma^k = {\mathbb{F}_2}^{k \log_2 n} \to \Sigma^n = {\mathbb{F}_2}^{n \log_2 n}$ a $[n \log_2 n, k \log_2 n, n-(k-1)]_2$-code. And if you put $n \log_2 n = N$ and $n-(k-1) = 5$ then this gives a $[N, N - 4 \log_2 N, 5]_2$-code. As we have already seen, Hamming's construction gave a $[N, N - \log_2 N, 3]_2$-code and Hamming's impossibility result said that for a $[N, k, 2t+1]_2$-code, $k \geq N - t \log_2 N$. Reed-Solomon code don't achieve this but give a fairly closer $k \geq N - 2t \log N$ (We will discuss later about BCH codes which match this bound).



Irving Stoy Reed (1923 – 2012) mathematician and engineer. He is best known for co-inventing a class of algebraic error-correcting and error-detecting codes known as Reed–Solomon codes in collaboration with Gustave Solomon. He also co-invented the Reed–Muller code.

Reed made many contributions to areas of electrical engineering including radar, signal processing, and image processing.

Gustave Solomon (1930 – 1996) was a mathematician and electrical engineer who was one of the founders of the algebraic theory of error detection and correction. Solomon was best known for developing, along with Irving S. Reed, the algebraic error correction and detection codes named the Reed-Solomon codes.

2.4 Multivariate Polynomial Codes

Multivariate Polynomial Codes

Here instead of considering polynomials over one variable (like in Reed-Solomon codes), we will consider multivariate polynomials. For example, let's see the following encoding similar to Reed-Solomon codes but using bivariate polynomials.

Definition 15. Let $\Sigma = \mathbb{F}_q$ and let $k = l^2$ and $n = q^2$. A message is typically $m = (m_{00}, m_{01}, \ldots, m_{ll})$ and is treated as the coefficients of a bivariate polynomial $p(X, Y) = \sum_{i=0}^{l} \sum_{j=0}^{l} m_{ij} X^i Y^j$ which has degree l in each variable. The encoding is just evaluating the polynomial over all the elements of $\mathbb{F}_q \times \mathbb{F}_q$. i.e. $C(m) = \langle p(x, y) \rangle_{(x,y) \in \mathbb{F}_q \times \mathbb{F}_q}$.

But what is the minimum distance of this code ? We will use the following lemma to find it.

Lemma 16 (Schwartz-Zippel lemma)**.** *A multivariate polynomial $Q(X_1, \ldots, X_m)$ (not identically 0) of total degree L is non-zero on at least $(1 - \frac{L}{|S|})$ fraction of points in S^m, where $S \subseteq \mathbb{F}_q$.*

proof-of-lemma .... Induction on the number of variables. For $m = 1$, it's easy as we know that $Q(X_1, \ldots, X_m)$ can have at most L zeros over the field $\mathbb{F}_q$. Now assume that the induction hypothesis is true for the a multivariate polynomial with upto $(m-1)$ variables, for $m > 1$. Consider

$$Q(X_1, \ldots, X_m) = \sum_{i=0}^{t} {X_1}^i Q_i(X_2, \ldots, X_m)$$

where $t \leq L$ is the largest exponent of X_1 in $Q(X_1, \ldots, X_m)$. So the total degree of $Q_t(X_2, \ldots, X_m)$ is at most $(L - t)$. □

... *proof-of-lemma.* Induction hypothesis $\Rightarrow Q_t(X_2, \ldots, X_m) \neq 0$ on at least $(1 - \frac{(L-t)}{|S|})$ points in $S^{(m-1)}$. But suppose $Q_t(s_2, \ldots, s_m) \neq 0$ then $Q(X_1, s_2, \ldots, s_m)$ is a not-identically-zero polynomial of degree t in X_1, and therefore is non-zero on at least $(1 - \frac{t}{|S|})$ fraction of choices for X_1. So putting it all together, $Q(X_1, \ldots, X_m)$ is non-zero on at least $(1 - \frac{(L-t)}{|S|})(1 - \frac{t}{|S|}) \geq (1 - \frac{L}{|S|})$ points in S^m. □

Using this, for $m = 2$ and $S = \mathbb{F}_q$, we get that the above bivariate polynomial code is a $[q^2, l^2, (1 - \frac{2l}{q})q^2]_q$-code.

Some interesting cases of multivariate polynomial codes include Reed-Solomon code ($l = k$ and $m = 1$), Bivariate polynomial code ($l = \sqrt{k}$ and $m = 2$) and Hadamard code ($l = 1$, $m = k$ and $q = 2$) !

2.5 BCH Codes

BCH Codes

- A class of important and widely used cyclic codes that can correct multiple errors, developed by R. C. Bose

  and D. K. Ray-Chaudhuri (1960)[4] and independently by A. Hocquenghem (1959)[5]

- Let C_1 be the code over $F_n = F_{2^t}$, $n = 2^t$ with the following parity check matrix:

$$H = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 \\ 1 & \alpha_2 & \alpha_2^2 & \alpha_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \alpha_n^3 \end{bmatrix}$$

- C_1 is a $[n, n-4, 5]_n$ code. The reason why the code has distance 5 is as follows. If the code has distance 4, there exists 4 rows of H that are linearly dependent. However, this cannot happen because the submatrix consisting of 4 rows of H is a Vandemonde matrix whose determinant is non-zero when the elements are distinct.
- Now consider $C_{BCH} = C \cap \{0,1\}^n$. Clearly, the length and the distance of the code do not change so $C_{BCH} = [n, ?, 5]_2$. The main question here is how many codewords there are in CBCH. We know that the all zero codeword is in C_{BCH} but is there any other codeword?
- Let's represent all entries in $\mathbb{F}n$ by vectors such that:

$$\begin{aligned} \mathbb{F}_{2^t} &\longleftrightarrow \mathbb{F}_2^t \\ \alpha &\longleftrightarrow V_\alpha \\ 1 &\longleftrightarrow (100\ldots0) \end{aligned}$$

- Apply the representation above to H and we get a new matrix:

$$H' = \begin{bmatrix} 1 & V_{\alpha_1} & V_{\alpha_1^2} & V_{\alpha_1^3} \\ 1 & V_{\alpha_2} & V_{\alpha_2^2} & V_{\alpha_2^3} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & V_{\alpha_n} & V_{\alpha_n^2} & V_{\alpha_n^3} \end{bmatrix}$$

- If a $\{0,1\}$ vector $X = [x_1 \ldots x_n]$ satisfies $XH' = 0$ then

$$\begin{aligned} \sum x_i V\alpha_i &= 0 \\ V_{\sum x_i\alpha_i} &= 0 \\ \sum x_i\alpha_i &= 0 \\ XH &= 0 \end{aligned}$$

- Consider a matrix $\widetilde{H}$ equal to H' with the third column removed:

$$\widetilde{H} = \begin{bmatrix} 1 & V_{\alpha_1} & V_{\alpha_1^3} \\ 1 & V_{\alpha_2} & V_{\alpha_2^3} \\ \vdots & \vdots & \vdots \\ 1 & V_{\alpha_n} & V_{\alpha_n^3} \end{bmatrix}$$

Proposition 17. *For any $X \in \{0,1\}^n$ such that $X\widetilde{H} = 0$, $XH = 0$.*

Proof. The only question is whether $\sum x_i\alpha_i^2 = 0$. Over $\mathbb{F}_{p^t}$, $(x+y)^p = x^p + y^p$ $\forall x, y$. Therefore,

$$\begin{aligned} \sum x_i\alpha_i^2 &= \sum x_i^2\alpha_i^2 \\ &= \left(\sum x_i\alpha_i\right)^2 \\ &= \sum x_i\alpha_i = 0 \end{aligned}$$

The second and third columns of H impose on the code $\log n$ linear constraints each so the dimension of the code is $n - 2\log n - 1$. Thus, C_{BCH} is a $[n, n - 2\log n - 1, 5]$ code. □

- In general, the BCH code of distance d has the following parity check matrix:

$$\begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{d-2} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{d-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^{d-2} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \alpha_1 & \alpha_1^3 & \dots & \alpha_1^{d-2} \\ 1 & \alpha_2 & \alpha_2^3 & \dots & \alpha_2^{d-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^3 & \dots & \alpha_n^{d-2} \end{bmatrix}$$

- The number of columns is $1 + \left\lceil \frac{d-2}{2} \right\rceil \log n$ so the binary code we get satisfies $n - k \leq \left\lceil \frac{d-2}{2} \right\rceil \log n$.

- By the Hamming bound for code of length n and distance d,

$$2^k \binom{n}{d/2} \leq 2^n \rightarrow n - k \geq \frac{d}{2} \log \frac{n}{d}$$

- In the case $d = n^{o(1)}$, $n - k \geq \frac{d}{2} \log n$. Thus, BCH is essentially optimal as long as d is small.

- The problem is more difficult for bigger alphabet. Consider code over the ternary alphabet $\{0, 1, 2\}$. The Hamming bound is $n - k \geq \frac{d}{2} \log_3 n$. BCH technique gives $n - k = \frac{2}{3} d \log_3 n$ and we do not know how to get a better code in this case.

Raj Chandra Bose (19 June 1901 – 31 October 1987) was an Indian American mathematician and statistician best known for his work in design theory and the theory of error-correcting codes in which the class of BCH codes is partly named after him.

Dwijendra Kumar Ray-Chaudhuri (Born November 1, 1933) a Bengali-born mathematician and a statistician is a professor emeritus at Ohio State University. He is best known for his work in design theory and the theory of error-correcting codes, in which the class of BCH codes is partly named after him and his Ph.D. advisor Bose.

Alexis Hocquenghem (1908?-1990) was a French mathematician. He is known for his discovery of Bose–Chaudhuri–Hocquenghem codes, today better known under the acronym BCH codes.

References

References

[1] Thomas M. Cover, Joy A. Thomas, *Elements of Information Theory*, 2nd edition, Wiley, 2006.

[2] David J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.

[3] Robert M. Gray, *Entropy and Information Theory*, Springer, 2009

[4] John C. Bowman, Coding Theory, University of Alberta, Edmonton, Canada, 2003

[5] D. G. Hoffmann, Coding Theory. The Essential, Marcel Deker, 1991

References

[1] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, 1948.

[2] R. V. Hamming, Error detecting and error correcting codes, *Bell System Technical Journal*, 29: 147-160, 1950

[3] Reed, Irving S.; Solomon, Gustave, Polynomial Codes over Certain Finite Fields, *Journal of the Society for Industrial and Applied Mathematics* (SIAM) 8 (2): 300–304, 1960

[4] Bose, R. C., Ray-Chaudhuri, D. K., On A Class of Error Correcting Binary Group Codes, *Information and Control* 3 (1): 68–79, 1960

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