Generalized BCH codes

A Generalized BCH code has the following *design parameters*:

n = the *base degree* of the code

m = the *extension degree* of the code

t = the *correction capability* of the code

n is the number of bits per symbol, while $2^{nm} - 1$ is the size of a block (in symbols) and t is the number of symbol errors that can be corrected per block.

Before we get started, a few special cases with special names: if $m = 1$ the code is called a Reed-Solomon code, if $n = 1$ it is a "binary BCH code."

In the discussion that follows, let m, n and t be fixed. If $\gamma \in GF(2^{mn})$, let $K(\gamma)$ denote the smallest positive integer such that $\gamma^{2^{nK(\gamma)}} = \gamma$. Also, let $P_\gamma(x)$ denote the unique monic polynomial of smallest degree with coefficients in $GF(2^n)$ such that $P_\gamma(\gamma) = 0$. It is reasonably easy to show with a bit of algebra (and we'll do so later) that

$$P_\gamma(x) = \prod_{i=0}^{K(\gamma)-1} (x + \gamma^{2^{ni}})$$

Let α be a generator for $GF(2^{mn})$ (that is, an element α such that the powers of α are all the nonzero elements of $GF(2^{mn})$ (such an element always exists by a fairly high-powered bit of mathematics which states that finite fields always have cyclic multiplicative groups). Then, compute the polynomial

$$G(x) = LCM(P_\alpha(x), P_{\alpha^2}(x), \ldots, P_{\alpha^{2t}}(x))$$

Then, $G(x)$ will have coefficients in $GF(2^n)$ and will be the generator of a Generalized BCH Code with the desired properties.

The Big Question, though, is: What is the degree of $G(x)$?

Some estimates can be given in the special cases - in the case of a Reed-Solomon code, $P_\gamma(x)$ always has degree 1 (since $K(\gamma) = 1$ because the multiplicative group of $GF(n)$ has order $2^n - 1$ and hence $\gamma^{2^n - 1} = 1$ so that $\gamma^{2^n} = \gamma$). Furthermore, all the factors in the definition of $G(x)$ are distinct, hence relatively prime, so that $G(x)$ has degree $2t$. An example of a code like this would come from $n = 8$, $m = 1$, $t = 3$. The degree of $G(x)$ would be 6 and the block would have 255 8-bit symbols, 6 of which would be generated, so that altogether we would have a $(2040, 1992)$ code, which would be able to correct any 3 8-bit symbols in error (if the errors were aligned just right, it would be able to correct up to 24 bits).

In the case of the binary BCH code, it's harder to say exactly what degrees the P_{α^j} have, but it *is* easy to see that half of them are redundant. This is because $n = 1$ so that $P_\gamma = P_{\gamma^2}$, thus all the even exponents from 2 up to $2t$ are redundant in the definition of $G(x)$. To get a bound on the degree of $G(x)$, we observe that, for all γ, $K(\gamma) \leq K(\alpha)$ and $K(\alpha) = m$ so that $\deg(G(x)) \leq mt$. An example of a code like this would correspond to $n = 1$, $m = 11$, $t = 4$ which would have a 2047 bit block size and (at most) 44 generated bits, giving (at worst) a $(2047, 2003)$ code, which would be able to correct any 4 bits in error.

I will calculate an example of each of these special cases and give one general case as well. For these cases, I will use arithmetic in $GF(16)$ which will be performed using the following scheme: each nonzero element in $GF(16)$ can be written in one of two ways, either as a nonzero 4-bit string (relative to which addition is easy) or as a power of α (a generator of $GF(16)$), relative to which multiplication is easy. All we need, then, is a means to translate between the two, which is provided by the following:

$$\begin{aligned}
&0001 = \alpha^0, 0010 = \alpha, 0011 = \alpha^4,\\
&0100 = \alpha^2, 0101 = \alpha^8, 0110 = \alpha^5, 0111 = \alpha^{10},\\
&1000 = \alpha^3, 1001 = \alpha^{14}, 1010 = \alpha^9, 1011 = \alpha^7,\\
&1100 = \alpha^6, 1101 = \alpha^{13}, 1110 = \alpha^{11}, 1111 = \alpha^{12}
\end{aligned}$$

Example 1: $m = 4, n = 1, t = 4$ (a 4-bit-correcting binary BCH code) We need to calculate P_α, P_{α^3}, P_{α^5} and P_{α^7}. We will do this by seeing what the zeroes of each of these polynomials are: P_α has roots α, α^2, α^4, and α^8. P_{α^3} has roots α^3, α^6, α^{12}, and α^9. P_{α^5} has roots α^5, and α^{10}. P_{α^7} has roots α^7, α^{14}, α^{13}, and α^{11}. None of these have factors in common, so that $G(x)$ has degree 14 and is equal to

$$\begin{aligned}
G(x) &= \prod_{i=1}^{14}(x + \alpha^i)\\
&= x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1
\end{aligned}$$

One might note that this is a $(15, 1)$ code and probably would never be used in practice!

Example 2: $m = 1, n = 4, t = 4$ (a 4-symbol correcting (60,28) Reed-Solomon code) Here, $G(x)$ is easy to write down (but rather tedious to calculate its coefficients):

$$\begin{aligned}
G(x) &= (x + \alpha)(x + \alpha^2)(x + \alpha^3)(x + \alpha^4)(x + \alpha^5)(x + \alpha^6)(x + \alpha^7)(x + \alpha^8)\\
&= x^8 + \alpha^{14}x^7 + \alpha^2x^6 + \alpha^4x^5 + \alpha^2x^4 + \alpha^{13}x^3 + \alpha^5x^2 + \alpha^{11}x + \alpha^6
\end{aligned}$$

Example 3: $m = 2, n = 2, t = 2$ (a 2-symbol correcting code with 2-bit symbols and a 15-symbol block size). Here, we need to calculate P_α, P_{α^2}, P_{α^3} and P_{α^4}. As in Example 1, we will do this by seeing what the zeroes of each of these polynomials are: P_α has roots α, α^4 (and hence $P_\alpha = P_{\alpha^4}$). P_{α^2} has roots α^2, α^8. P_{α^3} has roots α^3 and α^{12}. Except for the redundancy between P_α and P_{α^4}, none of these has common factors, so that

$$\begin{aligned}
G(x) &= (x + \alpha)(x + \alpha^2)(x + \alpha^3)(x + \alpha^4)(x + \alpha^8)(x + \alpha^{12})\\
&= x^6 + \beta^2x^5 + x^4 + x^3 + \beta x^2 + \beta x + 1
\end{aligned}$$

where β is a generator of $GF(4)$ (arithmetic is generated by $01 = \beta^0, 10 = \beta, 11 = \beta^2$). This is a (30,18) code.

As an exercise, work out the Reed-Solomon code corresponding to $m = 1, n = 5, t = 2$. In order to work out the arithmetic in $GF(32)$ you will need an irreducible polynomial over Z_2 of degree 5 - use

$$x^5 + x^2 + 1$$

With this polynomial, $\alpha = 00010$ works as a generator. If you can work this example out completely, then you really understand this stuff...

Why this works:

There are two aspects of all this to prove. First, we need to prove that the coefficients of $P_\gamma(x)$ are actually in $GF(2^n)$. Then, we also need to show that no multiples of $G(x)$ with degree less than $2^{mn}-1$ have fewer than $2t+1$ nonzero coefficients.

Why $P_\gamma(x)$ has coefficients in $GF(2^n)$:

We will need a couple of lemmas:

Lemma 1: Let x, m, n be positive integers with $x \geq 2$. Then, $(x^m-1)|(x^n-1)$ if and only if $m|n$.

Proof: First, the "if" part – suppose $m|n$, say, $n = mk$. Then, using the usual factorization of the difference of powers, $x^n - 1 = x^{mk} - 1^k = (x^m-1)\sum_{i=0}^{k-1} x^{mi}$. Thus, $(x^m-1)|(x^n-1)$.

To go the other direction, suppose that $(x^m-1)|(x^n-1)$. Use the integer division algorithm to write $n = mq + r$ where q, r are nonnegative integers, and $0 \leq r < m$. Then,

$$\begin{aligned} x^n - 1 &= x^{mq+r} - 1 \\ &= x^{mq+r} - x^r + x^r - 1 \\ &= x^r(x^{mq}-1) + (x^r - 1) \\ &= (x^m-1)x^r \sum_{i=0}^{q-1} x^{mi} + (x^r-1) \end{aligned}$$

Since $x^m - 1$ is a factor of the left-hand side and of the first term on the right, it must be a factor of the second term as well. So, $(x^m-1)|(x^r-1)$. Now, $r < m$ and $x \geq 2$ implies that $x^r - 1 < x^m - 1$ so that $x^r - 1$ is a nonnegative multiple of $x^m - 1$ that is less than $x^m - 1$. The only possibility is $x^r - 1 = 0$ which implies that $r = 0$ and hence $m|n$. **Q.E.D. (Lemma 1)**

Lemma 2: If K is a field of characteristic 2, then $f : K \to K$ given by $f(x) = x^2$ is a 1-1 field homomorphism. Furthermore, if K is finite, f is an isomorphism (i.e., it is actually onto K).

Proof: First, recall that characteristic 2 means that $x + x = 0$ for all $x \in K$. Next, recall that to show that f is a field homomorphism, we merely need to show that $f(xy) = f(x)f(y)$ and $f(x+y) = f(x) + f(y)$. The first equation is always true for the indicated f, so the second is the only one we need worry about. Using the "characteristic 2" assumption, we have:

$$\begin{aligned} f(x+y) &= (x+y)^2 \\ &= x^2 + xy + yx + y^2 \\ &= x^2 + y^2 + (xy + xy) \\ &= x^2 + y^2 \\ &= f(x) + f(y) \end{aligned}$$

To show that f is 1-1, we need to show that any element of K that has a square root has exactly **one** square root. Let a, b be in K and let b be a zero of of the polynomial $x^2 - a$. Since the linear term of this polynomial is zero, this implies that the "other" root

is the negative of b. But, again using characteristic 2, $-b = b$. Hence, this polynomial has one repeated root and thus f is 1-1.

If K is finite, then we simply observe that $f(K)$ has the same cardinality as K, so it must be all of K. **Q.E.D. (Lemma 2)**

Now, back to the subject at hand - when you have one field contained in another field, the easy way to show that an element is in the smaller field is to use **Galois theory**. The part of this theory that we will need for our purposes is that the subfield is always the *fixed field* of some group of transformations of the larger field, that is, the set of elements which are left intact by every transformation in the group. Such a transformation is completely determined by what it does to α, since α is a generator of the field. Now, the question is: where can α be sent under such a transformation? The short answer is that it has to go to another generator, that is, another element whose powers generate all the $2^{mn} - 1$ nonzero elements of $GF(2^{mn})$. In other words, α must go to α^k where k is relatively prime to $2^{mn} - 1$. In addition, it must go to another root of the *same* irreducible polynomial as α. This implies that there are at most mn possible values for k. Using Lemma 2, we see precisely what values of k work, since that lemma tells us that $k = 2$ always works. Repeating this transformation, we see that $k = 4$, $k = 8$, etc. work as well. This process may be repeated until we arrive at $k = 2^{mn}$ which is the identity map since $x^{2^{mn}} = x$ for all $x \in GF(2^{mn})$. Hence, this lemma provides us with all mn possible values for k.

Now, which of these transformations leave $GF(2^n)$ fixed? Well, which powers of α are in $GF(2^n)$? The easy way to see this is to use the "cyclic multiplicative group" theorem referred to before. Then, one sees that the nonzero elements in $GF(2^n)$ sit inside $GF(2^{mn})$ as the powers of $\beta = \alpha^{(2^{mn}-1)/(2^n-1)}$ (note that this exponent is an integer even though it doesn't look like one - it can also be written as $1 + 2^n + 2^{2n} + \ldots + 2^{(m-1)n}$). Hence, the transformations of $GF(2^{mn})$ which fix $GF(2^n)$ are precisely those for which α is taken to α^{2^k} and

$$2^k(2^{mn} - 1)/(2^n - 1) \equiv (2^{mn} - 1)/(2^n - 1) \pmod{2^{mn} - 1}$$

or, more helpfully,

$$(2^k - 1)(1 + 2^n + 2^{2n} + \ldots + 2^{(m-1)n}) \equiv 0 \pmod{2^{mn} - 1}$$

which means that $2^k - 1$ is a multiple of $2^n - 1$. Using Lemma 1 now gives us that k must be a multiple of n. Hence, the transformations that leave $GF(2^n)$ fixed are exactly the transformations generated by $\alpha \mapsto \alpha^{2^{nk}}$ for $k = 0, 1, \ldots, m - 1$. For our purposes, the key fact is that anything which is preserved by all of these transformations must be in $GF(2^n)$.

The rest of this result is easy, since, for any $\gamma \in GF(2^{mn})$, these transformations permute the roots of P_γ and the coefficients of P_γ are symmetric functions of the roots (sum, product, sum of products taken two at a time, etc.). Thus, these transformations fix all the coefficients of P_γ and therefore the coefficients are in $GF(2^n)$.

Why all nonzero multiples of G have at least $2t+1$ nonzero coefficients:
We will need a lemma for this one as well.

Lemma 3: For $n \geq 2$, let $\mathrm{VDM}(x_1, x_2, \ldots, x_n)$ denote the determinant of the matrix

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$$

Then, $\mathrm{VDM}(x_1, x_2, \ldots, x_n) = \prod_{1 \leq i < j \leq n}(x_j - x_i)$

Proof: This determinant is called the Van der Monde determinant. We will induct on n. First, the base case $n = 2$ is clearly true. Now, assume that the theorem is true for $n = 2, 3, \ldots, k-1$ and consider $n = k$. In particular, consider $\mathrm{VDM}(x_1, x_2, \ldots, x_k)$ as a function $f(x)$ where $x = x_k$. By expanding the determinant along the last column, we see that this is a polynomial of degree $k-1$. Furthermore, if $x_k = x_j$ for any $j < k$, the determinant is zero, hence $f(x) = C(x - x_1)(x - x_2)\ldots(x - x_{k-1})$ where C is constant with respect to x. Evaluating the coefficient of x_k^{k-1} in the expansion by minors, we see that $C = \mathrm{VDM}(x_1, x_2, \ldots, x_{k-1})$. Hence (using the induction hypothesis),

$$\begin{aligned} \mathrm{VDM}(x_1, x_2, \ldots, x_k) &= \mathrm{VDM}(x_1, x_2, \ldots, x_{k-1}) \prod_{i=1}^{k-1}(x_k - x_i) \\ &= \left(\prod_{1 \leq i < j \leq k-1} (x_j - x_i) \right) \left(\prod_{1 \leq i \leq k-1} (x_k - x_i) \right) \\ &= \prod_{1 \leq i < j \leq k} (x_j - x_i) \end{aligned}$$

Q.E.D. (Lemma 3)

The main proof of this section now begins with the first two words of all good proofs: "suppose not." That is, suppose we have a nonzero polynomial $H(x)$ such that:

(i) H has degree less than $2^{mn} - 1$,

(ii) H has coefficients in $GF(2^n)$, no more than $2t$ of which are nonzero, and

(iii) $H(x) = G(x)Q(x)$ for some $Q(x)$ with coefficients in $GF(2^n)$.

Believe it or not, the rest of this proof is just linear algebra (although it's linear algebra with all the coefficients in $GF(2^{mn})$ so it's a little different from what you're accustomed to).

Let's represent $H(x)$ as $H(x) = A_1 x^{k_1} + A_2 x^{k_2} + \cdots + A_{2t} x^{k_{2t}}$ where the A_i are in $GF(2^n)$, not all are zero, and the k_i are distinct integers between 0 and $2^{mn} - 2$, inclusive. Then, (iii) simply implies that $H(\alpha^i) = 0$ for all $1 \leq i \leq 2t$. This may be represented in

matrix form as $Hv = 0$ where

$$H = \begin{pmatrix} \alpha^{k_1} & \alpha^{k_2} & \cdots & \alpha^{k_{2t}} \\ \alpha^{2k_1} & \alpha^{2k_2} & \cdots & \alpha^{2k_{2t}} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{2tk_1} & \alpha^{2tk_2} & \cdots & \alpha^{2tk_{2t}} \end{pmatrix}$$

$$v = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{2t} \end{pmatrix}$$

Since H is a square matrix and not all the A_i are zero, the only way for this to be true is for $\det H = 0$. Now, factor α^{k_i} from column i of H to get

$$\begin{aligned} \det H &= \alpha^{k_1+k_2+\cdots+k_{2t}} \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{k_1} & \alpha^{k_2} & \cdots & \alpha^{k_{2t}} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(2t-1)k_1} & \alpha^{(2t-1)k_2} & \cdots & \alpha^{(2t-1)k_{2t}} \end{pmatrix} \\ &= \alpha^{k_1+k_2+\cdots+k_{2t}} \mathrm{VDM}(\alpha^{k_1}, \alpha^{k_2}, \ldots, \alpha^{k_{2t}}) \\ &= \alpha^{k_1+k_2+\cdots+k_{2t}} \prod_{1 \le i < j \le 2t} (\alpha^{k_j} - \alpha^{k_i}) \end{aligned}$$

Since all of the k_i are distinct and less than $2^{mn} - 1$, all of the α^{k_i} are distinct and thus $\det H$ is nonzero, contradicting our assumption.

Summary We have proven that $G(x)$ is a well-defined polynomial with coefficients in $GF(2^n)$ and that all its nonzero multiples of degree $< 2^{mn}$ must have at least $2t + 1$ nonzero coefficients. It follows that the cyclic code generated by $G(x)$ over $GF(2^n)$ is capable of correcting t symbol errors.

It is also worth noting that it is not, strictly speaking, necessary to use $\alpha, \alpha^2, \ldots, \alpha^{2t}$ as the generating roots for $G(x)$ – any sequence of $2t$ consecutive powers of α will work (the same proof applies, except that the factor in front of the Van der Monde determinant is different). Occasionally, this allows for a slightly more efficient code (lower degree for $G(x)$).