The Enumeration of Information Symbols in BCH Codes

By E. R. BERLEKAMP

(Manuscript received May 9, 1967)

This paper presents certain formulas for $I(q, n, d)$, the number of information symbols in the q-ary Bose-Chaudhuri-Hocquenghem code of block length $n = q^m - 1$ and designed distance d. By appropriate manipulations on the m-digit q-ary representation of d, we derive a simple linear recurrence for a sequence whose mth term is the number of information symbols in the BCH code.

In addition to an exact solution of all finite cases, we obtain exact asymptotic results, as n and d go to infinity while their ratio n/d remains fixed. In this limit, the number of information symbols increases as n^s. Specifically, we show that for fixed u, $0 \leqq u \leqq 1$,

$$\lim_{m \to \infty} q^{-ms} I(q, q^m - 1, uq^m) = 1,$$

where s is a singular function of u. The function s(u) is continuous and monotonic nonincreasing; it has derivative zero almost everywhere. Yet $s(0) = 1$ and $s(1) = 0$. For $q = 2$, s(u) is plotted in Fig. 1.

Any cyclic code of block length n over $GF(q)$ may be defined by its generator polynomial, $g(x)$, which is some factor of $x^n - 1$ over $GF(q)$, or by its check polynomial, $h(x) = (x^n - 1)/g(x)$. The number of check digits in the code is given by the degree of $g(x)$; the number of information digits, by the degree of $h(x)$. We assume that n and q are relatively prime. It is most convenient to work in a particular extension field of $GF(q)$, namely $GF(q^m)$, where m is the multiplicative order of q mod n. In this field, $x^n - 1$ factors into distinct linear factors:

$$x^n - 1 = \prod_{i=1}^{n} (x - \alpha^i).$$

Here α is any primitive nth root of unity in $GF(q^m)$; $\alpha^n = 1$. From the factorization $x^n - 1 = g(x)h(x)$, we see that every power of α is a root

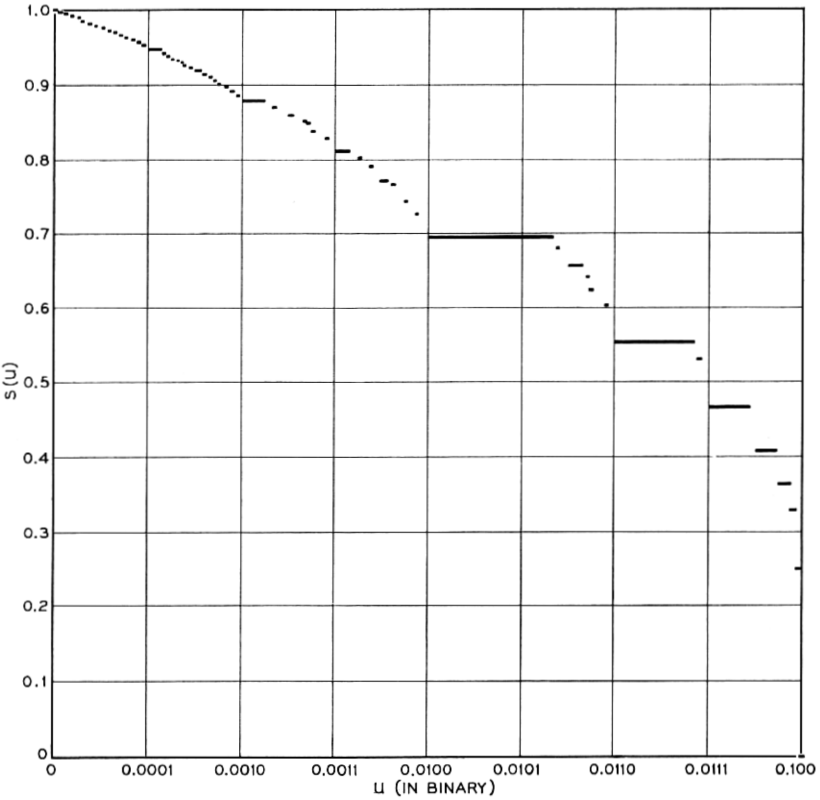

Fig. 1 — Graph of $s(u)$ vs. u.

of either $g(x)$ or of $h(x)$, but not both. Thus, a cyclic code partitions the powers of α into two sets: those powers which are roots of the generator polynomial, and those powers which are roots of the check polynomial. If $g(x)$ and $h(x)$ were permitted to have coefficients in $GF(q^m)$, then any partition of the powers of α would define a cyclic code. However, the coefficients of $g(x)$ and $h(x)$ must lie in the ground field $GF(q)$. Consequently, if α^i is a root of $g(x)$, then so are the conjugates of α^i, namely α^{iq}, α^{iq^2}, α^{iq^3}, $\cdots$. Conversely, if all conjugates of roots of $g(x)$ are also roots of $g(x)$, and all conjugates of roots of $h(x)$ are also roots of $h(x)$, then all of the coefficients of the polynomials $g(x)$ and $h(x)$ lie in $GF(q)$.

The previous remarks hold for all cyclic codes.

A q-ary BCH code of block length n over $GF(q)$ may be defined as the cyclic code whose generator's roots include only $\alpha, \alpha^2, \cdots, \alpha^{d-1}$ and their conjugates. This code is capable of correcting any combination of less than $d/2$ errors; (*cf.* Berlekamp[1]) the minimum Hamming distance of this code is at least d. For this reason, d is called the designed distance of the code.

The first result on the number of information symbols in BCH codes is the following lemma:

Classical Lemma I: Let $I(q, n, d)$ be the number of information symbols in the q-ary BCH code of block length n and designed distance d.

Define $\lceil i \rceil$ by the equations

$$i \equiv \lceil i \rceil \text{ mod } n \quad \text{and} \quad 1 \leqq \lceil i \rceil \leqq n.$$

Then $I(q, n, d)$ is the number of integers j, such that $1 \leqq j \leqq n$ and $\lceil jq^k \rceil \geqq d$ for all k.

Proof: α^i is a root of the generator polynomial of the BCH code iff there exists some $k(j)$ such that $\lceil jq^k \rceil < d$. Conversely, α^i is a root of the check polynomial iff $\lceil jq^k \rceil \geqq d$ for all k. Q.E.D.

The classical lemma enables one to compute the number of information symbols in any q-ary given BCH code without doing any calculations in $GF(q)$ or its extensions. One need only enumerate certain types of residue classes mod n. In practice, this enumeration is still often tedious, particularly when n and d are large.

In order to obtain more tractable results for large n and d, we prefer to start from an alternate form of the classical lemma:

Classical Lemma II: Let $I(q, n, d)$ be the number of information symbols in the q-ary BCH code of block length n and designed distance d.

Define $\lfloor i \rfloor$ by the equations

$$i \equiv \lfloor i \rfloor \text{ mod } n \quad \text{and} \quad 0 \leqq \lfloor i \rfloor \leqq n - 1.$$

Then, $I(q, n, d)$ is the number of integers i, such that $0 \leqq i \leqq n - 1$ and $\lfloor iq^k \rfloor < n + 1 - d$ for all k.

Proof: $1 \leqq j \leqq n$ and $\lceil jq^k \rceil \geqq d$ for all k iff $0 \leqq (n - j) \leqq n - 1$ and $\lfloor (n - j)q^k \rfloor \leqq n - d$ for all k. Let $i = n - j$. Q.E.D.

In the wide sense, BCH codes may be defined over any alphabet whose order, q, is a prime power, and for any block length, n, which is relatively prime to q. In the narrow sense, however, n is required to be one less than a power of q. For narrow sense codes, the smallest extension field of $GF(q)$ which contains the nth roots of unity is $GF(n+1)$. For wide sense codes, this extension field is always larger, usually much larger. Since the decoder must perform certain computations in this extension field, narrow sense BCH codes are more easily implemented than their more general wide sense counterparts.

We shall enumerate the information symbols in narrow sense BCH codes by reducing the problem to the enumeration of certain kinds of sequences over the alphabet consisting of the integers $0, 1, 2, \cdots, q-1$, as first suggested by Mann.[2] We begin by defining the appropriate manipulations with such sequences.

We shall always use *capital letters for sequences.* We let $(Q - 1)$ denote the sequence consisting of the single letter $q - 1$. Unless otherwise stated, we allow every sequence to be either finite or infinite.

Let $V = V_1V_2V_3 \cdots$ be any finite or infinite q-ary sequence (i.e., a sequence of numbers V_i, where V_i is an integer, $0 \leqq V_i \leqq q - 1$. We let $\bar{V} = \bar{V}_1\bar{V}_2\bar{V}_3 \cdots$ denote the *complement* of V, defined by $\bar{V}_i = (q-1) - V_i$ for all i. If $W = W_1W_2 \cdots W_k$ is a finite q-ary sequence, then we may form the *cyclic shifts* of W: $W_2W_3 \cdots W_kW_1$, $W_3W_4 \cdots W_kW_1W_2$, $\cdots$. If X is a finite q-ary sequence, $X = X_1X_2 \cdots X_j$, then we may form the *concatenation* $X * V = X_1X_2X_3 \cdots X_jV_1V_2V_3 \cdots$. This concatenation may be formed whenever V is a finite or infinite q-ary sequence. If V is a finite q-ary sequence, then $V * X$ is a cyclic shift of $X * V$.

The q-ary sequence Y is said to be a *prefix* of X iff $X = Y * Z$ for some Z; Y is said to be a *suffix* of X iff $X = Z * Y$ for some Z. A prefix must be a finite (or empty) sequence; a suffix may be empty, finite, or infinite. V is a *proper prefix* of X iff $X = V * Z$, and neither V nor Z is empty. Z is a *proper suffix* of X iff $X = V * Z$ and neither V nor Z is empty. If X is a finite q-ary sequence, $X = X_1X_2 \cdots X_k$, then we may form the *iterated concatenation* of X with itself, $\dot{X} = X_1X_2 \cdots X_kX_1X_2 \cdots X_k \cdots$. In particular, $(Q \dot{-} 1)$ denotes the infinite q-ary sequence all of whose letters are $q - 1$.

We say $X < Y$ iff there exists a j such that $X_i = Y_i$ for $i = 1, 2, \cdots, j-1$, but $X_j < Y_j$. If $X \not< Y$ and $Y \not< X$, then one is a prefix of the other.

This ordering is similar to the numerical ordering of q-ary fractions, but there are important differences. For example, $\frac{1}{4} = 0.01 < 0.0101 = 5/16$, but the sequences 01 and 0101 are incomparable, because one is a prefix of the other. On the other hand, $0.0111111 \cdots = 0.1 = \frac{1}{2}$, yet $01111 \cdots < 1$. This type of example may be excluded by writing all fractions in their terminating form if they have one. We may then assert the following:

Let

$$u = \sum_i U_i q^{-i}, \qquad v = \sum_i V_i q^{-i}, \qquad U = U_1 U_2 U_3 \cdots$$

and $V = V_1 V_2 V_3 \cdots$, *and suppose that* $(Q \dot{-} 1)$ *is not a suffix of* U *or* V. *Then*

$$U < V \Rightarrow u < v$$

$$u \leqq v \Leftrightarrow \begin{cases} U < V \\ \quad \textit{or} \\ U \textit{ is a prefix of } V. \end{cases}$$

We say that X is an *immediate subordinate* of Y iff X is a finite sequence, $X = X_1 X_2 X_3 \cdots X_k$, and $X_1 = Y_1$, $X_2 = Y_2$, $\cdots$, $X_{k-1} = Y_{k-1}$, but $X_k < Y_k$. The sequence Y has Y_1 immediate subordinates of length 1, Y_2 immediate subordinates of length 2, Y_3 immediate subordinates of length 3, $\cdots$ Y_k immediate subordinates of length k. If the sequence Y has only a finite number of nonzeros, then we may define the *greatest immediate subordinate* of Y. If the last nonzero in the sequence $Y = Y_1 Y_2 \cdots$ is Y_k, then the greatest immediate subordinate of Y is $Y_1 Y_2 \cdots Y_{k-1}(Y_k - 1)$. If the sequence Y contains an infinite number of nonzeros, then Y has infinitely many immediate subordinates. All of them are less than Y itself, but none of them is the greatest immediate subordinate.

Similarly, we say that Y is an *immediate superior* of X iff $Y = Y_1 Y_2 Y_3 \cdots Y_k$, where $Y_1 = X_1$, $Y_2 = X_2$, $\cdots$, $Y_{k-1} = X_{k-1}$ but $Y_k > X_k$. If $X = X_1 X_2 \cdots X_k$ and $X_k \neq (Q - 1)$, then the *least immediate superior* of X is $Y = Y_1 Y_2 \cdots Y_k$; $Y_i = X_i$ for $i = 1, 2, \cdots, k - 1$, and $Y_k = X_k + 1$. It should be evident that the least immediate superior is among the longest immediate superiors, and the greatest immediate subordinate is among the longest immediate subordinates.

Definition: If q is any integer, U is any infinite q-ary sequence and m is any integer, we define $J(q, U, m)$ to be the number of q-ary m-tuples all of whose cyclic shifts are less than U.

Lemma III: (Complemented form of Mann's Lemma)

If

$$n = q^m - 1, \qquad n + 1 - d = \sum_{i=1}^{m} U_i q^{m-i}, \qquad 0 \leqq U_i < q,$$

$U = U_1U_2 \cdots U_m$, *and* Y *is any* q*-ary sequence then*

$$I(q, n, d) = J(q, U * Y, m).$$

Proof: Lemma III reduces to Lemma II under the following correspondence: The q-ary m-tuple U corresponds to the integer $n + 1 - d$; another q-ary m-tuple $W = W_1W_2 \cdots W_m$ corresponds to the integer $w = \sum_{i=1}^{m} W_i q^{m-i}$. The first cyclic shift of W is the sequence $W_2W_3 \cdots W_mW_1$, which then corresponds to the integer

$$\sum_{i=1}^{m-1} W_i q^{m+1-i} + W_1 = qw - (q^m - 1)W_1 .$$

Modulo $n = q^m - 1$, the integer corresponding to the first cyclic shift of W is seen to be congruent to qw. Therefore, the successive cyclic shifts of an m-digit q-ary sequence W correspond to the integers $\lfloor w \rfloor$, $\lfloor wq \rfloor$, $\lfloor wq^2 \rfloor$, $\cdots$, $\lfloor wq^{m-1} \rfloor$. These integers are all $< n + 1 - d$ iff all cyclic shifts of W are $< U$, which is true iff all cyclic shifts of W are $< U * Y$, for any Y. Q.E.D.

The choice $Y = \dot{U}$ has an interesting interpretation:

$$\sum_{i=1}^{\infty} \dot{U}_i q^{-i} = \sum_{k=0}^{\infty} \sum_{i=1}^{m} U_i q^{-(i+mk)} = \left(\sum_{i=1}^{m} U_i q^{-i} \right) / (1 - q^{-m})$$

$$= \left(\sum_{i=1}^{m} U_i q^{m-i} \right) / (q^m - 1) = 1 - \frac{(d - 1)}{n}.$$

Thus, the sequence $\dot{U}$ is the q-ary expansion of $1 - (d - 1)/n$. For this reason, we may investigate the behavior of $I(q, n, d)$ for large n and d with a fixed fractional error correction capability, $(d - 1)/2n$, by studying $J(q, \dot{U}, m)$ as a function of m for fixed q and $\dot{U}$.

We shall temporarily ignore the periodicity of the $\dot{U}$ sequence, and consider the function $J(q, V, m)$ for an arbitrary q-ary sequence V. We assume only that the sequence V has no terminal zeros.

From the definition of the immediate subordinates of V, it is clear that *if an* m*-digit* q*-ary sequence* W *is less than* V*, then some immediate subordinate of* V *is a prefix of* W. For if W is less than V, then there exists a k such that $W_i = V_i$ for $i = 1, 2, \cdots, k - 1$, but $W_k < V_k$, and the sequence $W_1W_2 \cdots W_k$ is a prefix of W and an immediate subordinate of V.

Now suppose that some m-digit sequence W and all of its cyclic shifts are less than V. Since W itself is less than V, some prefix of W must be an immediate subordinate of V. Are all possible immediate subordinates of V possible prefixes of W? In general, they are not, for some immediate subordinates may have suffixes which are greater than V. If $X * Y$ is an immediate subordinate of V and Y is greater than V, then $X * Y$ cannot be a prefix of W. For, if $W = X * Y * Z$, then one of the cyclic shifts of W is $Y * Z * X$ which is greater than V.

For example, consider the ternary sequence $V = 20212$. Its immediate subordinates are 0, 1, 200, 201, 2020, 20210, and 20211. The immediate subordinate 20210 has the suffix 210 which is greater than V. Therefore, if 20210 is the prefix of W, then the second cyclic left shift of W is greater than V. Similarly, V's immediate subordinate 20211 has the suffix 211, which is also greater than V.

For some sequences V, this difficulty does not arise. If V exceeds all of its own proper suffixes, then we have the following theorem:

Theorem 1: Let V be a q-ary sequence which exceeds all of its own proper suffixes. Then:

(i) No immediate subordinate of V is a proper prefix of any other immediate subordinate of V.

(ii) Every suffix of every immediate subordinate of V is a concatenation of other immediate subordinates of V.

(iii) If W and all of its cyclic shifts are less than V, then W can be uniquely decomposed into a concatenation of immediate subordinates of V, including a (possibly empty) end-around immediate subordinate. Specifically $W = W^{(1)} * W^{(2)} * \cdots * W^{(i)} * W^{(i+1)} * W^{(i+2)} * \cdots * W^{(j-1)} * W^{(j)}$; $W^{(1)}, W^{(2)}, \cdots, W^{(j-1)}$ are immediate subordinates of V; $W^{(j)} * W^{(1)} * W^{(2)} * \cdots * W^{(i)}$ is the end-around immediate subordinate. The end-around immediate subordinate has a prefix, $W^{(j)}$, which is a suffix of W, and a suffix $W^{(1)} * W^{(2)} * \cdots * W^{(i)}$ which is a prefix of W, as well as a concatenation of the shorter immediate superiors $W^{(1)}, W^{(2)}, \cdots, W^{(i)}$.

(iv) Every concatenation of immediate subordinates of V, including a (possibly empty) end-around immediate subordinate yields a sequence which has the property that all of its cyclic shifts are less than V. No such sequence of length m can exceed the maximum m-digit concatenation of immediate subordinates of V. If Y is the maximum m-digit concatenation of immediate subordinates of V, and $Y \leqq U \leqq V$, then $J(q, V, m) = J(q, U, m)$.

(v)

$$J(q, V, m) = mV_m + \sum_{k=1}^{m-1} V_k J(q, V, m - k),$$

where V_j is taken as 0 if j exceeds the length of the sequence V.

(*vi*) *Let*

$$n = q^m - 1, \qquad d = \Sigma\, D_i q^{m-i}, \qquad 0 \leqq D_i < q,$$
$$D = D_1 D_2 \cdots D_m .$$

If

$$\bar{V} * (Q \dot{-} 1) < D \leqq$$
least m-digit concatenation of immediate superiors of V,

then,

$$I(q, n, d) = J(q, V, m).$$

Proofs:

(*i*) This property of immediate subordinates does not even depend on the suffix condition on V. From the definition of immediate subordinates, each immediate subordinate must disagree with V only in the immediate subordinate's last digit, and hence no immediate subordinate can be a prefix of any other.

(*ii*) Let us first prove the weaker assertion:

(*a*) Every proper suffix of every immediate subordinate of V has a prefix which is a shorter immediate subordinate of V.

Let S be an immediate subordinate of V, and let $S^{(2)}$ be a suffix of S. We may write $S = S^{(1)} * S^{(2)}$. Since S differs from V only in its last digit, $S^{(1)}$ is a prefix of V, and $V = S^{(1)} * V^{(2)}$. Since $S < V$, $S^{(2)} < V^{(2)}$. Since V exceeds all of its own proper suffixes, $V^{(2)} < V$. Therefore, $S^{(2)} < V$. Therefore, some prefix of $S^{(2)}$ is an immediate subordinate of V.

(*b*) If every suffix of an immediate subordinate has a prefix which is an immediate subordinate, then every suffix of every immediate subordinate is a concatenation of immediate subordinates.

For, suppose F is a suffix on an immediate subordinate, then $F = B^{(1)} * F^{(2)}$, where $B^{(1)}$ is an immediate subordinate. Since $F^{(2)}$ is a suffix of F, it is also a suffix of an immediate subordinate, and $F^{(2)} = B^{(2)} * F^{(3)}$, where $B^{(2)}$ is an immediate subordinate $\cdots$ $F = B^{(1)} * B^{(2)} * B^{(3)} * \cdots$.

(*iii*) Since $W < V$, it contains a prefix $W^{(1)}$ which is an immediate subordinate of V. After shifting this prefix around to the end, we may similarly identify $W^{(2)}, W^{(3)}, \cdots, W^{(i-1)}$, each of which is an immediate subordinate of V. The sequence $W^{(i)} * W^{(1)} * W^{(2)} * \cdots * W^{(i-1)}$ is a cyclic shift of W, and so it must have a prefix, P, which is an immediate subordinate of V. P is not a prefix of $W^{(i)}$, so $W^{(i)}$ must be a prefix of P. Suppose that $W^{(i)} * W^{(1)} * \cdots * W^{(i)}$ is a prefix of P, but that $W^{(j)} * W^{(1)} * \cdots * W^{(i+1)}$ is not a prefix of P. (This defines i.) Then $P = W^{(j)} * W^{(1)} * \cdots * W^{(i)} * S$, where the (possibly empty) sequence S is a proper prefix of $W^{(i+1)}$. Since S is a suffix of P, which is an immediate subordinate of V, S is itself a concatenation of immediate subordinates of V. But no immediate subordinate of V is a proper prefix of any other immediate subordinate of V, so S must be empty.

(*iv*) This is the converse of (*iii*). Suppose we are given the sequence $W = S^{(j)} * W^{(1)} * W^{(2)} * \cdots * W^{(j-1)} * P^{(j)}$, where $W^{(1)}, W^{(2)}, \cdots, W^{(j-1)}$ and $W^{(j)} = P^{(j)} * S^{(j)}$ are immediate subordinates of V. We must show that all cyclic shifts of W are less than V. Any cyclic shift is of the form $C = S^{(k)} * W^{(k+1)} * W^{(k+2)} * \cdots * W^{(j)} * W^{(1)} * W^{(2)} * \cdots * W^{(k+1)} * P^{(k)}$, where $W^{(k)} = P^{(k)} * S^{(k)}$. If $S^{(k)}$ is empty, C has the prefix $W^{(k+1)}$, which is an immediate subordinate of V. If $S^{(k)}$ is not empty, by (*ii*) it has a prefix which is an immediate subordinate of V, which is a prefix of C. In either case, C has a prefix which is an immediate subordinate of V. Therefore, $C < V$.

(*v*) V has V_m immediate subordinates of length m, each of which has m distinct cyclic shifts. Thus, W may be chosen as a single end-around immediate subordinate of V in mV_m ways.

If W is a concatenation of several immediate subordinates of V, $W = W^{(1)} * W^{(2)} * \cdots * W^{(j-1)} * W^{(j)}$ where $W^{(1)}, W^{(2)}, \cdots, W^{(j-1)}$ are immediate subordinates of V and $W^{(j)}$ is a (possibly empty) proper prefix of the immediate subordinate $W^{(j)} * W^{(1)} * \cdots * W^{(i)}$, then the length of W is the length of $W^{(j-1)}$ plus the length of $W^{(1)} * W^{(2)} * \cdots * W^{(j-2)} * W^{(j)}$. For each k, there are V_k choices of $W^{(j-1)}$ of length k, and $J(q, V, m - k)$ choices for $W^{(1)} * W^{(2)} * \cdots * W^{(j-2)} * W^{(j)}$.

(*vi*) *Least special case:* Suppose D is the least m-digit q-ary sequence greater than $\bar{V} * (Q \dot{-} 1)$. Letting

$$d = \Sigma D_i q^{m-i}, \qquad v = \Sigma V_i q^{m-i}, \qquad \bar{v} = \Sigma \bar{V}_i q^{m-i},$$

it is evident that $d + v = n + 1$ and $v = n + 1 - d$. According to Lemma III, $I(q, n, d) = J(q, V, m)$.

(*vi*) *Greatest special case:* Let D be the least m-digit concatenation of immediate superiors of $\bar{V}$. Complementing, $\bar{D}$ is the greatest m-digit concatenation of immediate subordinates of V. In the notation of part (*iv*), $\bar{D} = Y$. Letting $\bar{d} = \Sigma \bar{D}_i q^{m-i}$, $\bar{d} = n - d$. Letting $n + 1 - d = \Sigma U_i q^{m-i}$, $U > Y$ because $n + 1 - d > n - d$. Theorem follows from part (*iv*) and Lemma III.

(*vi*) *The general case* follows because $J(q, U, m)$ is a monotonic function of U. Q.E.D.

Example I: Let V be the binary sequence 1101. We compute

		Bose distance†		Designed distance	
m	$J(q, V, m)$	*Binary*	*Decimal*	*Binary*	*Decimal*
1	1	1	1	1	1
2	3	01	1	01	1
3	4	011	3	010	2
4	11	0011	3‡	0011‡	3
5	16	00111	7	00110	6
6	30	001101	13	001100	12
7	50	0011011	27	0011000	24
8	91	00110011	51	00110000	48
9	157	001100111	103	001100000	96
10	278	0011001101	205	0011000000	192
11	485	00110011011	411	00110000000	384
12	854	001100110011	819	001100000000	768
			⋮		⋮

Here $J(q, V, m)$ is computed by Theorem 1*v*. The designed distances are computed according to Theorem 1*vi*, using $\bar{V} = 0010$, with immediate superiors 1, 01, and 0011. $\bar{V} * (Q \dot{-} 1) = 001011111111 \cdots$.

Evidently, the binary BCH code of block length $2^{12} - 1$ and designed distance 768 is identical to the binary BCH code of block length $2^{12} - 1$ and designed distance 769 or 770 or $\cdots$ or 819. This code has 854 information symbols. This code is distinct from the binary BCH code of block length $2^{12} - 1$ and designed distance 820. This is true in general, because the least m-digit concatenation of immediate superiors of $\bar{V}$ is necessarily minimum among all of its own cyclic shifts. This "greatest designed distance" is called the *Bose distance*.

It happens that the binary BCH code of block length $2^{12} - 1$ and designed distance 768 is also distinct from the binary BCH code of block length $2^{12} - 1$ and designed distance 767, because the 12-digit binary expansion of 767 is minimum among all of its cyclic shifts. This, however, need not be true in general. For example, the binary BCH code of block length $2^4 - 1$ and designed distance 3 is not distinct from the binary BCH code of block length $2^4 - 1$ and designed distance 2, because the 4-digit binary expansion of $2 = 0010$ is not minimum among its cyclic shifts; the minimum is 0001.

† Defined later in the text.

‡ This code is identical to the binary BCH code of block length 15 and designed distance 2.

In general, we would like to determine the number of information digits in the q-ary BCH code of block length $n = q^m - 1$ and designed distance $d = \Sigma D_i q^{m-i}$. The previous theorem gives us a solution to this problem if we can find a sequence V which is greater than all of its own suffixes and has the property that

$$\bar{V} * (Q \mathbin{\dot{-}} 1) < D \leqq \text{least } m\text{-digit concatenation of immediate superiors of } \bar{V}.$$

Complementing this condition gives

$$V > \bar{D} \geqq \text{greatest } m\text{-digit concatenation of immediate subordinates of } V.$$

or

$$V > \bar{D} * (Q \mathbin{\dot{-}} 1) > \begin{pmatrix} \text{greatest } m\text{-digit concatenation} \\ \text{of immediate subordinates of } V \end{pmatrix} * \dot{0} > \dot{X},$$

where X is the greatest immediate subordinate of V. We may assume that V has no terminal zeros, and that the length of V does not exceed the length of $\bar{D}$. Since X and V have the same length, X is a prefix of $\bar{D}$.

Since V is the least immediate superior of X, the problem of finding V is reduced to the problem of finding X, which is a prefix of $\bar{D}$. The solution is as follows:

Theorem 2: Let X be the shortest prefix of $\bar{D}$ such that

$$\bar{D} = X * F, \qquad F * (Q \mathbin{\dot{-}} 1) \geqq \bar{D} * (Q \mathbin{\dot{-}} 1),$$

and let V be the least immediate superior of X. Then

(i) $$\bar{V} * (Q \mathbin{\dot{-}} 1) < D \leqq \textit{least } m\textit{-digit concatenation of immediate superiors of } \bar{V}.$$

(ii) V exceeds all of its own proper suffixes.

Proof of (i):

Since X is a prefix of $\bar{D}$ and V is an immediate superior of X, V is an immediate superior of $\bar{D}$. So $V > \bar{D}$ and $V > \bar{D} * (Q \mathbin{\dot{-}} 1)$.

Complementing gives $\bar{V} * (Q \mathbin{\dot{-}} 1) < D * \dot{0}$, so $V * (Q \mathbin{\dot{-}} 1) < D$.

Let $X^{(k)} = \overset{\longleftarrow\!\!-\!\!-\!\!-\; k \;\longrightarrow}{X * X * \cdots * X}$. Then $F * (Q \mathbin{\dot{-}} 1) \geqq \bar{D} * (Q \mathbin{\dot{-}} 1)$ is equivalent to $X^{(0)} * F * (Q \mathbin{\dot{-}} 1) \geqq X^{(1)} * F * (Q \mathbin{\dot{-}} 1)$. Therefore, $X * X^{(0)} * F * (Q \mathbin{\dot{-}} 1) \geqq X * X^{(1)} * F * (Q \mathbin{\dot{-}} 1)$ or $X^{(1)} * F * (Q \mathbin{\dot{-}} 1\ ) \geqq$

$X^{(2)} * F * (Q \dot{-} 1)$. By induction, $X^{(k)} * F * (Q \dot{-} 1) \geqq X^{(k+1)} * F * (Q \dot{-} 1)$ and $\bar{D} * (Q \dot{-} 1) \geqq X^{(k)} * F * (Q \dot{-} 1)$ for all k. Since this is true for arbitrarily large k, $\bar{D} * (Q \dot{-} 1) \geqq \dot{X}$. Complementing, $D * \dot{0} \leqq \dot{\bar{X}} \leqq$ any infinite concatenation of immediate superiors of $\bar{V}$. Therefore, $D \leqq$ any m-digit concatenation of immediate superiors of $\bar{V}$.

Proof of (ii):

Let $X = Y * Z * L$, where Y and Z are arbitrary (possibly empty) and L is the final digit of X. We have

$$V = Y * Z * (L + 1)$$

$$\bar{D} = Y * Z * L * F$$

$$F * (Q \dot{-} 1) \geqq Y * Z * L * F * (Q \dot{-} 1) = X * F * (Q \dot{-} 1);$$

$X * F * (Q \dot{-} 1) = Y * Z * L * F * (Q \dot{-} 1) > Z * L * F * (Q \dot{-} 1)$, else Y would be a shorter prefix than X which satisfied the same conditions.

No proper suffix of V can equal V, for the suffix must be shorter.

If some proper prefix of V, say $Z * (L + 1)$, (Z possibly empty) exceeds V, then

$$Z * (L + 1) > Y * Z * (L + 1) > Y * Z * L = X.$$

If $Z * L > X$, then $Z * L * F * (Q \dot{-} 1) > X * F * (Q \dot{-} 1)$, a contradiction. If $Z * L$ is a prefix of X, then $X = Z * L * G$ and from

$$X * F * (Q \dot{-} 1) > Z * L * F * (Q \dot{-} 1)$$

we have

$$Z * L * G * F * (Q \dot{-} 1) > Z * L * F * (Q \dot{-} 1)$$

$$G * F * (Q \dot{-} 1) > F * (Q \dot{-} 1) \geqq X * F * (Q \dot{-} 1).$$

Now $Z * L$ is a shorter prefix than X, a contradiction. Therefore, $Z * (L + 1) < Y * Z * (L + 1)$, i.e., V exceeds all of its own proper suffixes. Q.E.D.

Example II: Let $q = 9$, $n = 728 = 9^3 - 1$, $d = 217$. Then $D = 261$, $\bar{D} = 627$, $\bar{D} * \dot{8} = 627888 \cdots$, $X = 62$, $V = 63$, $\bar{V} = 25$.

		Bose distance	
m	*J*†	*q-ary*	*decimal*
1	6	3	3
2	42	26	24
3	270	263	219

† In this and subsequent tables we use the single later J as an abbreviation for $J(q, V, m)$.

Example III: Let $q = 2$, $n = 511$, $d = 185$. Then $D = 010111001$, $\bar{D} = 101000110$, $\bar{D} * 1 = 101000110111111 \cdots$, $X = 101000$, $V = 101001$, $\bar{V} = 010110$. Immediate superiors are 010111, 011, 1.

m	*J*	*Bose distance*	*Smaller designed distance*
1	1	1	1
2	1	11	11
3	4	011	011
4	5	0111	0111
5	6	01111	01111
6	16	010111	010111
7	22	0101111	0101110
8	29	01011111	01011100
9	49	010111011	010111000

Example IV: Let $q = 2$, $n = 511 = 2^9 - 1$, $d = 187$. Then $D = 010111011$, $\bar{D} = 101000100 = X$, $V = 101000101$, $\bar{V} = 010111010$.

m	*J*	*Bose distance*
1	1	1
2	1	11
3	4	011
4	5	0111
5	6	01111
6	10	011011
7	22	0101111
8	29	01011111
9	49	010111011

The answer agrees with Example III, although the recurrence is different. This illustrates the general nonuniqueness of V. Theorem 2 specifies one satisfactory method of finding V, but as seen from this example, this V need not be unique. The simplest recurrence rule generally arises from the shortest possible V, which corresponds to the greatest V, or the least D. This can generally be found by first reducing D insofar as permissible.

Example V: Let $n = 2^{11} - 1$, $d = 411$. Then $D = 00110011011$. We could take $\bar{D} = 11001100100 = X$ and proceed. However, we instead consider $d = 410$, $D = 00110011010$. Since D has a cyclic shift smaller than itself, the code is unchanged. But $\bar{D} = 11001100101$, $X = 1100110010$ does not look much easier, so we continue. Each prime marks the starting point of a smaller cyclic shift.

$$
\begin{array}{c}
D \\
0011001101'0 \\
00110011'001 \\
00110011'000 \\
0011'0010111 \\
0011'0010110 \\
\vdots \\
0011'0010000 \\
0011'0001111 \\
\vdots \\
0011'0000000 \\
00101111111
\end{array}
$$

Since 00101111111 has no cyclic shift less than itself, this designed distance is the Bose distance of a *different* BCH code. We must instead use $D = 00110000000$, $\bar{D} = 11001111111$, $X = 1100$, $V = 1101$. The recurrence is given in Example I: $I(2,\ 2^{11} - 1{,}411) = 485$. This same V is obtained if we started with $D = 00110011000$, or any D in the region

$$00110000000 \leqq D \leqq 00110011000$$

Example VI: Let $q = 2$, $n = 2^{15} - 1$, $D = 001010010100111$, $\bar{D} * \dot{1} = 110101101011000 1111 \cdots$, $X = 110101101011000$, $V = 110101101011001$, $\bar{V} = 001010010100110$

m	J	*Bose distance*
1	1	1
2	3	01
3	4	011
4	11	0011
5	16	00111
6	36	001011
7	64	0010101
8	115	00101011
9	211	001010011
10	378	0010100111
11	694	00101001011
12	1256	001010010101
13	2276	0010100101011
14	4112	00101001010111
15	7474	001010010100111

Although the brute force method just used gives the right answer, a more devious approach proves easier. Instead of $D = 001010010100111$, let us consider $D = 001010010100101$, $X = 11010$, $V = 11011$, $\bar{V} = 00100$. This yields a different set of codes, with a much simpler recurrence:

m	*J*	*Bose distance*
1	1	1
2	3	01
3	4	011
4	11	0011
5	21	00101
6	36	001011
7	64	0010101
8	115	00101011
9	211	001010011
10	383	0010100101
11	694	00101001011
12	1256	001010010101
13	2276	0010100101011
14	4126	00101001010011
15	7479	001010010100101

The code with $D = 001010010100111$ has 5 less information digits than the code with $D = 001010010100101$, corresponding to the 5 distinct cyclic shifts of 001010010100101.

I. ASYMPTOTIC RESULTS

Let us define the enumerator

$$J(q, U; z) = \sum_{m=1}^{\infty} J(q, U, m) z^m .$$

Given a sequence V which is less than all of its own proper suffixes, we may also define

$$V(z) = \sum_k v_k z^k ,$$

so that

$$zV'(z) = \sum_k k V_k z^k .$$

The recurrence

$$J(q, V, m) = mV_m + \sum_{k=1}^{m-1} V_k J(q, V, m - k)$$

becomes

$$J(q,\ V; z) = zV'(z) + V(z)J(q,\ V; z)$$

whose solution is

$$J(q,\ V; z) = \frac{zV'(z)}{1 - V(z)}.$$

Let ρ_1, ρ_2, $\cdots$ be the (not necessarily distinct) complex reciprocal roots of $1 - V(z)$. Then

$$1 - V(z) = \prod_i (1 - \rho_i z)$$

$$-V'(z) = -\sum_i \rho_i \prod_{j \neq i} (1 - \rho_j z)$$

$$\frac{zV'(z)}{1 - V(z)} = \sum_i \frac{\rho_i z}{1 - \rho_i z} = \sum_i \sum_{m=1}^{\infty} (\rho_i z)^m = \sum_{m=1}^{\infty} \sum_i \rho^m_i z^m.$$

Therefore,

$$J(q,\ V; z) = \sum_{m=1}^{\infty} \sum_i \rho^m_i z^m$$

so

$$J(q,\ V,\ m) = \sum_i \rho_i^m,$$

where ρ_i are the complex numbers defined by the equation

$$1 - V(z) = \prod_i (1 - \rho_i z).$$

Although this gives an explicit expression for $J(q, V, m)$, the expression depends upon the complex numbers ρ_i. For finite values of m, it is usually easier to compute $J(q, V, m)$ directly from the recurrence relation of the previous section, since these calculations involve only integers. For asymptotic results, however, the above equation is very useful.

Definition: Let $\rho = \max_i |\rho_i|$, let $s = \log_q \rho$.

Since all coefficients of the polynomial $V(z)$ must be nonnegative integers not exceeding $q - 1$, it is easily seen that the ρ_i with the maximum absolute value is real and positive, and $1 \leqq \rho \leqq q$. Clearly,

$$J(q,\ V,\ m) \approx \rho^m$$

for large m, in the sense that

$$\lim_{m\to\infty} \rho^{-m} J(q, V, m) = 1.$$

Similarly,

$$\log_q J(q, V, m) \approx m \log_q \rho = ms.$$

If

$$u = \sum_{i=1}^{\infty} U_i q^{-i},$$

and

$$\dot{X} \leqq \bar{U} \leqq V * \dot{0},$$

where V exceeds all of its own suffixes and X is the maximum subordinate of V, then

$$I(q, q^m - 1, uq^m) \approx q^{ms}.$$

In other words, if we fix the fraction $d/n = u$ and let n and d grow large, then

$$I \approx n^s$$

or, more precisely,

$$s(u) = \lim_{m\to\infty} \frac{\log_q I(q, q^m - 1, uq^m)}{m}.$$

For given q, the function $s(u)$ is a rather complicated animal. To compute it, one must first write u in q-ary. If $\bar{U}$ exceeds all of its proper suffixes, set $V = \bar{U}$; otherwise write $\bar{U} = X * F$ where X is the shortest prefix such that $\bar{U} \leqq F$. V is then taken to be the least immediate superior of X. Then, s is defined as the logarithm (base q) of the maximum reciprocal root of $1 - V(z)$.

It may easily be shown that s is a continuous, monotonic nonincreasing function of u. It may also be shown that the derivative of s with respect to u is either 0 or it is undefined. There are two kinds of points at which the derivative $s'(u)$ is undefined. First, there are the endpoints of the intervals on which $s(u)$ is constant. u is a lower endpoint of such an interval iff $\bar{U}$ is a finite sequence which exceeds all of its own proper suffixes; u is an upper endpoint of such an interval iff $\bar{U}$ is a periodic sequence, equal to some of its suffixes but not less than any others. At these endpoints, $s(u)$ is undifferentiable because it has only a right derivative or a left derivative, but not both. There is only a countable number of points of this type.

The more interesting points are those at which $s(u)$ has neither a right derivative nor a left derivative. This happens iff $\bar{U}$ is an infinite sequence which exceeds all of its own proper suffixes, and $\dot{0}$ is not a suffix of $\bar{U}$.

The set of points u such that $s(u)$ is not differentiable is uncountable, but it has measure 0. Professor T. Pitcher of the University of Southern California has also shown[3] that this set has Hausdorf dimension 1. This appears to be entirely due to the large density of these points in the vicinity of $u = 0$. In general, I conjecture that the Hausdorf dimension of the set of points in the interval $a \leqq u \leqq b$ [where $0 \leqq a$, $b \leqq 1$, $s(a) \neq s(b)$] is $s(a)$. In some sense, almost all of the nondifferentiable points in any interval seem to lie very near the leftmost cluster point of the interval.

When Mann[2] first obtained results identical to those here in the special cases $u = q^{-k}$, he also showed that ρ is the only reciprocal root of $1 - V(z)$ with magnitude greater than 1. Thus, not only is

$$I \approx \rho^m,$$

but in fact, for sufficiently large m,

$$I = \langle \rho^m \rangle,$$

where $\langle \cdot \rangle$ denotes the nearest integer to "$\cdot$". Unfortunately, this strengthened result is not true in general. For some values of u, $1 - V(z)$ has only one reciprocal root with magnitude greater than 1, but for other values of u, $1 - V(z)$ has many reciprocal roots with magnitude greater than 1. Little is known about the behavior of the smaller complex reciprocal roots of $1 - V(z)$ as a function of u, although B. F. Logan[4] has obtained a few preliminary results in this area.

II. ACTUAL DISTANCE

As one increases the designed distance, the number of information symbols in the resulting code must either remain constant or decrease. Thus,

$I(q, n, d)$ is the maximum number of information symbols in any of the q-ary BCH codes with designed distance $\geqq d$.

We must be careful to distinguish between I and $\hat{I}$, defined by

$\hat{I}(q, n, d)$ is the maximum number of information symbols in any of the q-ary BCH codes with actual distance $\geqq d$.

It is obvious that $\hat{I}(q, n, d) \geqq I(q, n, d)$.

For example, there are three binary BCH codes of block length 23, having 23, 12, and 1 information symbols. The code with 23 information digits has Bose distance = actual distance = 1, but the code with 12 information digits has Bose distance 5, actual distance 7. The code with 1 information digit has Bose distance = actual distance = 23. Therefore, $I(2, 23, 6) = I(2, 23, 7) = 1$, but $\hat{I}(2, 23, 6) = \hat{I}(2, 23, 7) = 12$. For all values of $d \neq 6$ or 7, $I(2, 23, d) = \hat{I}(2, 23, d)$.

The known cases in which $\hat{I}(q, n, d) > I(q, n, d)$ are relatively sparse. Peterson, Kasami, and Lin[5] and Berlekamp[6] have investigated this question for narrow sense binary BCH codes (where $q = 2$ and $n = 2^m - 1$). They proved that $\hat{I}(2, 2^m - 1, d) = I(2, 2^m - 1, d)$ if d divides $2^m - 1$, or if d is one less than a power of 2, or if m' divides m and $\hat{I}(2, 2^{m'} - 1, d) = I(2, 2^{m'} - 1, d) > I(2, 2^{m'} - 1, d + 1)$, or if m is sufficiently small, or if d is sufficiently small, or if d and/or m satisfy any of various other number theoretical constraints. More recently, Peterson and Lin[7] have shown that if $\hat{I}(2, 2^m - 1, d) = I(2, 2^m - 1, d) > I(2, 2^m - 1, d + 1)$, and $1 \leqq j \leqq m - d$ then $\hat{I}(2, 2^m - 1, 2^j d + 2^j - 1) = I(2, 2^m - 1, 2^j d + 2^j - 1)$. No examples are known in which $\hat{I}(2, 2^m - 1, d) > I(2, 2^m - 1, d)$, and it has been conjectured that $\hat{I}(2, 2^m - 1, d) = I(2, 2^m - 1, d)$ for all m and d.

Although this conjecture remains open, we can obtain certain results about the asymptotic behavior of $I(2, 2^m - 1, u2^m)$ from the known classes of special cases in which $\hat{I}(2, 2^m - 1, d) = I(2, 2^m - 1, d)$. We would like to define

$$\hat{s}(u) \overset{?}{=} \lim_{m\to\infty} \frac{\log_2 \hat{I}(2, 2^m - 1, u2^m)}{m}.$$

Unfortunately, however, we have no assurance that the limit exists. In order to discuss the asymptotic behavior of the best BCH codes, we define

$$\hat{s}(u) = \limsup_{m\to\infty} \frac{\log_2 \hat{I}(2, 2^m - 1, u2^m)}{m}.$$

Clearly $\hat{s}(u) \geqq s(u)$. Like $s(u)$, $\hat{s}(u)$ must be a monotonic nonincreasing function of u, because if $d' > d$, the codewords of the q-ary BCH code of distance d' are a subset of the codewords of the q-ary BCH code of distance d.

We can prove that $\hat{s}(u) = s(u)$ for certain values of u, as indicated by the following theorem:

If $u \geqq 2^{-k}$, *then* $\hat{s}(u) \leqq s(2^{-k})$

Proof: We know that if $u \geqq 2^{-k}$, then

$$\begin{aligned}\hat{I}(2, 2^{m-1}, u2^m) &\leqq \hat{I}(2, 2^m - 1, 2^{m-k} - 1)\\ &= I(2, 2^m - 1, 2^{m-k} - 1) \qquad m \geqq k\end{aligned}$$

Hence,

$$\frac{\log \hat{I}(2, 2^m - 1, u2^m)}{m} \leqq \frac{\log I(2, 2^m - 1, 2^{m-k} - 1)}{m}.$$

So

$$\hat{s}(u) \leqq \lim_{m\to\infty} \frac{\log I(2, 2^m - 1, 2^{m-k} - 1)}{m} = s(2^{-k})$$

because $s(u)$ is continuous. Q.E.D.

This shows that $\hat{s}(u) = s(u)$ if $u = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots$. Similarly, one can show from the recent theorem of Peterson and Lin that $\hat{s}(u) = s(u)$ for certain other values of u.

We conjecture that $\hat{s}(u) = s(u)$ for all u. This is a weakened form of Peterson's conjecture that $\hat{I}(2, 2^m - 1, d) = I(2, 2^m - 1, d)$ for all m and d.

REFERENCES

1. Berlekamp, E. R., *Algebraic Coding Theory,* McGraw-Hill Book Company, Inc., New York, 1968.
2. Mann, H. P., On the Number of Information Symbols in Bose-Chaudhuri Codes, Inform. Control *5,* 1962, pp. 153–162.
3. Pitcher, T., unpublished correspondence, 1966.
4. Logan, B. F., unpublished oral communication, 1966.
5. Kasami, J., Lin, S., and Peterson, W. W., Some Results on Weight Distributions of BCH Codes, presented at PGIT Symposium at UCLA, January, 1966.
6. Berlekamp, E. R., Practical BCH Decoders, unpublished work, 1966.
7. Peterson, W. W. and Lin, S., Some New Results on Finite Fields and Their Applications to the Theory of BCH Codes, presented at Conference on Combinatorial Mathematics and Its Applications at the University of North Carolina, April, 1967.
8. Peterson, W. W., *Error Correcting Codes,* MIT Press, 1961.