A Class of Binary Signaling Alphabets

By DAVID SLEPIAN

(Manuscript received September 27, 1955)

A class of binary signaling alphabets called "group alphabets" is described. The alphabets are generalizations of Hamming's error correcting codes and possess the following special features: (1) all letters are treated alike in transmission; (2) the encoding is simple to instrument; (3) maximum likelihood detection is relatively simple to instrument; and (4) in certain practical cases there exist no better alphabets. A compilation is given of group alphabets of length equal to or less than 10 binary digits.

INTRODUCTION

This paper is concerned with a class of signaling alphabets, called "group alphabets," for use on the symmetric binary channel. The class in question is sufficiently broad to include the error correcting codes of Hamming,[1] the Reed-Muller codes,[2] and all "systematic codes".[3] On the other hand, because they constitute a rather small subclass of the class of all binary alphabets, group alphabets possess many important special features of practical interest.

In particular, (1) all letters of the alphabets are treated alike under transmission; (2) the encoding scheme is particularly simple to instrument; (3) the decoder — a maximum likelihood detector — is the best possible theoretically and is relatively easy to instrument; and (4) in certain cases of practical interest the alphabets are the best possible theoretically.

It has very recently been proved by Peter Elias[4] that there exist group alphabets which signal at a rate arbitarily close to the capacity, C, of the symmetric binary channel with an arbitrarily small probability of error. Elias' demonstration is an existence proof in that it does not show *explicitly* how to construct a group alphabet signaling at a rate greater than $C - \varepsilon$ with a probability of error less than δ for arbitrary positive δ and ε. Unfortunately, in this respect and in many others, our understanding of group alphabets is still fragmentary.

In Part I, group alphabets are defined along with some related concepts necessary for their understanding. The main results obtained up to the present time are stated without proof. Examples of these concepts are given and a compilation of the best group alphabets of small size is presented and explained. This section is intended for the casual reader.

In Part II, proofs of the statements of Part I are given along with such theory as is needed for these proofs.

The reader is assumed to be familiar with the paper of Hamming,[1] the basic papers of Shannon[5] and the most elementary notions of the theory of finite groups.[6]

Part I — Group Alphabets and Their Properties

1.1 INTRODUCTION

We shall be concerned in all that follows with communication over the symmetric binary channel shown on Fig. 1. The channel can accept either of the two symbols 0 or 1. A transmitted 0 is received as a 0 with probability q and is received as a 1 with probability $p = 1 - q$: a transmitted 1 is received as a 1 with probability q and is received as a 0 with probability p. We assume $0 \leqq p \leqq \frac{1}{2}$. The "noise" on the channel operates independently on each symbol presented for transmission. The capacity of this channel is

$$C = 1 + p \log_2 p + q \log_2 q \text{ bits/symbol} \tag{1}$$

By a *K-letter, n-place binary signaling alphabet* we shall mean a collection of K distinct sequences of n binary digits. An individual sequence of the collection will be referred to as a *letter* of the alphabet. The integer K is called the size of the alphabet. A letter is transmitted over the channel by presenting in order to the channel input the sequence of n zeros and ones that comprise the letter. A *detection scheme* or *detector* for

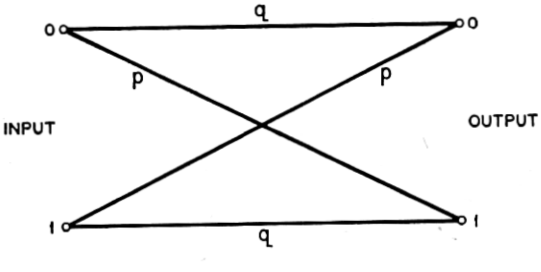


Fig. 1 — The symmetric binary channel.

a given K-letter, n-place alphabet is a procedure for producing a sequence of letters of the alphabet from the channel output.

Throughout this paper we shall assume that signaling is accomplished with a given K-letter, *n-place* alphabet by choosing the letters of the alphabet for transmission independently with equal probability $1/K$.

Shannon[5] has shown that for sufficiently large n, there exist K-letter, n-place alphabets and detection schemes that signal over the symmetric binary channel at a rate $R > C - \varepsilon$ for arbitrary $\varepsilon > 0$ and such that the probability of error in the letters of the detector output is less than any $\delta > 0$. Here C is given by (1) and is shown as a function of p in Fig. 2. No algorithm is known (other than exhaustvie procedures) for the construction of K-letter, n-place alphabets satisfying the above inequalities for arbitrary positive δ and ε except in the trivial cases $C = 0$ and $C = 1$.

1.2 THE GROUP B_n

There are a totality of 2^n different n-place binary sequences. It is frequently convenient to consider these sequences as the vertices of a cube of unit edge in a Euclidean space of n-dimensions. For example the 5-place sequence 0, 1, 0, 0, 1 is associated with the point in 5-space whose

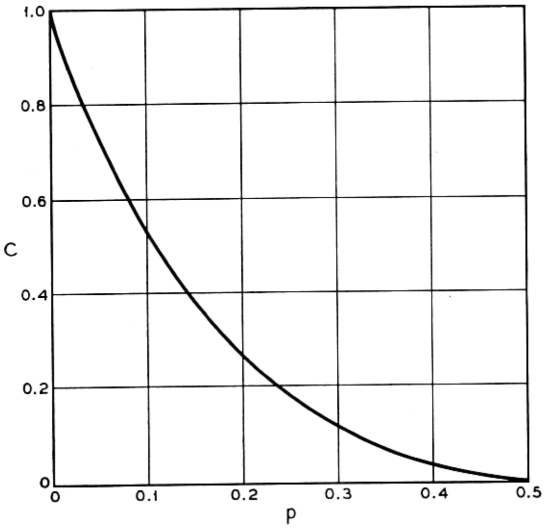


Fig. 2 — The capacity of the symmetric binary channel.
$C = 1 + p \log_2 p + (1 - p) \log_2 (1 - p)$

coordinates are (0, 1, 0, 0, 1). For convenience of notation we shall generally omit commas in writing a sequence. The above 5-place sequence will be written, for example, 01001.

We define the *product of two n-place binary sequences*, $a_1a_2 \cdots a_n$ and $b_1b_2 \cdots b_n$ as the n-place binary sequence

$$a_1 \dotplus b_1, \qquad a_2 \dotplus b_2, \cdots, a_n \dotplus b_n$$

Here the a's and b's are zero or one and the $\dotplus$ sign means addition modulo 2. (That is $0 \dotplus 0 = 1 \dotplus 1 = 0, \quad 0 \dotplus 1 = 1 \dotplus 0 = 1$) For example, (01101) (00111) = 01010. With this rule of multiplication the 2^n n-place binary sequences form an Abelian group of order 2^n. The elements of the group, denoted by $T_1, T_2, \cdots, T_{2^n}$, say, are the n-place binary sequences; the identity element I is the sequence $000 \cdots 0$ and

$$IT_i = T_iI = T_i; \qquad T_iT_j = T_jT_i; \qquad T_i(T_jT_k) = (T_iT_j)T_k;$$

the product of any number of elements is again an element; every element is its own reciprocal, $T_i = T_i^{-1}$, $T_i^2 = I$. We denote this group by B_n.

All subgroups of B_n are of order 2^k where k is an integer from the set $0, 1, 2, \cdots, n$. There are exactly

$$N(n, k) = \frac{(2^n - 2^0)(2^n - 2^1)(2^n - 2^2) \cdots (2^n - 2^{k-1})}{(2^k - 2^0)(2^k - 2^1)(2^k - 2^2) \cdots (2^k - 2^{k-1})} \tag{2}$$
$$= N(n, n - k)$$

distinct subgroups of B_n of order 2^k. Some values of $N(n, k)$ are given in Table I.

Table I — Some Values of $N(\mathrm{n}, k)$, the Number of Subgroups of B_n of Order 2^k. $N(n, k) = N(n, n - k)$

$n \backslash k$	0	1	2	3	4	5
2	1	3	1			
3	1	7	7	1		
4	1	15	35	15	1	
5	1	31	155	155	31	1
6	1	63	651	1395	651	63
7	1	127	2667	11811	11811	2667
8	1	255	10795	97155	200787	97155
9	1	511	43435	788035	3309747	3309747
10	1	1023	174251	6347715	53743987	109221651

1.3 GROUP ALPHABETS

An n-place *group alphabet* is a K-letter, n-place binary signaling alphabet whose letters form a subgroup of B_n. Of necessity the size of an n-place group alphabet is $K = 2^k$ where k is an integer satisfying $0 \leqq k \leqq n$. By an *(n, k)-alphabet* we shall mean an n-place group alphabet of size 2^k. Example: the $N(3, 2) = 7$ distinct $(3, 2)$-alphabets are given by the seven columns

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
000	000	000	000	000	000	000
100	100	100	010	010	001	110
010	001	011	001	101	110	011
110	101	111	011	111	111	101

(3)

1.4 STANDARD ARRAYS

Let the letters of a specific (n, k)-alphabet be $A_1 = I = 00 \cdots 0$, $A_2, A_3, \cdots, A_\mu$, where $\mu = 2^k$. The group B_n can be developed according to this subgroup and its cosets:

$$B_n = \begin{array}{llll} I, & A_2, & A_3, & \cdots, A_\mu \\ S_2, & S_2A_2, & S_2A_3, & \cdots, S_2A_\mu \\ S_3, & S_3A_2, & S_3A_3, & \cdots, S_3A_\mu \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ S_\nu, & S_\nu A_2, & S_\nu A_3, & \cdots, S_\nu A_\mu \\ \mu = 2^k, & \nu = 2^{n-k}. & & \end{array} \tag{4}$$

In this array every element of B_n appears once and only once. The collection of elements in any row of this array is called a *coset* of the (n, k)-alphabet. Here S_2 is any element of B_n not in the first row of the array, S_3 is any element of B_n not in the first two rows of the array, etc. The elements $S_2, S_3, \cdots, S_\nu$ appearing under I in such an array will be called the *coset leaders*.

If a coset leader is replaced by any element in the coset, the same coset will result. That is to say the two collections of elements

$$S_i, \quad S_iA_2, \quad S_iS_3, \cdots, S_iA_\mu$$

and

$$S_iA_k, \quad (S_iA_k)A_2, \quad (S_iA_k)A_3, \cdots (S_iA_k)A_\mu$$

are the same.

We define the *weight* $w_i = w(T_i)$ of an element, T_i, of B_n to be the number of ones in the n-place binary sequence T_i.

Henceforth, unless otherwise stated, we agree in dealing with an array such as (4) to adopt the following convention:

the leader of each coset shall be taken to be an element of minimal weight in that coset. (5)

Such a table will be called a *standard array*.

Example: B_4 can be developed according to the (4, 2)-alphabet 0000, 1100, 0011, 1111 as follows

$$\begin{array}{cccc} 0000 & 1100 & 0011 & 1111 \\ 1010 & 0110 & 1001 & 0101 \\ 1110 & 0010 & 1101 & 0001 \\ 1000 & 0100 & 1011 & 0111 \end{array} \tag{6}$$

According to (5), however, we should write, for example

$$\begin{array}{cccc} 0000 & 1100 & 0011 & 1111 \\ 1010 & 0110 & 1001 & 0101 \\ 0010 & 1110 & 0001 & 1101 \\ 1000 & 0100 & 1011 & 0111 \end{array} \tag{7}$$

The coset leader of the second coset of (6) can be taken as any element of that row since all are of weight 2. The leader of the third coset, however, should be either 0010 or 0001 since these are of weight one. The leader of the fourth coset should be either 1000 or 0100.

1.5 THE DETECTION SCHEME

Consider now communicating with an (n, k)-alphabet over the symmetric binary channel. When any letter, say A_j, of the alphabet is transmitted, the received sequence can be of any element of B_n. We agree to use the following detector:

if the received element of B_n lies in column i of the array (4), the detector prints the letter A_i, $i = 1, 2, \cdots, \mu$. The array (4) is to be constructed according to the convention (5). (8)

The following propositions and theorems can be proved concerning signaling with an (n, k)-alphabet and the detection scheme given by (8).

1.6 BEST DETECTOR AND SYMMETRIC SIGNALING

Define the *probability* $\ell_i = \ell(T_i)$ of an element T_i of B_n to be $\ell_i = p^{w_i}q^{n-w_i}$ where p and q are as in (1) and w_i is the weight of T_i. Let

Q_i, $i = 1, 2, \cdots, \mu$ be the sum of the probabilities of the elements in the ith column of the standard array (4).

Proposition 1. The probability that any transmitted letter of the (n, k)-alphabet be produced correctly by the detector is Q_1.

Proposition 2. The equivocation[5] per symbol is

$$H_y(x) = -\frac{1}{n} \sum_{i=1}^{\mu} Q_i \log_2 Q_i$$

Theorem 1. The detector (8) is a maximum likelihood detector. That is, for the given alphabet no other detection scheme has a greater average probability that a transmitted letter be produced correctly by the detector.

Let us return to the geometrical picture of n-place binary sequences as vertices of a unit cube in n-space. The choice of a K-letter, n-place alphabet corresponds to designating K particular vertices as letters. Since the binary sequence corresponding to any vertex can be produced by the channel output, any detector must consist of a set of rules that associates various vertices of the cube with the vertices designated as letters of the alphabet. We assume that every vertex is associated with some letter. The vertices of the cube are divided then into disjoint sets, $W_1, W_2, \cdots, W_K$ where W_i is the set of vertices associated with ith letter of the signaling alphabet. A maximum likelihood detector is characterized by the fact that every vertex in W_i is as close to or closer to the ith letter than to any other letter, $i = 1, 2, \cdots, K$. For group alphabets and the detector (8), this means that no element in the ith column of array (4) is closer to any other A than it is to A_i, $i = 1, 2, \cdots, \mu$.

Theorem 2. Associated with each (n, k)-alphabet considered as a point configuration in Euclidean n-space, there is a group of $n \times n$ orthogonal matrices which is transitive on the letters of the alphabet and which leaves the unit cube invariant. The maximum likelihood sets W_1, W_2, $\cdots$ W_μ are all geometrically similar.

Stated in loose terms, this theorem asserts that in an (n, k)-alphabet every letter is treated the same. Every two letters have the same number of nearest neighbors associated with them, the same number of next nearest neighbors, etc. The disposition of points in any two W regions is the same.

1.7 GROUP ALPHABETS AND PARITY CHECKS

Theorem 3. Every group alphabet is a systematic[3] code: every systematic code is a group alphabet.[7]

We prefer to use the word "alphabet" in place of "code" since the latter has many meanings. In a *systematic alphabet*, the places in any letter can be divided into two classes: the information places — k in number for an (n, k)-alphabet — and the check positions. All letters have the same information places and the same check places. If there are k information places, these may be occupied by any of the 2^k k-place binary sequences. The entries in the $n - k$ check positions are fixed linear (mod 2) combinations of the entries in the information positions. The rules by which the entries in the check places are determined are called *parity checks*. Examples: for the (4, 2)-alphabet of (6), namely 0000, 1100, 0011, 1111, positions 2 and 3 can be regarded as the information positions. If a letter of the alphabet is the sequence $a_1a_2a_3a_4$, then $a_1 = a_2$, $a_4 = a_3$ are the parity checks determining the check places 1 and 4. For the (5, 3)-alphabet 00000, 10001, 01011, 00111, 11010, 10110, 01100, 11101 places 1, 2, and 3 (numbered from the left) can be taken as the information places. If a general letter of the alphabet is $a_1a_2a_3a_4a_5$, then $a_4 = a_2 \dotplus a_3$, $a_5 = a_1 \dotplus a_2 \dotplus a_3$.

Two group alphabets are called *equivalent* if one can be obtained from the other by a permutation of places. Example: the 7 distinct (3, 2)-alphabets given in (3) separate into three equivalence classes. Alphabets (i), (ii), and (iv) are equivalent; alphabets (iii), (v), (vi), are equivalent; (vii) is in a class by itself.

Proposition 3. Equivalent (n, k)-alphabets have the same probability Q_1 of correct transmission for each letter.

Proposition 4. Every (n, k)-alphabet is equivalent to an (n, k)-alphabet whose first k places are information places and whose last $n - k$ places are determined by parity checks over the first k places.

Henceforth we shall be concerned only with (n, k)-alphabets whose first k places are information places. The parity check rules can then be written

$$a_i = \sum_{j=1}^{k} \gamma_{ij} a_j, \qquad i = k + 1, \cdots, n \tag{9}$$

where the sums are of course mod 2. Here, as before, a typical letter of the alphabet is the sequence $a_1a_2 \cdots a_n$. The γ_{ij} are $k(n - k)$ quantities, zero or one, that serve to define the particular (n, k)-alphabet in question.

1.8 MAXIMUM LIKELIHOOD DETECTION BY PARITY CHECKS

For any element, T, of B_n we can form the sum given on the right of (9). This sum may or may not agree with the symbol in the ith place of

T. If it does, we say T satisfies the ith-place parity check; otherwise T fails the ith-place parity check. When a set of parity check rules (9) is given, we can associate an $(n - k)$-place binary sequence, $R(T)$, with each element T of B_n. We examine each check place of T in order starting with the $(k + 1)$-*st* place of T. We write a zero if a place of T satisfies the parity check; we write a one if a place fails the parity check. The resultant sequence of zeros and ones, written from left to right is $R(T)$. We call $R(T)$ the *parity check sequence* of T. Example: with the parity rules $a_4 = a_2 \dotplus a_3$, $a_5 = a_1 \dotplus a_2 \dotplus a_3$ used to define the (5, 3)-alphabet in the examples of Theorem 3, we find $R(11000) = 10$ since the sum of the entries in the second and third places of 11001 is not the entry of the fourth place and since the sum of $a_1 = 1$, $a_2 = 1$, and $a_3 = 0$ is $0 = a_5$.

Theorem 4. Let $I, A_2, \cdots A_\mu$ be an (n, k)-alphabet. Let $R(T)$ be the parity check sequence of an element T of B_n formed in accordance with the parity check rules of the (n, k)-alphabet. Then $R(T_1) = R(T_2)$ if and only if T_1 and T_2 lie in the same row of array (4). The coset leaders can be ordered so that $R(S_i)$ is the binary symbol for the integer $i - 1$.

As an example of Theorem 4 consider the (4, 2)-alphabet shown with its cosets below

0000	1011	0101	1110
0100	1111	0001	1010
0010	1001	0111	1100
1000	0011	1101	0110

The parity check rules for this alphabet are $a_3 = a_1$, $a_4 = a_1 \dotplus a_2$. Every element of the second row of this array satisfies the parity check in the third place and fails the parity check in the 4th place. The parity check sequence for the second row is 01. The parity check for the third row is 10, and for the fourth row 11. Since every letter of the alphabet satisfies the parity checks, the parity check sequence for the first row is 00. We therefore make the following association between parity check sequences and coset leaders

$$\begin{array}{l}
00 \rightarrow 0000 = S_1 \\
01 \rightarrow 0100 = S_2 \\
10 \rightarrow 0010 = S_3 \\
11 \rightarrow 1000 = S_4
\end{array}$$

1.9 INSTRUMENTING A GROUP ALPHABET

Proposition 4 attests to the ease of the encoding operation involved with the use of an (n, k)-alphabet. If the original message is presented as a long sequence of zeros and ones, the sequence is broken into blocks of length k places. Each block is used as the first k places of a letter of the signaling alphabet. The last n-k places of the letter are determined by fixed parity checks over the first k places.

Theorem 4 demonstrates the relative ease of instrumenting the maximum likelihood detector (8) for use with an (n, k)-alphabet. When an element T of B_n is received at the channel output, it is subjected to the n-k parity checks of the alphabet being used. This results in a parity check sequence $R(T)$. $R(T)$ serves to identify a unique coset leader, say S_i. The product S_iT is then formed and produced as the detector output. The probability that this be the correct letter of the alphabet is Q_1.

1.10 BEST GROUP ALPHABETS

Two important questions regarding (n, k)-alphabets naturally arise. What is the maximum value of Q_1 possible for a given n and k and which of the $N(n, k)$ different subgroups give rise to this maximum Q_1? The answers to these questions for general n and k are not known. For many special values of n and k the answers are known. They are presented in Tables II, III and IV, which are explained below.

The probability Q_1 that a transmitted letter be produced correctly by the detector is the sum, $Q_1 = \sum_1^{\nu} \ell(S_i)$ of the probabilities of the coset leaders. This sum can be rewritten as $Q_1 = \sum_{i=0}^{n} \alpha_i \, p^i q^{n-i}$ where α_i is the number of coset leaders of weight i. One has, of course, $\sum \alpha_i = \nu = 2^{n-k}$ for an (n, k)-alphabet. Also $\alpha_i \leqq \binom{n}{i} = \frac{n!}{i!(n-i)}$! since this is the number of elements of B_n of weight i.

The α_i have a special physical significance. Due to the noise on the channel, a transmitted letter, A_i, of an (n, k)-alphabet will in general be received at the channel output as some element T of B_n different from A_i. If T differs from A_i in s places, i.e., if $w(A_iT) = s$, we say that an s-tuple error has occurred. For a given (n, k)-alphabet, α_i is the number of i-tuple errors which can be corrected by the alphabet in question, $i = 0, 1, 2, \cdots, n$.

Table II gives the α_i corresponding to the largest possible value of Q_1 for a given k and n for $k = 2, 3, \cdots n - 1$, $n = 4 \cdots, 10$ along with a few other scattered values of n and k. For reference the binomial coefficients $\binom{n}{i}$ are also listed. For example, we find from Table II that the best group alphabet with $2^4 = 16$ letters that uses $n = 10$ places has a probability of correct transmission $Q_1 = q^{10} + 10q^9p + 39q^8p^2 + 14q^7p^3$. The alphabet corrects all 10 possible single errors. It corrects 39 of the possible $\binom{10}{2} = 45$ double errors (second column of Table II) and in addition corrects 14 of the 120 possible triple errors. By adding an additional place to the alphabet one obtains with the best (11, 4)-alphabet an alphabet with 16 letters that corrects all 11 possible single errors and all 55 possible double errors as well as 61 triple errors. Such an alphabet might be useful in a computer representing decimal numbers in binary form.

For each set of α's listed in Table II, there is in Table III a set of parity check rules which determines an (n, k)-alphabet having the given α's. The notation used in Table III is best explained by an example. A (10, 4)-alphabet which realizes the α's discussed in the preceding paragraph can be obtained as follows. Places 1, 2, 3, 4 carry the information. Place 5 is determined to make the mod 2 sum of the entries in places 3, 4, and 5 equal to zero. Place 6 is determined by a similar parity check on places 1, 2, 3, and 6; place 7 by a check on places 1, 2, 4, and 7, etc.

It is a surprising fact that for all cases investigated thus far an (n, k)-alphabet best for a given value of p is uniformly best for all values of p, $0 \leqq p \leqq \frac{1}{2}$. It is of course conjectured that this is true for all n and k.

It is a further (perhaps) surprising fact that the best (n, k)-alphabets are not necessarily those with greatest nearest neighbor distance between letters when the alphabets are regarded as point configurations on the n-cube. For example, in the best (7, 3)-alphabet as listed in Table III, each letter has two nearest neighbors distant 3 edges away. On the other hand, in the (7, 3)-alphabet given by the parity check rules 413, 512, 623, 7123 each letter has its nearest neighbors 4 edges away. This latter alphabet does not have as large a value of Q_1, however, as does the (7, 3)-alphabet listed on Table III.

The cases $k = 0, 1, n - 1, n$ have not been listed in Tables II and III. The cases $k = 0$ and $k = n$ are completely trivial. For $k = 1$, all $n > 1$ the best alphabet is obtained using the parity rule $a_2 = a_3 = \cdots = a_n = a_1$. If $n = 2j$,

$$Q_1 = \sum_0^{j-1} \binom{n}{i} p^i q^{n-i} + \frac{1}{2} \binom{n}{j} p^j q^j. \text{ If } n = 2j + 1,\ Q_1 = \sum_0^{j} \binom{n}{i} p^i q^{n-i}.$$

For $k = n - 1$, $n > 1$, the maximum Q_1 is $Q_1 = q^{n-1}$ and a parity rule for an alphabet realizing this Q_1 is $a_n = a_1$.

If the α's of an (n, k)-alphabet are of the form $\alpha_i = \binom{n}{i}$, $i = 0, 1,$

TABLE II — PROBABILITY OF NO ERROR WITH BEST ALPHABETS, $Q_1 = \sum \alpha_i p^i q^{n-i}$

	i	$\binom{n}{i}$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
			a_i	a_i	a_i	a_i	a_i	a_i	a_i	a_i	a_i
$n = 4$	0	1	1								
	1	4	3								
$n = 5$	0	1	1	1							
	1	5	5	3							
	2	10	2								
$n = 6$	0	1	1	1	1						
	1	6	6	6	3						
	2	15	9	1							
$n = 7$	0	1	1	1	1	1					
	1	7	7	7	7	3					
	2	21	18	8							
	3	25	6								
$n = 8$	0	1	1	1	1	1	1				
	1	8	8	8	8	7	3				
	2	28	28	20	7						
	3	56	27	3							
$n = 9$	0	1	1	1	1	1	1	1			
	1	9	9	9	9	9	7	3			
	2	36	36	33	22	6					
	3	84	64	21							
	4	126	18								
$n = 10$	0	1	1	1	1	1	1	1	1		
	1	10	10	10	10	10	10	7	3		
	2	45	45	45	39	21	5				
	3	120	110	64	14						
	4	210	90	8							
$n = 11$	0	1	1	1	1		1	1	1	1	
	1	11	11	11	11		11	11	7	3	
	2	55	55	55	55		20	4			
	3	165	165	126	61						
	4	330	226	63							
	5	462	54								
$n = 12$	0	1	1	1				1	1	1	1
	1	12	12	12				12	12	7	3
	2	66	66	66				19	3		
	3	220	220	200							
	4	495	425	233							
	5	792	300								

$2, \cdots, j$, $\alpha_{j+1} = r$ some integer, $\alpha_{j+2} = \alpha_{j+3} = \cdots = \alpha_n = 0$, then there does not exist a 2^k-letter, n-place alphabet of any sort better than the given (n, k)-alphabet. It will be observed that many of the α's of Table II are of this form. It can be shown that

Proposition 5 if $n + \binom{n-k}{2} + \binom{n-k}{3} \geqq 2^{n-k} - 1$ there exists no 2^k-letter, n-place alphabet better than the best (n, k)-alphabet.

When the inequality of proposition 5 holds the $\boldsymbol{\alpha}$'s are either $\alpha_0 = 1$, $\alpha_1 = 2^{n-k} - 1$, all other $\alpha = 0$; or $\alpha_0 = 1$, $\alpha_1 = \binom{n}{1}$, $\alpha_2 = 2^{n-k} - 1 - \binom{n}{1}$ all other $\alpha = 0$; or the trivial $\alpha_0 = 1$ all other $\alpha = 0$ which holds when $k = n$. The region of the $n - k$ plane for which it is known that (n, k)-alphabets cannot be excelled by any other is shown in Table IV.

1.11 A DETAILED EXAMPLE

As an example of the use of (n, k)-alphabets consider the not unrealistic case of a channel with $p = 0.001$, i.e., on the average one binary digit per thousand is received incorrectly. Suppose we wish to transmit messages using 32 different letters. If we encode the letters into the 32 5-place binary sequences and transmit these sequences without further encoding, the probability that a received letter be in error is $1 - (1 - p)^5 = 0.00449$. If the best (10, 5)-alphabet as shown in Tables II and III is used, the probability that a letter be wrong is $1 - Q_1 = 1 - q^{10} - 10q^9p - 21q^8p^2 = 24p^2 - 72p^3 + \cdots = 0.000024$. Thus by reducing the signaling rate by ½, a more than *one hundredfold* reduction in probability of error is accomplished.

A (10, 5)-alphabet to achieve these results is given in Table III. Let a typical letter of the alphabet be the 10-place sequence of binary digits $a_1 a_2 \cdots a_9 a_{10}$. The symbols $a_1 a_2 a_3 a_4 a_5$ carry the information and can be any of 32 different arrangements of zeros and ones. The remaining places are determined by

$$\begin{aligned} a_6 &= a_1 \dotplus a_3 \dotplus a_4 \dotplus a_5 \\ a_7 &= a_1 \dotplus a_2 \dotplus a_4 \dotplus a_5 \\ a_8 &= a_1 \dotplus a_2 \dotplus a_3 \dotplus a_5 \\ a_9 &= a_1 \dotplus a_2 \dotplus a_3 \dotplus a_4 \\ a_{10} &= a_1 \dotplus a_2 \dotplus a_3 \dotplus a_4 \dotplus a_5 \end{aligned}$$

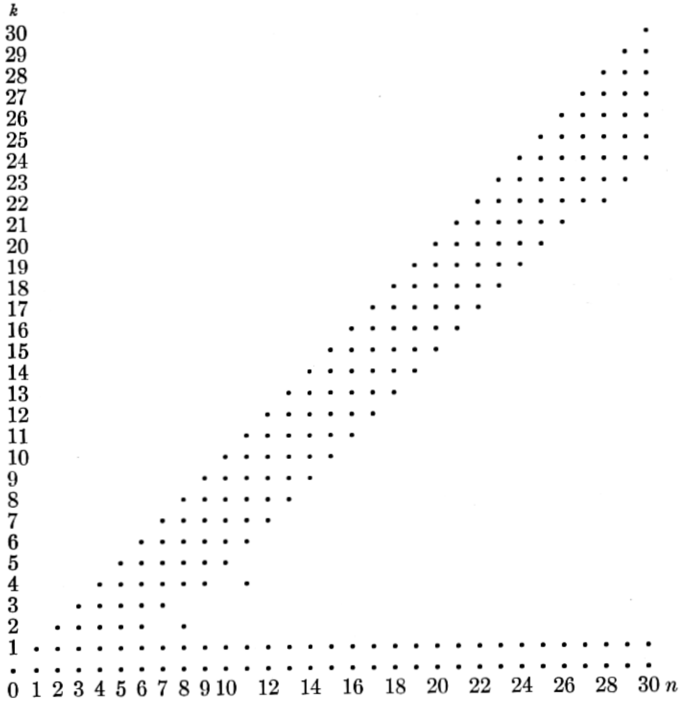
To design the detector for this alphabet, it is first necessary to determine the coset leaders for a standard array (4) formed for this alphabet.

Table III — Parity Check Rules for Best Alphabets

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
$n = 4$	3 2 4 1 2								
$n = 5$	3 1 2 4 2 5 1	4 1 2 5 1 3							
$n = 6$	3 2 4 1 2 5 1 6 1	4 1 2 5 1 3 6 2 3	5 1 2 3 6 1 2 4						
$n = 7$	3 1 4 1 5 1 6 1 2 7 2	4 1 3 5 1 2 6 1 2 3 7 1 2 3	5 1 3 4 6 1 2 4 7 1 2 3	6 1 7 1					
$n = 8$	3 1 4 1 5 2 6 2 7 1 2 8 1 2	4 1 5 1 2 6 1 3 7 2 3 8 1 2 3	5 1 3 4 6 1 2 4 7 1 2 3 8 1 2 3 4	6 1 3 4 7 1 2 4 8 1 2 3	7 1 8 1				
$n = 9$	3 1 4 1 5 1 6 2 7 2 8 1 2 9 1 2	4 1 5 2 6 1 2 7 1 3 8 2 3 9 1 2 3	5 1 3 4 6 1 2 4 7 1 2 3 8 1 2 3 9 1 2 3	6 1 3 4 5 7 1 2 4 5 8 1 2 3 5 9 1 2 3 4	7 1 3 4 8 1 2 4 9 1 2 3	8 1 9 1			

$n = 10$	3 1 4 1 5 1 6 2 7 2 8 1 2 9 1 2 10 1 2	4 1 5 2 6 3 7 1 2 8 1 3 9 2 3 10 1 2 3	5 3 4 6 1 2 3 7 1 2 4 8 1 3 4 9 2 3 4 10 1 2 3 4	6 1 3 4 5 7 1 2 4 5 8 1 2 3 5 9 1 2 3 4 10 1 2 3 4 5	7 1 3 4 5 8 1 2 4 5 9 1 2 3 5 6 10 1 2 3 4 6	8 1 3 4 9 1 2 4 10 1 2 3	9 1 10 1		
$n = 11$	3 1 4 1 5 1 6 2 7 2 8 2 9 1 2 10 1 2 11 1 2	4 3 5 3 6 2 7 1 3 8 1 3 9 1 2 10 1 2 3 11 1 2 3	5 1 3 6 2 4 7 1 4 8 2 3 9 1 3 4 10 2 3 4 11 1 2 3 4		7 1 3 4 5 6 8 1 2 4 5 6 9 1 2 3 5 6 10 1 2 3 4 6 11 1 2 3 4 5	8 1 3 4 5 9 1 2 4 5 7 10 1 2 3 5 6 11 1 2 3 4 6 7	9 1 3 4 10 1 2 4 11 1 2 3	10 1 11 1	
$n = 12$	3 1 4 1 5 1 6 1 7 2 8 2 9 2 10 2 11 1 2 12 1 2	4 1 5 2 6 3 7 1 2 8 1 2 9 1 3 10 2 3 11 1 2 3 12 1 2 3				8 1 3 4 5 6 9 1 2 4 5 6 10 1 2 3 5 6 7 11 1 2 3 4 6 7 12 1 2 3 4 5 7	9 1 2 3 5 6 7 8 10 1 2 3 4 6 11 1 2 4 5 7 12 1 3 4 5 8	10 1 2 3 11 1 2 4 12 1 3 4	11 1 12 1

TABLE IV — REGION OF THE n-k PLANE FOR WHICH IT IS KNOWN THAT (n, k)-ALPHABETS CANNOT BE EXCELLED



This can be done by a variety of special methods which considerably reduce the obvious labor of making such an array. A set of best S's along with their parity check symbols is given in Table V.

A maximum likelihood detector for the (10, 5)-alphabet in question forms from each received sequence $b_1 b_2 \cdots b_{10}$ the parity check symbol $c_1 c_2 c_3 c_4 c_5$ where

$$\begin{aligned} c_1 &= b_6 \dotplus b_1 \dotplus b_3 \dotplus b_4 \dotplus b_5 \\ c_2 &= b_7 \dotplus b_1 \dotplus b_2 \dotplus b_4 \dotplus b_5 \\ c_3 &= b_8 \dotplus b_1 \dotplus b_2 \dotplus b_3 \dotplus b_5 \\ c_4 &= b_9 \dotplus b_1 \dotplus b_2 \dotplus b_3 \dotplus b_4 \\ c_5 &= b_{10} \dotplus b_1 \dotplus b_2 \dotplus b_3 \dotplus b_4 \dotplus b_5 \end{aligned}$$

According to Table V, if $c_1 c_2 c_3 c_4 c_5$ contains less than three ones, the detector should brint $b_1 b_2 b_3 b_4 b_5$. The detector should print $(b_1 \dotplus 1) b_2 b_3 b_4 b_5$ if the parity check sequence $c_1 c_2 c_3 c_4 c_5$ is either 11111 or 11110; the de-

TABLE V — COSET LEADERS AND PARITY CHECK SEQUENCES FOR (10,5)-ALPHABET

$c_1c_2c_3c_4c_5$	↔	S	$c_1c_2c_3c_4c_5$	↔	S
00000		0000000000	11100		0000100001
10000		0000010000	11010		0001000001
01000		0000001000	11001		0001000010
00100		0000000100	10110		0010000001
00010		0000000010	10101		0010000010
00001		0000000001	10011		0010000100
11000		0000011000	01110		0100000001
10100		0000010100	01101		0100000010
10010		0000010010	01011		0100000100
10001		0000010001	00111		0100001000
01100		0000001100	11110		1000000001
01010		0000001010	11101		0000100000
01001		0000001001	11011		0001000000
00110		0000000110	10111		0010000000
00101		0000000101	01111		0100000000
00011		0000000011	11111		1000000000

tector should print $b_1(b_2 \dot{+} 1)b_3b_4b_5$ if the parity check sequence is 01111, 00111, 01011, 01101, or 01110; the detector should print $b_1b_2(b_3 \dot{+} 1)b_4b_5$ if the parity check sequence is 10111, 10011, 10101, or 10110; the detector should print $b_1b_2b_3(b_4 \dot{+} 1)b_5$ if the parity check sequence is 11011, 11001, 11010; and finally the detector should print $b_1b_2b_3b_4(b_5 \dot{+} 1)$ if the parity check sequence is 11101 or 11100.

Simpler rules of operation for the detector may possibly be obtained by choice of a different set of S's in Table V. These quantities in general are not unique. Also there may exist non-equivalent alphabets with simpler detector rules that achieve the same probability of error as the alphabet in question.

PART II — ADDITIONAL THEORY AND PROOFS OF THEOREMS OF PART I

2.1 THE ABSTRACT GROUP C_n

It will be helpful here to say a few more words about B_n, the group of n-place binary sequences under the operation of addition mod 2. This group is simply isomorphic with the abstract group C_n generated by n commuting elements of order two, say a_1, a_2, $\cdots$, a_n. Here $a_ia_j = a_ja_i$ and $a_i^2 = I$, $i, j = 1, 2, \cdots, n$, where I is the identity for the group. The eight distinct elements of C_3 are, for example, I, a_1, a_2, a_3, a_1a_2, a_1a_3, a_2a_3, $a_1a_2a_3$. The group C_n is easily seen to be isomorphic with the n-fold direct product of the group C_1 with itself.

It is a considerable saving in notation in dealing with C_n to omit the symbol "a" and write only the subscripts. In this notation for example, the elements of C_4 are I, 1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234. The product of two or more elements of C_n can readily be written down. Its symbol consists of those numerals that occur an odd number of times in the collection of numerals that comprise the symbols of the factors. Thus, $(12)(234)(123) = 24$.

The isomorphism between C_n and B_n can be established in many ways. The most convenient way, perhaps, is to associate with the element $i_1 i_2 i_3 \cdots i_k$ of C_n the element of B_n that has ones in places $i_1, i_2, \cdots, i_k$ and zeros in the remaining $n - k$ places. For example, one can associate 124 of C_4 with 1101 of B_4; 14 with 1001, etc. In fact, the numeral notation afforded by this isomorphism is a much neater notation for B_n than is afforded by the awkward strings of zeros and ones. There are, of course, other ways in which elements of C_n can be paired with elements of B_n so that group multiplication is preserved. The collection of all such "pairings" makes up the group of automorphisms of C_n. This group of automorphisms of C_n is isomorphic with the group of non-singular linear homogenous transformations in a field of characteristic 2.

An element T of C_n is said to be *dependent* upon the set of elements $T_1, T_2, \cdots, T_j$ of C_n if T can be expressed as a product of some elements of the set $T_1, T_2, \cdots, T_j$; otherwise, T is said to be *independent* of the set. A set of elements is said to be independent if no member can be expressed solely in terms of the other members of the set. For example, in C_8, 1, 2, 3, 4 form a set of independent elements as do likewise 2357, 12357, 14. However, 135 depends upon 145, 3457, 57 since $135 = (145)(3457)(57)$. Clearly any set of n independent elements of C_n can be taken as generators for the group. For example, all possible products formed of 12, 123, and 23 yield the elements of C_3.

Any k independent elements of C_n serve as generators for a subgroup of order 2^k. The subgroup so generated is clearly isomorphic with C_k. All subgroups of C_n of order 2^k can be obtained in this way.

The number of ways in which k independent elements can be chosen from the 2^n elements of C_n is

$$F(n, k) = (2^n - 2^0)(2^n - 2^1)(2^n - 2^2) \cdots (2^n - 2^{k-1})$$

For, the first element can be chosen in $2^n - 1$ ways (the identity cannot be included in a non-trivial set of independent elements) and the second element can be chosen in $2^n - 2$ ways. These two elements determine a subgroup of order 2^2. The third element can be chosen as any element of the remaining $2^n - 2^2$ elements. The 3 elements chosen determine a subgroup of order 2^3. A fourth independent element can be chosen as any of the remaining $2^n - 2^3$ elements, etc.

Each set of k independent elements serves to generate a subgroup of order 2^k. The quantity $F(n, k)$ is not, however, the number of distinct subgroups of C_n of this order, for, a given subgroup can be obtained from many different sets of generators. Indeed, the number of different sets of generators that can generate a given subgroup of order 2^k of C_n is just $F(k, k)$ since any such subgroup is isomorphic with C_k. Therefore the number of subgroups of C_n of order 2^k is $N(n, k) = F(n, k)/F(k, k)$ which is (2). A simple calculation gives $N(n, k) = N(n, n - k)$.

2.2 PROOF OF PROPOSITIONS 1 AND 2

After an element A of B_n has been presented for transmission over a noisy binary channel, an element T of B_n is produced at the channel output. The element $U = AT$ of B_n serves as a record of the noise during the transmission. U is an n-place binary sequence with a one at each place altered in A by the noise. The channel output, T, is obtained from the input A by multiplication by U: $T = UA$. For channels of the sort under consideration here, the probability that U be any particular element of B_n of weight w is $p^w q^{n-w}$.

Consider now signaling with a particular (n, k)-alphabet and consider the standard array (4) of the alphabet. If the detection scheme (8) is used, a transmitted letter A_i will be produced without error if and only if the received symbol is of the form $S_j A_i$. That is, there will be no error only if the noise in the channel during the transmission of A_i is represented by one of the coset leaders. (This applies for $i = 1, 2, \cdots, \mu = 2^k$). The probability of this event is Q_1 (Proposition 1, Section 1.6). The convention (5) makes Q_1 as large as is possible for the given alphabet.

Let X refer to transmitted letters and let Y refer to letters produced by the detector. We use a vertical bar to denote conditions when writing probabilities. The quantity to the right of the bar is the condition. We suppose the letters of the alphabet to be chosen independently with equal probability 2^{-k}.

The equivocation $h(X \mid Y)$ obtained when using an (n, k)-alphabet with the detector (8) can most easily be computed from the formula

$$h(X \mid Y) = h(X) - h(Y) + h(Y \mid X) \qquad (10)$$

The entropy of the source is $h(X) = k/n$ bits per symbol. The probability that the detector produce A_j when A_i was sent is the probability that the noise be represented by $A_i A_j S_\ell$, $\ell = 1, 2, \cdots, \nu$. In symbols,

$$Pr(Y \to A_j \mid X \to A_i) = \sum_{\ell} Pr(N \to A_iA_jS_\ell) = Q(A_iA_j)$$

where $Q(A_i)$ is the sum of the probabilities of the elements that are in the same column as A_i in the standard array. Therefore

$$Pr(Y \to A_j) = \sum_i Pr(Y \to A_j \mid X \to A_i)Pr(X \to A_i) = \frac{1}{2^k}\sum_i Q(A_iA_j)$$

$$= \frac{1}{2^k}, \qquad \text{since } \sum_i Q(A_iA_j) = \sum_i Q(A_i) = 1.$$

This last follows from the group property of the alphabet. Therefore

$$h(Y) = -\frac{1}{n}\sum Pr(Y \to A_j) \log Pr(Y \to A_j) = \frac{k}{n} \text{ bits/symbol.}$$

It follows then from (10) that

$$h(X \mid Y) = h(Y \mid X)$$

The computation of $h(Y \mid X)$ follows readily from its definition

$$\begin{aligned} h(Y \mid X) &= \sum_i Pr(X \to A_i)h(Y \mid X \to A_i) \\ &= -\sum_{ij} Pr(X \to A_i)Pr(Y \to A_j \mid X \to A_i) \\ &\qquad\qquad \log Pr(Y \to A_j \mid X \to A_i) \\ &= -\frac{1}{2^k}\sum_{ij}\sum_{\ell} Pr(N \to A_iS_\ell A_j) \log \sum_m Pr(N \to A_iS_mA_j) \\ &= -\frac{1}{2^k}\sum_{ij} Q(A_iA_j) \log Q(A_iA_j) \\ &= -\sum_i Q(A_i) \log Q(A_i) \end{aligned}$$

Each letter is n binary places. Proposition 2, then follows.

2.3 DISTANCE AND THE PROOF OF THEOREM 1

Let A and B be two elements of B_n. We define the *distance*, $d(A, B)$, between A and B to be the weight of their product,

$$d(A, B) = w(AB) \tag{11}$$

The distance between A and B is the number of places in which A and B differ and is just the "Hamming distance."[1] In terms of the n-cube, $d(A, B)$ is the minimum number of edges that must be traversed to go from vertex A to vertex B. The distance so defined is a monotone function of the Euclidean distance between vertices.

It follows from (11) that if C is any element of B_n then

$$d(A,\, B) = d(AC,\, BC) \tag{12}$$

This fact shows the detection scheme (8) to be a maximum likelihood detector. By definition of a standard array, one has

$$d(S_i\,,\, I) \leqq d(S_iA_j\,,\, I) \qquad \text{for all } i \text{ and } j$$

The coset leaders were chosen to make this true. From (12),

$$\begin{aligned} d(S_i\,,\, I) &= d(S_iA_mS_i\,,\, I\ A_mS_i) = d(S_iA_m\,,\, A_m) \\ d(S_iA_j\,,\, I) &= d(S_iA_jS_iA_m\,,\, I\ S_iA_m) = d(A_jA_m\,,\, S_iA_m) \\ &= d(S_iA_m\,,\, A_\ell) \end{aligned}$$

where $A_\ell = A_jA_m$. Substituting these expressions in the inequality above yields

$$d(S_iA_m\,,\, A_m) \leqq d(S_iA_m\,,\, A_\ell) \qquad \text{for all } i,\, m,\, \ell$$

This equation says that an arbitrary element in the array (4) is at least as close to the element at the top of its column as it is to any other letter of the alphabet. This is the maximum likelihood property.

2.4 PROOF OF THEOREM 2

Again consider an (n, k)-alphabet as a set of vertices of the unit n-cube. Consider also n mutually perpendicular hyperplanes through the centroid of the cube parallel to the coordinate planes. We call these planes "symmetry planes of the cube" and suppose the planes numbered in accordance with the corresponding parallel coordinate planes.

The reflection of the vertex with coordinates $(a_1\,, a_2\,, \cdots, a_i\,, \cdots, a_n)$ in symmetry plane i yields the vertex of the cube whose coordinates are $(a_1\,, a_2\,, \cdots, a_i \dotplus 1, \cdots, a_n)$. More generally, reflecting a given vertex successively in symmetry planes i, j, k, $\cdots$ yields a new vertex whose coordinates differ from the original vertex precisely in places i, j, k $\cdots$. Successive reflections in hyperplanes constitute a transformation that leaves distances between points unaltered and is therefore a "rotation." The rotation obtained by reflecting successively in symmetry planes i, j, k, etc. can be represented by an n-place symbol having a one in places i, j, k, etc. and a zero elsewhere.

We now regard a given (n, k)-alphabet as generated by operating on the vertex $(0, 0, \cdots, 0)$ of the cube with a certain collection of 2^k rotation operators. The symbols for these operators are identical with the sequences of zeros and ones that form the coordinates of the 2^k points. It is readily seen that these rotation operators form a group which is transitive on the letters of the alphabet and which leave the unit cube invariant. Theorem 2 then follows.

Theorem 2 also follows readily from consideration of the array (4). For example, the maximum likelihood region associated with I is the set of points I, S_2, S_3, $\cdots$, S_ν. The maximum likelihood region associated with A_i is the set of points A_i, A_iS_2, A_iS_3, $\cdots$, A_iS_ν. The rotation (successive reflections in symmetry planes of the cube) whose symbol is the same as the coordinate sequence of A_i sends the maximum likelihood region of I into the maximum likelihood region of A_i, $i = 1, 2, \cdots, \mu$.

2.5 PROOF OF THEOREM 3

That every systematic alphabet is a group alphabet follows trivially from the fact that the sum mod 2 of two letters satisfying parity checks is again a letter satisfying the parity checks. The totality of letters satisfying given parity checks thus constitutes a finite group.

To prove that every group alphabet is a systematic code, consider the letters of a given (n, k)-alphabet listed in a column. One obtains in this way a matrix with 2^k rows and n columns whose entries are zeros and ones. Because the rows are distinct and form a group isomorphic to C_k, there are k linearly independent rows (mod 2) and no set of more than k independent rows. The rank of the matrix is therefore k. The matrix therefore possesses k linearly independent (mod 2) columns and the remaining $n - k$ columns are linear combinations of these k. Maintaining only these k linearly independent columns, we obtain a matrix of k columns and 2^k rows with rank k. This matrix must, therefore, have k linearly independent rows. The rows, however, form a group under mod 2 addition and hence, since k are linearly independent, all 2^k rows must be distinct. The matrix contains only zeros and ones as entries; it has 2^k distinct rows of k entries each. The matrix must be a listing of the numbers from 0 to $2^k - 1$ in binary notation. The other $n - k$ columns of the original matrix considered are linear combinations of the columns of this matrix. This completes the proof of Theorem 3 and Proposition 4.

2.6 PROOF OF THEOREM 4

To prove Theorem 4 we first note that the parity check sequence of the product of two elements of B_n is the mod 2 sum of their separate parity check sequences. It follows then that all elements in a given coset have the same parity check sequence. For, let the coset be S_i, S_iA_2, S_iA_3, $\cdots$ S_iA_μ. Since the elements I, A_2, A_3, $\cdots$, A_μ all have parity check sequence $00 \cdots 0$, all elements of the coset have parity check $R(S_i)$.

In the array (4) there are 2^{n-k} cosets. We observe that there are 2^{n-k} elements of B_n that have zeros in their first k places. These elements have parity check symbols identical with the last $n - k$ places of their symbols. These elements therefore give rise to 2^{n-k} different parity check symbols. The elements must be distributed one per coset. This proves Theorem 4.

2.7 PROOF OF PROPOSITION 5

If

$$n \geqq 2^{n-k} - \binom{n-k}{2} - \binom{n-k}{3} - 1$$

we can explicity exhibit group alphabets having the property mentioned in the paragraph preceding Proposition 5. The notation of the demonstration is cumbersome, but the idea is relatively simple.

We shall use the notation of paragraph 2.1 for elements of B_n, i.e., an element of B_n will be given by a list of integers that specify what places of the sequence for the element contain ones. It will be convenient furthermore to designate the first k places of a sequence by the integers $1, 2, 3, \cdots, k$ and the remaining $n - k$ places by the "integers" $1', 2', 3', \cdots, \ell'$, where $\ell = n - k$. For example, if $n = 8$, $k = 5$, we have

$$\begin{gathered} 10111010 \leftrightarrow 13452' \\ 10000100 \leftrightarrow 11' \\ 00000101 \leftrightarrow 1'3' \end{gathered}$$

Consider the group generated by the elements $1', 2', 3', \cdots, \ell'$, i.e. the 2^ℓ elements $I, 1', 2', \cdots, \ell', 1'2', 1'3', \cdots, 1'2'3' \cdots \ell'$. Suppose these elements listed according to decreasing weight (say in decreasing order when regarded as numbers in the decimal system) and numbered consecutively. Let B_i be the ith element in the list. Example: if $\ell = 3$, $B_1 = 1'2'3'$, $B_2 = 2'3'$, $B_3 = 1'3'$, $B_4 = 1'2'$, $B_5 = 3'$, $B_6 = 2'$, $B_7 = 1'$.

Consider now the (n, k)-alphabet whose generators are

$$1B_1, 2B_2, 3B_3, \cdots, kB_k$$

We assert that if

$$n \geqq 2^{n-k} - \binom{n-k}{2} - \binom{n-k}{3} - 1$$

this alphabet is as good as any other alphabet of 2^k letters and n places.

In the first place, we observe that every letter of this (n, k)-alphabet (except I) has unprimed numbers in its symbols. It follows that each of the 2^ℓ letters $I, 1', 2', \cdots, \ell', 1'2', \cdots, 1'2' \cdots \ell'$ occurs in a different coset of the given (n, k)-alphabet. For, if two of these letters appeared in the same coset, their product (which contains only primed numbers) would have to be a letter of the (n, k) alphabet. This is impossible since every letter of the (n, k) alphabet has unprimed numbers in its symbol. Since there are precisely 2^ℓ cosets we can designate a coset by the single element of the list $B_1, B_2, \cdots, B_{2^\ell} = I$ which appears in the coset.

We next observe that the condition

$$n \geqq 2^{n-k} - \binom{n-k}{2} - \binom{n-k}{3} - 1$$

guarantees that B_{k+1} is of weight 3 or less. For, the given condition is equivalent to

$$k \geqq 2^\ell - \binom{\ell}{0} - \binom{\ell}{1} - \binom{\ell}{2} - \binom{\ell}{3}$$

We treat several cases depending on the weight of B_{k+1}.

If B_{k+1} is of weight 3, we note that for $i = 1, 2, \cdots, k$, the coset containing B_i also contains an element of weight one, namely the element i obtained as the product of B_i with the letter iB_i of the given (n, k)-alphabet. Of the remaining $(2^\ell - k)$ B's, one is of weight zero, ℓ are of weight one, $\binom{\ell}{2}$ are of weight 2 and the remaining are of weight 3. We have, then $\alpha_0 = 1$, $\alpha_1 = \ell + k = n$. Now every B of weight 4 occurs in the list of generators $1B_1, 2B_2, \cdots, kB_k$. It follows that on multiplying this list of generators by any B of weight 3, at least one element of weight two will result. (E.g., $(1'2'3')(j1'2'3'4') = j4'$) Thus every coset with a B of weight 2 or 3 contains an element of weight 2 and $\alpha_2 = 2^\ell - \alpha_0 - \alpha_1$.

The argument in case B_{k+1} is of weight two or one is similar.

2.8 MODULAR REPRESENTATIONS OF C_n

In order to explain one of the methods used to obtain the best (n, k)-alphabets listed in Tables II and III, it is necessary to digress here to present additional theory.

It has been remarked that every (n, k)-alphabet is isomorphic with C_k. Let us suppose the elements of C_k listed in a column starting with I and proceeding in order $I, 1, 2, 3, \cdots, k, 12, 13, \cdots, (k-1)k, 123, \cdots \cdots, 123 \cdots k$. The elements of a given (n, k)-alphabet can be paired off with these abstract elements so as to preserve group multiplication. This can be done in many different ways. The result is a matrix with elements zero and one with n columns and 2^k rows, these latter being labelled by the symbols $I, 1, 2, \cdots$ etc. What can be said about the columns of this matrix? How many different columns are possible when all (n, k)-alphabets and all methods of establishing isomorphism with C_k are considered?

In a given column, once the entries in rows $1, 2, \cdots, k$ are known, the entire column is determined by the group property. There are therefore only 2^k possible different columns for such a matrix. A table showing these 2^k possible columns of zeros and ones will be called a *modular representation* table for C_k. An example of such a table is shown for $k = 4$ in Table VI.

It is clear that the columns of a modular representation table can also be labelled by the elements of C_k, and that group multiplication of these column labels is isomorphic with mod 2 addition of the columns. The table is a symmetric matrix. The element with row label A and column label B is one if the symbols A and B have an odd number of different numerals in common and is zero otherwise.

Every (n, k)-alphabet can be made from a modular representation table by choosing n columns of the table (with possible repetitions) at least k of which form an independent set.

TABLE VI — MODULAR REPRESENTATION TABLE FOR GROUP C_4

	I	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	1	1	1	0	0	0	1	1	1	0	1
2	0	0	1	0	0	1	0	0	1	1	0	1	1	0	1	1
3	0	0	0	1	0	0	1	0	1	0	1	1	0	1	1	1
4	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1
12	0	1	1	0	0	0	1	1	1	1	0	0	0	1	1	0
13	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0
14	0	1	0	0	1	1	1	0	0	1	1	1	0	0	1	0
23	0	0	1	1	0	1	1	0	0	1	1	0	1	1	0	0
24	0	0	1	0	1	1	0	1	1	0	1	1	0	1	0	0
34	0	0	0	1	1	0	1	1	1	1	0	1	1	0	0	0
123	0	1	1	1	0	0	0	1	0	1	1	1	0	0	0	1
124	0	1	1	0	1	0	1	0	1	0	1	0	1	0	0	1
134	0	1	0	1	1	1	0	0	1	1	0	0	0	1	0	1
234	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1
1234	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0

We henceforth exclude consideration of the column I of a modular representation table. Its inclusion in an (n, k)-alphabet is clearly a waste of 1 binary digit.

It is easy to show that every column of a modular representation table for C_k contains exactly 2^{k-1} ones. Since an (n, k)-alphabet is made from n such columns the alphabet contains a total of $n2^{k-1}$ ones and we have

Proposition 6. The weights of an (n, k)-alphabet form a partition of $n2^{k-1}$ into $2^k - 1$ non-zero parts, each part being an integer from the set $1, 2, \cdots, n$.

The identity element always has weight zero, of course.

It is readily established that the product of two elements of even weight is again an element of even weight as is the product of two elements of odd weight. The product of an element of even weight with an element of odd weight yields an element of odd weight.

The elements of even weight of an (n, k)-alphabet form a subgroup and the preceding argument shows that this subgroup must be of order 2^k or 2^{k-1}. If the group of even elements is of order 2^{k-1}, then the collection of even elements is a possible $(n, k - 1)$-alphabet. This $(n, k - 1)$ alphabet may, however, contain the column I of the modular representation table of C_{k-1}. We therefore have

Proposition 7. The partition of Proposition 6 must be either into $2^k - 1$ even parts or else into 2^{k-1} odd parts and $2^{k-1} - 1$ even parts. In the latter case, the even parts form a partition of $\alpha 2^{k-2}$ where α is some integer of the set $k - 1, k, \cdots, n$ and each of the parts is an integer from the set $1, 2, \cdots, n$.

2.9 THE CHARACTERS OF C_k

Let us replace the elements of B_n (each of which is a sequence of zeros and ones) by sequences of $+1$'s and -1's by means of the following substitution

$$\begin{aligned} 0 &\leftrightarrow 1 \\ 1 &\leftrightarrow -1. \end{aligned} \tag{13}$$

The multiplicative properties of elements of B_n can be preserved in this new notation if we define the product of two $+1$, -1 symbols to be the symbol whose ith component is the ordinary product of the ith components of the two factors. For example, 1011 and 0110 become respectively $-11\ -1\ -1$ and $1\ -1\ -11$. We have

$$(-11\ -1\ -1)(1\ -1\ -11) = (-1\ -11\ -1)$$

corresponding to the fact that

$$(1011)\ (0110) = (1101)$$

If the $+1$, -1 symbols are regarded as shorthand for diagonal matrices, so that for example

$$-1 1 -1 -1 \leftrightarrow \begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

then group multiplication corresponds to matrix multiplication.

(While much of what follows here can be established in an elementary way for the simple group at hand, it is convenient to fall back upon the established general theory of group representations[8] for several propositions.

The substitution (13) converts a modular representation table (column I included) into a square array of $+1$'s and -1's. Each column (or row) of this array is clearly an irreducible representation of C_k. Since C_k is Abelian it has precisely 2^k irreducible representations each of degree one. These are furnished by the converted modular table. This table also furnishes then the characters of the irreducible representations of C_k and we refer to it henceforth as a *character table.*

Let $\chi^{\alpha}(A)$ be the entry of the character table in the row labelled A and column labelled α. The orthogonality relationship for characters gives

$$\sum_{A \subset C_k} \chi^{\alpha}(A)\chi^{\beta}(A) = 2^k \delta_{\alpha\beta}$$

$$\sum_{\alpha \subset C_k} \chi^{\alpha}(A)\chi^{\alpha}(B) = 2^k \delta_{AB}$$

where δ is the usual Kronecker symbol. In particular

$$\sum_{A \subset C_k} \chi^{I}(A)\chi^{\beta}(A) = \sum_{A \subset C_k} \chi^{\beta}(A) = 0, \qquad \beta \neq I$$

Since each $\chi^{\beta}(A)$ is $+1$ or -1, these must occur in equal numbers in any column $\beta \neq I$. This implies that each column except I of the modular representation table contains 2^{k-1} ones, a fact used earlier.

Every matrix representation of C_k can be reduced to its irreducible components. If the trace of the matrix representing the element A in an arbitrary matrix representation of C_k is $\chi(A)$, then this representation contains the irreducible representation having label β in the character table d_{β} times where

$$d_\beta = \frac{1}{2^k} \sum_{A \subset C_k} \chi(A)\chi^\beta(A) \tag{14}$$

Every (n, k)-alphabet furnishes us with a matrix representation of C_k by means of (13) and the procedure outlined below (13). The trace $\chi(A)$ of the matrix representing the element A of C_k is related to the weight of the letter by

$$\chi(A) = n - 2w(A) \tag{15}$$

Equations (14) and (15) permit us to compute from the weights of an (n, k)-alphabet what irreducible representations are present in the alphabet and how many times each is contained. It is assumed here that the given alphabet has been made isomorphic to C_k and that the weights are labelled by elements of C_k .

Consider the converse problem. Given a set of numbers w_1 , w_2 , $\cdots$, w_{2^k} that satisfy Propositions 6 and 7. From these we can compute quantities $\chi_i = n - 2w_i$ as in (15). It is clear that the given w's will constitute the weights of an (n, k)-alphabet if and only if the 2^k χ_i can be labelled with elements of C_k so that the 2^k sums (14) (β ranges over all elements of C_k) are non-negative integers. The integers d_β tell what representations to choose to construct an (n, k)-alphabet with the given weights w_1 .

2.10 CONSTRUCTION OF BEST ALPHABETS

A great many different techniques were used to construct the group alphabets listed in Tables II and III and to show that for each n and k there are no group alphabets with smaller probability of error. Space prohibits the exhibition of proofs for all the alphabets listed. We content ourselves here with a sample argument and treat the case $n = 10$, $k = 4$ in detail.

According to (2) there are $N(10, 4) = 53{,}743{,}987$ different (10, 4)-alphabets. We now show that none is better than the one given in Table III. The letters of this alphabet and weights of the letters are

I	0
1 6 7 8 10	5
2 6 7 9 10	5
3 5 6 8 9 10	6
4 5 7 8 9 10	6
1 2 8 9	4
1 3 5 7 9	5

1 4 5 6 9	5
2 3 5 7 8	5
2 4 5 6 8	5
3 4 6 7	4
1 2 3 5 7 9	6
1 2 4 5 7 10	6
1 3 4 8 10	5
2 3 4 9 10	5
1 2 3 4 6 7 8 9	8

The notation is that of Section 2.1. By actually forming the standard array of this alphabet, it is verified that

$$\alpha_0 = 1, \qquad \alpha_1 = 10, \qquad \alpha_2 = 39, \qquad \alpha_3 = 14.$$

Table II shows $\binom{10}{2} = 45$, whereas $\alpha_2 = 39$, so the given alphabet does not correct all possible double errors. In the standard array for the alphabet, 39 coset leaders are of weight 2. Of these 39 cosets, 33 have only one element of weight 2; the remaining 6 cosets each contain two elements of weight 2. This is due to the two elements of weight 4 in the given group, namely 1289 and 3467. A portion of the standard array that demonstrates these points is

I	1289	3467
·	·	·
·	·	·
12	89	·
18	29	·
19	28	·
34	·	67
36	·	47
37	·	46
·	·	·
·	·	·

In order to have a smaller probability of error than the exhibited alphabet, it is necessary that a (10, 4)-alphabet have an $\alpha_2 > 39$. We proceed to show that this is impossible by consideration of the weights of the letters of possible (10, 4)-alphabets.

We first show that every (10, 4)-alphabet must have at least one element (other than the identity, I) of weight less than 5. By Propositions 6 and 7, Section 2.8, the weights must form a partition of $10 \cdot 8 = 80$ into 15 positive parts. If the weights are all even, at least two must be less than 6 since $14 \cdot 6 = 84 > 80$. If eight of the weights are odd, we see from $8 \cdot 5 + 7 \cdot 6 = 82 > 80$ that at least one weight must be less than 5.

An alphabet with one or more elements of weight 1 must have an $\alpha_2 \leqq 36$, for there are nine elements of weight 2 which cannot possibly be coset leaders. To see this, suppose (without loss of generality) that the alphabet contains the letter 1. The elements 12, 13, 14, $\cdots$ 1 10 cannot possibly be coset leaders since the product of any one of them with the letter 1 yields an element of weight 1.

An alphabet with one or more elements of weight 2 must have an $\alpha_2 \leqq 37$. Suppose for example, the alphabet contained the letter 12. Then 13 and 23 must be in the same coset, 14 and 24 must be in the same coset, $\cdots$, 1 10 and 2 10 must be in the same coset. There are at least eight elements of weight two which are not coset leaders.

Each element of weight 3 in the alphabet prevents three elements of weight 2 from being coset leaders. For example, if the alphabet contains 123, then 12, 13, and 23 cannot be coset leaders. We say that the three elements of weight 2 are "blocked" by the letter of weight 3. Suppose an alphabet contains at least three letters of weight three. There are several cases: (A) if three letters have no numerals in common, e.g., 123, 456, 789, then nine distinct elements of weight 2 are blocked and $\alpha_2 \leqq 36$; (B) if no two of the letters have more than a single numeral in common, e.g., 123, 345, 789, then again nine elements of weight 2 are blocked and $\alpha_2 \leqq 36$; and (C) if two of the letters of weight 3 have two numerals in common, e.g., 123, 234, then their product is a letter of weight 2 and by the preceding paragraph $\alpha_2 \leqq 37$. If an alphabet contains exactly two elements of weight 3 and no elements of weight 2, the elements of weight 3 block six elements of weight 2 and $\alpha_2 \leqq 39$.

The preceding argument shows that to be better than the exhibited alphabet a (10, 4)-alphabet with letters of weight 3 must have just one such letter. A similar argument (omitted here) shows that to be better than the exhibited alphabet, a (10, 4)-alphabet cannot contain more than one element of weight 4. Furthermore, it is easily seen that an alphabet containing one element of weight 3 and one element of weight 4 must have an $\alpha_2 \leqq 39$.

The only new contenders for best (10, 4)-alphabet are, therefore, alphabets with a single letter other than I of weight less than 5, and this letter must have weight 3 or 4. Application of Propositions 6 and 7 show that the only possible weights for alphabets of this sort are: 35^76^7 and 5^846^6 where 5^7 means seven letters of weight 5, etc. We next show that there do not exist (10, 4)-alphabets having these weights.

Consider first the suggested alphabet with weights 35^76^7. As explained in Section 2.9, from such an alphabet we can construct a matrix representation of C_4 having the character $\chi(I) = 10$, one matrix of trace 4, seven of trace 0 and seven of trace -2. The latter seven matrices correspond to elements of even weight and together with I must represent a subgroup of order 8. We associate them with the subgroup generated by the elements 2, 3, and 4. We have therefore

$$\chi(I) = 10, \qquad \chi(2) = \chi(3) = \chi(4) = \chi(23)$$
$$= \chi(24) = \chi(34) = \chi(234) = -2.$$

Examination of the symmetries involved shows that it doesn't matter how the remaining χ_i are associated with the remaining group elements. We take, for example

$$\chi(1) = 4, \qquad \chi(12) = \chi(13) = \chi(14) = \chi(123)$$
$$= \chi(124) = \chi(134) = \chi(1234) = 0.$$

Now form the sum shown in equation (14) with $\beta = 1234$ (i.e., with the character χ^{1234} obtained from column 1234 of the Table VI by means of substitution (13). There results $d_{1234} = \frac{1}{2}$ which is impossible. Therefore there does not exist a (10, 4)-alphabet with weights $35^7 6^7$.

The weights $5^8 46^6$ correspond to a representation of C_4 with character $\chi(I) = 10, 0^8, 2, (-2)^6$. We take the subgroup of elements of even weight to be generated by 2, 3, and 4. Except for the identity, it is clearly immaterial to which of these elements we assign the character 2. We make the following assignment: $\chi(I) = 10$, $\chi(2) = 2$, $\chi(3) = \chi(4) = \chi(23) = \chi(24) = \chi(34) = \chi(234) = -2$, $\quad \chi(1) = \chi(12) = \chi(13) = \chi(14) = \chi(123) = \chi(124) = \chi(134) = \chi(1234) = 0$. The use of equation (14) shows that $d_2 = \frac{1}{2}$ which is impossible.

It follows that of the 53,743,987 (10, 4)-alphabets, none is better than the one listed on Table III.

Not all the entries of Table III were established in the manner just demonstrated for the (10, 4)-alphabet. In many cases the search for a best alphabet was narrowed down to a few alphabets by simple arguments. The standard arrays for the alphabets were constructed and the best alphabet chosen. For large n the labor in making such a table can be considerable and the operations involved are highly liable to error when performed by hand.

I am deeply indebted to V. M. Wolontis who programmed the IBM CPC computer to determine the α's of a given alphabet and who patiently ran off many such alphabets in course of the construction of Tables II and III. I am also indebted to Mrs. D. R. Fursdon who evaluated many of the smaller alphabets by hand.

REFERENCES

1. R. W. Hamming, B.S.T.J., **29,** pp. 147–160, 1950.
2. I. S. Reed, Transactions of the Professional Group on Information Theory, PGIT-4, pp. 38–49, 1954.
3. See section 7 of R. W. Hamming's paper, loc. cit.
4. I.R.E. Convention Record, Part 4, pp. 37–45, 1955 National Convention, March, 1955.
5. C. E. Shannon, B.S.T.J., **27,** pp. 379–423 and pp. 623–656, 1948.
6. Birkhoff and MacLane, A Survey of Modern Algebra, Macmillan Co., New York, 1941. Van der Waerden, Modern Algebra, Ungar Co., New York, 1953. Miller, Blichfeldt, and Dickson, Finite Groups, Stechert, New York, 1938.
7. This theorem has been previously noted in the literature by Kiyasu-Zen'iti, Research and Development Data No. 4, Ele. Comm. Lab., Nippon Tele. Corp. Tokyo, Aug., 1953.
8. F. D. Murnaghan, Theory of Group Representations, Johns Hopkins Press, Baltimore, 1938. E. Wigner, Gruppentheorie, Edwards Brothers, Ann Arbor, Michigan, 1944.