THE DISTANCE OF ARITHMETIC CODES

MUNEHIRO GOTO* and TERUO FUKUMURA

Department of Electrical Engineering

(Received October 30, 1968)

1. Introduction

Arithmetic codes are useful for error control in both arithmetic operations and data transmission. When used in data transmission, they do not require special coding equipment, as encoding and decoding operations may be easily carried out in a general-purpose computer.

An arithmetic code is a code of the form AN ($N=0, 1, \dots, B$-1) where A is an odd integer. As all code words here are numerals, the distance of this code can not be defined by Hamming distance. This makes it difficult to evaluate the distance of the arithmetic code. To the best of our knowledge, there are no papers discussing the evaluation of the large distance of the arithmetic code expect Mandelbaum's[1]. In his recent work, he has found the upper and lower bounds on the distance of a certain class of these codes.

In this paper we discuss the distance of the arithmetic code generated by an odd integer A. After defining some symbols, a minimal representation for the binary number and its property are introduced. By using this representation, a distance of the code is defined and the main theorem in this paper is stated in advance, in section 4. The proof is given in section 5. This main theorem gives the formula to evaluate the distance of the code and its corollary gives the exact value, not the upper or the lower bounds, of the distance of the Mandelboum's code. Theorem 2 in section 6 shows that the formula similar to that in the case of symmetric error-detecting codes holds also for the asymmetric case.

2. Definitions of symbols

We use below the following symbols.

A; code generator, *i.e.* an odd integer.

$e(A)$; the minimum positive integer which satisfies

$$2^{e(A)} - 1 \equiv 0 \qquad (\text{mod } A)$$

B; an odd integer defined by

$$B = (2^{e(A)} - 1)/A \tag{1}$$

J; an integer.

$R_0 = \{J \mid 1 \leqq J \leqq B-1\}$

$R_1 = \{J \mid 1 \leqq J \leqq B/3\}$

* The author is with Gifu University.

$P_0 = \{\alpha \mid -B/3 < \alpha \leqq B/3, \text{ odd integer}\}$
$P_1 = \{J \mid J = \alpha \pmod{B},\ J \in R_0,\ \alpha \in P_0\}$
$Q_1 = \{J \mid J \in R_0, \text{ odd integer}\}$
$H_1 = \{J \mid J \in R_1 \text{ or twice of } J \in R_1\}$
$H_2 = \{J \mid B/3 < J \leqq 2B/3, \text{ odd integer}\}$
$H_3 = \{J \mid 2B/3 < J < B\}$
$a \in Z$; a is in the set Z.
$Z_1 \cup Z_2$; union of Z_1 and Z_2
$Z_1 \cap Z_2$; intersection of Z_1 and Z_2

$\sim$; equivalence relation defined on integers, *i.e.*, $J_1 \sim J_2$ if and only if J_1, $J_2 \in R_0$ and there exists j such that

$$J_1 \equiv J_2 2^j \quad (\mathrm{mod}\ B).$$

z_k; an equivalence class on R_0 defined by the relation $\sim$. For any two distinct classes z_k, $z_{k'}$,

$$z_k \cap z_{k'} = \phi$$

L; number of z_k's, *i.e.*,

$$R_0 = \bigcup_{k=1}^{L} z_k$$

$\#(S)$; number of members in a set S.

$$w_k = \#(z_k \cap P_1) e(A) / \#(z_k)$$
$$W_k = \#(z_k \cap Q_1) e(A) / \#(z_k)$$

Note that in general

$$e(A) = e(B)\alpha_1 \quad \text{and} \quad e(B) = \#(z_k)\alpha_2,$$

where α_1 and α_2 are positive integers respectively. Thus, from Eq. (1),

$$A = (2^{e(B)\alpha_1} - 1)/B. \tag{2}$$

3. Minimal representation

The binary minimal representation for any integer J is the expression of J having the least number of nonzero terms among all expressions of the form

$$J = b_0 2^0 + b_1 2^1 + \cdots + b_m 2^m, \tag{3}$$

where $b_i = 0$ or ± 1 $(i = 0, 1, \ldots, m)$. The least number of nonzero terms is denoted by $w(J)$ referring to the arithmetic weight of J.

P.1. (Reitwiesner[2]) If the expression in the form of Eq. (3) satisfies

$$b_i b_{i+1} = 0 \quad (i = 0, 1, \ldots, m-1),$$

then the expression is the minimal representation for J.

4. Distance of codes and main theorem

In the arithmetic codes the messages N are presented by the integers in R_0 and by zero, and the corresponding code words are the binary representations for $A \times N$. If we discuss the symmetric errors where each digit may change both from zero to one and from one to zero, then the distance between two code words AN_1 and AN_2 must be defined by

$$w(AN_1 - AN_2).$$

In general the distance of code, d_A, is defined to be the minimum distance between distinct code words. In this coding, it follows that

$$d_A = \min_{N \in R_0} (w(AN)), \tag{4}$$

because the difference $(AN_1 - AN_2)$ is also a word of the code generated by A.

Massey[6] proves the following. The arithmetic code correct all patterns of t or fewer errors (or detect all patterns of $2t$ or fewer errors) if and only if its distance d_A is at least $2t+1$. In what follows we regard d_A as an ability to detect the symmetric errors.

One of the purposes of this note is to prove the following theorem.

Theorem 1. When the generator A is an odd integer,

$$d_A = \min_{1 \leqq k \leqq L} (w_k), \tag{5}$$

where the length of code word is $e(A)$ and the number of messages is B, that is, $N = 0, 1, \ldots, B-1$.

Our plan of the proof of this theorem is as follows. First the non-restoring binary division[3], of which quotient digits satisfy P. 1, is introduced. Then it is shown that if the divisor is B and the dividend is $N \in z_k$, then the arithmetic weight for the sequence of quotient digits $(q_1 q_2 \cdots q_{e(A)})_2$ is w_k. Using this result and

$$AN = \frac{N}{B} 2^{e(A)} - \frac{N}{B}, \tag{6}$$

the arithmetic weight of $(AN)_2$ is evaluated. Finally a simple number theoretic consideration leads us to Theorem 1.

5. Proof of Theorem 1

Preparation 1. Non-restoring binary division and its quotients

We assume in what follows that the divisor is B and the dividend X_0 is the number in R_0. Then each step of the non-restoring binary division is performed in accordance with the equations

$$\left.\begin{aligned} X_i' &= X_i - q_i B \\ X_{i+1} &= 2X_i' \end{aligned}\right\} \tag{7}$$

where X_i is the partial remainder and i represents the step of division. The quotient digit q_i is selected by the rules;*

$$\begin{array}{lll} \text{If} & |X_i| \leqq 2B/3, & \text{then } q_i = 0. \\ \text{If} & 2B/3 < X_i \leqq 4B/3, & \text{then } q_i = 1. \\ \text{If} & -4B/3 \leqq X_i < -2B/3, & \text{then } q_i = -1. \end{array} \tag{8}$$

As B is an odd integer, this division produces the infinite sequence of the quotient digits, *i.e.*, for any integer N

$$\frac{N}{B} = \sum_{i=0}^{\infty} q_i 2^{-i}. \tag{9}$$

From the quotient selection rule, it can be easily shown that
P. 2.

$$q_i q_{i+1} = 0, \quad (i = 0, 1, 2, \ldots) \tag{10}$$

that is, the sequence of the quotient digits has the property stated in P. 1.

Since, in ordinary binary division, the partial remainders Y_i always remain positive and that

$$Y_i = Y_{i+e(B)} \quad (i = 0, 1, 2, \ldots), \tag{11}$$

the corresponding quotients have also the same relation

$$q_i = q_{i+e(B)} \quad (i = 0, 1, 2, \ldots). \tag{12}$$

However, in the non-restoring division, Eq. (12) does not always hold, because the partial remainders are allowed to take negative values. But it is not difficult to show that once the nonzero quotient, say $q_j = 0$, is obtained,

$$q_i = q_{i+e(B)} = q_{i+e(A)} \quad (i = j+1,\ j = 2, \ldots). \tag{13}$$

Thus, from $e(A) = e(B)\alpha_1$,

$$\begin{aligned} X_0/B &= (q_0 \cdot q_1 \cdots q_j \dot{q}_{j+1} \cdots \dot{q}_{j+e(B)})_2 \\ &= (q_0 \cdot q_1 \cdots q_j \dot{q}_{j+1} \cdots \dot{q}_{j+e(A)})_2 \end{aligned} \tag{14}$$

where $(\dot{q}_j \cdots \dot{q}_k)$ means infinite recurrences of the sequence $q_j \cdots q_k$.

We will now consider the properties associated with the sequence of q_i's. From Eqs. (7) and (8), the partial remainders X_i $(i \neq 0)$ are always even, and X_i's are even or odd according as $q_i = 0$ or not. If X_i is odd, its absolute value is less than or equal to $B/3$, that is, otherwise $X_i' \notin P_0$ noting that $X_0 > 0$.

P. 3. In the non-restoring binary division, $q_i = 0$ $(i \neq 0)$ if and only if $X_i' \in P_0$.

In other words, as $X_i \equiv X_i' \equiv Y_i$ (mod B), $q_i \neq 0$ if and only if $Y \in P_1$. On the other hand, since each number in z_k is contained $e(A)/\#(z_k)$ times in the sequence $(Y_1 Y_2 \cdots Y_{e(A)})$, we have the following.

P. 4. The number of nonzero digits in the sequence $(q_1 q_2 \cdots q_{e(A)})$ is equal to w_k, where the divisor $X_0 \in z_k$.

---

* The selection rules are made to have the quotient digits possess property P. 1. Here the word non-restoring division means the division by these rules.

Since $(AN)_2$ is not always $(q_1 \cdots q_{e(A)})_2$, three cases must be considered in order to evaluate $w(AN)$ using Eq. (6). First, if $X_0 \in H_1$, then the dividend X_0 might be regarded as the virtual partial remainder obtained at some step in a fictious division process where $q_i \neq 0$ for some integer i and $q_{i+1}=q_{i+2}=\cdots=q_0=0$ or $q_0 \neq 0$. Thus j in Eq. (13) might be regarded as negative or zero. Hence

P. 5-1. If $X_0 \in H_1$, then

$$X_0/B = (0.\dot{q}_1 q_2 \cdots q_{e(A)})_2.$$

Next we consider the case such that $X_0 \in H_2$.

P. 5-2. If $X_0 \in H_2$, then

$$X_0/B = (0.1\dot{0} \cdots q_{e(A)-1} 0 \dot{\bar{1}})_2,$$

where $\bar{1} = -1$.

Poof. As X_0 is an odd integer greater than $B/3$, it is not allowed to regard X_0 as a partial remainder in non-restoring division process as before. From the quotient selection rules, $q_0=0$. Since $2B/3<2X_0=X_1<4B/3$, $q_1=1$. Thus Eq. (13) shows

$$q_i = q_{i+e(A)} \qquad (i=2, 3, \ldots).$$

Now, as $X_{e(A)} \equiv X_0 \pmod{B}$ but X_0 is an odd number, $X_{e(A)}$ given in the non-restoring division process must be a negative even integer

$$X_{e(A)} = X_0 - B.$$

This means that $q_{e(A)+1} = -1 = \bar{1}$, and by using P. 2 $q_2=q_{e(A)}=0$. (QED)

Finally, when $X_0 \in H_3$, $q_0=1$ according to the selection rules. Thus Eq. (14) shows that

P. 5-3. If $X_0 \in H_3$, then

$$X_0/B = (1.\dot{q}_1 \cdots \dot{q}_{e(A)})_2$$

We can obtain the following expressions from P. 5 and Eq. (6).

$$AN = (q_1 \cdots q_{e(A)})_2, \quad \text{if } N \in H_1. \tag{15}$$

$$AN = (10\, q_3 \cdots q_{e(A)-1} 0)_2 - 1, \quad \text{if } N \in H_2. \tag{16}$$

$$AN = (10\, q_2 \cdots q_{e(A)})_2 - 1, \quad \text{if } N \in H_3. \tag{17}$$

It follows from P. 2 that Eq. (15) is a minimal representation for AN. Thus, by virtue of P. 4,

P. 6-1. If $N \in H_1 \cap z_k$, then

$$w(AN) = w_k$$

For $N \in H_2 \cup H_3$, Eqs. (16) and (17) are not minimal representations due to the last term -1. In order to obtain the minimal representation for Eqs. (16) and (17), the following problem must be considered:

Calculate the weight of the number $(q_1 \cdots q_{e(A)})_2 - 1$ under the conditions that $(q_1 \cdots q_{e(A)})_2$ is minimally represented and its weight is w_k. To solve this problem, it is sufficient to consider the seven cases.

1 $(\cdots 00)_2 - 1 = (\cdots 0\bar{1})_2$

2 $(\cdots 10)_2 - 1 = (\cdots 01)_2$

3 $(\cdots 01)_2 - 1 = (\cdots 00)_2$

4 $(\cdots 10\bar{1}0\bar{1}0 \cdots 010)_2 - 1 = (\cdots 0101 \cdots 101)_2$

5 $(\cdots 00\bar{1}0\bar{1} \cdots 0\bar{1}0)_2 - 1 = (\cdots 0\bar{1}010 \cdots 101)_2$

6 $(\cdots 10\bar{1}0 \cdots \bar{1}0\bar{1})_2 - 1 = (\cdots 0101 \cdots 010)_2$

7 $(\cdots 00\bar{1}0\bar{1}0 \cdots \bar{1}0\bar{1})_2 - 1 = (\cdots 0\bar{1}0\bar{1}0\bar{1} \cdots 0\bar{1}0)_2$

Clearly, all expressions in the right hand sides of above equations are the minimal representation satisfying P. 1. Then the arithmetic weights are invariably w_k in the cases 2, 4 and 7, (w_k-1) in 3 and 6, and (w_k+1) in 1 and 5 respectively. These results lead us to P. 6-2 and P. 6-3.

P. 6-2. If $N \in H_2 \cap z_k$, then

$$w(AN) = w_k \text{ or } w_k + 1.$$

Proof. From Eq. (16), $q_{e(A)} = 0$. Then the cases 3 and 6 do not occur. (QED)

P. 6-3. If $N \in H_3 \cap z_k$, then

$$w(AN) = w_k,\ w_k + 1 \text{ or } w_k + 2.$$

Proof. Since $N \in H_3$, $q_0 = 1$. Thus $w((q_0 \cdots q_{e(A)})_2) = w_k + 1$. In the cases 1, 2, ..., and 7, the weights change at most by 1. Therefore $w(AN) = (w_k+1) - 1$, w_k+1 or w_k+1+1. (QED)

Preparation 2

The next lemma is useful for the proof of Theorem 1.

Lemma 1. If $z_k \cap H_1 = \phi$, there exists $z_{k'} \neq z_k$ such that $z_k \cap H_1 \neq \phi$ and $w_{k'} = w_k$.

Proof. Suppose z_k is composed of only the member in H_2. Then, for some number $N_1 \in z_k$, $2B/3 < 2N_1 < 4B/3$. Thus $J \equiv 2N_1 \pmod{B}$ $(J \in R_0)$ is the number included in $H_1 \cup H_3$. As $z_k \cap H_1 = \phi$, J must be in H_3. Hence

$$z_k \cap H_3 \neq \phi \qquad \text{and} \qquad z_k \cap H_2 \neq 0.$$

Next suppose that both N and $B\text{-}N$ are contained in z_k. Then, for some $N \in z_k \cap H_3$, $B\text{-}N \in H_1 \cap z_k$. This contradicts that $z_k \cap H_1 = \phi$. Therefore, if $z_k \cap H_1 = \phi$, then we find the other equivalence class $z_{k'}$ such that

$$z_{k'} = \{N' \mid N' = B\text{-}N,\ N \in z_k\}.$$

Then, from a simple number theoretic consideration, it is known that z_k is also one of the equivalence class defined by $\sim$. By the definition of w_k, we can say that w_k is the sum of the numbers of odd integer in $z_k \cap H_1$ and the number of even integers in $z_k \cap H_3$. Since B is odd, the number of even integers in $z_k \cap H_3 (z_{k'} \cap H_1)$ is equal to the number of the odd ones in $z_k \cap H_1 (z_k \cap H_3)$. Now, $z_k \cap H_1 = \phi$. Hence $w_k = w_{k'}$.

Finally $z_k \cap H_3 \neq \phi$ means $z_k \cap H_1 \neq \phi$. (QED)

Proof of Theorem 1

From Eq. (4) and P. 6, obviously

$$\min_{1\leqq k\leqq L}(w_k)\leqq d_A\leqq\min_{1\leqq k\leqq L}(w_k)+2.$$

If $z_k\cap H_1\neq\phi$ then from P. 6–1

$$\min_{N\in z_k}(w(AN))=w_k,$$

otherwise, from Lemma 1

$$\min_{N\in z_k\cup z_{k'}}(w(AN))=w_k.$$

Thus we can obtain Theorem 1. (QED)

Corollary 1. If B is an odd prime and has 2 as a primitive root, then

$$d_A=2\left[\frac{B+3}{6}\right],$$

where $e(A)=e(B)$.

Proof. Under the conditions stated in the corollary, $R_0=z_1$. Thus

$$d_A=w_1=\#(R_0\cap P_1)=2\left[\frac{B+3}{6}\right]. \quad \text{(QED)}$$

The generator $A=(2^{B-1}-1)/B$ in this case coincides with Mandelbaum's[1].

Corollary 2. If B is a prime and has -2 as a primitive root (but not 2), then

$$d_A=\left[\frac{B+3}{6}\right],$$

where $e(A)=e(B)$.

Proof. Since -2 is a primitive root of B, $e(A)=e(B)=\dfrac{B-1}{2}$.

Hence the member of R_0 is devided into two equivalence classes z_1 and z_2 which correspond to z_k and $z_{k'}$ mentioned in Lemma 1 respectively[4]. Thus $w_1=w_2$. On the other hand, $R_0=z_1\cup z_2$ and $z_1\cap z_2=\phi$ mean $2\left[\dfrac{B+3}{6}\right]=w_1+w_2$. (QED)

For example, B=23 is a prime and has -2 as a primitive root. Then, by virtue of Eq. (1), $A=89$. Thus from Corollary 2 $d_A=4$. Table 1 shows the code words and their minimal representations. The arithmetic weights presented in the last column show that $d_A=4$.

6. Detection of asymmetric errors

When we discuss the asymmetric errors, where all digits may change only from zero to one or only from one to zero, we must define the distance between two code words by

$$W(AN_1-AN_2),$$

TABLE 1. $B=23$ and $A=89$. The symbol * means -1.

N_i	$(AN)_2$	minimal representation	weight
0	00000000000	00000000000	0
1	00001011001	00010*0*001	4
2	00010110010	0010*0*0010	4
3	00100001011	00100010*0*	4
4	00101100100	010*0*00100	4
5	00110111101	0100*000*01	4
6	01000010110	0100010*0*0	4
7	01001101111	010100*000*	4
8	01011001000	10*0*001000	4
9	01100100001	10*00100001	4
10	01101111010	100*000*010	4
11	01111010011	1000*01010*	5
12	10000101100	100010*0*00	4
13	10010000101	10010000101	4
14	10011011110	10100*000*0	4
15	10100110111	1010100*00*	5
16	10110010000	10*0*0010000	4
17	10111101001	10*000*01001	5
18	11001000010	10*001000010	4
19	11010011011	10*010100*0*	6
20	11011110100	100*000*0100	4
21	11101001101	100*01010*01	6
22	11110100110	1000*01010*0	5

where $W(J)$ denotes the number of nonzero digits in the ordinary binary representation for J. Except this point, definition of the distance of code is same as that in symmetric case, that is, the distance of code D_A for the asymmetric error detection is the minimum distance among the distinct code words. The all patterns of $2t$ or fewer errors are detected, if $D_A=2t+1$. However, it is not always true that all patterns of t or fewer errors are corrected.

By noting that for any N_1, $N_2 \in z_k$, the code word $(AN_1)_2$ is equal to $(AN_2)_2$ shifted cyclically,

$$W(AN_1) = W(AN_2) = W_k.$$

It follows that

$$D_A = \min_{N \in K_0} (W(AN)) = \min_{1 \leqq k \leqq L} (W_k).$$

However, in this note, we will derive more strong result, that is,

Theorem 2. When the generator A is given by Eq. (1),

$$D_A = \min_{N \geqq 1} (W(AN)) = \min_{1 \leqq k \leqq L} (W_k).$$

As the number of messages N_0 is arbitrary in this coding, so the length of the code words must be long enough to represent the longest code word $A(N_0-1)$.

Proof. For any integer $N \geqq 1$, there exist integers α and γ such that $N=\alpha B+\gamma$ ($\alpha=0, 1, 2, \ldots, 0 \leqq \gamma \leqq B-1$). Then, using the relation

$$W(J) + W(K) \geqq W(J+K)^{5)},$$

where J and K are any integers,

$$W(AN) = W(A\alpha B + A\gamma) = W(\alpha 2^{e(A)} + A\gamma - \alpha)$$
$$\geqq W(\alpha 2^{e(A)} + A\gamma) - W(\alpha) = W(A\gamma).$$

Hence, if $\gamma \neq 0$ there exists some z_k such that $\gamma \in z_k$, thus $W(AN) \geqq W_k$, otherwise $W(AN) = W(\alpha(2^{e(A)} - 1)) \geqq e(A)$[5]. Clearly $e(A)$ is greater than each of W_k's. (QED) From the definition of W_k, we get

Corollary 1. If B is a prime and has 2 as a primitive root,

$$D_A = \frac{B-1}{2}$$

where $e(A) = e(B)$.

7. Conclusion

Theorem 1 and 2 derived here are most useful for the evaluation of distances of arithmetic codes. Especially, when B is a prime and $e(A) = e(B)$, the equivalence classes z_k's are the residue classes of the Abelian group R_0 derived by a subgroup $H = \{J \mid J \equiv 2^j \pmod{B},\ j = 0, 1, \ldots, e(B)\}$. We are also interested in that d_A relates to the number theoretic property of R_0.

Acknowledgement

The authers would like to acknowledge the continuing guidance and encouragement of Pro. N. Honda and Pro. K. Ikegaya and also to express their gratitude to many colleagues for frequent, stimulating, and helpful discussion.

References

1) D. Mandelbaum, "Arithmetic codes with large distance," IEEE Trans. on Information Theory, vol. IT-13, p. 238, April 1967.
2) G. W. Reitwiesner, "Binary arithmetic", in Advances in Computers, vol. 1, F. L. Alt, ed., New York, Academic Press, 1960.
3) K. Udagawa and M. Goto, "Generalized minimal representation of binary numbers and its application to a non-restoring binary division", Electronics and Commun. in Japan, vol. 48, p. 79, May 1965.
4) W. W. Peterson, "Error-Correcting Codes", New York, Wiley, 1961, ch. 13, p. 240.
5) D. Mandelbaum, "Arithmetic error detection codes for communications links involving computers". IEEE Trans. on COM, vol. COM-13, p. 165, Jun. 1965.
6) J. L. Massey, "Survey of residue coding for arithmetic errors", in Internat'l computation Center Bull. Rome, Italy: UNESCO, Sept. 1964.