CYBERNETICS AND CONTROL THEORY

R. R. Varshamov

1965

SovietRxiv

Abstract

Full Text

CYBERNETICS AND CONTROL THEORY

R. R. Varshamov

ON THE THEORY OF ASYMMETRIC CODES

(Presented by Academician V. S. Kulebakin, 15 January 1965)

Recently, attention has been attracted by the problem of synthesizing codes used for transmission over an asymmetric channel ([1,2]). In coding theory it is important to obtain upper and lower bounds for the maximum possible number of signals in error-correcting codes, since these estimates can serve as a good basis for conclusions about the effectiveness of individual codes. Nevertheless, in the theory of asymmetric codes nothing is known about these important characteristics of correcting codes. Their mathematical structure has also been poorly studied, knowledge of which greatly facilitates the study of their properties and regularities. Meanwhile, the theory of asymmetric codes has a great future. It finds broad application in various areas of technology, for example, in the creation of reliable information devices whose elements are distinguished by great complexity.

The purpose of the present note is to reveal, as far as possible, the mathematical structure and characteristic features of asymmetric codes, and also to establish lower and upper bounds for the maximum possible number of signals in codes correcting asymmetric errors.

The set $G_{n,q}$ of q^n sequences of the form $x = (x_1, x_2, \dots, x_n)$, where each symbol x_i can take a value from the Galois field $GF(q)$, is naturally regarded as a linear algebra of rank n over the field $GF(q)$, if the product of the elements x and y is defined as follows: $x \cdot y = (x_1 y_1, x_2 y_2, \dots, x_{ny} n)$.

In the case of nonbinary codes, in principle several different definitions of errors are possible; for example, unrestricted errors, which change symbols to any other value, or small errors with a limited range of distortion. In the case of an unrestricted error, we define the distance between signals x and y as the number of symbols in them that do not coincide, $\rho(x, y) = |x - y|$, where $|x|$ is the total number of nonzero coordinates of the vector x.

Mathematically, by a distortion of a signal x is meant the usual addition of it with some n-dimensional vector $\varepsilon_x \in G_{n,q}$. However, in the case of asymmetric errors, not every symbol in the signal is subject to distortion. Consequently, not every element of the algebra $G_{n,q}$ can be chosen as an admissible interference. Clearly, the vector ε_x is chosen from some subset $E_x \subset G_{n,q}$. Let Ω be the set

of ω of those, and only those, elements of the field $GF(q)$ that are subject to channel distortions. Errors of this kind we shall call Ω-asymmetric.

It is of interest to investigate a correcting code resistant to a prescribed number r of $\Omega(q)$-asymmetric errors, i.e., a code correcting any r or fewer $\Omega(\theta)$ $(=\Omega+\theta)$-asymmetric errors, where θ may take any value from the field $GF(q)$. As is easy to understand, for any admissible interference ε_x under $\Omega(\theta)$-asymmetric distortions the equality $x^{(\theta)}\varepsilon_x = \varepsilon_x$ holds, which permits one to conclude that

$$E_x = x^{(\theta)} G_{n,q}, \tag{1}$$

where each sign $x_i^{(\theta)}$ of the vector $x^{(\theta)}$ is equal to one if $x_i \in \Omega(\theta)$, or is equal to zero if $x_i \notin \Omega(\theta)$. Let $\bar{\rho}(x,y) = |x-y| + ||x|-|y||$. We introduce the functions

$$\alpha_\theta(x,y) = \bar{x}^{(\theta)} y^{(\theta)}(x-y), \quad \beta_\theta(x,y) = x^{(\theta)} y^{(\theta)}(x-y), \quad \gamma_\theta(x,y) = \bar{x}^{(\theta)} \bar{y}^{(\theta)}(x-y),$$

taking $\bar{x}^{(\theta)} = e - x^{(\theta)}$ $(|e| = n,\ e^{q-1} = e)$, and note two interesting facts:

$$\bar{\rho}(x^{(\theta)}, y^{(\theta)}) = 2\max(|\alpha_\theta(x,y)|,\ |\alpha_\theta(y,x)|); \tag{2}$$

$$\rho(x,y) = |\alpha_\theta(x,y)| + |\alpha_\theta(y,x)| + |\beta_\theta(x,y)| + |\gamma_\theta(x,y)|. \tag{3}$$

The first expression follows from the relation

$$|\alpha_\theta(x,y)| = |y^{(\theta)}| - |x^{(\theta)} y^{(\theta)}|,$$

taking into account that

$$|x^{(\theta)} - y^{(\theta)}| = |x^{(\theta)}| + |y^{(\theta)}| - 2|x^{(\theta)} y^{(\theta)}|.$$

The second is obtained from the equality

$$x - y = \alpha_\theta(x,y) - \alpha_\theta(y,x) + \beta_\theta(x,y) + \gamma_\theta(x,y), \tag{4}$$

taking into account that the vectors $\alpha_\theta(x,y)$, $\alpha_\theta(y,x)$, $\beta_\theta(x,y)$, and $\gamma_\theta(x,y)$ are pairwise orthogonal.

We shall take the quantity

$$\rho_\theta(x,y) = \max(\rho(x,y), \bar{\rho}(x^{(\theta)}, y^{(\theta)}))$$

as the $\Omega(\theta)$-distance between the signals x and y.

Definition. Two distinct vectors x and y satisfying the condition

$$\gamma_\theta(x,y) = \bar{0} \quad (|\bar{0}| = 0)$$

will be called $\Omega(\theta)$-mutually representable.

Lemma. *For it to be possible to correct r $\Omega(q)$-asymmetric errors, it is necessary and sufficient that the $\Omega(\theta)$-distance between any two used $\Omega(\theta)$-mutually representable signals x, y be not less than $d = 2r + 1$, i.e.*

$$\rho_\theta(x, y) \geq 2r + 1.$$

Proof. Necessity. Suppose the set of signals under consideration is resistant to r $\Omega(q)$-asymmetric errors. In the case $\gamma_\theta(x, y) = \bar{0}$ and $\rho_\theta(x, y) \leq 2r$ for at least one distinct pair of signals x and y, in view of $\rho'(x, y) \leq 2r$, $\bar{\rho}(x^{(\theta)}, y^{(\theta)}) \leq 2r$, and according to (1), (2), (3), it is clear that, as their admissible errors ε_x and ε_y, the vectors

$$\alpha_\theta(y, x) + e_1\beta_\theta(x, y), \quad \alpha_\theta(x, y) + e_2\beta_\theta(x, y) \quad (-e_1 + e_2 = e),$$

may serve, with

$$r - |\alpha_\theta(y, x)| \leq |e_1\beta_\theta(x, y)|, \quad r - |\alpha_\theta(x, y)| \leq e_2\beta_\theta(x, y).$$

Putting

$$\varepsilon_x = \alpha_\theta(y, x) + e_1\beta_\theta(x, y), \quad \varepsilon_y = \alpha_\theta(x, y) + e_2\beta_\theta(x, y),$$

by (4) we shall have the relation

$$x + \varepsilon_x = y + \varepsilon_y,$$

which contradicts the fundamental principle

$$(x + \varepsilon_x \neq y + \varepsilon_y),$$

characterizing error-correcting codes.

Sufficiency. If $\gamma_\theta(x, y) = \bar{0}$, then, according to (4) and the evident identity

$$\gamma_\theta(x, y)x^{(\theta)} = \gamma_\theta(x, y)y^{(\theta)} = \bar{0},$$

we shall have the expression

$$x - y \in x^{(\theta)}G_{n,q} + y^{(\theta)}G_{n,q},$$

equivalent to the basic relation for correcting codes $(x + \varepsilon_x = y + \varepsilon_y)$. If $\rho_\theta(x, y) \geq 2r + 1$, then at least one of the two inequalities

$$\rho(x, y) \geq 2r + 1$$

or

$$\rho(x^{(\theta)}, y^{(\theta)}) \geq 2r + 1$$

holds.

In the case $\rho(x, y) \geq 2r + 1$, in view of

$$\tau(x, y) = \max(|\varepsilon_x|, |\varepsilon_y|) \leq r,$$

it is clear that

$$x + \varepsilon_x \neq y + \varepsilon_y.$$

If $\rho(x^{(\theta)}, y^{(\theta)}) \geq 2r + 1$, then, arguing by contradiction, putting $x + \varepsilon_x = y + \varepsilon_y$, by virtue of (4) we shall have

$$\varepsilon_x = \alpha_\theta(y, x) + c_1(x, y), \quad \varepsilon_y = \alpha_\theta(x, y) + c_2(x, y),$$

$$c_1(x, y)\alpha_\theta(y, x) = c_2(x, y)\alpha_\theta(x, y) = \bar{0}.$$

Therefore

$$|\varepsilon_x| \geq |\alpha_\theta(y, x)|, \quad |\varepsilon_y| \geq |\alpha_\theta(x, y)|,$$

which, together with (2), gives the contradictory expression

$$\tau(x, y) \geq r + 1.$$

Thus, the lemma is fully proved.

As a particular case of the proved lemma, one can obtain some earlier known results. Thus, for example, putting $\omega = q$ (the case of a symmetric channel), in view of

$$x^{(\theta)} \equiv e, \quad \gamma_\theta(x, y) \equiv \bar{0}, \quad \rho(x^{(\theta)}, y^{(\theta)}) \equiv 0$$

and

$$\rho_\theta(x, y) = \rho(x, y),$$

we obtain the well-known proposition of the theory of correcting codes:

For the possible correction of r symmetric errors, it is necessary and sufficient that the code distance of the set of used signals be not less than

$$d = 2r + 1.$$

In the case $\omega = q - 1$ we shall have $x^{(\theta)} = (x - \lambda e)^\omega$, where $\lambda = GF(q) \setminus \Omega(\theta)$, and since $\gamma_\theta(x, y) = 0$, for $q = 2$, in view of $\rho(x^{(\theta)}, y^{(\theta)}) \geq \rho(x, y)$, we obtain, according to the lemma, the known result ([2]):

In order that a binary code correct r asymmetric errors*, it is necessary and sufficient that, for any distinct pair of signals used x and y, the condition $\rho(x, y) \geq 2r + 1$ be satisfied.

Theorem 1. *The number of signals used by any code correcting r $\Omega(q)$-asymmetric errors is at most*

$$q^{(n+1)} \left(\sum_{\nu=0}^{q-1} S^{(r)}_{[(n\omega+\nu)q^{-1}],q} \right)^{-1}, \tag{5}$$

where

$$S_{a,q}^{b} = \sum_{i=0}^{b} (q-1)^i C_a^i, \qquad C_a^b = \frac{a!}{b!(a-b)!}, \qquad C_a^b = 0 \quad (b > a).$$

Proof. Let a code G be given that is resistant to r $\Omega(q)$-asymmetric errors. Then, as is easy to see, for any element $\theta \in GF(q)$ we shall have

$$\sum_{y \in G} S_{|x^{(\theta)}|,q}^{r} \leq q^n.$$

Hence, putting

$$\sigma(x, \theta) = \sum_{g \in \Omega(\theta)} \Delta(x, g),$$

where $\Delta(x, g)$ is the total number of coordinates of the vector x corresponding to the value g, we obtain

$$\sum_{x \in G} \sum_{\theta \in GF(q)} S_{\sigma(x,\theta),q}^{r} \leq q^{n+1}. \tag{6}$$

We shall now show that, whatever the sequence of integers $\Delta(g) \geq 0^{**}$ ($g \in GF(q)$),

$$\sum_{\theta \in GF(q)} S_{\sigma(\theta),q}^{r} \geq \sum_{\nu=0}^{q-1} S_{[(n\omega+\nu)q^{-1}],q}^{r}, \tag{7}$$

where

$$n = \sum_{g \in GF(q)} \Delta(g)$$

and

$$\sigma(\theta) = \sum_{g \in \Omega(\theta)} \Delta(g),$$

assuming $C_a^0 = 1$, $C_a^b = 0$ ($a \leq 0$, $b < a$). For this, it is evidently sufficient to prove the validity of the inequality

$$I_m(\Delta) = \sum_{\theta \in GF(q)} C^m_{\sigma(\theta)} \geq I^{(q-1)}_m(n) \tag{8}$$

for any $m \geq 0$, where

$$I^{(k)}_m(n) = \sum_{\nu=0}^{k} C^m_{[(n\omega+\nu)q^{-1}]}.$$

For $n < (1-q)/\omega$ the validity of inequality (8) is obvious, since $I^{(q-1)}_0(n) = q$, $I^{(q-1)}_m(n) = 0$ $(m \neq 0)$. Suppose that it has already been proved for all numbers $< n$ $(n > (1-q)/\omega)$. Putting

$$I_u(\Delta, h) = \sum_{\theta \in \Omega(h)} C^u_{\sigma(\theta)} \left(\sum_{g \in GF(q)} \Delta(g) = n - 1 \right),$$

we note that among the elements of $GF(q)$ one can find at least one element ξ satisfying the condition

$$I_u(\Delta, \xi) \geq I^{(\omega)}_u(n-1) \qquad (u \geq 0). \tag{9}$$

Indeed, otherwise

$$\sum_{g \in GF(q)} I_u(\Delta, g) = \omega I_u(\Delta) < q I^{(\omega)}_u(n-1),$$

and since

$$q I^{(\omega)}_u(n-1) \leq \omega I^{(q-1)}_u(n-1),$$

then $I_u(\Delta) < I^{(q-1)}_u(n-1)$, which would contradict our assumption.

Now let $\Delta'(g) = \Delta(g)$ $(g \in GF(q) \setminus \xi)$ and $\Delta'(\xi) = \Delta(\xi) + 1$. But then

$$\sum_{g \in GF(q)} \Delta'(g) = n,$$

and, by virtue of $C^b_a = C^b_{a-1} + C^{b-1}_{a-1}$, we shall have $I_m(\Delta') =$

* Distortions of the form $(1 \to 0)$ or $(0 \to 1)$.

** For convenience, abstracting from the physical side.

$$= I_m(\Delta) + I_{m-1}(\Delta, \xi) \quad \text{and} \quad I_m^{(q-1)}(n) = I_m^{(q-1)}(n-1) + I_{m-1}^{(\omega)}(n-1).$$

And this, in turn, in view of (9), gives the relation of interest to us

$$I_m(\Delta') \geq I_m^{(q-1)}(n).$$

Thus, inequality (8), and consequently also inequality (7), hold for any integer n. From (6) and (7), in view of $\sigma(x, \theta) \geq 0$, (5) follows automatically. The theorem is proved.

In the case of symmetric errors ($\omega = q$), Theorem 1, since

$$[(n\omega + \nu)q^{-1}] = n \qquad (0 \leq \nu \leq q-1),$$

gives the Hamming upper bound [3]. If we put $q = 2,\ \omega = 1$, we obtain

$$M \leq 2^{n+1}\left(S^r_{[n/2],2} + S^r_{[(n+1)/2],2}\right)^{-1},$$

where M is the maximum possible number of signals of a binary code with correction of r asymmetric errors—a result obtained earlier [4].

Theorem 2. *The number of working signals in the best code correcting r $\Omega(\theta)$-asymmetric errors is equal to the very least*

$$q^n \left(\sum_{t=1}^{rsg(q-\omega)} L_{n,\omega}(t) + \sum_{v=0}^{r} \sum_{u=0}^{2r-v} (q-\omega)^v (\omega-1)^u C_n^{v+u} \right)^{-1},$$

where

$$L_{n,\omega}(t) = \sum_{v=0}^{r} \sum_{u=0}^{2r-v-t} \omega^t (q-\omega)^v (\omega-1)^u C^t_{n-[(n+1)(v+u)/(v+u+t)]} C^{v+u}_{[(n+1)(v+u)/(v+u+t)]}$$

and, according to the definition, $0^0 = 1$.

Proof. Let a code G be given with the maximum possible number N of code words, correcting r $\Omega(\theta)$-asymmetric errors. Then, according to the lemma, we shall have

$$\sum_{x \in G} \sum_{\substack{v+u+t \leq 2r \\ v \leq r,\ t \leq r}} \omega^t (q-\omega)^v (\omega-1)^u C^t_{n-|x^{(\theta)}|} C^{v+u}_{|x^{(\theta)}|} \geq q^n. \tag{10}$$

Meanwhile, one can show that for any positive k and nonnegative integers $l,\ p \leq n$ the inequality

$$C_{n-p}^{k} C_{p}^{l} \leq C_{n-[(n+1)l/(k+l)]}^{k} C_{[(n+1)l/(k+l)]}^{l}$$

holds. Therefore, for any $x \in G$, putting $p = |x^{(\theta)}|,\ k = t,\ l = v+u$ and taking into account that for $\omega = q$, $x^{(\theta)} \equiv e$, we may write the relation

$$\sum_{\substack{v+u+t \leq 2r \\ v \leq r,\ t \leq r}} \omega^{t} (q-\omega)^{v} (\omega-1)^{u} C_{n-|x^{(\theta)}|}^{t} C_{|x^{(\theta)}|}^{v+u} \leq$$

$$\leq \sum_{t=1}^{rsg(q-\omega)} L_{n,\omega}(t) + \sum_{v=0}^{r} \sum_{u=0}^{2r-v} (q-\omega)^{v} (\omega-1)^{u} C_{n}^{v+u},$$

or, by virtue of (10),

$$N \geq q^{n} \left(\sum_{t=1}^{rsg(q-\omega)} L_{n,\omega}(t) + \sum_{v=0}^{r} \sum_{u=0}^{2r-v} (q-\omega)^{v} (\omega-1)^{u} C_{n}^{v+u} \right)^{-1},$$

which is the content of Theorem 2. In the case $q = \omega$, from Theorem 2, in view of $sg(q-\omega) = 0$, it follows that

$$N \geq q^{n} \left(\sum_{u=0}^{2r} (q-1)^{u} C_{n}^{u} \right)^{-1},$$

i.e. the Hamming lower bound.

Institute of Automation
and Telemechanics

Received
14 I 1965

CITED LITERATURE

1. W. H. Kim, C. V. Freiman, *TRE Trans. Inform. Theory*, **5**, 2, 62 (1959).

2. R. R. Varshamov, *DAN*, **157**, No. 3 (1964).

3. W. W. Peterson, *Error-Correcting Codes*, N. Y.–London, 1961.

4. R. R. Varshamov, *Automation and Telemechanics*, **25**, No. 11 (1964).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.