A Study of Varshamov Codes for Asymmetric Channels

R. P. Stanley[1] and M. F. Yoder[2]
Communications Systems Research Section

An important class of single-error-correcting codes for binary and nonbinary discrete asymmetric channels recently discovered by Varshamov is studied. Among other things, a wide generalization of Varshamov's construction is given, and the complete weight distribution of Varshamov's codes is calculated.

I. Introduction

Recently Varshamov (Ref. 1) discovered an impressive class of single-error-correcting codes for the binary asymmetric, or "Z" channel. (The reason for the letter "Z" appears in Fig. 1.)

In the Z channel, a 0 is always transmitted reliably but 1 may be received as either 1 or 0. Actual physical channels, in particular the Ground Communications Facility (GCF), usually exhibit some degree of asymmetry, and so a study of the Z channel provides insight into the effects of asymmetry on practical data-processing systems.

It is the object of this paper to extend Varshamov's work in several directions. In Subsection II, Varshamov's codes will be introduced, and a larger class of single-error-correcting codes will be described that contain Varshamov codes as a proper subset. Estimates on the number of codewords in these codes will be obtained, and a general upper bound on the number of words in *any* single-error-correcting code for the Z channel will be obtained. In Subsection III, the exact number of codewords in each of Varshamov's codes will be calculated; indeed the complete weight distribution of each code will be found. (In fact in Subsection III we will consider q-ary, rather than binary, codes, where q is an arbitrary integer.)

[1]Consultant, University of California, Berkeley, Calif.
[2]Student at the California Institute of Technology, Pasadena, Calif.

II. A Generalization of the Binary Varshamov Codes

Varshamov's single-error correcting codes for the binary Z channel may be described as the set of all vectors $(e_1, e_2, \cdots, e_n)$ with $e_i = 0$ or 1, such that

$$\sum_{i=1}^{n} ie_i \equiv d \,(\mathrm{mod}\, n+1)$$

for a fixed d. There are, then, $n + 1$ distinct Varshamov codes of length n, one for each choice of d. A generalization of this construction that immediately suggests itself is the following: let G be an arbitrary group of order $n + 1$, and let $g_1, g_2, \cdots, g_n$ be an ordering of the nonidentity elements of g. For a fixed $d \in G$ consider the set of $\{0, 1\}$ vectors $(e_1, e_2, \cdots, e_n)$ such that

$$\prod_{i=1}^{n} g_i^{e_i} = d \tag{1}$$

Varshamov's codes are the special case where G is a cyclic group. Unfortunately in this generality the sets of $\{0,1\}$ vectors thus formed are not necessarily single-error correcting codes for the Z channel. We must restrict both the group G and the ordering of the elements of G, as in Theorem 1.

Theorem 1

Let G be a group of order $n+1$ such that every element commutes with all of its conjugates (e.g., if G is abelian or nilpotent of rank 2 this condition is satisfied). Let $g_1, g_2, \cdots, g_n$ be an ordering of the nonidentity elements of G with the property that the conjugacy classes appear serially; i.e., every conjugacy class appears as a set of consecutive elements $g_m, g_{m+1}, \cdots, g_{m+k}$ in the ordering. Then for every $d \in G$, the set of $\{0,1\}$ vectors $(e_1, e_2, \cdots, e_n)$ which satisfy Eq. (1) is a single-error correcting code for the Z channel.

Proof:

We first observe that no two vectors satisfying Eq. (1) differ in only one position; for if $(e_1, \cdots, e_i, \cdots, e_n)$ and $(e_1, \cdots, \bar{e}_i, \cdots, e_n)$ both satisfy Eq. (1), we would have

$$w_1 g_i w_2 = d = w_1 w_2$$

where

$$w_1 = g_i^{e_1} \cdots g_{i-1}^{e_{i-1}}, \qquad w_2 = g_{i+1}^{e_{i+1}} \cdots g_n^{e_n}$$

But then $g_i = 1$, a contradiction. Also, it is easy to prove that there cannot be two vectors such that a single error in one produces the same result as a single error in the other; for in such a case there would result an equation

$$w_1 g_i w_2 w_3 = d = w_1 w_2 g_j w_3$$

Then $g_j = w_2^{-1} g_i w_2$, and so w_2, being a product of elements that lie between the conjugates g_i and g_j commutes with g_i and so $g_i = g_j$, another contradiction.

Of course we would like to know the number of codewords in each of the codes constructed in Theorem 1. Unfortunately this is a very difficult problem, for which we have only partial solutions. If G is cyclic, i.e., for Varshamov's original codes, a complete solution will be given in Subsection III. In the general case, we make the observation that the 2^n $\{0,1\}$ vectors are distributed into $n+1$ codes and so at least one such code contains at least $2^n/(n+1)$ codewords. On the other hand, Hamming's bound says that a single-error-correcting code for the binary *symmetric* channel has at *most* $2^n/(n+1)$ codewords. Thus unless $n = 2^m - 1$ for some m the asymmetric channel will support a larger single-error-correcting code than the symmetric channel will. If $n = 2^m - 1$ for some m the assertion very probably remains true, but the codes of Theorem 1 cannot be used to demonstrate that fact, because of Theorem 2.

Theorem 2

If $|G| = 2^m$, then all of the codes defined in Theorem 1 contain 2^{2^m-1-m} codewords; i.e. no more than the Hamming codes of the same length.

Proof:

It is easier to prove the more general statement that if $(g_1, g_2, \cdots, g_r)$ is a sequence of elements of G such that every element with the possible exception of 1 appears at least once, then the number of $\{0,1\}$ vectors $(e_1, e_2, \cdots, e_r)$ such that

$$\prod_{i=1}^{r} g_i^{e_i} = d \tag{2}$$

is independent of d. To prove this fact we induct on m, the case $m = 1$ being easily treated. For $m > 1$, let $z \neq 1$ be an element of order 2 in the center of G. For convenience we assume $g_1 = z$. Then if ϕ is the homomorphism from $G \to G/\{z\}$, and if Eq. (2) holds, we have

$$\prod_{i=2}^{r} \phi(g_i)^{e_i} = \phi(d) \tag{3}$$

Furthermore, every element $\neq 1$ of $G/\{z\}$ occurs among the $\phi(g_i)$, and so by induction the number of vectors $(e_2, \cdots, e_r)$ which satisfy Eq. (3) is independent of d. But if Eq. (3) is satisfied it follows that

$$\prod_{i=2}^{r} g_i^{e_i} = d \quad \text{or} \quad dz$$

and so there is a unique choice of e_1 that forces Eq. (2) to hold.

Although Theorem 2 shows that the codes of Theorem 1 are unimpressive when $n = 2^m - 1$, there is good reason to believe that codes of these lengths do exist with more than $2^n/(n+1)$ codewords, if $m \geqq 3$. For $n = 7$, a code with 18 codewords exists:

0 0 0 0 0 0 0	1 1 0 0 1 0 0	0 1 0 1 1 0 1
0 1 0 0 0 1 0	1 0 1 0 0 1 0	0 1 1 0 1 1 0
0 0 1 0 0 0 1	1 0 0 1 0 0 1	1 1 0 0 1 1 1
0 0 0 1 1 0 0	1 1 1 0 0 0 1	1 0 1 1 1 0 1
0 1 1 1 0 0 0	1 0 0 1 1 1 0	1 1 1 1 0 1 0
0 0 0 0 1 1 1	0 0 1 1 0 1 1	1 1 1 1 1 1 1

It is in fact possible to show that no single-error-correcting code of length 7 for the Z channel can have 19 words. The above code was found ad hoc by hand calculation. A computer search might yield an $n = 15$ code with more than 2^{11} words, but a general construction is desirable.

We conclude this section with a general upper bound on the number M_n of codewords in a single-error-correcting code for the Z channel.

Theorem 3

$$M_n \leq B_{n+1}$$

where B_{n+1} is the maximum number of words possible for a single-error correcting code for the symmetric channel.

Proof:

There are two asymmetric binary channels: one that changes 0's to 1's, and one that changes 1's to 0's. It is an odd but easily checked fact that a code which corrects t errors on one of these channels will also correct t errors on the other. We use this fact to obtain the upper bound of Theorem 3.

For a given code of length n for the Z channel, construct a new code of length $n + 1$ by adding a "parity" bit to each codeword that is 0 if the weight of the codeword is congruent to 0 or 1 (mod 4) and is 1 if the weight is $\equiv 2$ or 3 (mod 4). Now this extended code will correct 1 error on the *symmetric* channel, since an error in the parity bit will be obvious (the first n bits will be a codeword from the original code, but the parity bit will not check), and if an error occurs elsewhere the parity bit will indicate whether it was a $0 \to 1$ or a $1 \to 0$ transition, and thus the error can be corrected. Thus $M_n \leq B_{n+1}$, the maximum number of words in a single-error-correcting code for the symmetric channel.

Corollary

$$M_n \leq \frac{2^{n+1}}{n+2}$$

Proof:

$$B_{n+1} \leq \frac{2^{n+1}}{n+2}$$

by Hamming's bond.

Remark:

It is very probable that Theorem 3 is quite weak, and that

$$M_n \leq \sim \frac{2^n}{n+1}$$

for large n.

III. The Weight Distribution of Varshamov's Codes

Let q, m, and d be natural numbers satisfying $q > 1$, $m > 1$, $1 \leq d \leq m$. Set $n = m - 1$. Let $C(q, n, d)$ be the set of all n-tuples (*vectors*) $e = (e_1, e_2, \cdots, e_n)$, where $e_i \in \{0, 1, 2, \cdots, q-1\}$, and

$$\sum_{1}^{n} ie_i \equiv d \pmod{m} \tag{4}$$

Then $C(q, n, d)$ is a single-error-correcting asymmetric code in the sense of Varshamov. Since multiplying Eq. (4) by a unit modulo m merely permutes the e_i's, from now on we assume without loss of generality that d divides m (written $d \mid m$).

If $e = (e_1, e_2, \cdots, e_n) \in C(q, n, d)$, define the weight $|e|$ by

$$|e| = e_1 + e_2 + \cdots + e_n \text{ (real addition)}$$

Although this definition of weight differs from the usual Hamming or Lee weights (except when $q = 2$), it is in accordance with Varshamov's usage.

Let $c_i = c_i(q, n, d)$ be the number of vectors in $C(q, n, d)$ of weight i, and define the *weight enumerator* $W(y) = W(q, n, d; y)$ by

$$W(y) = \sum_{i=0}^{\infty} c_i y^i$$

$W(y)$ is actually a polynomial since $c_i = 0$ for $i > (q-1)n$. Finally let $c = c(q, n, d) = |C(q, n, d)|$, so $c = \sum c_i = W(1)$. Our object is to obtain an expression for $W(y)$ and for $c(q, n, d)$.

Theorem 4

We have

$$W(q, n, d; y) = \frac{1-y}{m(1-y^q)} \sum_{f \mid d} f \sum_{g \mid \frac{m}{f}} \mu(g) \frac{(1 - y^{fgq/(fg,q)})^{m(fg,q)/fg}}{(1 - y^{fg})^{m/fg}}$$

where (fg, q) is the g.c.d. of fg and q.

Before proving Theorem 4, we first discuss some consequences.

Corollary 1

Let k be the largest factor of m relatively prime to q. Then

$$c(q, n, d) = \frac{1}{m} \sum_{f|(k,d)} f \sum_{g|\frac{k}{f}} \mu(g)\, q^{(m/fg)-1}$$

Proof of Corollary 1:

Set $y = 1$ in Theorem 4. For a given choice of f and g, the factor $1 - y$ appears $1 + m(fg, q)/fg$ times in the numerator and $1 + m/fg$ times in the denominator. Hence the term corresponding to f, g will be 0 unless $(fg, q) = 1$. Hence we may assume $f|(k, d)$ and

$$g \left| \frac{k}{f} \right.$$

so

$$c(q, n, d) = \frac{1 - y}{m(1 - y^q)} \times \sum_{f|(k,d)} f \sum_{g|\frac{k}{f}} \mu(g) \left[\frac{(1 - y^{fgq})^{m/fg}}{(1 - y^{fg})^{m/fg}}\right]_{y=1}$$

$$= \frac{1}{m} \sum_{f|(k,d)} f \sum_{g|\frac{k}{f}} \mu(g)\, q^{(m/fg)-1}$$

Remark:

Let $M(q, r)$ be the number of q-symbol "necklaces" with r beads and with no symmetry. As is well-known

$$M(q, r) = \frac{1}{r} \sum_{d|r} \mu(d)\, q^{r/d}$$

Hence, by Corollary 1,

$$c(q, n, d) = \frac{k}{mq} \sum_{f|(k,d)} M\left(q^{m/k}, \frac{k}{f}\right)$$

where $M\left(q^{m/k}, \frac{k}{f}\right) > 0$. Hence we have:

Corollary 2

If $e|d|m$, then

$$c(q, n, d) \geqq c(q, n, e)$$

with equality if and only if every prime dividing d/e also divides q. In particular, $c(q, n, d)$ is maximized (for fixed q, n) at precisely those $d|m$ such that every prime divisor of m/d also divides q, and therefore for $d = m$.

Corollary 3

For fixed q, n, we have

$$\max_{d|m} c(q, n, d) = c(q, n, m) = \frac{1}{m} \sum_{h|k} \phi(h)\, q^{(m/h)-1}$$

where k as usual is the largest factor of m relatively prime to q.

Proof of Corollary 3:

By Corollary 2, $\max c(q, n, d) = c(q, n, m)$. By Corollary 1,

$$c(q, n, m) = \frac{1}{m} \sum_{f|k} f \sum_{g|\frac{k}{f}} \mu(g)\, q^{(m/pg)-1}$$

$$= \frac{1}{m} \sum_{h|k} q^{(m/h)-1} \sum_{f|h} f \mu\left(\frac{h}{f}\right) \qquad (h = fg)$$

But

$$\sum_{f|h} f \mu\left(\frac{h}{f}\right) = \phi(h)$$

so the proof follows.

Remark 1:

The number $N(t, r)$ of inequivalent t-symbol necklaces with r beads is

$$\frac{1}{r} \sum_{h|r} \phi(h)\, t^{r/h}$$

Hence

$$c(q, n, m) = \frac{k}{qm} N(q^{m/k}, k)$$

This suggests that a combinatorial proof of Corollary 3 may be possible, especially in the case $(m, q) = 1$ (so $k = m$), but we have been unable to find one. More generally, if $(m, q) = 1$ and n_i is the number of q-symbol necklaces with m beads summing to i (where the symbols are $0, 1, \cdots, q-1$), then it follows from Theorem 4 that

$$n_i = c_i + c_{i-1} + \cdots + c_{i-q+1}$$

since

$$\sum n_i y^i = \frac{1}{m} \sum_{f|m} \phi(f) (1 + y^f + y^{2f} + \cdots + y^{(q-1)f})^{m/f}$$

This suggests that with each word $e \in C(q, n, m)$ of weight i, one can associate a q-symbol necklace with m beads of weight $i + j$ for each $j = 0, 1, \cdots, q-1$, but we have been unable to find such a correspondence.

Remark 2:

The Hamming bound for *symmetric* q-ary single-error correcting codes of length $n = m - 1$ is q^{m-1}/m. Hence by Corollary 3, Varshamov's code in the optimum case $d = m$ does better than any symmetric code as long as m has a prime divisor not dividing q. As remarked in Subsection II, we have been unable to do better than the Hamming bound when every prime divisor of m divides q, except for the special cases for $q = 2$ listed there. The largest code has 18 elements (though $c(2, 7, 8) = 16$). There is also a 12-element binary code of length 6 and a 32-element binary code of length 8, both exceeding the cardinalities given by Corollary 3 of $c(2, 6, 7) = 10$ and $c(2, 8, 9) = 30$. These codes are:

$n = 6$:

0 0 0 0 0 0	0 1 1 0 0 1
1 1 0 0 0 0	0 1 0 1 1 0
0 0 1 1 0 0	1 1 1 1 0 0
0 0 0 0 1 1	1 1 0 0 1 1
1 0 1 0 1 0	0 0 1 1 1 1
1 0 0 1 0 1	1 1 1 1 1 1

$n = 8$:

0 0 0 0 0 0 0 0	1 0 1 0 1 0 0 0
1 1 0 0 0 0 0 0	1 0 0 1 0 0 1 0
0 0 1 1 0 0 0 0	1 0 0 0 0 1 0 1
0 0 0 0 1 1 0 0	0 1 1 0 0 0 0 1
0 0 0 0 0 0 1 1	0 1 0 1 0 1 0 0
1 1 1 1 0 0 0 0	0 1 0 0 1 0 1 0
1 1 0 0 1 1 0 0	0 0 1 0 0 1 1 0
1 1 0 0 0 0 1 1	0 0 0 1 1 0 0 1
0 0 1 1 1 1 0 0	0 1 0 1 0 1 1 1

0 0 1 1 0 0 1 1	0 1 1 0 1 1 0 1
0 0 0 0 1 1 1 1	0 1 1 1 1 0 1 0
1 1 1 1 1 1 0 0	1 0 0 1 1 1 1 0
1 1 1 1 0 0 1 1	1 0 1 0 1 0 1 1
1 1 0 0 1 1 1 1	1 0 1 1 0 1 0 1
0 0 1 1 1 1 1 1	1 1 0 1 1 0 0 1
1 1 1 1 1 1 1 1	1 1 1 0 0 1 1 0

Proof of Theorem 4:

Set

$$\begin{aligned} F(z) &= (1 + yz + y^2z^2 + \cdots + y^{q-1}z^{q-1}) \\ &\quad \times (1 + yz^2 + y^2z^4 + \cdots + y^{q-1}z^{2(q-1)}) \\ &\quad \cdots (1 + yz^n + y^2z^{2n} + \cdots + y^{q-1}z^{n(q-1)}) \\ &= \prod_{i=1}^{n} (1 - \omega z^i y)(1 - \omega^2 z^i y) \cdots (1 - \omega^{q-1} z^i y) \end{aligned}$$

where ω is a primitive q-th root of 1. Let $G(z)$ be the unique polynomial in z of degree $< m$ such that

$$F(z) \equiv G(z) \,(\mathrm{mod}\, z^m - 1)$$

Then the coefficient of z^d in $G(z)$ is $W(g, n, d; y)$, since choosing a term $y^i x^{ij}$ from the i-th factor

$$1 + yz^i + y^2z^{2i} + \cdots + y^{q-1}z^{(q-1)i}$$

of $F(z)$ corresponds to choosing $e_i = j$ in Eq. (4).

Now $G(z)$ is the unique polynomial of degree $< m$ satisfying $F(\zeta) = G(\zeta)$ for every root ζ of $z^m - 1 = 0$ i.e., for every m-th root of unity ζ. We shall therefore now evaluate $F(\zeta)$. Suppose $e|m$ and ζ is a primitive e-th root of 1. Then

$$\begin{aligned} F(\zeta) &= \prod_{i=1}^{n} \prod_{j=1}^{q-1} (1 - \omega^j \zeta^i y) \\ &= \left[\prod_{j=1}^{q-1} (1 - \omega^i y)^{-1} \right] \cdot \prod_{j=1}^{q-1} \prod_{k=0}^{(m/e)-1} \prod_{i=1}^{e} (1 - \omega^i \zeta^{ke+i} y) \\ &= \frac{1 - y}{1 - y^q} \prod_{j=1}^{q-1} (1 - \omega^{je} y^e)^{m/e} \\ &= \frac{(1 - y)(1 - y^{eq/(e,q)})^{mq(e,q)/e}}{(1 - y^q)(1 - y^e)^{m/e}} \end{aligned}$$

since ω^e is a primitive $q/(e, q)$ root of 1.

We therefore have

$$G(z) = \sum_{e|m} \frac{(1 - y)(1 - y^{eq/(e,q)})^{m(e,q)/e}}{(1 - y^q)(1 - y^e)^{m/e}} G_e(z)$$

where

$$G_e(\zeta) = \begin{cases} 1, \text{ if } \zeta \text{ is a primitive } e\text{-th root of } 1 \\ 0, \text{ if } \zeta^m = 1 \text{ but } \zeta \text{ is not a primitive} \\ \quad e\text{-th root of } 1. \end{cases}$$

We claim

$$G_e(z) = \frac{1}{m}(z^m - 1)\sum \frac{\zeta}{z - \zeta}$$

where the sum is over all primitive e-th roots of 1. Let

$$H(z) = \frac{z^m - 1}{z - \zeta_0}$$

(where ζ_0 is a primitive e-th root of 1). If $\zeta^m = 1$, $\zeta \neq \zeta_0$, then $H(\zeta) = 0$. Also $H(\zeta_0) = H'(\zeta_0) = m\,\zeta_0^{m-1} = m\,\zeta_0^{-1}$. Hence $G_e(\zeta_0) = (1/m)(m\,\zeta_0^{-1}\zeta_0) = 1$, while $G_e(\zeta) = 0$ if $\zeta^m = 1$ and ζ is not a primitive e-th root of 1. This proves the claim.

Summing a geometric series, we have

$$\frac{\zeta(z^m - 1)}{z - \zeta} = \frac{1 - z^m}{1 - \zeta^{-1}z}$$
$$= 1 + \zeta^{-1}z + \zeta^{-2}z^2 + \cdots + \zeta^{-m}z^m$$

Interchanging ζ with ζ^{-1} in the sum for $G_e(z)$ gives

$$G(z) = \sum_{e|m} \frac{1}{m}\frac{(1-y)}{(1-y^q)}\frac{(1 - y^{eq/(e,q)})^{m(e,q)/e}}{(1-y^e)^{m/e}}$$
$$\times \sum_{\substack{\zeta = \text{primitive} \\ e\text{-th root} \\ \text{of } 1}} (1 + \zeta z + \cdots + \zeta^n z^n)$$

Hence

$$W(q,n,d;y) = \text{coefficient of } z^d \text{ in } G(z)$$
$$= \frac{1}{m}\frac{(1-y)}{(1-y^q)}\sum_{e|m}\frac{(1 - y^{eq/(e,q)})^{m(e,q)/e}}{(1-y^e)^{m/e}}$$
$$\times \sum_{\substack{\zeta = \text{primitive} \\ e\text{-th root} \\ \text{of } 1}} \zeta^d$$

It is well-known that $\sum \zeta = \mu(e)$, where the sum ranges over all primitive e-th roots of 1. Now ζ^d is a primitive $e/(d,e)$ root of 1, so

$$\sum_{\substack{\zeta = \text{primitive} \\ e\text{-th root} \\ \text{of } 1}} \zeta^d = \mu\frac{e}{(d,e)}\frac{\phi(e)}{\phi\dfrac{e}{(d,e)}}$$

Hence

$$W(q,n,d;y) = \frac{1}{m}\frac{(1-y)}{(1-y^q)}\sum_{e|m}\frac{(1 - y^{eq/(e,q)})^{m(e,q)/e}}{(1-y^e)^{m/e}}$$
$$\times \mu\frac{e}{(d,e)}\frac{\phi(e)}{\phi\dfrac{e}{(d,e)}}$$

To complete the proof we need the following result:

Lemma: (Brauer-Rademacher):

For all positive integers e, d,

$$\sum_{f|(e,d)} f\,\mu\left(\frac{e}{f}\right) = \mu\left(\frac{e}{(e,d)}\right)\frac{\phi(e)}{\phi(e/(e,d))}$$

Proof:

See Ref. 2.

References

1. Varshamov, R. R., "A Class of Codes for Asymmetric Channels and a Problem from the additive Theory of Numbers," *IEEE Transactions on Information Theory*, IT-19, pp. 92–95, Jan. 1973.
2. Subbarao, M. V., "The Brauer-Rademacher Identity," *Amer. Math. Monthly*, Vol. 72, pp. 135–138, 1965.

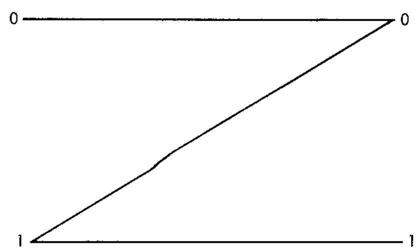

Fig. 1. The Z channel