The largest number of weights in cyclic codes

Minjia Shi, Patrick Solé

School of Mathematical Sciences, Anhui University

Shanghai, 2018

Outline

- Motivation
- Definitions and Notation
- Upper bound
- Lower bound
- Asymptotic
- Main results

Motivation

The largest number of nonzero weights **a linear code** of given length and dimension has been studied [1].

We are now ready to address the same type of questions for **cyclic codes**. Thus, we study the function $\Gamma(k, q)$, the largest number of nonzero weights a cyclic code of dimension k over $\mathbb{F}_q$ can have.

[1] M. Shi, H. Zhu, P. Solé, G.D.Cohen, How many weights can a linear code have? , *Des. Codes and Crypt.* To appear (2018).

Definitions and Notation

2.1 Linear codes

A **(linear) code** C of length n over a finite field $\mathbb{F}_q$ is a $\mathbb{F}_q$ vector subspace of $\mathbb{F}_q^n$. The dimension of the code is its dimension as a $\mathbb{F}_q$ vector space, and is denoted by k. The elements of C are called **codewords**.

The dual $C^\perp$ of a linear code C is understood w.r.t. the standard inner product. The minimum nonzero weight d of a linear code is called the minimum distance.

Every linear code satisfies the Singleton bound on its parameters $d \leq n - k + 1$. A code meeting that bound is called MDS. See [2] for general knowledge on this family of codes.

[2] F.J. MacWilliams, N.J.A. Sloane, The theory of error correcting code, *North Holland*, Amsterdam (1977).

2.2 Cyclic codes

A **cyclic** code of length n over a finite field $\mathbb{F}_q$ is a $\mathbb{F}_q$ linear code of length n invariant under the coordinate shift. Under the polynomial correspondence a cyclic code of length n can be regarded as an ideal in the ring $\mathbb{F}_q[x]/(x^n - 1)$.

It can be shown that this ideal is principal, with a unique monic generator $g(x)$, called **the generator polynomial** of the code. The **check polynomial** $h(x)$ is then defined as the quotient $(x^n - 1)/g(x)$.

A well-known fact is that the codewords are the periods of the linear recurrence of characteristic polynomial the reciprocal polynomial of $h(x)$. Thus any codeword c can be continued into an infinite periodic sequence $\widehat{c}$ which is periodic of period n.

The **period** of a codeword c is understood to be the smallest integer T such $\widehat{c}_{i+T} = \widehat{c}_i$ for all integers i. Thus the period is always a divisor of n.

A cyclic code is **irreducible** over $\mathbb{F}_q$ if its check polynomial $h(x)$ is irreducible over $\mathbb{F}_q[x]$.

The **period** of a polynomial $h(x) \in \mathbb{F}_q[x]$, such that $h(0) \neq 0$ is the smallest integer T such that $h(x)$ divides $x^T - 1$ over $\mathbb{F}_q[x]$.

If C is a cyclic code, its codewords are partitioned into orbits under the action of the shift. We call these orbits the **cyclic classes** of C.

2.3 Combinatorial functions

Define $\Gamma(k, q)$ as the largest number of nonzero weights of a cyclic code of dimension k over $\mathbb{F}_q$.

Define $\Gamma(n, k, q)$ as the largest number of nonzero weights of a cyclic code of length n and dimension k over $\mathbb{F}_q$, if such a code exists, and by zero otherwise.

The same functions for **strongly cyclic codes** (to be defined below) are denoted by $\Gamma^0(k, q)$, and $\Gamma^0(n, k, q)$, respectively.

3.1 Cyclic Structure

If C is a cyclic code, denote by B_t the number of nonzero codewords of period t it contains. A cyclic code such that $B_t = 0$ for $1 \leq t < n$ shall be called **strongly cyclic**.

Lemma 1

If C is an $[n, k]_q$ cyclic code with s nonzero weights, then

$$s \leq \sum_{t|n} \frac{B_t}{t}.$$

Proof The number of cyclic classes of codewords of period t is at most $\frac{B_t}{t}$. All codewords in the same class share the same weight. ■

Example

Consider the code of dimension 2 over $\mathbb{F}_5$, with length 20 and check polynomial $x^2 + x - 1$. This code contains the Fibonacci numbers mod 5 [3][A082116]. It can be checked to contain 4 codewords of period 4 (namely $1, 3, 4, 2$, repeated five times) and 20 codewords of period 20. Thus, it is a two-weight code satisfying $\frac{B_4}{4} = \frac{B_{20}}{20} = 1$. The bound of Lemma 1 is met with equality.

[3] Online Encyclopedia of Integer Sequences www.oeis.org

{(0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0), (1 0 1 4 2 2 0 2 3 4 4 0 4 1 3 3 0 3 2 1), (2 0 2 3 4 4 0 4 1 3 3 0 3 2 1 1 0 1 4 2), (3 0 3 2 1 1 0 1 4 2 2 0 2 3 4 4 0 4 1 3), (4 0 4 1 3 3 0 3 2 1 1 0 1 4 2 2 0 2 3 4), (4 1 3 3 0 3 2 1 1 0 1 4 2 2 0 2 3 4 4 0), (0 1 4 2 2 0 2 3 4 4 0 4 1 3 3 0 3 2 1 1), (1 1 0 1 4 2 2 0 2 3 4 4 0 4 1 3 3 0 3 2), (2 1 1 0 1 4 2 2 0 2 3 4 4 0 4 1 3 3 0 3), (3 1 2 4 3 1 2 4 3 1 2 4 3 1 2 4 3 1 2 4), (3 2 1 1 0 1 4 2 2 0 2 3 4 4 0 4 1 3 3 0), (4 2 2 0 2 3 4 4 0 4 1 3 3 0 3 2 1 1 0 1), (0 2 3 4 4 0 4 1 3 3 0 3 2 1 1 0 1 4 2 2), (1 2 4 3 1 2 4 3 1 2 4 3 1 2 4 3 1 2 4 3), (2 2 0 2 3 4 4 0 4 1 3 3 0 3 2 1 1 0 1 4), (2 3 4 4 0 4 1 3 3 0 3 2 1 1 0 1 4 2 2 0), (3 3 0 3 2 1 1 0 1 4 2 2 0 2 3 4 4 0 4 1), (4 3 1 2 4 3 1 2 4 3 1 2 4 3 1 2 4 3 1 2), (0 3 2 1 1 0 1 4 2 2 0 2 3 4 4 0 4 1 3 3), (1 3 3 0 3 2 1 1 0 1 4 2 2 0 2 3 4 4 0 4), (1 4 2 2 0 2 3 4 4 0 4 1 3 3 0 3 2 1 1 0), (2 4 3 1 2 4 3 1 2 4 3 1 2 4 3 1 2 4 3 1), (3 4 4 0 4 1 3 3 0 3 2 1 1 0 1 4 2 2 0 2), (4 4 0 4 1 3 3 0 3 2 1 1 0 1 4 2 2 0 2 3), (0 4 1 3 3 0 3 2 1 1 0 1 4 2 2 0 2 3 4 4)}

This simple counting lemma has two important applications. First, we improve the upper bound on $L(k,q) \leq \frac{q^k-1}{q-1}$ of [1, Prop. 2] by a factor $\frac{n}{q-1}$ for some large class of cyclic codes.

Theorem 2

If C is a $[n,k]_q$ strongly cyclic code with s nonzero weights, then

$$s \leq \frac{q^k-1}{n}.$$

Thus $\Gamma^0(n,k,q) \leq \frac{q^k-1}{n}$.

Proof We apply the lemma when $B_t = 0$ for $t < n$, so that the sum in the right handside contains only one summand. ■

Theorem 3

If C is an $[n,k]_q$ cyclic code with s nonzero weights, not containing the all-one codeword, then

$$s^2 \leq (q^k-1)^2 \sum_{1<t|n} \frac{1}{t^2}.$$

Proof. Note that, by hypothesis, $B_1 = 0$. Squaring the bound in the lemma, and applying Cauchy-Schwarz inequality we obtain

$$s^2 \leq \sum_{1<t|n} B_t^2 \sum_{1<t|n} \frac{1}{t^2}.$$

By definition of the B_t's note that $\sum_{t|n} B_t = q^k - 1$, implying $\sum_{t|n} B_t^2 \leq (q^k-1)^2$. The result follows. ■

Remark : Trivially, $s \leq q^k - 1$ for all linear codes, so we avoid $B_1 > 0$ and the summation on t to be > 1.

3.2 Character sums

The following result can be derived by using the character sums techniques of [4].

Theorem 4

If C is an $[n,k]_q$ strongly cyclic code with s weights, then $s \leq 2(1-\frac{1}{q})q^{k/2}$. Thus $\Gamma^0(n,k,q) \leq 2(1-\frac{1}{q})q^{k/2}$.

Proof By [4, Cor. 8.83] we know that the weights w of C lie in the range

$$|n(1-\frac{1}{q})-w| \leq (1-\frac{1}{q})q^{k/2}.$$

The result follows by computing the length of that interval. ■

[4] R. Lidl, H. Niederreiter, *Finite fields*, Encycl. of Math and Appl., vol. 20, Cambridge (1997).

3.3 Irreducible cyclic codes

The weight structure of irreducible cyclic codes has been a research topic since the first works of McEliece and others [4-6] due to their connection to Gauss sums and L-functions, and its intrinsic complexity.

[5] C. Ding, J. Yang, Hamming weights of irreducible cyclic codes, Discr. Math., **313** (4), (2013), 434–446.

[6] R.J. MacEliece, Irreducible Cyclic Codes and Gauss Sums, Combinatorica, **16**, (1975), 185–202.

Theorem 5

If C is an $[n = \frac{q^k-1}{N}, k]_q$ irreducible cyclic code with a check polynomial of period n, and s nonzero weights, then $s \leq N$.

Proof. Since the check polynomial $h(x)$ is irreducible it generates the annihilating ideal of each sequence attached to a codeword. If the period of such a sequence were $T < n$, then $h(x)$ would divide $x^T - 1$, contradicting the hypothesis on the period of $h(x)$. Hence C is strongly cyclic, and we can apply Theorem 2. The result follows. ■

Example : Consider the case of $N = 2$, and $q = p$ an odd prime. Such a code is well-known to be a two-weight code [6].

A slightly sharper bound can be derived using the results in [[5], Ding and Yang 2013], DM.

Theorem 6

If C is an $[n = \frac{q^k-1}{N}, k]_q$ irreducible cyclic code with a check polynomial of period n, and s nonzero weights then $s \leq N_k = GCD(N, \frac{q^k-1}{q-1})$.

Proof Follows by [5, (12)] which involves Gaussian periods of order N_k. ■

Theorem 7

If C is an $[n = \frac{q^k-1}{N}, k]_q$ irreducible cyclic code with a check polynomial of period n, and s nonzero weights then $s \leq 2(1 - \frac{1}{q})\sqrt{1 + nN}$.

Proof. As explained in the proof of Theorem 5 we know that all nonzero codewords have period n. Thus the code C is strongly cyclic, and we can apply Theorem 4. We get rid of k in Theorem 4 by writing $q^k = 1 + nN$. ■

Remark : Depending on the relative values of n and N, either Theorem 7, or Theorem 6 is sharper than the other.

A slight improvement on Theorem 7 can be derived for irreducible cyclic codes.

Theorem 8

If C is an $[n = \frac{q^k-1}{N}, k]_q$ irreducible cyclic code with a check polynomial of period n, and s nonzero weights then $s \leq 2(1-\frac{1}{q})(\frac{n}{h}-\frac{1}{N})\sqrt{1+nN}$ where $h = LCM(n, q-1)$.

Proof The proof follows the lines of Theorem 7 with [4, Th. 8.84, (8.37)] replacing [4, Cor. 8.83]. We get rid of k by writing $q^k = 1 + nN$. ■

4 Lower bounds

4.1 Special values

Proposition 1

For all prime powers q, we have $\Gamma(k,q) \geq k$.

Proof The **universe code**, the cyclic $[k,k]_q$ code with generator the zero polynomial, has k nonzero weights. This shows that $\Gamma(k,k,q) \geq k$. The result follows by $\Gamma(k,k,q) \leq \Gamma(k,q)$. ■

The following result is immediate by [[1], Shi, Zhu, Solé, 2018, DCC]. The proof is omitted.

Proposition 2

For all prime powers q, we have $\Gamma(2,q) = 2$.

The **repetition code** $R(n, q)$ is the ideal of $\mathbb{F}_q[x]/(x^n - 1)$ with generator $\frac{x^n-1}{x-1}$. Its dual is $P(n, q) = \langle (x - 1) \rangle$.

The **Hamming code** $\mathcal{H}_m$ is the binary cyclic code of length $n = 2^m - 1$ with generator any primitive irreducible polynomial of $\mathbb{F}_2[x]$ of degree m. Its dual the **simplex code** S_m is a one-weight code.

Theorem 9

For all integers $n \geq 1$ and all prime powers q with $(n, q) = 1$, we have that $\Gamma(n, 1, q) = 1$, and that $\Gamma(n, n - 1, q)$ is the number of nonzero weights in $P(n, q)$. For all primes $m \geq 2$, we have $\Gamma(n, n - m, 2) = n - 4$, and $\Gamma(n, m, 2) = 1$, where $n = 2^m - 1$.

The next two theorems rely on some deep algebraic geometric enumeration of cyclic codes [7,8].

Theorem 10

For all integers $m \geq 3$, we have $\Gamma(2^m - 1, 2m, 2) \geq \lceil 2^{m/2} \rceil$, and $\Gamma(2^m + 1, 2m, 2) \geq \lceil 2^{m/2} \rceil$.

Proof. The dual of the binary Melas code is cyclic of parameters $[2^m - 1, 2m]$. It is proved in [[7, Th. 6.3]] that its nonzero weights are all the even integers w in the range

$$|w - \frac{2^m - 1}{2}| \leq 2^{m/2}.$$

Similarly, the dual of the Zetterberg code is an irreducible cyclic code of parameters $[2^m + 1, 2m]$. It is proved in [[7, Th. 6.6]] that its nonzero weights are all the even integers w in the range

$$|w - \frac{2^m + 1}{2}| \leq 2^{m/2}.$$

The result follows after elementary calculations. ■

[7] G. Lachaud, J. Wolfmann, The weights of the orthogonals of the extended quadratic binary Goppa codes, IEEE Trans. on Information Theory **36**, (1990) 686–692.

Theorem 11

For all integers $m \geq 2$ we have $\Gamma(3^m - 1, 2m, 3) \geq \lceil 4 \times 3^{\frac{m-2}{2}} \rceil$.

Proof. The dual of the ternary Melas code is cyclic of parameters $[3^m - 1, 3m]$. It is proved in [8] that its nonzero weights are of the form $\frac{3^m - 1 + t}{3}$ with $t \in \mathbb{Z}$, satisfying $t \equiv 1 \pmod 3$, and $t^2 < 3^m$. The result follows after elementary calculations.
It is remarkable that the last two theorems imply lower bounds on $\Gamma(k, 2)$ and $\Gamma(k, 3)$ that are exponential in the dimension. It would be desirable to extend these results to $\Gamma(k, q)$ with q a prime power > 3. ■

[8] G. Van der Geer, M. van der Vlugt, Artin ?Schreier curves and codes, J. Algebra **139** (1991), 256–272.

4.2 Covering radius

Recall that the **covering radius** $\rho(C)$ of a code C is the smallest integer t such that every point in $\mathbb{F}_q^n$ is at distance at most t from some codeword of C. A combinatorial function that is, as far as we know, new, is $T[n, k, q]$, the largest covering radius of a cyclic code of length n and dimension k over $\mathbb{F}_q$. Note that the closest classical function in that context is, for $q = 2$, the quantity $t[n, k]$, the smallest covering radius of a binary linear code of length n and dimension k. Trivially $t[n, k] \leq T[n, k, 2]$. The Delsarte bound [9], stated for the dual of a linear code C, is $\rho(C^{\perp}) \leq s(C)$ [9][Chap. 6, Th. 21].

With the above definitions, we can state the following result.

Proposition 3

For all integers n, k with $1 \leq k \leq n$, we have

$$\Gamma(n, k, q) \geq T[n, n-k, q].$$

Proof. Upon using Delsarte bound for the dual of an $[n, k]_q$ code with $\Gamma(n, k, q)$ nonzero weights, which is, in particular, an $[n, n-k]_q$ code we see that $\Gamma(n, k, q) \geq T[n, n-k, q]$. ■

[9] F.J. MacWilliams, N.J.A. Sloane, *The theory of error correcting codes*, North Holland, Amsterdam (1977).

5 Asymptotics

The q-ary **entropy function** $H_q(\cdot)$ is defined for $0 < y < \frac{q-1}{q}$ by

$$H_q(y) = y \log_q(q-1) - y \log_q(y) - (1-y) \log_q(1-y).$$

To consider the number of weights of long codes of given rate, we study the behavior of $\gamma_q(R)$ defined for $0 < R < 1$ as

$$\gamma_q(R) = \limsup_{n \to \infty} \Gamma(n, \lfloor Rn \rfloor, q).$$

Theorem 12

For all rates $R \in (0,1)$ we have $H_q^{-1}(R) \leq \gamma_q(R) \leq R$. In particular, $\gamma_q(R) \leq t(q)$, the unique solution in $(0, \frac{q-1}{q})$ of the equation $H_q(x) = x$.

Proof The upper bound comes from the immediate inequalities $\Gamma(n,k,q) \leq \Gamma(k,q) \leq q^k - 1$. The lower bound follows by combining the sphere-covering bound [10] with Proposition 3. ■

[10] G.D. Cohen, I. Honkala, S. Litsyn, A. Lobstein, *Covering codes*, North-Holland, Amsterdam (1997).

Similarly for strongly cyclic codes we define

$$\gamma_q^0(R) = \limsup_{n\to\infty} \Gamma^0(n, \lfloor Rn \rfloor, q).$$

We obtain a different upper bound.

Theorem 13

For all rates $R \in (0,1)$ we have $H_q^{-1}(R) \leq \gamma_q^0(R) \leq \frac{R}{2}$. In particular, $\gamma_q(R) \leq t^0(q)$, the unique solution in $(0, \frac{q-1}{q})$ of the equation $H_q(x) = \frac{x}{2}$.

Main results

• Upper and lower bounds on the largest number of weights in a cyclic code of given length, dimension and alphabet are given.

• An application to irreducible cyclic codes is considered.

• Sharper upper bounds are given for cyclic codes (called here strongly cyclic), all codewords of which have period the length.

• Asymptotics are derived on the function $\Gamma(k, q)$, the largest number of nonzero weights a cyclic code of dimension k over $\mathbb{F}_q$ can have.

Thanks for your attention !