On quadratic residue codes and hyperelliptic curves

David Joyner*

8-28-2006

A long standing problem has been to develop "good" binary linear codes to be used for error-correction. This paper investigates in some detail an attack on this problem using a connection between quadratic residue codes and hyperelliptic curves. Codes with this kind of relationship have been investigated in Helleseth [H], Bazzi-Mitter [BM], Voloch [V1], and Helleseth-Voloch [HV]. This rest of this introduction is devoted to explaining in more detail the ideas discussed in later sections.

Let $\mathbb{F} = GF(2)$ be the field with two elements and $C \subset \mathbb{F}^n$ denote a binary block code of length n. For any two $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, let $d(\mathbf{x}, \mathbf{y})$ denote the *Hamming metric*:

$$d(\mathbf{x}, \mathbf{y}) = |\{1 \leq i \leq n \mid x_i \neq y_i\}|. \tag{1}$$

The *weight* w of $\mathbf{v}$ is the number of non-zero entries of $\mathbf{v}$. The smallest weight of any non-zero codeword is denoted d - the minimum distance if C is linear. When C is linear, denote the dimension of C by k and call C an $[n, k, d]_2$-code.

Denoting the volume of a Hamming sphere of radius r in $\mathbb{F}^n$ by $V(n, r)$, the binary version of the *Gilbert-Varshamov bound* asserts that (given n and d) there is an $[n, k, d]_2$ code C satisfying $k \geq \log_2(\frac{2^n}{V(n,d-1)})$ [HP].

Conjecture 1 *(Goppa's conjecture [JV],[G]) The binary version of the Gilbert-Varshamov bound is asymptotically exact.*

For each odd prime $p > 5$, a QQR code[1] is a linear code of length $2p$. Like the quadratic residue codes, the length and dimension are easy to determine but the minimum distance is more mysterious. In fact, the weight of each codeword can be explicitly computed in terms of the number of solutions in integers mod p to a certain type of ("hyperelliptic") polynomial equation. To explain the results better, some more notation is needed.

For our purposes, a *hyperelliptic curve* X over $GF(p)$ is a polynomial equation of the form $y^2 = h(x)$, where $h(x)$ is a polynomial with coefficients in $GF(p)$ with

*Department of Mathematics, United States Naval Academy, wdj@usna.edu

[1]This code is defined in §2 below.

distinct roots[2]. The number of solutions to $y^2 = h(x) \bmod p$, plus the number of "points at infinity" on X, will be denoted $|X(GF(p))|$. This quantity can be related to a sum of Legendre characters (see Lemma 4 below), thanks to classical work of Artin, Hasse, and Weil. This formula yields good estimates for $|X(GF(p))|$ in many cases (especially when p is large compared to the degree of h). A long-standing open problem has been to improve on the trivial estimate when p is small compared to the degree of h.

For each non-empty subset $S \subset GF(p)$, consider the hyperelliptic curve X_S defined by $y^2 = f_S(x)$, where $f_S(x) = \prod_{a\in S}(x-a)$. Let $B(c,p)$ be the statement: *For all subsets $S \subset GF(p)$, $|X_S(GF(p))| \le c \cdot p$ holds.* Note that $B(2,p)$ is trivially true, so the statement $B(2-\epsilon,p)$, for some fixed $\epsilon > 0$, might not be horribly unreasonable.

Conjecture 2 *("Bazzi-Mitter conjecture" [BM]) There is a $c \in (0,2)$ such that, for an infinite number of primes p the statement $B(c,p)$ holds.*

It is remarkable that these two conjectures are related. In fact, using QQR codes we show that if, for an infinite number of primes p with $p \equiv 1 \pmod 4$, $B(1.77,p)$ holds then Goppa's conjecture is false. Although this is a new result, it turns out that it's an easy consequences of the QQR construction given in [BM] if you think about things in the right way. Using LQR codes[3] we will remove the condition $p \equiv 1 \pmod 4$ at a cost of slightly weakening the constant 1.77 (see Corollary 2).

Though of secondary importance here, the Duursma zeta function of these QQR codes is discussed in a brief section below and some examples are given (with the help of the software package `SAGE` [S]).

We close this introduction with a few open questions which, on the basis of this result, seem natural.

Question 1: For each prime $p > 5$ is there an effectively computable subset $S \subset GF(p)$ such that $|X_S(GF(p))|$ is "large"?

Here "large" is left vague but what is intended is some quantity which is unusual. By Weil's estimate (valid for "small"-sized subsets S), we could expect about p points to belong to $|X_S(GF(p))|$. Thus "large" could mean, say, $> c\cdot p$, for some fixed $c > 1$.

The next question is a strong version of the Bazzi-Mitter conjecture.

Question 2: Does there exist a $c < 2$ such that, for all sufficiently large p and all $S \subset GF(p)$, we have $|X_S(GF(p))| < c \cdot p$?

In the direction of these questions, for Question 1, a coding theory bound of McEliese-Rumsey-Rodemich-Welsh allows one to establish the following result (see Theorem 2): *There exists a constant p_0 having the following property: if $p > p_0$ then there exists a subset $S \subset GF(p)$ for which the bound $|X_S(GF(p))| > 1.62p$ holds.* Unfortunately, the method of proof gives no clue how to compute p_0 or S. Using the theory of self-dual codes, we prove the following lower (Corollary 1): *If $p \equiv 3 \pmod 4$ then there exists a subset $S \subset GF(p)$ for which the bound $|X_S(GF(p))| > \frac{5}{3}p - 4$ holds.* Again, we do not know what S is.

[2]This overly simplistic definition brings to mind the famous Felix Klein quote: "Everyone knows what a curve is, until he has studied enough mathematics to become confused through the countless number of possible exceptions." Please see Tsafsman-Vladut [TV] or Schmidt [Sc] for a rigorous treatment.

[3]These codes will be defined in §4 below.

Finally, Felipe Voloch [V2] has kindly allowed the author to include some interesting explicit constructions (which do not use any theory of error-correcting codes) in this paper (see §5 below). First, he shows the following result: *If* $p \equiv 1 \pmod 8$ *then there exists an explicit subset* $S \subset GF(p)$ *for which the bound* $|X_S(GF(p))| > 1.5p$ *holds.* A similar result holds for $p \equiv 3, 7 \pmod 8$. Second, he gives a construction which answers Question 2 in the negative.

1 Cyclotomic arithmetic mod 2

Let $R = \mathbb{F}[x]/(x^p - 1)$ and $r_S \in R$ denotes the polynomial

$$r_S(x) = \sum_{i \in S} x^i,$$

where $S \subseteq GF(p)$. By convention, if $S = \emptyset$ is the empty set, $r_S = 0$. We call $|S|$ the *weight* of r_S. Note that $r_S^2 = r_{2S}$, where $2S$ is the set of elements $2s \in GF(p)$, for $s \in S$. In particular, since $Q \subseteq GF(p)^\times$ is a subgroup, $r_Q^2 = r_Q$ if and only if $2 \in Q$ if and only if $p \equiv \pm 1 \pmod 8$ (by the quadratic reciprocity law). Moreover, if $2 \in N$ then $r_Q^2 = r_N$.

For $S_1, S_2 \subset GF(p)$, $a \in GF(p)$, let

$$H(S_1, S_2, a) = \{(s_1, s_2) \in S_1 \times S_2 \mid s_1 + s_2 \equiv a \pmod p\}.$$

In particular,

- $H(S_1, S_2, a) = H(S_2, S_1, a)$,
- there is a natural bijection $H(GF(p), S, a) \cong S$,
- if $S_1 \cap S_1' = \emptyset$ then $H(S_1, S_2, a) + H(S_1', S_2, a) = H(S_1 + S_1', S_2, a)$.

Let

$$h(S_1, S_2, a) = |H(S_1, S_2, a)| \pmod 2.$$

Adding $|H(S_1, S_2, a)| + |H(S_1^c, S_2, a)| = |S_2|$ to $|H(S_1^c, S_2^c, a)| + |H(S_1^c, S_2, a)| = |S_1^c|$, we obtain

$$h(S_1, S_2, a) \equiv h(S_1^c, S_2^c, a) + |S_1^c| + |S_2| \pmod 2, \tag{2}$$

so

$$r_{S_1} r_{S_2} = \sum_{a \in GF(p)} h(S_1, S_2, a) x^a.$$

Let $S^c = GF(p) - S$ and let $* : R \to R$ denote the involution defined by $(r_S)^* = r_{S^c} = r_S + r_{GF(p)}$. We shall see below that this is not an algebra involution.

Lemma 1 *For all* $S_1, S_2 \subset GF(p)$*, we have*

- $|S_1|$ *odd,* $|S_2|$ *even:* $r_{S_1}r_{S_2} = r^*_{S_1}r^*_{S_2}$ *has even weight.*
- $|S_1|$ *even,* $|S_2|$ *even:* $(r_{S_1}r_{S_2})^* = r^*_{S_1}r^*_{S_2}$ *has even weight.*
- $|S_1|$ *even,* $|S_2|$ *odd:* $r_{S_1}r_{S_2} = r^*_{S_1}r^*_{S_2}$ *has even weight.*
- $|S_1|$ *odd,* $|S_2|$ *odd:* $(r_{S_1}r_{S_2})^* = r^*_{S_1}r^*_{S_2}$ *has odd weight.*

This lemma follows from the discussion above by a straightforward argument.

Note that $R_{even} = \{r_S \mid |S| \text{ even}\}$, is a subring of R and, by the previous lemma, $*$ is an algebra involution on R_{even}.

2 QQR Codes

These are some observations on the interesting paper by Bazzi and Mitter [BM]. We shall need to remove the assumption $p \equiv 3 \pmod 8$ (which they make in their paper) below.

If $S \subseteq GF(p)$, let $f_S(x) = \prod_{a \in S}(x-a) \in GF(p)[x]$. Let $\chi = (\frac{\cdot}{p})$ be the quadratic residue character, which is 1 on the set Q quadratic residues in $GF(p)^\times$, -1 on the set N non-quadratic residues, and is 0 on $0 \in GF(p)$.

Define

$$C_{NQ} = \{(r_N r_S, r_Q r_S) \mid S \subseteq GF(p)\},$$

where N, Q are as above. (When S is the empty set, we associate to $(r_N r_S, r_Q r_S)$ the zero codeword.) We call this a *QQR code* (or a *quasi-quadratic residue code*). These are binary codes of length $2p$ and dimension

$$k = \begin{cases} p, & \text{if } p \equiv 3 \pmod 4, \\ p-1, & \text{if } p \equiv 1 \pmod 4. \end{cases}$$

This code has no codewords of odd weight.

Remark 1 *If* $p \equiv \pm 1 \pmod 8$ *then* C_{NQ} *contains a binary quadratic residue code. For such primes* p*, the minimum distance satisfies the well-known square-root lower bound,* $d \geq \sqrt{p}$*.*

Lemma 2 *For any* p*, the associated QQR code is self-orthogonal:* $C^\perp_{NQ} \subset C_{NQ}$*. If* $p \equiv 3 \pmod 4$ *then the associated QQR code is self-dual:* $C^\perp_{NQ} = C_{NQ}$*.*

Recall a code C is *self-orthogonal* if C is a subcode of its dual code $C^\perp$.

Proof: By Lemma 1 (and the fact that $|Q|$, $|N|$ are either both even or both odd), we know that each codeword in C_{NQ} has even weight. (In fact, when $p \equiv 1 \pmod 4$, each codeword in C_{NQ} has doubly even weight, i.e., all codewords have weight divisible by 4.) Therefore, $C^\perp_{NQ} \subset C_{NQ}$. However, if $p \equiv 3 \pmod 4$ then $\dim(C_{NQ})$ is half its length, so $\dim(C^\perp_{NQ}) = \dim(C_{NQ})$. This forces them to be equal. $\Box$

The self-dual binary codes have useful upper bounds on their minimum distance (for example Theorem 9.3.5 in [HP]). Combining this with the lower bound mentioned above, we have

Lemma 3 *If* $p \equiv 3 \pmod 4$ *then*

$$d \leq 4 \cdot [p/12] + 6.$$

If $p \equiv -1 \pmod 8$ *then*

$$\sqrt{p} \leq d \leq 4 \cdot [p/12] + 6.$$

Note that these upper bounds (in the cases they are valid) are better than the asymptotic bounds of McEliese-Rumsey-Rodemich-Welsh for rate $1/2$ codes.

Example 1 *When* $p = 5$*,* C_{NQ} *has weight distribution*

$$[1, 0, 0, 0, 5, 0, 10, 0, 0, 0, 0].$$

When $p = 7$*,* C_{NQ} *has weight distribution*

$$[1, 0, 0, 0, 14, 0, 49, 0, 49, 0, 14, 0, 0, 0, 1].$$

When $p = 11$*,* C_{NQ} *has weight distribution*

$$[1, 0, 0, 0, 0, 0, 77, 0, 330, 0, 616, 0, 616, 0, 330, 0, 77, 0, 0, 0, 0, 0, 1].$$

When $p = 13$*,* C_{NQ} *has weight distribution*

$$[1, 0, 0, 0, 0, 0, 0, 0, 273, 0, 598, 0, 1105, 0, 1300, 0, 598, 0, 182, 0, 39, 0, 0, 0, 0, 0, 0].$$

The following well-known result[4] shall be used to estimate the weights of codewords of QQR codes.

Lemma 4 *(Artin, Hasse, Weil) Assume* $S \subset GF(p)$ *is non-empty.*

- $|S|$ *even:*

$$\sum_{a \in GF(p)} \chi(f_S(a)) = -p - 2 + |X_S(GF(p))|.$$

- $|S|$ *odd:*

$$\sum_{a \in GF(p)} \chi(f_S(a)) = -p - 1 + |X_S(GF(p))|.$$

- $|S|$ *odd: The genus of the (smooth projective model of the) curve* $y^2 = f_S(x)$ *is* $g = \frac{|S|-1}{2}$ *and*

$$|\sum_{a \in GF(p)} \chi(f_S(a))| \leq (|S| - 1)p^{1/2} + 1.$$

[4]See for example Weil [W] or Schmidt [Sc], Lemma 2.11.2.

- $|S|$ *even: The genus of the (smooth projective model of the) curve* $y^2 = f_S(x)$ *is* $g = \frac{|S|-2}{2}$ *and*

$$|\sum_{a \in GF(p)} \chi(f_S(a))| \leq (|S|-2)p^{1/2} + 1.$$

Obviously, the last two estimates are only non-trivial for S "small" (e.g., $|S| < p^{1/2}$).

Lemma 5 *(Bazzi-Mitter [BM], Proposition 3.3) Assume* 2 *and* -1 *are non-quadratic residues mod* p *(i.e.* $p \equiv 3 \pmod 8$*).*

If $c = (r_N r_S, r_Q r_S)$ *is a nonzero codeword of the* $[2p, p]$ *binary code* C_{NQ} *then the weight of this codeword can be expressed in terms of a character sum as*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_S(a)),$$

if $|S|$ *is even, and*

$$wt(c) = p + \sum_{a \in GF(p)} \chi(f_{S^c}(a)),$$

if $|S|$ *is odd.*

In fact, looking carefully at their proof, one finds the following result:

Proposition 1 *Let* $c = (r_N r_S, r_Q r_S)$ *be a nonzero codeword of* C_{NQ}*.*

(a) If $|S|$ *is even*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_S(a)) = 2p + 2 - |X_S(GF(p))|.$$

(b) If $|S|$ *is odd and* $p \equiv 1 \pmod 4$ *then the weight is*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_{S^c}(a)) = 2p + 2 - |X_{S^c}(GF(p))|.$$

(c) If $|S|$ *is odd and* $p \equiv 3 \pmod 4$ *then*

$$wt(c) = p + \sum_{a \in GF(p)} \chi(f_{S^c}(a)) = |X_{S^c}(GF(p))| - 2.$$

Proof: If $A, B \subseteq GF(p)$ then the discussion in §1 implies

$$wt(r_A r_B) = \sum_{k \in GF(p)} \text{parity } |A \cap (k - B)|, \tag{3}$$

where $k - B = \{k - b \mid b \in B\}$. Let $S \subseteq GF(p)$, then we have

$$p - wt(r_Q r_S) - wt(r_N r_S) = \sum_{a \in GF(p)} 1 - \text{ parity } |Q \cap (a - S)| - \text{ parity } |N \cap (a - S)|.$$

Let

$$T_a(S) = 1 - \text{ parity } |Q \cap (a - S)| - \text{ parity } |N \cap (a - S)|.$$

Case 1. If $|S|$ is even and $a \in S$ then $0 \in a - S$ so $|Q \cap (a - S)|$ odd implies that $|N \cap (a - S)|$ is even, since 0 is not included in $Q \cap (a - S)$ or $N \cap (a - S)$. Likewise, $|Q \cap (a - S)|$ even implies that $|N \cap (a - S)|$ is odd. Therefore $T_a(S) = 0$.

Case 2. If $|S|$ is even and $a \notin S$ then parity $|Q \cap (a - S)|$ =parity$|N \cap (a - S)|$. If $|Q \cap (a - S)|$ is even then $T_a(S) = 1$ and if $|Q \cap (a - S)|$ is odd then $T_a(S) = -1$.

Case 3. $|S|$ is odd. We claim that $(a - S)^c = a - S^c$. (Proof: Let $s \in S$ and $\bar{s} \in S^c$. Then $a - s = a - \bar{s} \implies s = \bar{s}$, which is obviously a contradiction. Therefore $(a - S) \cap (a - S^c) = \emptyset$, so $(a - S)^c \supseteq (a - S^c)$. Replace S by S^c to prove the claim.) Also note that

$$Q \cap (a - S) \sqcup Q \cap (a - S^c) = GF(p) \cap Q = Q$$

has $|Q| = \frac{p-1}{2}$ elements ($\sqcup$ denotes disjoint union). So

$$\text{parity } |Q \cap (a - S)| = \text{parity } |Q \cap (a - S^c)|$$

if and only if $|Q|$ is even and

$$\text{parity } |Q \cap (a - S)| \neq \text{parity } |Q \cap (a - S^c)|$$

if and only if and only if $|Q|$ is odd.

Conclusion.

$$|S| \text{ even}: \; T_a(S) = \prod_{x \in a - S} \left(\frac{x}{p}\right)$$

$$|S| \text{ odd and } p \equiv 3 \pmod 4): \; T_a(S) = -T_a(S^c)$$

$$|S| \text{ odd and } p \equiv 1 \pmod 4): \; T_a(S) = T_a(S^c)$$

The relation between $wt(c)$ and the character sum follows from this. For the remaining part of the equation, use Lemma 4. $\Box$

Remark 2 *It can be shown, using the coding-theoretic results above, that if* $p \equiv -1 \pmod 8$ *then (for non-empty S)* $X_S(GF(p))$ *contains at least* $\sqrt{p} + 1$ *points. This also follows from Weil's estimate, but since the proof is short, it is given below.*

What part (c) of Proposition 1 gives is that If $p \equiv -1 \pmod 8$ *and* $|S|$ *is odd then* $X_S(GF(p))$ *contains at least* $\sqrt{p} + 2$ *points. If* $|S|$ *is even then perform the substitution* $x = a + 1/\overline{x}$, $y = \overline{y}/\overline{x}^{|S|}$ *on the equation* $y^2 = f_S(x)$*. This creates a hyperelliptic curve X in* $(\overline{x}, \overline{y})$ *for which* $|X(GF(p))| = |X_S(GF(p))|$ *and* $X \cong X_{S'}$, *where* $|S'| = |S| - 1$ *is odd. Now apply part (c) of the above proposition and Remark 1 to* $X_{S'}$. $\Box$

As a consequence of this Proposition and Lemma 3, we have the following result.

Corollary 1 *If $p \equiv 3 \pmod 4$ then $\max_S |X_S(GF(p))| > \frac{5}{3}p - 4$.*

Example 2 *The following examples were computed with the help of* `SAGE`*.*

If $p = 11$ and $S = \{1, 2, 3, 4\}$ then

$$(r_S(x)r_N(x), r_S(x)r_Q(x))$$
$$= (x^{10} + x^9 + x^7 + x^6 + x^5 + x^4 + x^2 + 1, x^{10} + x^9 + x^7 + x^6 + x^5 + x^3 + x + 1),$$

corresponds to the codeword $(1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1)$ of weight 16. An explicit computation shows that the character sum $\sum_{a\in GF(11)} \chi(f_S(a))$ is -5, as expected.

If $p = 11$ and $S = \{1, 2, 3\}$ then

$$(r_S(x)r_N(x), r_S(x)r_Q(x)) = (x^9+x^7+x^5+x^4+x^3+x^2+x, x^{10}+x^8+x^6+x^3+x^2+x+1).$$

corresponds to the codeword $(0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1)$ of weight 14. An explicit computation shows that the character sum $\sum_{a\in GF(11)} \chi(f_{S^c}(a))$ is 3, as predicted.

Recall $B(c, p)$ is the statement: $|X_S(GF(p))| \leq c \cdot p$ for all $S \subset GF(p)$.

Theorem 1 *(Bazzi-Mitter) Fix $c \in (0, 2)$. If $B(c, p)$ holds for infinitely many p with $p \equiv 1 \pmod 4$ then there exists an infinitely family of binary codes with asymptotic rate $R = 1/2$ and relative distance $\delta \geq 1 - \frac{c}{2}$.*

This is an easy consequence of the above results and is essentially in [BM] (though they assume $p \equiv 3 \pmod 8$).

Theorem 2 *If the $B(1.77, p)$ is true for infinitely many primes p with $p \equiv 1 \pmod 4$ then Goppa's conjecture is false.*

Proof: Recall Goppa's conjecture is that the binary asymptotic Gilbert-Varshamov bound is best possible for any family of binary codes. The asymptotic GV bound states that the rate R is greater than or equal to $1 - H(\delta)$, where

$$H(\delta) = \delta - \delta \log_2(\delta) - (1 - \delta) \log_2(1 - \delta)$$

is the entropy function (for a binary channel). Therefore, according to Goppa's conjecture, if $R = \frac{1}{2}$ then the best possible δ is $\delta_0 = .11$. Assume $p \equiv 1 \pmod 4$. Goppa's conjecture implies that the minimum distance of our QQR code with rate $R = \frac{1}{2}$ satisfies $d < \delta_0 \cdot 2p = .22p$, for sufficiently large p. Recall that the weight of a codeword c_S in this QQR code is given by Lemma 4, for $S \subset GF(p)$. $B(1.77, p)$ (with $p \equiv 1 \pmod 4$) implies (for all $S \subset GF(p)$) $wt(c_S) \geq 2p - |X_S(GF(p))| \geq 0.23p$. In other words, for $p \equiv 1 \pmod 4$, all nonzero codewords have weight at least $0.23p$. This contradicts the estimate above.$\Box$

Using the same argument and the first McEliese-Rumsey-Rodemich-Welsh bound ([HP], Theorem 2.10.6), we prove the following unconditional result.

Theorem 3 *For all sufficiently large primes p for which $p \equiv 1 \pmod 4$, the statement $B(1.62, p)$ is false.*

Proof: If a prime p satisfies $B(1.62, p)$ then we shall call it "admissible." We show that the statement "$B(1.62, p)$ holds for all sufficiently large primes p for which $p \equiv 1 \pmod 4$" contradicts the first asymptotic MRRW bound. Indeed, this MRRW bound states that the rate R is less than or equal to $h(\delta) = H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$. This, and the fact that $R = \frac{1}{2}$ for our QQR codes (with $p \equiv 1 \pmod 4$), imply $\delta \leq \delta_0 = h^{-1}(1/2) \cong 0.187$. Therefore, for all large p (admissible or not), $d \leq \delta_0 \cdot 2p$. On the other hand, if p is admissible and $|X_S(GF(p))| \leq c \cdot p$ (where $c = 1.62$) then by the above argument, $d \geq 2 \cdot (p - \frac{c}{2}p)$. Together, we obtain $1 - \frac{c}{2} \leq \delta_0$, so $c \geq 2 \cdot (1 - h^{-1}(1/2)) \cong 1.626$. This is a contradiction. $\Box$

Corollary 2 *There is a constant p_0 (ineffectively computable) having the following property: if $p > p_0$ then there is a subset $S \subset GF(p)$ for which the bound $|X_S(GF(p))| > 1.62p$ holds.*

This is of course the same as the above theorem, except that we have used Corollary 1 to remove the hypothesis $p \equiv 1 \pmod 4$.

3 Duursma zeta functions

In [D1] Iwan Duursma associates to a linear code C over $GF(q)$ a *zeta function* $Z = Z_C$ of the form

$$Z(T) = \frac{P(T)}{(1-T)(1-qT)},$$

where $P(T)$ is a "zeta polynomial" of degree $n + 2 - d - d^{\perp}$ which only depends on C through its weight enumerator polynomial (here d is the minimum distance of C and $d^{\perp}$ is the minimum distance of its dual code $C^{\perp}$). If $\gamma = \gamma(C) = n + k + 1 - d$ and $\xi_C(T) = Z_C(T)T^{1-\gamma}$ then the functional equation in [D1] can be written in the form $\xi_{C^{\perp}}(T) = \xi_C(1/qT)$. If we let $\zeta_C(s) = \xi_C(q^{-s})$ then it can be written

$$\zeta_{C^{\perp}}(s) = \zeta_C(1-s).$$

In fact, if ρ_i denotes the i-th zero of $Z(T)$ then equation (4.1) of [D2] implies (for the even weight codes we are considering here) the relation

$$d = 2 - \sum_i \rho_i^{-1}.$$

Therefore, further knowledge of the zeros of $Z(T)$ could be very useful.

If C is self-dual (or actually only formally self-dual) then the zeros of the ζ-function occur in pairs about the "critical line" $Re(s) = \frac{1}{2}$. Following Duursma, we say (for formally self-dual codes C) the "normalized" zeta function ζ_C satisfies the *Riemann hypothesis* if all its zeros occur on the "critical line". (The conjecture is that, for all extremal self-dual codes C, the ζ-function satisfies the Riemann hypothesis.)

Example 3 *If $p = 7$ then the $[14, 7, 4]$ code C_{NQ} has zeta polynomial*

$$P(T) = \frac{2}{143} + \frac{4}{143}T + \frac{19}{429}T^2 + \frac{28}{429}T^3 + \frac{40}{429}T^4 + \frac{56}{429}T^5 + \frac{76}{429}T^6 + \frac{32}{143}T^7 + \frac{32}{143}T^8.$$

Using SAGE, *it can be checked that all the roots ρ of this polynomial have $|\rho| = \sqrt{2}$.*

It would be interesting to know if the Duursma zeta function $Z(T)$ of C_{NQ} (for $p \equiv 3 \pmod 4$) satisfies the Riemann hypothesis.

If $p \equiv 1 \pmod 4$ then we conjecture that the code C' spanned by C_{NQ} and the all ones codeword (i.e., the smallest code containing C_{NQ} and all its complementary codewords) is a formally self-dual code of dimension p. Moreover, we if $A = [A_0, A_1, ..., A_n]$ is the weight distribution vector of C_{NQ} then we conjecture that the weight distribution vector of C' is $A + A^*$, where $A^* = [A_n, ..., A_1, A_0]$. However, the Riemann hypothesis is not valid for these codes in general.

Example 4 *If $p = 13$ then C' is a $[26, 13, 6]$ code with weight distribution*

$$[1, 0, 0, 0, 0, 0, 39, 0, 455, 0, 1196, 0, 2405, 0, 2405, 0, 1196, 0, 455, 0, 39, 0, 0, 0, 0, 0, 1]$$

and zeta polynomial

$$\begin{aligned} P(T) \quad &= \tfrac{3}{17710} + \tfrac{6}{8855}T + \tfrac{611}{336490}T^2 + \tfrac{9}{2185}T^3 + \tfrac{3441}{408595}T^4 + \tfrac{6448}{408595}T^5 + \tfrac{44499}{1634380}T^6 \\ &+ \tfrac{22539}{520030}T^7 + \tfrac{66303}{1040060}T^8 + \tfrac{22539}{260015}T^9 + \tfrac{44499}{408595}T^{10} + \tfrac{51584}{408595}T^{11} \\ &+ \tfrac{55056}{408595}T^{12} + \tfrac{288}{2185}T^{13} + \tfrac{19552}{168245}T^{14} + \tfrac{768}{8855}T^{15} + \tfrac{384}{8855}T^{16}. \end{aligned}$$

Using SAGE, *it can be checked that only 8 of the 12 zeros of this function have absolute value $\sqrt{2}$.*

4 Long Quadratic Residue Codes

We now introduce a new code, constructed similarly to the QQR codes discussed above:

$$C = \{(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*) \mid S \subseteq GF(p)\}.$$

We call this a *long quadratic residue code* or *LQR code* for short.

For any $S \subseteq GF(p)$, let

$$c_S = (r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*)$$

and let

$$v_S = (r_N r_S, r_Q r_S, r_N r_S, r_Q r_S).$$

Observe that this code is non-linear. If $S_1 \Delta S_2$ denotes the symmetric difference between S_1 and S_2 then it is easy to check that

$$c_{S_1} + c_{S_2} = v_{S_1 \Delta S_2}. \tag{4}$$

We now compute the size of C. We now prove the *claim*: if $p \equiv 3 \pmod 4$ then the map that sends S to the codeword $(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*)$ is injective. This implies $|C| = 2^p$. Suppose not, then there are two subsets $S_1,\ S_2 \subseteq GF(p)$ that are mapped to the same codeword. Subtracting, the subset $T = S_1 \Delta S_2$ satisfies $r_Q r_T = r_N r_T = r_Q r_{T^c} = r_N r_{T^c} = 0$. If $|T|$ is even then $0 = (r_Q + r_N) r_T = (r_{GF(p)} - 1) r_T = r_T$. This forces T to be the empty set, so $S_1 = S_2$. Now if $|T|$ is odd then similar reasoning implies that T^c is the empty set. Therefore, $S_1 = \emptyset$ and $S_2 = GF(p)$ or vice versa. This proves the claim.

In case $p \equiv 1 \pmod 4$, *claim*: $|C| = 2^{p-1}$. Again, suppose there are two subsets $S_1,\ S_2 \subseteq GF(p)$ that are mapped to the same codeword. Then the subset $T = S_1 \Delta S_2$ for which $r_Q r_T = r_N r_T = r_Q r_{T^c} = r_N r_{T^c} = 0$. This implies either $T = \emptyset$ or $T = GF(p)$. Therefore, either $S_1 = S_2$ or $S_1 = S_2^c$.

Combining this discussion with Lemma 4, we have proven the following result.

Theorem 4 *The code C has length $n = 4p$ and has size $M = 2^{p-1}$ if $p \equiv 1 \pmod 4$, and size $M = 2^p$ if $p \equiv 3 \pmod 4$. If $p \equiv 3 \pmod 4$ then the minimum non-zero weight is $2p$ and the minimum distance is at least*

$$d_p = 4p - 2 \max_{S \subset GF(p)} |X_S(GF(p))|.$$

If $p \equiv 1 \pmod 4$ then C is a binary $[4p, p-1, d_p]$-code.

Remark 3 *If $p \equiv 3 \pmod 4$, there is no simple reason I can think of why the minimum distance should actually be less than the minimum non-zero weight.*

Lemma 6 *If $p \equiv 1 \pmod 4$ then*

- $v_S = c_S$,
- $c_{S_1} + c_{S_2} = c_{S_1 \Delta S_2}$,
- *the code C is isomorphic to the QQR code C_{NQ}.*

In particular, C is linear, dimension $p - 1$, and self-orthogonal (by Lemma 2).

Proof: We shall show later (see the proof of Theorem 4) that if $p \equiv 1 \pmod 4$ then $r_N r_{S_1} = r_N r_{S_2}$ and $r_Q r_{S_1} = r_Q r_{S_2}$ if and only if $S_2 = S_1^c$. The lemma follows rather easily as a consquence of this and (4). □

Assume $p \equiv 3 \pmod 4$. Let

$$V = \{v_S \mid S \subset GF(p)\}$$

and let

$$\overline{C} = C \cup V.$$

Lemma 7 *The code $\overline{C}$ is*

- *the smallest linear subcode of* $\mathbb{F}^{4p}$ *containing* C,
- *dimension* $p+1$,
- *minimum distance* $\min(d_p, 2p)$
- *self-orthogonal.*

By abuse of terminology, we call $\overline{C}$ an *LQR code*.

Proof: The first part follows from (4). The second part follows from a counting argument (as in the proof of Theorem 4 below). The third part is a corollary of Theorem 4 below. The fourth part follows easily from Lemma 1. $\square$

We know that

$$wt(r_N r_S, r_Q r_S) = \begin{cases} p - \sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right), & |S| \text{ even (any } p), \\ p - \sum_{a\in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right), & |S| \text{ odd and } p \equiv 1 \pmod 4, \\ p + \sum_{a\in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right), & |S| \text{ odd and } p \equiv 3 \pmod 4, \end{cases}$$

by Proposition 1.

Lemma 8 *For each p, the codeword* $c_S = (r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*)$ *of* C *has weight*

$$wt(c_S) = \begin{cases} 2p - 2\sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right), & p \equiv 1 \pmod 4, \\ 2p, & p \equiv 3 \pmod 4. \end{cases}$$

In other words, if $p \equiv 3 \pmod 4$ then C is a constant weight code.

Proof: Indeed, Proposition 1 implies if $p \equiv 1 \pmod 4$ then

$$\begin{aligned} wt\left(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*\right) &= wt\left(r_N r_S, r_Q r_S\right) + wt\left(r_N r_S^*, r_Q r_S^*\right) \\ &= 2 \cdot wt\left(r_N r_S, r_Q r_S\right) \\ &= 2p - 2\sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right), \end{aligned} \tag{5}$$

if $p \equiv 3 \pmod 4$ and $|S|$ is even then

$$\begin{aligned} wt\left(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*\right) &= wt\left(r_N r_S, r_Q r_S\right) + wt\left(r_N r_S^*, r_Q r_S^*\right) \\ &= p - \sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right) + p + \sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right) \\ &= 2p, \end{aligned} \tag{6}$$

and if $p \equiv 3 \pmod 4$ and $|S|$ is odd then

$$\begin{aligned} wt\left(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*\right) &= wt\left(r_N r_S, r_Q r_S\right) + wt\left(r_N r_S^*, r_Q r_S^*\right) \\ &= p + \sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right) + p - \sum_{a\in GF(p)} \left(\frac{f_S(a)}{p}\right) \\ &= 2p. \end{aligned} \tag{7}$$

$\square$

Example 5 *The following examples were computed with the help of* `SAGE`*. When $p = 11$ and $S = \{1, 2, 3, 4\}$*

$$\left(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*\right) = (x^9+x^7+x^5+x^4+x^3+x^2+x,\ x^{10}+x^8+x^6+x^3+x^2+x+1,\ x^{10}+x^8+x^6+1,\ x^9+x^7+x^5+x^4)$$

corresponds to the codeword

$$(0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0)$$

of weight 22*. When $p = 11$ and $S = \{1, 2, 3\}$*

$$\left(r_N r_S, r_Q r_S, r_N r_S^*, r_Q r_S^*\right) = (x^{10}+x^9+x^7+x^6+x^5+x^4+x^2+1,\ x^{10}+x^9+x^7+x^6+x^5+x^3+x+1,\ x^8+x^3+x,\ x^8+x^4+x^2)$$

corresponds to the codeword

$$(1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0)$$

of weight 22*.*

Lemma 7 allows us to remove the condition $p \equiv 1 \pmod 4$ in one of the results in §2. The next subsection is devoted to this goal.

4.1 Goppa's conjecture revisited

We shall now remove the condition $p \equiv 1 \pmod 4$ in one of the results in §2, at a cost of weakening the constants involved.

Assuming $B(c, p)$ holds, we have that the minimum distance of $\overline{C}$ is $\geq \min(d_p, 2p) \geq 4p(1 - \frac{c}{2})$ and the information rate is $R = \frac{1}{4} + \frac{1}{4p}$. When $R = 1/4$, Goppa's conjecture gives $\delta = 0.214...$. So Goppa's conjecture will be false if $1 - \frac{c}{2} = 0.215$, or $c = 1.57$. We have the following improvement of Theorem 2.

Theorem 5 *If the $B(1.57, p)$ is true for infinitely many primes p then Goppa's conjecture is false.*

5 Some results of Voloch

Lemma 9 *(Voloch) If $p \equiv 1, 3 \pmod 8$ then $|X_Q(GF(p))| = 1.5p + a$, where Q is the set of quadratic residues and a is a small constant, $-\frac{1}{2} \leq a \leq \frac{5}{2}$.*

A similar bound holds if X_Q is replaced by X_N and $p \equiv 1, 3 \pmod 8$ is replaced by $p \equiv 7 \pmod 8$ (in which case 2 is a quadratic residue).

Proof: By Lemma 4, we know that if $p \equiv 3 \pmod 8$ (so $|Q|$ is odd):

$$\sum_{a \in GF(p)} \chi(f_Q(a)) = -p - 1 + |X_Q(GF(p))|.$$

Similarly, if $p \equiv 1 \pmod 8$ (so $|Q|$ is even):

$$\sum_{a \in GF(p)} \chi(f_Q(a)) = -p - 2 + |X_Q(GF(p))|.$$

Since $b^{\frac{p-1}{2}} \equiv \chi(b) \pmod p$, we have

$$x^{\frac{p-1}{2}} - 1 = \prod_{a \in Q}(x - a) = f_Q(x), \quad x^{\frac{p-1}{2}} + 1 = \prod_{a \in N}(x - a).$$

In particular, for all $n \in N$,

$$f_Q(n) = \prod_{a \in Q}(n - a) = n^{\frac{p-1}{2}} - 1 \equiv -2 \pmod p.$$

Since $p \equiv 1, 3 \pmod 8$, we have $\chi(-2) = 1$, so $\chi(f_Q(n)) = 1$ for all $n \in N$. It follows that $|X_Q(GF(p))| = \frac{3}{2}p + \chi(f_Q(0)) + \frac{1}{2}$ (if $p \equiv 3 \pmod 8$) or $|X_Q(GF(p))| = \frac{3}{2}p + \chi(f_Q(0)) + \frac{3}{2}$ (if $p \equiv 1 \pmod 8$). $\Box$

Here is an extension of the idea in the above proof. Fix an integer $\ell > 2$. Assuming ℓ divides $p - 1$, there are distinct ℓ-th roots $r_1 = 1$, r_2, ..., r_ℓ in $GF(p)$ for which $x^{p-1} - 1 = \prod_{i=1}^{\ell}(x^{\frac{p-1}{\ell}} - r_i)$. Also, $x^{\frac{p-1}{\ell}} - 1 = \prod_{a \in P_\ell}(x - a) = f_{P_\ell}(x)$, where P_ℓ denotes the set of non-zero ℓ-th powers in $GF(p)$.

Claim: It is possible to find an infinite sequence of primes p satisfying $p \equiv 1 \pmod \ell$ and $\chi(r_i - 1) = 1$, for all $2 \leq i \leq \ell$ (where χ denotes the Legendre character mod p). If the claim is true then we will have a lower bound for $|X_{P_\ell}(GF(p))|$ on the order of $(2 - \frac{1}{\ell})p$, along the lines above, by Lemma 4. Proof of claim: It's a well-known fact in algebraic number theory that $p \equiv 1 \pmod \ell$ implies the prime p splits completely in the cyclotomic field $\mathbb{Q}_\ell$ generated by the ℓ-th roots of unity in $\mathbb{C}$, denoted $\tilde{r}_1 = 1, \tilde{r}_2, ..., \tilde{r}_\ell$. The condition $\chi(r_i - 1) = 1$ means that p splits in the extension of $\mathbb{Q}_\ell$ obtained by adjoining $\sqrt{\tilde{r}_i - 1}$ (here $i = 2, ..., \ell$). By Chebotarev's density theorem there exists infinitely many such p, as claimed. In fact, there are effective versions which give explicit information on computing such p [LO], [Se].

This, together with the previous lemma, proves the following result.

Theorem 6 *(Voloch) If $\ell \geq 2$ is any fixed integer then there exist primes p and subsets $S \subset GF(p)$ for which $|X_S(GF(p))| = (2 - \frac{1}{\ell})p + a$, where a is a small constant, $-\frac{1}{2} \leq a \leq \frac{5}{2}$.*

Acknowledgements: I thank Prof. Amin Shokrollahi of the Ecole Polytechnique Fédérale de Lausanne for helpful advice and Prof. Felipe Voloch of the University of Texas for allowing his construction to be included above. Parts of this note (such as Prop. 1) can be found in the honors thesis [C] of my former student Greg Coy, who was a pleasure to work with.

References

[BM] L. Bazzi and S. Mitter, *Some constructions of codes from group actions*, preprint, 2001.

[C] G. Coy, *Long Quadratic Residue Codes*, USNA Mathematics Dept. Honors Project 2005-2006 (advisor Prof. Joyner), available: `http://cadigweb.ew.usna.edu/~wdj/coy/`

[D1] I. Duursma, *Weight distributions of geometric Goppa codes*, Transactions of the AMS, vol. 351, pp. 3609-3639, September 1999.

[D2] ——, *A Riemann hypothesis analogue for self-dual codes*, In: **Codes and Association schemes**, Eds. Barg and Litsyn, AMS Dimacs Series, vol. 56, pp. 115-124, 2001.

[G] V. D. Goppa, *Bounds for codes*, Dokl. Acad. Nauk. 333(1993)423.

[H] T. Helleseth, *Legendre sums and codes related to QR codes*, Discrete Applied Math., vol. 35, pp. 107-113, 1992.

[HP] W. C. Huffman and V. Pless, **Fundamentals of error-correcting codes**, Cambridge Univ. Press, 2003.

[HV] T. Helleseth and J. F. Voloch, *Double circulant quadratic residue codes*, IEEE Transactions in Information Theory, Volume 50, Issue 9, Sept. 2004 Page(s): 2154 - 2155. Available:
`http://www.ma.utexas.edu/users/voloch/preprint.html`

[JV] T. Jiang and A. Vardy, *Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes*, IEEE Trans Info Theory, to appear.

[LO] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in **Algebraic number fields (L- functions and Galois theory)** A. Fröhlich, ed., Academic Press 1977, 409-464.

[MS] F. MacWilliams and N. Sloane, **The theory of error-correcting codes**, North-Holland, 1977.

[Sc] W. Schmidt, **Equations over finite fields: an elementary approach**, 2nd ed., Kendrick Press, 2004.

[Se] J.-P. Serre, *Quelques applications du théorème de densité de Chebotarev*, Publications Mathmatiques de l'IHS, 54 (1981), p. 123-201
`http://www.numdam.org/item?id=PMIHES_1981__54__123_0`

[S] W. Stein, `SAGE`,
`http://sage.scipy.org/`
`http://modular.math.washington.edu/sage/`

[TV] M. A. Tsfasman and S. G. Vladut, **Algebraic-geometric codes**, Mathematics and its Applications, Kluwer Academic Publishers, Dordrechet 1991.

[V1] F. Voloch, *Asymptotics of the minimal distance of quadratic residue codes*, preprint. available: `http://www.ma.utexas.edu/users/voloch/preprint.html`

[V2] ——, email communications, 5-2006.

[W] A. Weil, *On some exponential sums*, Proceedings of the National Academy of Sciences, volume 34 1948)204-207.