ON QUADRATIC RESIDUES AND NONRESIDUES IN DIFFERENCE SETS MODULO m

J. FABRYKOWSKI

(Communicated by William Adams)

Abstract. Let $m > 1$, and consider a set $\mathcal{A} = \{a_i\}$ of residues modulo $m$ such that $a_i$ and $a_i - a_j$ for all $i$ and $j$ with $i \neq j$ are quadratic residues (nonresidues) modulo $m$. We investigate the estimation of the maximal cardinality of such a set $\mathcal{A}$ for various moduli $m$.

1. Introduction

Let $m > 1$ be an integer, and set

$$\mathcal{R}(m) = \{r; 1 \leq r \leq m-1, \text{r-quadratic residue modulo } m\},$$

$$\mathcal{N}(m) = \{n; 1 \leq n \leq m-1, \text{n-quadratic nonresidue modulo } m\}.$$  

We say that a set $\mathcal{A}$ of residues $a_1, a_2, \ldots, a_k$ modulo $m$ satisfies

(C1) if $a_i \in \mathcal{R}(m)$ for all $1 \leq i \leq k$,

(C2) if $a_i - a_j \in \mathcal{R}(m)$ for all $1 \leq i, j \leq k, i \neq j$.

Further, let $R_1(m), R_2(m),$ and $R_{1,2}(m)$ denote the maximal cardinality of a set $\mathcal{A}$ under conditions (C1), (C2), and both (C1) and (C2) respectively. Similarly, we define the functions $N_1(m), N_2(m),$ and $N_{1,2}(m)$, where in conditions (C1) and (C2) the set $\mathcal{R}(m)$ is replaced by the set $\mathcal{N}(m)$. We also introduce the set $\mathcal{R}^*(m) = \mathcal{R}(m) \cup \{0\}$ with mutatis mutandis counting functions $R_1^*(m), R_2^*(m),$ and $R_{1,2}^*(m)$.

The problem of estimating the above functions has already been investigated. The exact value of $R_1(m)$ and thus of $N_1(m)$, since $R_1(m) + N_1(m) = m - 1$, has been determined in [7]. Imposing condition (C2) in both cases of $\mathcal{R}(m)$ or $\mathcal{N}(m)$ makes the problem highly nontrivial.

It is easy to see that if $p$ is a prime $3 \pmod{4}$ then $R_2(p) = R_{1,2}(p) = N_2(p) = N_{1,2}(p) = 1$, and if $p$ is a prime $1 \pmod{4}$ then $R_2(p) = N_2(p) \geq R_{1,2}(p) = N_{1,2}(p)$.

It is also obvious that $R_2(m) \leq N_2(m), R_{1,2}(m) \leq N_{1,2}(m), R_{1,2}(m) \leq R_2(m),$ and $N_1(m) \leq N_2(m)$ for all integers $m$. If $p \equiv 3 \pmod{4}$ then $R_1^*(p) = R_1(p), R_2^*(p) = R_2(p),$ and $R_{1,2}^*(p) = R_{1,2}(p),$ and $R_1^*(m) =$
If \( p \) is a prime \( 1 \pmod{4} \) then Buell and Williams [2] have proved

\begin{align*}
(1) & \quad R_{1,2}(p) > \frac{1}{2} \log p \quad \text{for all } p, \\
(2) & \quad R_{1,2}(p) < p^{1/2} \log p \quad \text{for all } p, \\
(3) & \quad R_{1,2}(p) < (1 + \varepsilon) p^{1/2} \log p / 4 \log 2 \quad \text{for } p > p_0(\varepsilon),
\end{align*}

where both the upper and lower bounds have been obtained using an estimate for \( \sum_{x=0}^{p-1} \left( f(x)/p \right) \) as a consequence of a deep result of A. Weil related to his proof of the Riemann hypothesis for curves over finite fields.

Sarközy [10] made ingenious use of a certain result (see Lemma) of Erdös and Szekeres on extremal graph theory and derived

\begin{align*}
(4) & \quad N_2(p) > \log \left( \log \frac{p-1}{\log 4} + 1 \right) \quad \text{for all } p \equiv 1 \pmod{4}.
\end{align*}

A slight modification of his method leads to

\begin{align*}
(5) & \quad R_{1,2}(p) = N_{1,2}(p) > \left\lfloor \frac{\log(p-1)}{\log 4} + 1 \right\rfloor.
\end{align*}

Ruzsa [8, 9] proved that if \( m \) is a squarefree integer consisting exclusively of primes \( 1 \pmod{4} \) then

\begin{align*}
(6) & \quad N_2(m) \leq \sqrt{m},
\end{align*}

and conjectured that it holds for all squarefree integers.

Ebert [5] asked what is the value of \( R_{1,2}(p^n) \), where \( p \) is a prime and \( p^n \equiv 1 \pmod{4} \).

Recently the author proved [4] that in the case of primes \( 1 \pmod{4} \)

\begin{align*}
(7) & \quad R_{1,2}(p) \leq \sqrt{p} \quad \text{for all } p,
\end{align*}

and

\begin{align*}
(8) & \quad R_{1,2}(p) > \frac{1 - \varepsilon}{\log 4} \log p \quad \text{for } p > p_0(\varepsilon).
\end{align*}

In this paper the following will be proved.

**Theorem 1.** If \( p \equiv 1 \pmod{4} \) and \( n \geq 1 \) then \( R_{1,2}^*(p^n) \geq (R_{1,2}^*(p))^{\lfloor (n-1)/2 \rfloor + 1} \).

**Theorem 2.** If \( m \) is a squarefree integer then \( N_2(m) \ll m^\tau \) with \( \tau = \frac{1}{2} + 1/\log \log m \), the implied constant being absolute.

**Theorem 3.** If \( p \equiv 1 \pmod{4} \) and \( p \geq 29 \) then \( R_{1,2}(p) > \log p / \log 4. \)

**Corollary 1.** If \( m \) is a squarefree integer with all prime divisors \( 1 \pmod{4} \) then \( R_{1,2}(m) > C(m) \prod_{p \mid m} \log p \), where \( C(m) = (\log 4)^{-\omega(m)} \) and \( \omega(m) \) is the number of distinct prime factors of \( m \).

**Corollary 2.** \( \limsup_{m \to \infty} R_{1,2}(m)/m^{(\log \log \log m)/(3 \log \log m)} > 0. \)

### 2. Proof of Theorem 1

Let \( p \equiv 1 \pmod{4} \) and \( \mathcal{A}(p) = \{ \bar{r}_1, \bar{r}_2, \ldots, \bar{r}_k \} \) be the maximal set of residues modulo \( p \) satisfying conditions \((C_1)\) and \((C_2)\) with respect to the set \( R^*(p) \), so \( s = R_{1,2}^*(p) \) and \( 0 \in \mathcal{A}(p) \).
An easy calculation shows that the elements of the form
\[ \tilde{r}_0 + \tilde{r}_1 p^2 + \tilde{r}_2 p^4 + \cdots + \tilde{r}_{k} p^{2k}, \]
where \( \tilde{r}_i \in \mathcal{A}(p) \) and \( k = [(n-1)/2] \), are all distinct residues modulo \( p^n \) and satisfy \((C_1)\) and \((C_2)\). Counting the number of elements of the above form we see that \( R_{1, 2}(p^n) \geq (R_{1, 2}(p))^{k+1} \).

To prove the second part of the theorem we recall that if \( p \) is an odd prime and \((a, p) = 1\) then the number of solutions to \( x^2 \equiv a \pmod{p^n} \) is \( 1 + (a/p) \).

Suppose now \( p \equiv 3 \pmod{4} \), and assume \( R_{1, 2}(p^n) \geq 2 \). Let \( r_1 \) and \( r_2 \) be the residues in question such that \( r_1 \not\equiv r_2 \pmod{p^n} \). Let \( r_1 - r_2 \equiv p^\alpha c \pmod{p^n} \), where \( 0 \leq \alpha \leq n-1 \) and \((c, p) = 1\). By condition \((C_2)\), \( p^{\alpha} c \) is a quadratic residue modulo \( p^n \), so \( \alpha \) must be even, say \( \alpha = 2\beta \), and \( c \) must be a quadratic residue modulo \( p^{n-2\beta} \). Similar consideration of the difference \( r_2 - r_1 \) yields that \(-c\) is also a quadratic residue modulo \( p^{n-2\beta} \). Thus, by the above remark, both \( c \) and \(-c\) are quadratic residues modulo \( p \), which is impossible.

### 3. Proof of Theorem 2

Let \( m \) be squarefree, and set \( G_m(x) = \sum_{j=1}^{m-1} e(j^2 x / m) \) and \( g_m(x) = \sum_{a \in \mathcal{A}} e(ax / m) \), where \( \mathcal{A} \) is any set of residues modulo \( m \) that satisfies condition \((C_2)\) with respect to \( \mathcal{M}(m) \).

First we shall prove that \( \sum_{x=0}^{m-1} |g_m(x)|^2 G_m(x) = 0 \). We have
\[
\sum_{x=0}^{m-1} |g_m(x)|^2 G_m(x) = \sum_{x=0}^{m-1} \sum_{j=1}^{m-1} \sum_{a, a'} e\left(\frac{(a - a' + j^2 x)}{m}\right)
= \sum_{a \in \mathcal{A}} \sum_{j=1}^{m-1}\sum_{x=0}^{m-1} e\left(\frac{j^2 x}{m}\right) + \sum_{a \neq a'} \sum_{j=1}^{m-1}\sum_{x=0}^{m-1} e\left(\frac{(a - a' + j^2 x)}{m}\right)
= \Sigma_1 + \Sigma_2, \text{ say.}
\]

Since \( 1 \leq j \leq m-1 \) and \( m \) is squarefree, \( m \nmid j^2 \), so \( \Sigma_1 = 0 \). Also \( a - a' + j^2 \not\equiv 0 \pmod{m} \) for all \( a \neq a' \) and \( j \); thus, \( \Sigma_2 = 0 \). It follows that
\[
|g_m(0)|^2 G_m(0) = -\sum_{x=1}^{m-1} |g_m(x)|^2 G_m(x).
\]

Hence
\[
|\mathcal{A}|^2 (m - 1) \leq \sum_{x=1}^{m-1} |g_m(x)|^2 G_m(x).
\]

Since
\[
\sum_{j=1}^{m/(m, x)} e\left(\frac{j^2 x}{m}\right) = \sum_{j=1}^{m/(m, x)} e\left(\frac{j^2 x/(m, x)}{m/(m, x)}\right) = O\left(\left(\frac{m}{(m, x)}\right)^{1/2}\right),
\]
it follows that
\[
\sum_{j=1}^{m} e\left(\frac{j^2 x}{m}\right) = O(m^{1/2}(m, x)^{1/2}).
\]
From (9) and (10) it follows that

\[(11) \quad |\mathcal{A}|^2 m \ll m^{1/2} \sum_{x=1}^{m-1} |g_m(x)|^2 (x, m)^{1/2}.\]

We also have

\[(12) \quad \sum_{x=1}^{m-1} |g_m(x)|^2 = \sum_{x=1}^{m-1} \sum_{a, a' \in \mathcal{A}} e\left(\frac{(a - a')x}{m}\right) = \sum_{x=1}^{m-1} 1 + \sum_{a \neq a' x=1}^{m-1} e\left(\frac{(a - a')x}{m}\right) = |\mathcal{A}| (m - 1) - (|\mathcal{A}|^2 - |\mathcal{A}|) = |\mathcal{A}| m - |\mathcal{A}|^2 \leq |\mathcal{A}| m.

Further, since \( m \) is squarefree, we have

\[(13) \quad \sum_{x=1}^{m-1} |g_m(x)|^2 (x, m)^{1/2} = \sum_{x=1}^{m-1} |g_m(x)|^2 + \sum_{d|m} (d^{1/2} - 1) \sum_{(x, m) = d} |g_m(x)|^2
\]
\[= \sum_{x=1}^{m-1} |g_m(x)|^2 + \sum_{d|m} (d^{1/2} - 1) \sum_{x=1}^{m/d-1} |g_{m/d}(x)|^2\]
\[\leq m|\mathcal{A}| + \sum_{d|m} d^{1/2} \frac{m}{d} |\mathcal{A}|\]
\[\leq m|\mathcal{A}| \left(1 + \sum_{d|m} d^{-1/2}\right) \ll m|\mathcal{A}| 2^{\omega(m)}.\]

Combining (11) and (13) we obtain

\[|\mathcal{A}|^2 m \ll m^{1/2} m|\mathcal{A}| 2^{\omega(m)},\]

so \(|\mathcal{A}| \ll m^{1/2} 2^{\omega(m)}\). Since for a squarefree integer \( m \), one has \( \omega(m) \ll \log m / \log \log m \); finally

\[N_2(m) \ll m^\tau \quad \text{with} \quad \tau = \frac{1}{2} + \frac{1}{\log \log m}.\]

4. PROOF OF THEOREM 3

Here we follow the method of Sarkozy. We need the following lemma due to Erdos and Szekeres [3].

**Lemma.** Let \( K \) and \( N \) be positive integers and \( G_K \) a graph of \( K \) vertices. If

\[\left(\begin{array}{c} 2N - 2 \\ N - 1 \end{array}\right) \leq K\]

then either \( G_K \) or the complement of \( G_K \) contains a complete subgraph (clique) of \( N \) vertices. Let \( N \geq 12 \). It is known that

\[\left(\begin{array}{c} 2N - 2 \\ N - 1 \end{array}\right) < \frac{2^{2N-2}}{\sqrt{(N-1)\pi}} e^{1/24(N-1)} = f(N).\]
Let \( N_0 = \left[ \log K/ \log 4 + \log \log K/ \log 16 + 1 \right] \). With this choice
\[
f(N_0) \leq \frac{e^{1/24} \sqrt{\log 4}}{\sqrt{\pi}} \frac{K \sqrt{\log K}}{\sqrt{\log K + (\log \log K - \log 4)/2}} < K.
\]

Thus, it follows from the lemma that each graph \( G_K \) of \( K \) vertices or its complement contains a complete subgraph of \( \left[ \log K/ \log 4 + \log \log K/ \log 16 + 1 \right] \) vertices. Now let \( p \equiv 1 \pmod{4} \) be a prime. Define the graph \( G(p - 1)/2 \) of \( (p - 1)/2 \) vertices \( P_1, P_2, \ldots, P(p - 1)/2 \) that corresponds to all quadratic residues \( r_1, \ldots, r(p - 1)/2 \) modulo \( p \), \( 1 \leq r \leq p - 1 \). We connect vertices \( P_i \) and \( P_j \) \((i \neq j)\) iff \( (r_i - r_j)p \equiv 1 \). Since \( p \equiv 1 \pmod{4} \), the graph is not directed.

By the above either \( G(p - 1)/2 \) or its complement contains a complete graph of
\[
T = \left[ \frac{\log((p - 1)/2)}{\log 4} + \frac{\log\log((p - 1)/2)}{\log 16} + 1 \right]
\]
vertices, say \( P_1, P_2, \ldots, P_T \). Accordingly, either the set \( \mathscr{A} = \{ r_1, \ldots, r_T \} \) or \( \mathscr{B} = \{ sr_1, \ldots, sr_T \} \), where \( s \) is any fixed quadratic nonresidue modulo \( p \), satisfies both conditions \((C_1)\) and \((C_2)\). It follows that \( R_{1,2}(p) \geq T \), where
\[
T > \frac{\log(p - 1)}{\log 4} + \frac{\log\log((p - 1)/2)}{\log 16} - \frac{1}{2}.
\]

The inequality
\[
\frac{\log(p - 1)}{\log 4} + \frac{\log\log((p - 1)/2)}{\log 16} - \frac{1}{2} > \frac{\log p}{\log 4}
\]
is equivalent to
\[
(14) \quad \log \left( \frac{p - 1}{2} \right) > 4 \left( \frac{p}{p - 1} \right)^2.
\]

If \( p \geq 129 \) then \( 4[p/(p - 1)]^2 \leq 4(129/128)^2 < 4.07 \) and \( \log((p - 1)/2) \geq \log 64 > 4.15 \); thus, \((14)\) follows. The inequality \( T > \log p/\log 4 \) can easily be checked for all \( p \equiv 1 \pmod{4} \), \( 29 \leq p \leq 129 \).

5. Proof of Corollaries 1 and 2

Let \( m \) be a squarefree integer, \( m = p_1p_2 \cdots p_r \), \( p_i \equiv 1 \pmod{4} \), \( 1 \leq i \leq r \). For each \( p_i \) let \( \mathscr{A}^*(p_i) \subset \mathscr{A}^*(p_i) \) be a set of maximal cardinality satisfying \((C_1)\) and \((C_2)\) so that \( |\mathscr{A}^*(p_i)| = R^*_{1,2}(p_i) \). Every integer \( x \equiv r_i \pmod{p_i} \) for all \( 1 \leq i \leq r \) and all \( r_i \in \mathscr{A}^*(p_i) \) belongs to \( \mathscr{A}^*(m) \), where \( |\mathscr{A}^*(m)| = R^*_{1,2}(m) \). Thus Corollary 1 follows from the Chinese Remainder Theorem.

If \( m = p_1p_2 \cdots p_r \) is the product of the first \( r \) primes \( 1 \pmod{4} \) then \( \omega(m) \sim (\log m)/(2 \log \log m) \). Therefore,
\[
\sum_{i \leq \omega(m)} \log \log p_i \sim \frac{\log m \log \log m}{2(\log \log m)}.
\]
and by Corollary 1 for such integers \( m \)

\[
R_{1,2}(m) \gg \left( \frac{1}{\log 4} \right)^{(\log m)/(2 \log \log m)} \exp \left( \frac{\log m \log \log \log m}{2(\log \log m)} \right) \\
> \exp \left( \frac{\log m}{2 \log \log m} \left( \log \log \log m - \frac{1}{4} \right) \right) \\
> \exp \left( \frac{\log m \log \log \log m}{3 \log \log m} \right),
\]

where the implied constant is absolute and positive. This proves Corollary 2.

6. Final remarks

Based on the empirical data, Buell and Williams suggest \( R_{1,2}(p) \) \( (p \equiv 1 \text{ mod } 4) \) looking like \( c \log((p - 1)/2) \) for \( c \) ranging from 1.42 to 1.94. As we saw in the proof of Theorem 3, the value of \( R_{1,2}(p) \) is related to the size of a clique of the graph \( G_{(p-1)/2} \). Under the assumption that the graphs behave as if they were random graphs we can apply \([1, \text{Theorem 4, p. 257}]\) to get a good estimate of the expected sizes of cliques in the above graphs. Thus for almost all primes \( p \) these estimates have a constant \( c \) ranging from 1.44 to 2.16 on the low end and from 2.16 to 2.88 on the high end.

On the other hand, it is possible to have infinitely many exceptional primes for which the above estimations do not hold. To see this, let \( n(p) \) denote the smallest positive integer which is not a quadratic residue modulo a prime \( p \). Montgomery \([6]\) showed that if the Riemann hypothesis is true for all \( L \)-functions of real characters \( \chi \) then there is an \( \varepsilon > 0 \) such that \( n(p) > \varepsilon \log(p)(\log \log p) \) for infinitely many primes. A slight modification of his argument leads to the analog statement for primes \( p \equiv 1 \pmod{4} \). In such case, however, the set \( \{0, 1, 2, \ldots, n(p) - 1\} \) satisfies both conditions \((C_1)\) and \((C_2)\), giving \( R_{1,2}^*(p) \geq n(p) > \varepsilon'(\log(p)(\log \log p)) \) for infinitely many “exceptional” primes \( p \equiv 1 \pmod{4} \). This lower bound contradicts the conjecture of Buell and Williams.

We also point out that two different methods have been employed to calculate the lower bound for \( R_{1,2}(p) \)—one based on the Weil estimation \((9)\) and the other on the extremal graph theory \((\text{Theorem 3})\). Amazingly both yield almost the same result. It would be quite interesting to improve the constant \((\log 4)^{-1}\).

Finally we note that Theorem 2 can be generalized to the case of \( k \)-th power residues. Thus if \( N_{2,k}(m) \) denotes the maximal number of residues modulo a squarefree \( m \) that can be selected so that no difference between them is a \( k \)-th power residue then

\[
N_{2,k}(m) \ll m^\tau \quad \text{with} \quad m = 1 - \frac{1}{k} + \frac{1}{\log \log m}.
\]

Acknowledgment

The author thanks the referee for his helpful comments.
References

9. _____, private communication.

Department of Mathematics and Astronomy, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2
E-mail address: fabryko@ccu.umanitoba.ca