ON THE LEAST PAIR OF CONSECUTIVE QUADRATIC NON-RESIDUES

Adolf Hildebrand

1. Introduction. Let p be an odd prime and denote by $n_1(p)$ the least quadratic non-residue modulo p. It is a well-known consequence of Burgess' character sum estimate [1] that

$$n_1(p) \le p^{\theta+\epsilon} \quad (p \ge p_0(\epsilon)) \tag{1}$$

holds with $\theta = \theta_0 := 1/(4\sqrt{e})$ for every $\epsilon > 0$. (Here and in the sequel, $p_0(\epsilon)$ denotes a sufficiently large constant depending on ϵ, not necessarily the same at each occurrence.) A long-standing conjecture of Vinogradov asserts that one may take $\theta = 0$ in (1), but this seems to be very difficult, and up to date the exponent $\theta = \theta_0$ resulting from Burgess' estimate has not been improved upon.

One may ask whether a similar bound holds for $n_2(p)$, defined as the least positive integer n for which n and $n+1$ are both quadratic non-residues modulo p. It is easy to see that for all sufficiently large primes p such a pair $(n, n+1)$ exists, so that $n_2(p)$ is well-defined. Using Burgess' estimate, Elliott [2, Lemma 13] established the bound

$$n_2(p) \le p^{\theta+\epsilon} \quad (p \le p_0(\epsilon)) \tag{2}$$

with $\theta = 1/4$, and in a later paper [3] improved the exponent slightly to $\theta = 1/4 - e^{-10}/8$. We shall here prove the following.

THEOREM. (2) *holds with* $\theta = \theta_0 = 1/(4\sqrt{e})$.

Thus, we have the same upper bound for $n_2(p)$ as for $n_1(p)$. Since (trivially) $n_2(p) \ge n_1(p)$, any further improvement on the exponent in (2) would imply an improvement in the bound for $n_1(p)$ and would therefore seem to be very difficult.

2. A lemma. The proof of the theorem rests on Burgess' character sum estimate and an argument drawn from the author's paper [5]. In this paper, a sufficient condition on a set A of positive integers was given, which implies that A contains infinitely many pairs of consecutive integers. We require here a "finite" version of this result.

LEMMA. *For every* $\epsilon \in (0, 1]$ *there exist positive integers* $N_0(\epsilon)$ *and* $k_0(\epsilon)$ *with the following property: If* $N \ge N_0(\epsilon)$ *and* A *is a set of positive integers satisfying*

Received November 15, 1985.
The author was supported by NSF grant MCS 8108814(A04).
Michigan Math. J. 34 (1987).

(i) $$\sum_{\substack{n\le N\\ n\equiv l \bmod k\\ n\in A}} 1 \ge \epsilon\frac{N}{k} \quad (1\le k\le k_0(\epsilon),\ 0\le l\le k-1)$$

and

(ii) $$\left\{\frac{n}{d} : n\in A,\ d\mid n,\ d\le k_0(\epsilon)\right\}\subset A,$$

then A contains a pair of consecutive integers $\le N$.

Proof. Given $0\le\epsilon\le 1$, put $r=[2/\epsilon]+1$ and fix positive integers $t_1<t_2<\cdots<t_r$ satisfying

(3) $$t_j-t_i=(t_i,t_j) \quad (1\le i<j\le r).$$

The existence of such integers for every $r\ge 2$ has been first proved by Heath-Brown [4]; a simple construction is given in [5]. We shall prove the lemma with

(4) $$N_0(\epsilon)=\left[\frac{2t}{\epsilon}\right]+1, \quad k_0(\epsilon)=t,$$

where

$$t=\prod_{i=1}^{r} t_i.$$

Let $N\ge N_0(\epsilon)$ and a set $A\subset\mathbf{N}$ be given, such that hypotheses (i) and (ii) of the lemma are satisfied with $N_0(\epsilon)$ and $k_0(\epsilon)$ defined by (4). Consider the sets

(5) $$B_i=\{n\le N: t\mid n,\ n-t_i\in A\}, \quad 1\le i\le r.$$

If for some $i<j$, $B_i\cap B_j\ne\varnothing$, then there exists an integer $n\le N$ divisible by t and such that the numbers $n-t_i$ and $n-t_j$ belong to the set A. Since $(t_i,t_j)\mid t\mid n$, these two numbers are both divisible by (t_i,t_j). By (3) and hypothesis (ii) of the lemma, it follows that $(n-t_i)/(t_i,t_j)$ and $(n-t_j)/(t_i,t_j)$ are consecutive integers $\le N$ and both contained in A.

Thus, the conclusion of the lemma holds under the assumption that the sets (5) are not pairwise disjoint. To complete the proof of the lemma, we shall show by contradiction that this assumption is always satisfied.

Suppose that the sets (5) are pairwise disjoint. Denoting by $|B_i|$ the cardinality of B_i, we then have

(6) $$\sum_{i=1}^{r}|B_i|=\left|\bigcup_{i=1}^{r}B_i\right|\le\sum_{\substack{n\le N\\ t\mid n}}1\le\frac{N}{t}.$$

On the other hand, by (i) we have, for each $i\le r$,

$$|B_i|=\sum_{\substack{n\le N-t_i\\ n\equiv -t_i \bmod t\\ n\in A}}1\ge\sum_{\substack{n\le N\\ n\equiv -t_i \bmod t\\ n\in A}}1-1\ge\epsilon\frac{N}{t}-1,$$

and hence, by (4) and the hypothesis $N\ge N_0(\epsilon)$,

$$\sum_{i=1}^{r} |B_i| \geq r\left(\epsilon\frac{N}{t}-1\right) \geq \frac{r\epsilon N}{2t} > \frac{N}{t}.$$

This yields the desired contradiction to (6).

3. Proof of the theorem. Let $\epsilon > 0$ be given and p be an odd prime, which we shall later assume to be sufficiently large. For the proof of the theorem, we may clearly assume

$$0 < \epsilon \leq 1/100. \tag{7}$$

Let

$$N_1 = [p^{(1+\epsilon)/4}], \qquad N_2 = [p^{\theta_0+\epsilon}].$$

By Burgess' estimate [1] we have

$$\left| \sum_{n \leq N_1} \left(\frac{n}{p}\right) \right| \leq \epsilon N_1 \tag{8}$$

for $p \geq p_0(\epsilon)$, where (n/p) is the Legendre symbol. Using (8), we shall show that the hypotheses of the lemma are satisfied for the set $A = \{n \geq 1 : (n/p) = -1\}$ and a suitable $N \leq N_2$. The conclusion of the theorem then follows from that of the lemma.

For $x \geq 1$ let

$$S(x) = \sum_{\substack{q \leq x \\ q \in A}} \frac{1}{q},$$

where q denotes a generic prime. Obviously,

$$\sum_{\substack{n \leq N_1 \\ n \in A}} 1 \leq \sum_{\substack{q \leq N_1 \\ q \in A}} \sum_{\substack{n \leq N_1 \\ q \mid n}} 1 \leq N_1 S(N_1).$$

On the other hand, (8) implies

$$\sum_{\substack{n \leq N_1 \\ n \in A}} 1 = \frac{1}{2} \sum_{n \leq N_1} \left(1 - \left(\frac{n}{p}\right)\right) \geq \frac{1}{2} N_1 (1-\epsilon).$$

Hence we conclude

$$S(N_1) \geq \tfrac{1}{2}(1-\epsilon). \tag{9}$$

A standard prime number estimate yields

$$\begin{aligned} S(N_1) - S(N_2) \leq \sum_{N_2 < q \leq N_1} \frac{1}{q} &= \log \frac{\log N_1}{\log N_2} + O\left(\frac{1}{\log N_2}\right) \\ &= \log \frac{(1+\epsilon)/4}{\theta_0+\epsilon} + O\left(\frac{1}{\log p}\right) \\ &= \frac{1}{2} - \log \frac{1+4\epsilon\sqrt{e}}{1+\epsilon} + O\left(\frac{1}{\log p}\right). \end{aligned}$$

In view of (7), it follows that for sufficiently large p

$$S(N_1)-S(N_2)\le \tfrac{1}{2}-3\epsilon.$$

Combining this with (9), we obtain

$$S(N_2)\ge S(N_1)-(\tfrac{1}{2}-3\epsilon)\ge \tfrac{5}{2}\epsilon. \tag{10}$$

At this point, it is convenient to assume that

$$n_1(p)>p^{\epsilon/2}. \tag{11}$$

We shall deal with the remaining case, which is much easier, later on. (11) implies that the set A contains no integer $\le p^{\epsilon/2}$, so that $S(p^{\epsilon/2})=0$. From this and (10) we conclude that there exists an integer N satisfying $p^{\epsilon/2}\le N\le N_2$, for which

$$\tfrac{5}{2}\epsilon\le S(N)\le \tfrac{5}{2}\epsilon+\tfrac{1}{2}<\tfrac{4}{7},$$

where the last inequality follows from (7). We fix such an integer N.

For $k\le p^{\epsilon/2}$ and $0\le l\le k-1$ we have

$$\begin{aligned}
\sum_{\substack{n\le N\\ n\equiv l \bmod k\\ n\in A}} 1 &\ge \sum_{\substack{q\le N\\ q\in A}}\ \sum_{\substack{n\le N\\ n\equiv l \bmod k\\ q\mid n}} 1-\sum_{\substack{q,q'\le N\\ q,q'\in A}}\ \sum_{\substack{n\le N\\ n\equiv l \bmod k\\ qq'\mid n}} 1\\
&\ge \frac{N}{k}S(N)(1-S(N))-\sum_{q\le N} 1-\sum_{\substack{q,q'\le N\\ qq'\le N}} 1\\
&\ge \frac{N}{k}\,\frac{5}{2}\epsilon\left(1-\frac{4}{7}\right)+O\left(N\frac{\log\log(N+2)}{\log N}\right)\\
&= N\left(\frac{15}{14}\cdot\frac{\epsilon}{k}+O_\epsilon\left(\frac{\log\log(p+2)}{\log p}\right)\right).
\end{aligned}$$

If now $k\le k_0(\epsilon)$ and p is sufficiently large, then the last expression is $\ge \epsilon N/k$. Thus the first hypothesis of the lemma is fulfilled, whenever p is sufficiently large. The second hypothesis holds for $p\ge k_0(\epsilon)^{2/\epsilon}$ because of our assumption (11) and the multiplicativity of the Legendre symbol. Applying the lemma, we conclude that for $p\ge p_0(\epsilon)$ there exists a pair of consecutive integers $\le N$ belonging to A, that is, a pair of consecutive quadratic non-residues (modulo p) $\le N\le N_1\le p^{\theta_0+\epsilon}$, as asserted in the theorem.

It remains to deal with the case where (11) is not satisfied. We appeal to an elementary bound of Hudson [6], namely $n_2(p)\le (q_1-1)q_2$, where q_1 and q_2 denote the smallest and second smallest primes which are quadratic non-residues modulo p. If (11) is not satisfied, then $q_1=n_1(p)\le p^{\epsilon/2}$, and to obtain the desired bound for $n_2(p)$ it therefore suffices to show that

$$q_2\le N_3:=[p^{\theta_0+\epsilon/2}]. \tag{12}$$

Assume, to get a contradiction, that (12) does not hold. Then the set A contains, apart from q_1, no other prime $q\le N_3$. Therefore we have

$$S(N_1)=\sum_{\substack{q\le N_1\\ q\in A}}\frac{1}{q}=\frac{1}{q_1}+\sum_{\substack{N_3<q\le N_1\\ q\in A}}\frac{1}{q}=\frac{1}{q_1}+S_1,$$

say. As before, we see that for sufficiently large p

$$S_1\le\sum_{N_3<q\le N_1}\frac{1}{q}\le\frac{1}{2}-\epsilon. \tag{13}$$

In view of (9), it follows that

$$\frac{1}{q_1}=S(N_1)-S_1\ge\frac{\epsilon}{2};$$

that is, we have

$$2\le q_1\le 2/\epsilon. \tag{14}$$

Now, note that an integer $n\le N_1$ can have at most one prime factor $>N_3$ (since $N_3>\sqrt{N_1}$); hence n belongs to A (i.e., is a quadratic non-residue) if and only if it has the form

$$n=q_1^{2m}qn_1,\quad m\ge 0,\ N_3<q\le N_1,\ q\in A,\ (n_1,q_1)=1$$

or

$$n=q_1^{2m+1}n_1,\quad m\ge 0,\ \Big(n_1,\prod_{q\in A}q\Big)=1.$$

Using the sieve of Eratosthenes and the bound (14), it is easily seen that the number of such integers equals

$$\begin{aligned}\sum_{\substack{n<N_1\\ n\in A}}1&=N_1\Big(1-\frac{1}{q_1}\Big)\Big\{S_1\sum_{m\ge 0}q_1^{-2m}+(1-S_1)\sum_{m\ge 0}q_1^{-2m-1}\Big\}+O_\epsilon\Big(\frac{N_1}{\log N_1}\Big)\\&=N_1\Big\{S_1\Big(1-\frac{2}{q_1+1}\Big)+\frac{1}{q_1+1}\Big\}+O_\epsilon\Big(\frac{N_1}{\log p}\Big).\end{aligned}$$

By (13), the last expression is

$$\begin{aligned}&\le N_1\Big\{\Big(\frac{1}{2}-\epsilon\Big)\Big(1-\frac{2}{q_1+1}\Big)+\frac{1}{q_1+1}\Big\}+O_\epsilon\Big(\frac{N_1}{\log p}\Big)\\&=N_1\Big\{\Big(\frac{1}{2}-\epsilon\Big)+\frac{2\epsilon}{q_1+1}\Big\}+O_\epsilon\Big(\frac{N_1}{\log p}\Big)\le N_1\Big(\frac{1}{2}-\frac{\epsilon}{4}\Big)\end{aligned}$$

for sufficiently large p. It follows that

$$\sum_{n\le N_1}\Big(\frac{n}{p}\Big)=N_1-2\sum_{\substack{n\le N_1\\ n\in A}}1\ge\frac{\epsilon}{2}N_1.$$

But this contradicts Burgess' estimate, if p is sufficiently large. Hence (12) holds for $p\ge p_0(\epsilon)$, say, and the proof of the theorem is complete. □

REFERENCES

1. D. A. Burgess, *The distribution of quadratic residues and non-residues,* Mathematika 4 (1957), 106–112.
2. P. D. T. A. Elliott, *On the mean value of $f(p)$*, Proc. London Math. Soc. (3) 21 (1970), 28–96.
3. ———, *On the least pair of consecutive quadratic non-residues* (mod p). Proceedings of the number theory conference (Boulder, Colo., 1972), 75–79, Univ. Colorado, Boulder, Colo., 1972.
4. R. Heath-Brown, *The divisor function at consecutive integers,* Mathematika 31 (1984), 141–149.
5. A. Hildebrand, *On a conjecture of Balog,* Proc. Amer. Math. Soc. 95 (1985), 517–523.
6. R. Hudson, *The least pair of consecutive character non-residues,* J. Reine Angew. Math. 281 (1976), 219–220.

School of Mathematics
Institute for Advanced Study
Princeton, New Jersey 08540

Current address:
Department of Mathematics
University of Illinois
Urbana, Illinois 61801