THE DISTRIBUTION OF PRIMITIVE ROOTS

P. D. T. A. ELLIOTT

Notation. p and q are generic symbols for prime numbers.

$N(H, p)$ denotes the number of primes q, not exceeding H, which are primitive roots (mod p).

$g(p)$ denotes the least positive primitive root (mod p).

$g^*(p)$ is the least *prime* primitive root (mod p).

$\nu(m)$ denotes the number of distinct prime divisors of the integer m.

$\tau_k(m)$ is the number of ways of representing the integer m as the product of k integers, order being important.

$\pi(x, k, r)$ is the number of primes p, not exceeding x, which satisfy $p \equiv r \pmod{k}$; while $\pi(x)$ denotes the total number of $p \leqq x$.

$\log_m x$ denotes the mth iterated logarithmic function which is defined, when possible, by

$$\log_m x = \log(\log_{m-1} x), \quad m = 1, 2, \ldots, \qquad \log_0 x = x.$$

$[y]$ denotes the greatest integer not exceeding y.

$(n; \ldots)$ is the set of integers n which have the property $\ldots$, and

$\mathrm{Card}(n; \ldots)$ denotes its cardinality.

A, B, C will denote either sets, or sequences of integers. Occasionally, they will denote constants. For any real value of x, $A(x)$ denotes the number of integers in the sequence A which do not exceed x.

$c_1, c_2, \ldots$ will denote positive constants, and usually they will be absolute. It will be convenient to renumber constants from time to time.

1. Introduction. In a recent paper of Burgess and Elliott (**6**) it was shown that $g(p)$ is on average $O((\log p)^{2+\epsilon})$ for any fixed $\epsilon > 0$. Our aim is to investigate more closely the distribution of primitive roots to the various moduli. We essentially prove two theorems, and consider these in turn, beginning with the study of the function $N(H, p)$ introduced earlier.

A natural estimate for $N(H, p)$ is suggested by the following argument:

Let Γ be the group of reduced residue classes (mod p), and for each divisor d of $(p - 1)$ let Γ_d denote the subgroup consisting of those classes which are dth-powers. There is a natural homomorphism of the ring of integers onto the group Γ, and so onto the group Γ/Γ_d. For at any rate large values of H it seems reasonable to expect that the primes $q \leqq H$ are equidistributed in the

Received December 1, 1967.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

classes Γ/Γ_d. In terms of indices to a fixed primitive root (mod p) we write

$$\operatorname{Card}(q; q \leqq H, d|\operatorname{ind} q) \sim \pi(H)/d.$$

Thus, we see that it is likely that

$$N(H, p) = \sum_{d|(p-1)} \mu(d) \operatorname{Card}(q; q \leqq H, d|\operatorname{ind} q)$$
$$\sim \pi(H) \sum_{d|(p-1)} \frac{\mu(d)}{d} \sim \frac{\phi(p-1)}{p-1} \pi(H).$$

In order to consider $N(H, p)$, it is in fact more convenient to use Selberg's sieve method. Moreover, we can only prove the distribution for almost all (in a certain sense) prime moduli. More exactly, we prove the following theorem.

THEOREM 1. *Let ϵ and B be arbitrary positive constants. Then there is a set of primes E, and a positive constant $F = F(\epsilon, B)$, so that for all p not in E the estimate*

$$N(H, p) = \frac{\phi(p-1)}{p-1} \pi(H)\left\{1 + O\left(\frac{1}{(\log H)^B}\right)\right\}$$

holds uniformly for

$$H \geqq \exp(F \log_2 p \log_3 p).$$

Moreover, the sequence E satisfies

$$E(x) = O(x^\epsilon)$$

for all large values of x.

We can vary our conditions on H and E so as to lay more emphasis on obtaining a "thin" set E. For example, we have the following result.

THEOREM 2. *If we demand that the estimate for $N(H, p)$ in Theorem* 1 *holds in the range $H \geqq p^\epsilon$, then we can find a positive constant G such that*

$$E(x) = O((\log x)^G).$$

It seems very likely that in the second of these results the set E is actually empty. In particular, there is a long-standing conjecture that

$$g(p) = O(p^\epsilon)$$

for any fixed $\epsilon > 0$. The first non-trivial estimate in this direction was that of I. M. Vinogradov (see **18**), who obtained

$$g(p) = O(p^{1/2+\epsilon}), \qquad \epsilon > 0.$$

This exponent was improved by Burgess (**5**) to $\frac{1}{4} + \epsilon$, which is the best to date.

There is, on the other hand, a famous conjecture of Artin (**2**) concerning those primes for which a given integer a is a primitive root. In particular, he conjectured that the number of primes not exceeding x for which 2 is a primitive root is

$$(1 + o(1))A\pi(x) \qquad (x \to \infty),$$

where the constant A (known as Artin's constant) is defined by

$$A = \prod_p \left(1 - \frac{1}{p(p-1)}\right).$$

A modified constant is suggested for a general integer a. It was shown by Hooley (**16**) that this conjecture is certainly satisfied if a hypothesis of Riemann type concerning abelian algebraic number fields is satisfied.

By also using a form of extended Riemann hypothesis, namely that certain L-series formed with Dirichlet characters have all their zeros in the critical strip on the line $s = \frac{1}{2}$, but by an altogether different method, Ankeny (**1**) showed that

$$g^*(p) \leqq c_1(2^{\nu(p-1)} \log p \log(2^{\nu(p-1)}\log p))^2.$$

This was later improved by Wang (**22**) to

$$g(p) \leqq c_2\nu(p-1)^6\log^2 p.$$

It seems reasonable from these remarks to conjecture that $g^*(p) = 2$ infinitely often. The result of Burgess and Elliott (**6**) shows that a result of this type is true if we demand only that

$$g^*(p) = O((\log p)^{2+\epsilon}), \quad \epsilon > 0.$$

Our second main aim will be to improve this to the *unconditional*.

THEOREM 3. *For infinitely many primes p the inequality*

$$g^*(p) < 475(\log p)^{8/5}$$

is satisfied.

It will be convenient to prove somewhat more. Let α be a real number satisfying $0 < \alpha < 1$. Then for each $x \geqq 3$ we define

$$T_\alpha(x) = (p; p \leqq x, q|(p-1) \Rightarrow q = 2 \text{ or } q > x^\alpha).$$

It is well known that for certain values of α we can find a constant $D > 0$ so that

$$\operatorname{Card} T_\alpha(x) > x(\log x)^{-D}.$$

A result of this type was first proved by Rényi (**21**). For our present purposes we do not need explicit values. However, we can appeal to the recent result of Halberstam, Jurkat, and Richert (**15**). By using a result of Bombieri (**4**), they showed that one may take the values $\alpha = \frac{1}{4} - \epsilon$, $D = 2 + \epsilon$; see (**15**, Theorem 1). Indeed, they showed slightly more. By a system of $T_\alpha(x)$ we shall understand an infinite sequence of values of x, which are unbounded, so that $T_\alpha(x)$ exists for each value of x, but with the same values of α and D.

THEOREM 4. *Let α be associated with a system of $T_\alpha(x)$. Then the inequality*

$$\liminf g^*(p)(\log p)^{-2/(1+\alpha)} \leqq \left(\frac{8}{e}\left[\frac{1}{\alpha}\right]^2\right)^{2/(1+\alpha)}$$

is satisfied.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

COROLLARY. *For any* $\epsilon > 0$ *the inequality*

$$g^*(p) < (\log p)^{8/5+\epsilon}$$

is satisfied infinitely often.

The above corollary is obtained by using the values of α and D mentioned earlier. We would obtain Theorem 3 if we could choose $\alpha = \frac{1}{4}$, for then

$$(128e^{-1})^{8/5} = 475\cdot 97\ldots.$$

We shall show that we can effectively do this.

Finally, we note that the proofs of these theorems depend upon the large sieve in certain forms. We shall need to augment this in various ways, and it will sometimes be applied more than once in the same problem. We confine ourselves here to the remark that in the proof of Theorem 3 we use the sequence

Large Sieve (Bombieri) → Small Sieve (Halberstam,
Richert, Jurkat) → Large Sieve,

each of these steps being intrinsic in the method. We operate upon the sequence of primes, and consider $g^*(p)$ rather than $g(p)$, since this leads to simpler and stronger results. More general problems can be dealt with by the methods given.

My thanks are due to the referee whose comments enabled certain details to be simplified, and for his short proof of Lemma 10, which is given.

2. Various lemmas.

LEMMA 1. *Let* $a_1, a_2, \ldots$ *be a sequence of complex numbers. Then we have the inequality*

$$\sum_{p\leqq x}\ \sum_{\chi\neq\chi_0 \pmod p} \left|\sum_{n\leqq H} a_n\chi(n)\right|^2 \leqq c_3(x^2+H)\sum_{n\leqq H}|a_n|^2,$$

where χ *runs over all non-principal characters* (mod p).

Proof. This result follows from any of the well-known forms of the large sieve. A short proof is given in (**11**).

LEMMA 2. *Let* $\{\beta_{d,p}\}$ *be a double sequence of real numbers satisfying*

$$0 \leqq \beta_{d,p} \leqq \phi(d)^{-1}.$$

Let

$$T_p = \sum_{\substack{d\mid(p-1);\\ d>1}} \beta_{d,p}\sum_{\chi_d}\left|\sum_{q\leqq H}\chi_d(q)\right|,$$

where χ_d *runs through all characters* (mod p) *which are of order d. Furthermore, we set*

$$\rho(p) = \sum_{\substack{d\mid(p-1);\\ \beta_{d,p}>0}} 1,$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

and for real $\lambda > 0$, $R > 0$, $x \geqq 3$,

$$S = S(\lambda, R) = (p; p \leqq x, \rho(p) < R, T_p > \lambda^{-1}\pi(H)).$$

Then if $2 \leqq H \leqq x^2$, *the inequality*

$$\operatorname{Card} S \leqq c_4\left(\frac{\log x}{\log H}\right)^{1/2} \exp\left\{\frac{\log(x^2H)\log(\lambda^2R^2\log x)}{\log H}\right\}$$

is satisfied.

Proof. Lemma 2 has been proved by Burgess and Elliott (**6**, Lemma 2), under the condition $2 \leqq H \leqq x^{1/2}$. The same proof deals with the cases $2 \leqq H \leqq x^2$.

LEMMA 3. *Let* $a_1, a_2, \ldots$ *be complex numbers, and let* $x \geqq 2$, $L \geqq 2$ *hold. Then in terms of the Legendre symbol, the inequality*

$$\sum_{p \leqq x}\left|\sum_{n \leqq L} a_n\left(\frac{n}{p}\right)\right|^2 \leqq 4x \sum_{\substack{m,n \leqq L;\\ mn = t^2, 2t^2}}\sum |a_m a_n| + O\left(L \log L\left(\sum_{n \leqq L} |a_n|\right)^2\right)$$

is satisfied.

Proof. This result essentially replaces the factor $(x^2 + L)$ of Lemma 1 by $(x + L^2 \log L)$ for the particular character in question, and hence enables us to deal with it more effectively. For a proof of Lemma 3 and various generalizations, we refer the reader to (**9**). We need this lemma for Theorems 3 and 4 only.

LEMMA 4. *Let* $a_1, a_2, \ldots, a_N$ *be a sequence of positive integers, and let* $q_1, \ldots, q_s$ *be a sequence of primes which satisfy*

$$q_1 < q_2 < \ldots < q_s \leqq y, \qquad q_1 \ldots q_s = Q,$$

for a real number y. *Let* $f(d)$ *be a multiplicative function of* d, *satisfying*

$$p/f(p) = O(1)$$

uniformly for all primes p, *so that if* $d|Q$, *then*

$$\operatorname{Card}(a_n; a_n \equiv 0 \pmod{d}) = N/f(d) + R_d,$$

where R_d *may depend upon* N. *Define* $I(N, y)$ *to be the number of members* a_n *of the sequence which are not divisible by any of the primes* q_j. *Then we have* (*for any* $z \geqq y$ *and* $\epsilon > 0$) *the estimate*

$$I(N, y) = N\prod_{p|Q}\left(1 - \frac{1}{f(p)}\right)[1 + O(\exp(-c_5 \log z/\log y))] + O\left((\log z)^\epsilon \sum_{\substack{d \leqq z^3;\\ d|Q}} \tau_3(d)|R_d|\right).$$

Proof. This result proves valuable in the probabilistic theory of numbers, and has become known as the *first fundamental lemma.* It is derivable from any of the local sieve methods, and in particular from that of A. Selberg (for an account of his method, see Halberstam and Roth (**14**)). It is possible to improve the term $\Delta = \log z/\log y$ to $\Delta \log \Delta$, and this would lead to a marginal weakening of the lower bound on H needed in Theorem 1. To do this would introduce complications however, therefore we use the more readily obtainable result stated above. A detailed proof was given by Barban (**3**), and under slightly stronger conditions, by Kubilius (**17**, Lemma 4).

We shall use Lemma 4 in the proof of Theorems 1 and 2. For the remaining theorems we need a stronger form of the sieve result, and for completeness we define some functions anew.

From the previous lemma we keep the notation

$$a_1, a_2, \ldots, a_N,$$

for a sequence of positive integers. For each positive integer k, and each real value of z satisfying $z \geqq 2$ we set

$$P_k(z) = \prod_{\substack{p<z;\\ p\nmid k}} p.$$

Let $\gamma(d)$ be a function (corresponding to $d/f(d)$ in Lemma 5) which is multiplicative, and which satisfies the following four conditions:

(i) $1 \leqq \gamma(p) < p$, $(p \nmid k)$;

(ii) $\sum_{p \geqq z} (\gamma(p) - 1)/p = O(1/\log z)$;

(iii) there is a function $\eta(x, d) \geqq 0$ defined for $x > 1$, and a real number $X > 1$, such that

$$\left|\operatorname{Card}(a_n; a_n \equiv 0 (\operatorname{mod} d)) - \frac{\gamma(d)}{d} X\right| \leqq \eta(X, d) \quad \text{if } (d, k) = 1;$$

(iv) There is a constant α satisfying $0 < \alpha \leqq 1$, and for each real number $U > 15/14$ a function $\beta(X)$ satisfying $0 < \beta(X) = O((\log x)^{1/2})$ so that the inequality

$$\sum_{d \leqq X^\alpha/\beta(X)} \mu^2(d) 3^{\nu(d)} \eta(X, d) = O(X(\log X)^{-U})$$

is satisfied. We can then state the following lemma.

LEMMA 5. *Let*

$$\Gamma_k(z) = \prod_{p<z, p\nmid k} \left(1 - \frac{\gamma(p)}{p}\right),$$

and let $S_k(z)$ denote the number of integers a_i which are prime to $P_k(z)$. Define $\omega(u)$, $\rho(u)$ to be the solutions of the differential difference equations

$$\omega(u) = u^{-1}, \quad \rho(u) = 1, \qquad (0 < u \leqq 2),$$

$$(u\omega(u))' = \omega(u-1), \quad (u-1)\rho'(u) = -\rho(u-1), \qquad (u \geqq 2).$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

Then we can find a constant $c > 0$, and a real function f, so that the inequality

$$\frac{S_k(z)}{X\Gamma_k(z)} \geqq f\left(\alpha \frac{\log X}{\log z}\right) - c\frac{\log\log 3k}{(\log X)^{1/14}}$$

holds for all $z \leqq X$. In particular, we can take

$$f(u) = e^{\gamma}(\omega(u) - u^{-1}\rho(u)),$$

where γ denotes Euler's constant. This function is weakly increasing in u, and satisfies

$$f(u) = \begin{cases} 0 & \text{if } 0 < u \leqq 2, \\ \dfrac{2e^{\gamma}}{u}\log(u-1) & \text{if } 2 < u \leqq 3. \end{cases}$$

Proof. This result is one half of (**15**, Theorem 1). No details are given in that paper, but they are to appear in a monograph by Halberstam and Richert (**13**) dealing with sieve methods.

LEMMA 6 (*Siegel-Walfisz*). *Let a number $A > 0$ be given. Then there is a positive constant $c = c(A)$, so that the estimate*

$$\pi(x, k, l) = \frac{\operatorname{li}(x)}{\phi(k)}\{1 + O(e^{-c\sqrt{\log x}})\}$$

holds uniformly for all l prime to k, for all $k \leqq (\log x)^A$.

Proof. A proof of this well-known result is given in (**20**, Chapter IV, *Satz* 8.3).

LEMMA 7. *For any number $U > 0$ there is a number $V = V(U) > 0$, so that the inequality*

$$\sum_{d \leqq x^{1/2}(\log x)^{-V}} \max_{y \leqq x} \max_{(l,d)=1} \left| \pi(y, d, l) - \frac{\operatorname{li}(y)}{\phi(d)} \right| = O(x(\log x)^{-U})$$

is satisfied for all $x \geqq 3$.

Proof. This inequality, with $V = 4A + 40$, is a theorem of Bombieri (**4**). The proof depends upon the large sieve. Recently, a proof has been given by Gallagher (**12**) with $V = 16A + 103$. His proof also uses the large sieve, but is much simpler than that of Bombieri.

LEMMA 8. *Let ϵ and A be given positive constants. Then there is a (possibly empty) sequence of primes P so that if $x \geqq 3$, and q is a prime number not lying in P, then*

$$\pi(x, q, r) = \frac{\operatorname{li}(x)}{\phi(q)}\left\{1 + O\left(\frac{1}{(\log x)^A}\right)\right\}$$

holds for all r prime to q, and all $x \geqq q^{3+\epsilon}$. Moreover, there is a further constant D, depending upon ϵ and A, so that for all $y \geqq 3$ the inequality

$$P(y) = O((\log y)^D)$$

is satisfied.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

Proof. We can easily deduce the present result from (**8**, Theorem 2). There, a similar theorem is proved but with P replaced by a finite set P' of primes, none of which exceeds x, and for which

$$P'(x) = O((\log x)^{D-1})$$

holds. If we denote by P_j the set P' obtained for the value $x = 2^j$, and by P_0 the empty set, we need only take

$$P = \bigcup_{j=0}^{\infty} (P_{j+1} - P_j)$$

to obtain Lemma 8. We use this result only in the proof of Theorem 2.

LEMMA 9. *There are positive constants c_6, c_7, so that for any coprime integers k, r, and all real x satisfying $k^{c_6} \leqq x$, the lower bound*

$$\pi(x, k, l) \geqq k^{-c_7}\pi(x)$$

is satisfied.

Proof. This theorem is due to Fogels (**10**). The proof is complicated, developing some considerations of Linnik (**19**) in which he was concerned with the size of the least prime in an arithmetic progression. We use it only in some discussion at the end of this paper.

LEMMA 10. *Let S be a set of n distinct elements. For any positive integer m let $T(n, m)$ denote the number of ordered $2m$-tuples of elements of S with the property that any element which occurs in a tuple occurs an even number of times. Then the upper bound*

$$T(n, m) \leqq (nm)^n$$

is satisfied.

Proof. Since the proof of this result of Davenport and Erdős (**7**) is simple we give it here. More exactly, we give the variant suggested by the referee.

We begin by noting that if $(y_1, \ldots, y_{2m})$ is a $2m$-tuple with each element occurring an even number of times, then there is an integer k satisfying $1 \leqq k \leqq 2m - 1$ so that $y_k = y_{2m}$. If we denote by $\hat{y}_j$ the removal of the jth coordinate element, it is clear that

$$(y_1, \ldots, \hat{y}_k, \ldots, \hat{y}_{2m})$$

is a $2(m - 1)$-tuple of the same type. Since y_{2m} and k can be chosen in at most $n(2m - 1)$ ways, we conclude by induction that

$$T(n, m) \leqq n(2m - 1)T(n, m - 1) \leqq n^m(2m - 1)(2m - 3) \ldots 1$$
$$= n^m(2m)!/2^m m! \leqq (nm)^m,$$

as required.

LEMMA 11. *Let G be a fixed positive constant. Define a sequence of real numbers by*

$$u_1 = 2, \qquad u_{n+1} = u_n(1 + (\log u_n)^{-G}).$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

Then there is a constant $c_8 > 0$ so that for all integers $n \geqq 1$,

$$u_n > c_6 \exp(\tfrac{1}{2} n^{1/(G+1)}).$$

Proof. We assume n large in what follows. The result can then be obtained for small values of n by adjusting the value of the constant c_8.

The sequence $\{u_n\}$ is an increasing one for $n \geqq 2$, so that if then

$$\max_{1 \leqq m \leqq n-1} u_m > \exp(n^{1/(G+1)}),$$

the desired inequality is immediate. Otherwise, when $m = 1, 2, \ldots, n-1$, $(\log u_m)^{-G} \geqq n^{-G/(G+1)}$ and we conclude that

$$u_n \geqq \prod_{m=1}^{n-1} (1 + n^{-G/(G+1)}) \geqq \exp(\tfrac{1}{2} n^{1-G/(G+1)}) = \exp(\tfrac{1}{2} n^{1/(G+1)}).$$

Using these lemmas we can prove our main theorems.

3. Proof of Theorem 1. To begin with we shall keep p and H fixed and let $q_1, \ldots, q_s$ be the prime divisors of $(p-1)$ which satisfy

$$q_1 < q_2 < \ldots < q_s \leqq y$$

for a real number y to be specified later. We shall use the function "ind" taken with respect to a fixed primitive root (mod p), and χ will denote a typical character (mod p). Finally, if $p \leqq H$, we do not count p in $\pi(H)$. This clearly does not affect the statement of the theorem.

We mimic our introductory remarks. For each divisor d of $(p-1)$, and any character (mod p) of order d, we have the estimate

$$\operatorname{Card}(q; q \leqq H, d|\operatorname{ind} q) = \frac{\pi(H)}{d} + \frac{1}{d} \sum_{r=1}^{d-1} \sum_{q \leqq H} \chi^r(q).$$

If we set $a_j = \operatorname{ind} q_j$, q_j the jth natural prime number, we can regard the present situation as an example of that described in Lemma 4. We have $N = \pi(H)$ and $f(d) = d$, while

$$R_d = \frac{1}{d} \sum_{\substack{m|d;\\ m<d}} \sum_{\substack{r=1;\\ (r,d)=m}}^{d} \sum_{q \leqq H} \chi^r(q),$$

so that

$$|R_d| \leqq \frac{1}{d} \sum_{\substack{t|d;\\ t>1}} \sum_{\chi_t} \left| \sum_{q \leqq H} \chi(q) \right|,$$

where the sum over χ_t runs over all the characters (mod p) which have order t. These estimates are then in a form suitable for application, and we deduce that

$$(1) \quad I(\pi(H), y) = \pi(H) \prod_{\substack{q|(p-1);\\ q \leqq y}} \left(1 - \frac{1}{q}\right)[1 + O(\exp(-c_5 \log \pi(H)/\log y))] + O\left((\log z)^{\epsilon} \sum_{\substack{d \leqq z^3;\\ d|Q}} \tau_3(d)|R_d|\right).$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

Next, we simplify the last of these error terms by noting that it is majorized by

$$(\log z)^{\epsilon} \sum_{\substack{d \leqq z^3; \\ d \mid Q}} \tau_3(d) \frac{1}{d} \sum_{\substack{t \mid d; \\ t>1}} \sum_{\chi_t} \left| \sum_{q \leqq H} \chi(q) \right|$$

$$\leqq (\log z)^{\epsilon} \sum_{\substack{1 < t \leqq z^3; \\ t \mid Q}} \frac{\tau_3(t)}{t} \sum_{\chi_t} \left| \sum_{q \leqq H} \chi(q) \right| \sum_{m \leqq z^3 t^{-1}} \frac{\tau_3(m)}{m},$$

and, since the innermost sum is

$$O\left(\frac{z^3}{t} (\log z)^3\right),$$

by

$$(\log z)^4 \sum_{\substack{1 < t \leqq z^3; \\ t \mid Q}} \frac{1}{t} \sum_{\chi_t} \left| \sum_{q \leqq H} \chi(q) \right| . \tag{2}$$

For fixed positive real numbers K, L we define

$$J(H, x) = \left(p; \tfrac{1}{2}x < p \leqq x, \sum_{t \mid (p-1)}' \frac{1}{t} \sum_{\chi_t} \left| \sum_{q \leqq H} \chi(q) \right| > H(\log H)^{-1}\right).$$

Note that H is regarded as fixed. The values of t in the outer summation are restricted by $t|Q$ and

$$1 < t \leqq \max\{(\log H)^K, (\log x)^K\},$$

where Q has the appropriate value for each prime p. We shall estimate the cardinality of this set in order to prove our theorem, and investigate several cases.

Case 1. $\exp(\frac{1}{2}F \log_2 x \log_3 x) < H \leqq x^2$.

We apply Lemma 2 with

$$\beta_{d,p} = \begin{cases} 1/d & \text{if } 1 < d \leqq \max\{(\log H)^K, (\log x)^K\}, d|Q, \\ 0 & \text{otherwise,} \end{cases}$$

$$R = 2 \max\{(\log H)^K, (\log x)^K\},$$

$$\lambda = (\log H)^L,$$

and thus obtain the estimate

$$\text{Card } J(H, x) \leqq c_4 \left(\frac{\log x}{\log H}\right)^{1/2} \exp\left(\frac{4(K + L + 2) \log x \log_2 x}{\log_2 x \log_3 x}\right) < x^{\epsilon/2}. \tag{3}$$

Case 2. $x^2 < H \leqq \exp(x^{\delta})$, $6(L + 1)\delta = 1$.

In this case, Lemma 2 is of no value, being dependent for its success upon H being small. However, we can use the ideas involved in the proof, together with Lemma 1 as follows:

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

When p does not exceed x,

$$\sum_{t|(p-1)} \sum_{\chi_t} \beta_{t,p}^2 \leqq \sum_{t|(p-1)} \frac{1}{\phi(t)} \leqq \prod_{q|(p-1)} \left(1 + \frac{1}{q-1} + \frac{1}{q(q-1)} + \dots\right)$$
$$\leqq c_9 \prod_{2<q\leqq x} \left(1 - \frac{1}{q-1}\right)^{-1} \leqq c_{10} \log x,$$

and thus in the notation of Lemma 2,

$$\sum_{p\leqq x} T_p^2 \leqq c_{10} \log x \sum_{p\leqq x} \sum_{\substack{t|(p-1);\\ t>1}} \sum_{\chi_t} \left| \sum_{q\leqq H} \chi(q) \right|^2.$$

Hence, for any positive values of λ, R, and any $H \geqq 2$,

$$\text{Card}\, J(H, x) \leqq c_{11}\lambda^{-2}(x^2H^{-1} + 1) \log x \cdot \log H,$$

and in the present circumstances,

$$\text{Card}\, J(H, x) = O(x^{(2L+1)\delta} \log x) < x^{\epsilon/2}.$$

Case 3. $\exp(x^\delta) < H$.

In this case we shall show that $J(H, x)$ is empty. Indeed, since

$$p \leqq x < (\log H)^{1/\delta}$$

is satisfied, we may apply the Siegel-Walfisz theorem, which shows that

$$\sum_{q\leqq H} \chi(q) = \sum_{\substack{r=1;\\ (r,q)=1}}^{q} \chi(r) \frac{\text{li}(H)}{\phi(q)} [1 + O(e^{-c\sqrt{\log H}})]$$
$$= O(He^{-c\sqrt{\log H}}).$$

It is therefore clear that for any prime $p \leqq x$,

$$\sum_{t|(p-1)}{}' \frac{1}{t} \sum_{\chi_t} \left| \sum_{p\leqq H} \chi(p) \right| = O(H(\log H)^D e^{-c\sqrt{\log H}}) < H(\log H)^{-L},$$

so that $J(H, x)$ is empty.

In an analogous manner we define sets

$$V(H, x) = \left(p; \tfrac{1}{2}x < p \leqq x, \sum_{l|(p-1)}{}'' \frac{1}{l} \sum_{\chi_l} \left| \sum_{q\leqq H} \chi(q) \right| > H(\log H)^{-L}\right),$$

where l runs through primes, and satisfies a condition complementary to that on t, namely

$$l > \max\{(\log H)^K, (\log x)^K\}.$$

Now if $K > 1$, and we adopt the same values for R and λ, and define

$$\beta_{d,p} = \begin{cases} 1/d & \textit{if } d|(p-1) \textit{ and } d \textit{ is prime and satisfies} \\ & \qquad d > \max\{(\log H)^K, (\log x)^K\}, \\ 0 & \textit{otherwise}, \end{cases}$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

then, when p satisfies $p \leqq x$,

$$\rho(p-1) \leqq \sum_{l|(p-1)} 1 \leqq \frac{\log x}{\log 2} < (\log x)^K \leqq R$$

is satisfied. We can therefore apply the above arguments to show that whatever the value of $H > \exp(\frac{1}{2}F \log_2 x \log_3 x)$,

$$\text{(4)} \qquad \operatorname{Card} V(H, x) < x^{\epsilon/2}.$$

We define the above sets for each member of the sequence $\{u_r\}$ constructed in Lemma 11 with $G = B + 2$ which satisfies

$$\exp(\tfrac{1}{2}F \log_2 x \log_3 x) < u_r,$$

and set

$$W(x) = \cup_r J(u_r, x) \cup V(u_r, x), \qquad W^* = (p; p \leqq 2^k),$$

where k is chosen so that the sets J and V are well-defined for $x \geqq 2^k$. Finally we define

$$E = W^* \cup \bigcup_{m=k}^{\infty} W(2^m).$$

This, we maintain, is a set which has the properties stated in Theorem 1.

We obtain an estimate for $E(t)$ by means of (3), (4), and Lemma 11. Since $J(u_r, x)$ and $V(u_r, x)$ are empty when u_r exceeds $\exp(x^\delta)$ we see that

$$\operatorname{Card} W(x) \leqq 2x^{\epsilon/2} \sum_{u_r \leqq \exp(x^\delta)} 1 \leqq 2x^{\epsilon/2} \sum_{n^{1/(G+1)} \leqq x^\delta} 1 < x^\epsilon (\log x)^{-1}$$

provided $4\delta(B + 3) < \epsilon$. It follows immediately that for any $t \geqq 2$,

$$E(t) = O\left(\sum_{2^{m-1} \leqq t} \frac{2^{m\epsilon}}{m}\right) = O(t^\epsilon),$$

as desired.

It is our next step to show that when p does not lie in E, and H has one of the values u_r, then $N(H, p)$ satisfies an estimate of the type stated in Theorem 1. We can assume without loss of generality that $x = 2^j$ is large, and that $\frac{1}{2}x < p \leqq x$, $p \notin W(x)$.

We see from our definition of $W(x)$, and the estimate (1) with

$$z^3 = \max\{(\log H)^K, (\log x)^K\}, \qquad y = \exp(\log H \cdot (M \log_2 H)^{-1})$$

that

$$\text{(5)} \qquad I(\pi(H), y) = \pi(H) \prod_{\substack{q|(p-1);\\ q \leqq y}} \left(1 - \frac{1}{q}\right)[1 + O((\log H)^{-B})]$$

provided $L + 6 \geqq B$, and M is sufficiently large but fixed. Moreover, for the same primes,

$$\sum_{l|(p-1)}'' \left| \operatorname{Card}(q; q \leqq H, l|\operatorname{ind} q) - \frac{\pi(H)}{l} \right| \leqq \sum_{l|(p-1)}'' \frac{1}{l} \sum_{\chi_l} \left| \sum_{q \leqq H} \chi(q) \right|$$

$$\leqq \sum_{l|(p-1)}'' \frac{1}{l} H(\log H)^{-L} \leqq 2H(\log H)^{-L}$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

if $D \geqq 1$. We therefore see that

$$\sum_{l|(p-1)}'' \operatorname{Card}(q; q \leqq H, l|\operatorname{ind} q) < 2H(\log H)^{-L} + \pi(H) \sum_{l|(p-1)}'' \frac{1}{l} < 3H(\log H)^{-L}$$

provided $H > \exp(\frac{1}{2}F \log_2 x \log_3 x)$ for a sufficiently large, but otherwise fixed, value of F.

Putting these results together we deduce:

$$\left| I(\pi(H), p-1) - \pi(H) \prod_{q|(p-1)} \left(1 - \frac{1}{q}\right) \right|$$
$$\leqq |I(\pi(H), p-1) - I(\pi(H), y)| + \left| I(\pi(H), y) - \pi(H) \prod_{\substack{q|(p-1);\\ q \leqq y}} \left(1 - \frac{1}{q}\right) \right|$$
$$+ \left| \prod_{\substack{q|(p-1);\\ q \leqq y}} \left(1 - \frac{1}{q}\right) - \prod_{q|(p-1)} \left(1 - \frac{1}{q}\right) \right| \pi(H)$$
$$= \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3,$$

where

$$\sum\nolimits_1 \leqq \sum_{l|(p-1)}'' \operatorname{Card}(q; q \leqq H, l|\operatorname{ind} q) < H(\log H)^{-L},$$
$$\sum\nolimits_2 = O(H(\log H)^{-L}) \quad \text{by (5)},$$
$$\sum\nolimits_3 \leqq \pi(H) \prod_{\substack{q|(p-1);\\ q \leqq y}} \left(1 - \frac{1}{q}\right) \sum_{l|(p-1)}'' \frac{1}{l} < H(\log H)^{-L}.$$

Thus, reverting to the notation $H = u_r$ we have shown that

$$N(u_r, p) = \pi(u_r) \frac{\phi(p-1)}{p-1} \left[1 + O\left(\frac{1}{(\log u_r)^B}\right)\right]$$

for all $u_r \geqq \exp(\frac{2}{3}F \log_2 p \log_3 p)$. (The replacing of x by p in this condition is valid for all large values of x, since we have replaced $\frac{1}{2}F$ by $\frac{2}{3}F$, and since $p > \frac{1}{2}x$.)

We can now complete the proof of Theorem 1. For any

$$H > \exp(F \log_2 p \log_3 p)$$

there is a unique value of u_r satisfying

$$u_r < H \leqq u_{r+1},$$

for which the desired estimate holds. However, by the construction of the sequence $\{u_r\}$

$$N(H, p) - N(u_r, p) \leqq u_{r+1} - u_r = \frac{u_r}{(\log u_r)^G} < \frac{H}{(\log H)^{B+2}},$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

so that

$$N(H,p) = \pi(H)\,\frac{\phi(p-1)}{p-1}\left[1 + O\left(\frac{1}{(\log H)^B}\right)\right],$$

as required.

This completes the proof of Theorem 1, and under the more restrictive condition, $H \geqq p^{\epsilon}$, a modified form suffices. We give an outline of the necessary changes.

Proof of Theorem 2 (*outline*). We adopt similar definitions for the sets J, V, W, and E, but in place of the condition

$$u_r > \exp(\tfrac{1}{2}F \log_2 x \log_3 x)$$

we take

$$u_r > x^{\epsilon/2}.$$

The proof proceeds on exactly the same lines, and we need only prove the sharper estimates for $J(x)$ and $V(x)$. Once again we have three cases.

Case 1. $x^{\epsilon/2} < H \leqq x^2$.

In view of the above remarks, the application of Lemma 2 yields

$$\begin{aligned}\text{Card } J(H,x) &\leqq c_4\left(\frac{\log x}{\log H}\right)^{1/2} \exp\left(\frac{2\log x}{\log H}\right)\lambda^2 R^2 \log x\\ &= O((\log x)^{\eta})\end{aligned}$$

with $\eta = \frac{3}{2} + 2(K + L) + 4\epsilon^{-1}$.

Case 2. $x^2 < H \leqq x^4$.

We can apply the proof of the previous Case 2, obtaining exactly the same bound

$$\text{Card } J(H,x) \leqq c_{11}\lambda^{-2}(x^2H^{-1} + 1)\log x \cdot \log H;$$

however, in view of our more severe restriction on the size of H this now yields

$$\text{Card } J(H,x) = O((\log x)^{\mu})$$

with $\mu = 2(L + 1)$.

Case 3. $x^4 < H$.

In place of the Siegel-Walfisz theorem we now use Lemma 8. This guarantees that $J(H,x)$ is empty when $H > x^{3+\epsilon}$ is satisfied, save possibly for the set of primes P. Since, however,

$$P(t) = O((\log t)^D), \qquad t \geqq 3,$$

we can safely add this sequence to the sequence E already constructed and Theorem 2 follows easily.

Let us now consider Theorem 4. We shall indicate at the end of our proof what more is needed in order to prove Theorem 3.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

Proof of Theorem 4. Let a $T_\alpha(x)$-system with associated parameter D (see the introduction) be given. We shall consider a typical member $T_\alpha(x)$ for a large value of x, and begin by outlining how we modify the procedure for estimating the function $N(H, p)$. This time we settle for a non-zero lower bound only, and this, moreover, only for primes contained in $T_\alpha(x)$. For convenience during the present proof we use l to denote a typical prime number satisfying $l > x^\alpha$.

For any p satisfying $H < p \leqq x$, $p \in T_\alpha(x)$, there are no residue classes which have orders lying between 2 and x^α, so that

$$\begin{aligned}(6)\quad N(H,p) &\geqq \pi(H) - \mathrm{Card}(q\,;q \leqq H, 2|\mathrm{ind}\, q)\\ &\qquad\qquad - \sum_{l|(p-1)} \mathrm{Card}(q\,; q \leqq H, l|\mathrm{ind}\, q)\\ &= \pi(H) - N_1 - N_2,\end{aligned}$$

say. For any prime k and character χ of order k (mod p) we see that

$$\left|\mathrm{Card}(q\,; q \leqq H, k|\mathrm{ind}\, q) - \frac{\pi(H)}{k}\right| \leqq \frac{1}{k}\sum_{\chi_k}\left|\sum_{q \leqq H}\chi(q)\right|,$$

where the sum over χ_k runs through all the kth order characters (mod p). We deduce that

$$(7)\qquad \left|N_2 - \sum_{l|(p-1)}\frac{\pi(H)}{l}\right| \leqq \sum_{l|(p-1)}\frac{1}{l}\sum_{\chi_l}\left|\sum_{q \leqq H}\chi(q)\right|.$$

Let the number η satisfy $0 < 4\eta < 1$ and define H_0 by

$$2\log\left(\frac{2\log x}{e(\frac{1}{2} - 2\eta)^2}\cdot\frac{1}{(1-\eta)}\cdot\left[\frac{1}{\alpha}\right]^2\right) = \left(1 + \alpha - (D+4)\frac{\log\log x}{\log x}\right)\log H_0.$$

We shall show that for an appreciable number of primes $p \leqq x$ the inequality $g^*(p) \leqq H_0$ is satisfied. This then quickly leads to the desired result. The proof is in various stages.

LEMMA 12. *Let us define the set*

$$Y(\eta, H) = \left(p\,; p \leqq H, p \in T_\alpha(x), \sum_{l|(p-1)}\frac{1}{l}\sum_{\chi_l}\left|\sum_{q \leqq H}\chi(q)\right| > (\tfrac{1}{2} - 2\eta)\pi(H)\right).$$

Then the inequality

$$\mathrm{Card}\ Y(\eta, H_0) < \frac{\mathrm{Card}\ T_\alpha(x)}{\log x}$$

is satisfied for all large values of x.

Proof. To begin with, when $m \geqq 1$, $p \leqq x$ hold, the inequalities

$$\begin{aligned}\left(\sum_{l|(p-1)}\left(\frac{1}{l}\right)^{2m/(2m-1)}\sum_{\chi_l}1\right)^{2m-1} &= \left(\sum_{l|(p-1)} l^{-1/(2m-1)}\right)^{2m-1}\\ &\leqq x^{-\alpha}\left(\sum_{l|(p-1)}1\right)^{2m-1} \leqq x^{-\alpha}\left[\frac{1}{\alpha}\right]^{2m}\end{aligned}$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

are satisfied. By means of Hölder's inequality we then deduce that

$$(8) \quad \sum_{\substack{p \leqq x; \\ p \in T_\alpha(x)}} \left(\sum_{l \mid (p-1)} \frac{1}{l} \sum_{\chi_l} \left| \sum_{q \leqq H} \chi(q) \right| \right)^{2m} \leqq x^{-\alpha} \left[\frac{1}{\alpha} \right]^{2m} \sum_{p \leqq x} \sum_{\chi \neq \chi_0} \left| \sum_{q \leqq H} \chi(q) \right|^{2m} .$$

If we define the sequence of integers $a_1, a_2, \ldots$ in a natural way so that

$$(9) \quad \left(\sum_{q \leqq H} \chi(q) \right)^m = \sum_{r \leqq H^m} a_r \chi(r),$$

we can then apply Lemma 1 to the right-hand sum in the inequality (8) and obtain for it the upper bound

$$c_3(x^2 + H^m) \sum_{r \leqq H^m} a_r^2.$$

However, $a_r \leqq m!$ $(r = 1, 2, \ldots)$ so that this expression cannot exceed

$$c_3(x^2 + H^m) m! \{\pi(H)\}^m.$$

From these remarks we see that

$$(10) \quad \text{Card } Y(\eta, H) \leqq c_3(x^2 + H^m) x^{-\alpha} m! \left((\tfrac{1}{2} - 2\eta)^2 \left[\frac{1}{\alpha} \right]^{-2} \pi(H) \right)^{-m}$$

for $m = 1, 2, \ldots$.

If H satisfies the inequality $H \leqq x^2$, we define m by

$$m = \left[\frac{2 \log x}{\log H} \right]$$

so that $m \geqq 1$ holds. If, moreover, H is so large that we can apply the prime number theorem in the form

$$\pi(H) > (1 - \eta) H (\log H)^{-1},$$

we can apply Stirling's formula in the inequality (10), and deduce that

$$\text{Card } Y(\eta, H) \leqq c_{12} x^{-\alpha} H \left(\frac{\log x}{\log H} \right)^{1/2} \exp\left\{ m \log \left(\frac{m \log H}{e(\frac{1}{2} - 2\eta)^2 (1 - \eta)} \left[\frac{1}{\alpha} \right]^2 \right) \right\}$$

$$\leqq c_{13} x^{-\alpha} H \left(\frac{\log x}{\log H} \right)^{1/2} \exp\left\{ \frac{2 \log x}{\log H} \log \left(\frac{2 \log x}{e(\frac{1}{2} - 2\eta)^2} \cdot \frac{1}{(1 - \eta)} \cdot \left[\frac{1}{\alpha} \right]^2 \right) \right\} .$$

Remembering that x is large we see that

$$\text{Card } Y(\eta, H_0) \leqq c_{14} x^{-\alpha} (\log x)^{5/2} \exp\{ (1 + \alpha) \log x - (D + 4) \log \log x \}$$

$$< \frac{\text{Card } T_\alpha(x)}{\log x},$$

and our proof is complete.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

By means of this estimate we shall be able to show that N_2 is very often "small". In order to show that the same is true for N_1 we adopt a similar procedure but confine ourselves to quadratic characters. However, there is only one non-trivial quadratic character to a prime modulus, namely the Legendre symbol.

LEMMA 13. *On analogy with the set Y, we define*

$$Z(\eta,H) = \left(p\,;p \leqq H, \left|\sum_{q\leqq H}\left(\frac{q}{p}\right)\right| > \tfrac{1}{2}\eta\pi(H)\right).$$

Then for each constant $c_{15} > 0$ we can find constants d_1 and d_2 depending upon C_{15} and η so that the inequality

$$\text{Card } Z(\eta,H) = O\left(x \exp\left(-\frac{d_1 \log x}{\log\log x}\right)\right)$$

is satisfied for each value of H lying in the interval

$$d_2 \log x \leqq H \leqq (\log x)^{c_{15}}.$$

Proof. With the sequence $a_1, a_2, \ldots$ defined in (9) we can apply Lemma 3 with $L = H^m$ for a positive integer m, to show that

$$\sum_{p\leqq x}\left|\sum_{q\leqq H}\left(\frac{q}{p}\right)\right|^{2m} \leqq 4x \sum_{\substack{\nu,\mu\leqq H^m;\\ \nu\mu=t^2,\,2t^2}}\sum a_\nu a_\mu + O\left(H^m \log H^m\left(\sum_{r\leqq H^m} a_r\right)^2\right).$$

The sum of the a_r is once again $(\pi(H))^m$. If we express the integers a_ν, a_μ as products of primes we note that

$$\sum_{\nu\mu=t^2,2t^2}\sum a_\nu a_\mu \leqq \sum_t \tau(t^2) \sum_{q_1\cdots q_{2m}=t^2,q_i\leqq H}\cdots\sum 1$$

since $q_1 \ldots q_{2m} = 2t^2$ cannot occur. However, t^2 has $2m$ prime factors so that $\tau(t^2) \leqq 4^m$, and, in the notation of Lemma 10, this last multiple sum does not exceed

$$4^m T(\pi(H), m) \leqq (4m\pi(H))^m.$$

These inequalities together show that

$$\text{Card } Z(\eta, H) \leqq 4x(16m\eta^{-2})^m\{\pi(H)\}^{-m} + O((4\eta^{-2}H)^m \log H^m)$$

for $m = 1, 2, \ldots$.

Finally, we define m by

$$m = \left[\frac{3 \log x}{4 \log H}\right]$$

so that if H is sufficiently large depending upon η,

$$\begin{aligned}\text{Card } Z(\eta,H) &\leqq 4x\left(\frac{12 \log x}{\eta^2 \pi(H) \log H}\right)^m + O(x^{7/8})\\ &= O\left(x \exp\left\{-\frac{1}{2}\frac{\log x}{\log H}\log\left(c_{16}\frac{H}{\log x}\right)\right\}\right) + O(x^{7/8})\end{aligned}$$

from which the desired result is immediate.

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

In particular, setting $c_{15} = 2$ we can take $H = H_0$ in this result, and thus have proved that

$$\text{Card } Z(\eta, H_0) < \frac{\text{Card } T_\alpha(x)}{\log x}.$$

Completion of the proof of Theorem 4. Our hypothesis on $T_\alpha(x)$ included the estimate

$$\text{Card } T_\alpha(x) > x(\log x)^{-D}.$$

Let us define a set S (depending upon x, η) by

$$S = T_\alpha(x) - Y(\eta, H_0) - Z(\eta, H_0) - (p; p \leqq x(\log x)^{-D-1}).$$

Then in view of Lemmas 12 and 13, and the above remark,

$$\text{Card } S \geqq x(\log x)^{-D}\left(1 - \frac{3}{\log x}\right) > 0,$$

so that S is non-empty. Moreover, for any prime p in S the inequality (7) shows that

$$N_2 \leqq \pi(H_0) \sum_{l|(p-1)} \frac{1}{l} + (\tfrac{1}{2} - 2\eta)\pi(H_0) \leqq \pi(H_0)(x^{-\alpha}\alpha^{-1} + \tfrac{1}{2} - 2\eta)$$
$$< (\tfrac{1}{2} - \eta)\pi(H_0),$$

while

$$N_1 \leqq \tfrac{1}{2}\pi(H_0) + \tfrac{1}{2}\eta\pi(H_0).$$

These results and the inequality (6) together show that

$$N(H_0, p) > \pi(H_0)(1 - \tfrac{1}{2} + \eta - \tfrac{1}{2} - \tfrac{1}{2}\eta) = \tfrac{1}{2}\eta\pi(H_0) > 0,$$

so that

$$(11) \quad g^*(p) \leqq H_0$$
$$< \left(\frac{8}{e(1-4\eta)^2} \cdot \frac{1}{(1-\eta)} \cdot \left[\frac{1}{\alpha}\right]^2 \log x\right)^{2/(1+\alpha)} \exp\left(c_{17} \frac{(\log \log x)^2}{\log x}\right)$$
$$< \left(\frac{8}{e(1-4\eta)^2} \cdot \frac{1+\eta}{1-\eta} \cdot \left[\frac{1}{\alpha}\right]^2 \log p\right)^{2/(1+\alpha)} (1 + \eta).$$

Since $\eta < \frac{1}{4}$ can be taken arbitrarily small, and there exist sets $T_\alpha(x)$, and hence S, for arbitrarily large values of x, Theorem 4 is proved.

Proof of Theorem 3. Almost the same proof applies. In place of the sets $T_\alpha(x)$ we now use the sets

$$T_\alpha^*(x) = (p; p \leqq x, q|(p-1) \Rightarrow q = 2 \text{ or } q > x^{1/4}\exp(-(\log x)^{12/13})).$$

We maintain that these sets contain

$$\text{Card } T_\alpha^*(x) > x(\log x)^{-3}$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

primes for all large values of x. In order to prove this, we apply Lemma 7 to Lemma 9 for the sequence $(p - 1, p \leqq x)$ with

$$\alpha = \tfrac{1}{2}, \qquad \beta(x) = (\log X)^{4U+40},$$
$$X = \pi(x), \quad k = 2, \quad \gamma(d) = d/\phi(d), \quad z = x^{1/4}\exp(-(\log x)^{12/13}).$$

In the notation of Lemma 9, this yields

$$S_2(z) \geqq \pi(x)\Gamma_2(z)\left[f\left(\frac{\log \pi(x)}{2\log z}\right) - c\frac{\log\log 6}{(\log \pi(x))^{1/14}}\right],$$

where

$$f(u) = 2e^{\gamma}u^{-1}\log(u-1) \quad \text{for } 2 < u \leqq 3.$$

However, our choice of z guarantees that

$$\frac{\log \pi(x)}{2\log z} \geqq 2 + \tfrac{1}{2}(\log x)^{-1/13} + O((\log x)^{-2/13})$$

so that

$$f\left(\frac{\log \pi(x)}{2\log z}\right) \geqq \tfrac{1}{2}e^{\gamma}(\log x)^{-1/13} + O((\log x)^{-2/13}).$$

In other words,

$$\text{Card } T_\alpha^*(x) \geqq \pi(x)\prod_{2<p<z}\left(1 - \frac{1}{p-1}\right)[\tfrac{1}{2}e^{\gamma}(\log x)^{-1/13} + O((\log x)^{-2/13})]$$
$$> x(\log x)^{-3},$$

as stated earlier.

After changes, the exponent $c_{17}(\log_2 x)^2(\log x)^{-1}$ on the extreme right-hand side of the inequality (11) becomes $c_{18}(\log_2 x)^2(\log x)^{-1/13}$; however, this clearly does not affect the final result.

It is perhaps of interest to note that for each $\epsilon > 0$ we can find a positive constant c, depending upon ϵ, so that for at least $x^{1-\epsilon}$ primes $p \leqq x$, the least quadratic non-residue satisfies

$$n_2(p) > c\log p.$$

For, let $q_1, \ldots, q_r, \ldots$ denote the sequence of rational prime numbers. Then $n_2(p) > q_r$ holds if p belongs to certain reduced classes (mod $8q_1 \ldots q_r$). If we choose the constant δ to be suitably small we can ensure that

$$8q_1 \ldots q_r = 8\exp\left(\sum_{q \leqq \delta \log x}\log q\right) < x^{\eta}, \quad \text{with } \eta = \min(\tfrac{1}{2}\epsilon c^{-1}, c_6^{-1}).$$

We can then apply the result of Fogels (**10**) (Lemma 9 with $k = 8q_1 \ldots q_r$) and deduce that there are at least

$$(8q_1 \ldots q_r)^{-c_7}\pi(x) > x^{1-\epsilon}$$

primes $p \leqq x$ for which

$$n_2(p) \geqq q_r > \tfrac{1}{2}\delta\log x \geqq c\log p \qquad (c = \tfrac{1}{2}\delta).$$

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press

On the other hand, if now d is chosen to be a large constant, Lemma 13 shows that the inequality

$$n_2(p) \leqq d \log p$$

holds for all, save at most

$$O\left(x \exp\left(-c_{16}\frac{\log x}{\log\log x}\right)\right)$$

primes not exceeding x. Indeed, for any $\delta > 0$, there is a constant $\mu = \mu(\delta) > 0$ so that all but $O(x^{1-\mu})$ primes $p \leqq x$ have

$$n_2(p) \leqq (\log p)^{1+\delta}.$$

In view of this result, we conjecture that

$$n_2(p) = O((\log p)^{1+\epsilon})$$

for any fixed $\epsilon > 0$. Similar remarks can be made about the least prime quadratic residue (mod p).

References

1. N. C. Ankeny, *The least quadratic non-residue*, Ann. of Math. (2) *55* (1952), 65–71.
2. E. Artin, *Collected papers* (Addison-Wesley, Reading, Massachusetts, 1965).
3. M. B. Barban, *On a theorem of I. P. Kubilius*, Izv. Akad. Nauk USSR Ser. Fiz.-Mat. Nauk *1961* (5), 3–9. (Russian)
4. E. Bombieri, *On the large sieve*, Mathematika *12* (1965), 201–225.
5. D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. (3) *12* (1962), 179–192.
6. D. A. Burgess and P. D. T. A. Elliott, *On the average value of the least primitive root*, Mathematika *15* (1968), 39–50.
7. H. Davenport and P. Erdős, *The distribution of quadratic and higher residues*, Publ. Math. (Debrecen) *2* (1952), 252–265.
8. P. D. T. A. Elliott, *On the size of* $L(1, \chi)$, J. Reine Angew. Math. *236* (1969), 26–36.
9. ——— *On the mean value of* $f(p)$ (to appear).
10. E. Fogels, *On the distribution of prime ideals*, Acta Arithmetica VII (1962), 255–269.
11. P. X. Gallagher, *The large sieve*, Mathematika *14* (1967), 14–20.
12. ——— *Bombieri's mean-value theorem*, Mathematika *15* (1968), 1–6.
13. H. Halberstam and H. E. Richert, *Sieve methods* (to be published by Markham, Chicago).
14. H. Halberstam and K. F. Roth, *Sequences*, Vol. 1 (Oxford Univ. Press, Oxford, 1968).
15. H. Halberstam, W. Jurkat, and H. E. Richert, *Un nouveau résultat de la méthode du crible*, C. R. Acad. Sci. Paris *264* (1967), 920–923.
16. C. Hooley, *On Artin's conjecture*, J. Reine Angew. Math. *225* (1967), 209–220.
17. I. P. Kubilius, *Probabilistic methods in the theory of numbers*, Amer. Math. Soc. Transl. Vol. 11 (Amer. Math. Soc., Providence, R. I., 1964).
18. E. Landau, *Vorlesungen über Zahlentheorie*, Band 2, Teil VII, Kap. 14 (Leipzig, Berlin, 1927).
19. U. V. Linnik, *On the least prime in arithmetic progression*. I. *The basic theorem*; II. *The Deuring-Heilbronn phenomenon*, Mat. Sb. (N.S.) *15* (*57*) (1944), 139–178; 347–367.
20. K. Prachar, *Primzahlverteilung*, (Springer, Berlin, 1957).
21. A. Rényi, *On the representation of an even number as the sum of a prime and an almost prime*, Izv. Akad. Nauk SSSR Ser. Mat. *12* (1948), 57–78.
22. Y. Wang, *On the least primitive root of a prime*, Sci. Sinica *X* (*1*) (1961), 1–14.

University of Nottingham,
Nottingham, England

https://doi.org/10.4153/CJM-1969-092-6 Published online by Cambridge University Press