Publ. Math. d'Orsay,
Journees Arithmetiques,
Colloque Hubert Delange,
1982, 8-19.
(No reprints available)

MODULI r FOR WHICH THERE ARE MANY SMALL PRIMES CONGRUENT TO A MODULO r

Paul T. BATEMAN and Carl POMERANCE*

§1. Introduction.

Suppose a is a given non-zero integer. In this note we show that there is an infinitude of positive integers r (relatively prime to a) for each of which there are many primes p not much larger than r and congruent to a modulo r. We prove two theorems of this type, in the second of which we impose the additional requirement that the ratios $(p-a)/r$ are also prime. The proof of each theorem consists of a straightforward application of the pigeon-hole principle and is based only on classical results in multiplicative number theory. Arguments of the sort given here were used in an essential way in [1].

Our theorems are as follows. Both theorems remain true when $k = 0$ but are trivial in that case.

THEOREM 1. *If a is a given non-zero integer and k is a given positive integer, there are infinitely many positive integers r coprime to a for which we can find $k+1$ distinct primes $p_0, p_1, \ldots, p_k$ satisfying*

$$p_i \equiv a \pmod{r}, \quad p_i < e\, k\, r \log r$$

for $i = 0, 1, \ldots, k$.

* The research of the second-named author was supported by a grant from the National Science Foundation.

THEOREM 2. Suppose $\lambda > e^{\gamma}$ If a is a given non-zero integer and k is a given positive integer, there are infinitely many positive integers r coprime to a for which we can find $k+1$ distinct primes $p_0, p_1, \ldots, p_k$ satisfying

$$p_i \equiv a \pmod{r},\ (p_i - a)/r \text{ is prime},\ p_i < \lambda k r \log r \operatorname{loglog} r$$

for $i = 0, 1, \ldots, k$.

It is not hard to see that the prime k-tuples conjecture would imply stronger results than these theorems, for example the assertion in which the inequalities for the primes p_i in Theorem 1 are replaced by the inequalities of the form $p_i < Akr$, where A is a constant depending only on a. However, the inequalities for the primes p_i given in these theorems are about the best that can be expected by simple averaging arguments, aside possibly for constant factors (possibly depending on a.) This optimality follows from the fact that the relative frequency of primes around r is about $1/\log r$ and the relative frequency of primes around $\log r$ is about $1/\operatorname{loglog} r$. More specifically, in the case of Theorem 1 the early primes congruent to a modulo r could be expected to lie about $r \log r$ apart, so that we could not expect to find k such primes until we reach numbers of the order of magnitude $k r \log r$. In the case of Theorem 2 the ratios $(p_i - a)/r$ are $O((\log r)^{1+\varepsilon})$ and so the extra condition that these ratios are primes introduces an extra factor $\operatorname{loglog} r$ in the preceding discussion. In the case of both Theorems 1 and 2 we do not guarantee that the results cannot be improved by a

numerical factor (possibly depending on a.) In fact, in the case of Theorem 2, the condition $\lambda > e$ could be replaced by the condition $\log \lambda > \phi(a)/a$, where, as usual, $\phi(a)$ denotes the number of positive integers not exceeding $|a|$ and coprime to a.

The authors would like to thank Paul Erdős for his interest in these theorems.

§2. Necessary Lemmas.

We require the following classical results from multiplicative number theory. As usual $\pi(y;m,\ell)$ denotes the number of primes not exceeding y which are congruent to ℓ modulo m. In Lemmas B1 and B2 (as in our two theorems) the letter a stands for a given non-zero integer.

LEMMA A. *If* $y \geq 3$, *then*

$$\pi(y;m,\ell) = \frac{1}{\phi(m)} \int_2^y \frac{du}{\log u} + O\left(\frac{y}{(\log y)^{100}}\right)$$

for all m *less than* $(\log y)^{3/2}$ *and all* ℓ *relatively prime to* m, *where the constant implied by the* O *symbol is absolute and effectively computable.*

The result of Lemma A follows from equation (36) of (3). The exponent 3/2 could be replaced by any number less than 2 and the exponent 100 could be replaced by any positive constant whatever.

LEMMA B1. *If* ρ *is a fixed number greater than* 1, *then for* $y \geq 3$ *we have*

$$\sum_{y<q\leq \rho y,\ (q,a)=1} \frac{1}{\phi(q)} = C_a \log \rho + O\left(\frac{\log y}{y}\right),$$

where

$$C_a = \frac{\phi(a)}{a} \prod_{p\nmid a} \left(1 + \frac{1}{p(p-1)}\right) > \frac{\phi(a)}{a}$$

and the constant implied by the O*-symbol depends on* a.

PROOF. By [2] we have

$$\sum_{1\le q\le y,\ (q,a)=1} \frac{1}{\phi(q)} = C_a \log y + D_a + O\left(\frac{\log y}{y}\right),$$

where C_a is as above, D_a is another constant depending on a, and the constant implied by the O-symbol depends on a. The stated result follows by subtraction.

LEMMA B2. *If ρ is a fixed number greater than 1, then for $y \ge 3$ we have*

$$\sum_{y<q\le \rho y,\ q \text{ prime}} \frac{1}{\phi(q)} = \frac{\log \rho}{\log y} + O\left(\frac{1}{(\log y)^2}\right),$$

where the constant implied by the O-symbol depends on ρ.

PROOF. From Lemma A with $m = 1$ we readily obtain by partial summation

$$\sum_{1\le q\le y,\ q \text{ prime}} \frac{1}{\phi(q)} = \sum_{1\le q\le y,\ q \text{ prime}} \frac{1}{q-1}$$

$$= \log\log y + b + O\left(\frac{1}{(\log y)^{99}}\right),$$

where b is a certain absolute constant. By subtraction we obtain

$$\sum_{y<q\le \rho y,\ q \text{ prime}} \frac{1}{\phi(q)} = \log\left(1 + \frac{\log \rho}{\log y}\right) + O\left(\frac{1}{(\log y)^{99}}\right),$$

from which the conclusion of the lemma follows.

§3. Proof of Theorem 1.

Let K be a large positive constant, to be specified later in terms of a. For large positive x let P be the set of primes p such that $x<p\le(K+1)x$, let Q be the set of integers q such that

$$(q,a) = 1,\ (1+\frac{1}{K})\frac{k\,\phi(a)}{C_a\,a}\log x < q \le (1+\frac{1}{K})\frac{ek\phi(a)}{C_a\,a}\log x,$$

and let M be the set of pairs (p,q) with $p\in P$, $q\in Q$, and $p\equiv a \pmod q$. We define a function f on M by putting $f(p,q) = (p-a)/q$. In view of the definitions of P and Q, the range of f is contained in the set R consisting of the integers r satisfying

$$(r,a) = 1,\ \frac{C_a\,a\,(x-a)}{(1+K^{-1})\,e\,k\,\phi(a)\log x} < r < \frac{C_a\,a\{(K+1)x-a\}}{(1+K^{-1})k\,\phi(a)\log x}.$$

By Lemma A the cardinality of M is given by

$$\begin{aligned}|M| &= \sum_{q\in Q}\{\pi(\,(K+1)\,x;\,q,a) - \pi(x;q,a)\}\\ &= \sum_{q\in Q}\{\frac{1}{\phi(q)}\int_x^{(K+1)x}\frac{du}{\log u} + O(\frac{x}{(\log x)^{100}})\}\\ &= \sum_{q\in Q}\frac{1}{\phi(q)}\{\frac{Kx}{\log x} + O(\frac{x}{(\log x)^2}\} + O(\frac{x}{(\log x)^{99}}).\end{aligned}$$

By Lemma B 1

$$\sum_{q\in Q}\frac{1}{\phi(q)} = C_a + O(\frac{\log\log x}{\log x}),$$

so that

$$|M| = C_a K \frac{x}{\log x} + O\left(\frac{x \operatorname{loglog} x}{(\log x)^2}\right).$$

On the other hand the cardinality of R satisfies

$$|R| = C_a K\{1 - \frac{1}{(K+1)e}\} \frac{x}{k \log x} + O(1).$$

Hence for large K we have $k|R| < |M|$, so that the function f must take on some value at least $k + 1$ times. Thus for sufficiently large x there exists an element r of R and $k + 1$ distinct pairs $(p_0,q_0), (p_1,q_1),\ldots,(p_k,q_k)$ in M such that

$$\frac{p_0 - a}{q_0} = \frac{p_1 - a}{q_1} = \ldots = \frac{p_k - a}{q_k} = r.$$

Clearly the primes p_i are distinct and $p_i \equiv a \pmod r$ for each i. Further

$$p_i = q_i r + a \leq (1 + \frac{1}{K}) \frac{e\phi(a)}{C_a a} k r \log x + a.$$

Since $\log r = \log x + O(\operatorname{loglog} x)$, we have

$$p_i < (1 + \frac{2}{K}) \frac{e\phi(a)}{C_a a} k r \log r$$

if x is sufficiently large. Since $C_a > \frac{\phi(a)}{a}$, we may take K large enough so that

$$(1 + \frac{2}{K}) \frac{1}{C_a} \leq \frac{a}{\phi(a)}.$$

Then the primes p_i satisfy the inequality of the theorem, provided of course that x is sufficiently large. Since r tends to infinity with x, there are infinitely many positive integers r for which the conclusion of the theorem holds.

§4. Proof of Theorem 2.

Let K be a large positive constant and let ρ be a constant greater than 1, both of which will be specified later. For large positive x let P be the set of primes p such that $x < p \leq (K + 1)\, x$, let Q be the set of primes q such that

$$k \log x \operatorname{loglog} x < q \leq \rho\, k \log x \operatorname{loglog} x,$$

and let M be the set of pairs (p,q) with $p \in P$, $q \in Q$, and $p \equiv a \pmod q$. We define a function f on M by putting $f(p,q) = (p - a)/q$. Clearly the range of f is contained in the set R consisting of the integers r satisfying

$$\frac{x - a}{k\, \rho \log x \operatorname{loglog} x} < r < \frac{(K + 1)\, x - a}{k \log x \operatorname{loglog} x}.$$

As in the proof of Theorem 1 the cardinality of M is given by

$$|M| = \sum_{q \in Q} \frac{1}{\phi(q)} \left\{ \frac{K\, x}{\log x} + O\left(\frac{x}{(\log x)^2}\right) \right\} + O\left(\frac{x}{(\log x)^{99}}\right).$$

By Lemma B2

$$\sum_{q \in Q} \frac{1}{\phi(q)} = \frac{\log \rho}{\log(k \log x \operatorname{loglog} x)} + O\left(\frac{1}{(\operatorname{loglog} x)^2}\right),$$

so that

$$|M| = K \log \rho \, \frac{x}{\log x \operatorname{loglog} x} + O\left(\frac{x \operatorname{logloglog} x}{\log x\, (\operatorname{loglog} x)^2}\right).$$

On the other hand the cardinality of R satisfies

$$|R| = \left(K + 1 - \frac{1}{\rho}\right) \frac{x}{k \log x \operatorname{loglog} x} + O(1).$$

Hence for large x we have $k|R| < |M|$, provided that $\rho > e$ and we choose K large enough so that

$$K(\log \rho - 1) > 1 - 1/\rho.$$

Accordingly for sufficiently large x there exists an element r of R and $k+1$ distinct pairs $(p_0,q_0), (p_1,q_1),\dots,(p_k,q_k)$ in M such that

$$\frac{p_0 - a}{q_0} = \frac{p_1 - a}{q_1} = \dots = \frac{p_k - a}{q_k} = r.$$

Clearly the primes p_i are distinct, $p_i \equiv a \pmod r$ for each i, and $(p_i - a)/r = q_i$ is prime for each i. Further

$$p_i = q_i\, r + a \le \rho\, k\, r \log x \operatorname{loglog} x + a.$$

Since $\log r = \log x + O(\operatorname{loglog} x)$, we have

$$p_j < \rho\, k\, r \log r \operatorname{loglog} r + O(r\, (\operatorname{loglog} r)^2)$$

if x is sufficiently large. If we now choose ρ so that $e < \rho < \lambda$, say, $\rho = (e + \lambda)/2$, and choose $K > (1 - \rho^{-1})/(\log \rho - 1)$, we have

$$p_i < \lambda\, k\, r \log r \operatorname{loglog} r,$$

provided of course that x is sufficiently large x. Since r tends to infinity with x, there are infinitely many positive integers r for which the conclusion of the theorem holds.

By redefining the set R to include only integers co-prime to a, we could replace the condition $\lambda > e$ in Theorem 2 by the condition $\log \lambda > \phi(a)/a$.

§5. Some Related Conjectures.

If a is a non-zero integer and r is a positive integer coprime to a, let $p_k(r,a)$ denote the k-th prime number congruent to a modulo r and greater than r. Put

$$c(r,a) = \sup_k \frac{p_k(r,a)}{k\, r \log(k\, r)}.$$

Since $p_k(r,a) = O(k \log k)$ for fixed r and a by Lemma A, clearly $c(r,a)$ exists.

Theorem 1 asserts that if a and k are given, there are infinitely many r coprime to a for which

$$p_k(r,a)/\{k\, r \log(k\, r)\} < e.$$

However $c(r,a)$ considers the somewhat deeper question of bounding the ratio $p_k(r,a)/\{k\, r \log(k\, r)\}$ for all positive integral values of k while r and a remain fixed. The following conjectures about $c(r,a)$ seem reasonable to us.

CONJECTURE 1. *There is an absolute constant c for which there are infinitely many pairs of integers r,a such that*

$$0 < |a| < r, \quad (r,a) = 1, \quad \text{and} \quad c(r,a) \leq c.$$

Conjecture 2 is the more specific form of Conjecture 1 in which a is specified in advance.

CONJECTURE 2. *There is an absolute constant c such that, for any given nonzero integer a, there are infinitely many positive integers r coprime to a for which*

$$c(r,a) \leq c.$$

Conjecture 3 is a quantitative form of Conjecture 2 in which we assert (1) that, for any positive c, a positive fraction of the positive integers coprime to a have the property $c(r,a) \leq c$ and also (2) that, if c is large, the vast majority of the positive integers coprime to a have the property $c(r,a) \leq c$. In order to state this formally, we need the following notation. If a is a non-zero integer and $c \geq 0$, we let $N(a,c,x)$ denote the number of integers r satisfying

$$(r,a) = 1, \quad 1 \leq r \leq x, \quad c(r,a) \leq c.$$

CONJECTURE 3. *If a is a given non-zero integer, then*

$$f_a(c) = \lim_{x \to +\infty} x^{-1} N(a,c,x)$$

exists for every $c \geq 0$ and is a continuous function of c. Moreover $f_a(c) > 0$ for $c > 0$ and

$$\lim_{c \to +\infty} f_a(c) = \phi(a)/a.$$

Clearly Conjecture 3 implies Conjecture 2, which in turn implies Conjecture 1. While Conjecture 3 seems difficult, Conjectures 1 and 2 may be assailable.

REFERENCES

1. Paul T. Bateman, Carl Pomerance, and Robert C. Vaughan, On the size of the coefficients of the cyclotomic polynomial, to be published in the Proceedings of the Colloquium on Number Theory held in Budapest, Hungary, June 20-26, 1981.

2. E. Landau, On a Titchmarsh-Estermann sum, J. London Math. Soc. 11, 242-245 (1936).

3. A Page, On the number of primes in an arithmetic progression, Proc. London Math. Soc. (2) 39, 116-141 (1935).

University of Illinois at Urbana-Champaign

University of Georgia