ON THE NUMBER OF PRIMES p FOR WHICH $p+a$ HAS A LARGE PRIME FACTOR

MORRIS GOLDFELD

1. *Introduction.* For any fixed integer a and real variables x, y with $y < x$ let $N_a(x, y)$ denote the number of primes $p \leqslant x$ for which $p + a$ has at least one prime factor greater than y. As an elementary application of the following deep theorem of Bombieri on arithmetic progressions,

THEOREM (Bombieri, [1]). *For any constant $A > 0$, there exists a positive constant B such that if $M = x^{\frac{1}{2}} l^{-B}$ with $l = \log x$, then for $x > 1$*

$$\sum_{m \leqslant M} \max_{(a, m)=1} \left| \Pi(x;\ m, a) - \frac{\operatorname{Li}(x)}{\phi(m)} \right| \ll x l^{-A},$$

where† *$\Pi(x;\ m, a)$ denotes the number of primes less than x which are congruent to $a \bmod m$;*

we shall prove the following theorem:

THEOREM 1. *Let a be any fixed integer and let $x > e$. We then have*

$$\sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p+a}} \log q = x/2 + O\left(\frac{x \log\log x}{\log x}\right),$$

where the double sum is taken over primes p and q.

The transition from Theorem 1 to information about $N_a(x, x^{\frac{1}{2}})$ is relatively easy, for

$$N_a(x, x^{\frac{1}{2}}) \geqslant \sum_{p \leqslant x-a} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p+a}} 1 = \sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p+a}} 1 + O(\log x)$$

$$\geqslant \frac{1}{\log x} \sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p+a}} \log q + O(\log x);$$

so that

$$N_a(x, x^{\frac{1}{2}}) \geqslant \frac{1}{2} \frac{x}{\log x} + O\left(\frac{x \log\log x}{\log^2 x}\right).$$

2. *Proof of Theorem* 1. Let $\Lambda(n)$ denote the von Mangoldt function $\Lambda(n) = \log p$ if n is a power of the prime p, and otherwise $\Lambda(n) = 0$. It follows that

$$\sum_{m \leqslant x} \Pi(x;\ m, -a)\, \Lambda(m) = \sum_{m \leqslant x} \sum_{\substack{p \leqslant x \\ m \mid p+a}} \Lambda(m) = \sum_{p \leqslant x} \sum_{\substack{m \leqslant x \\ m \mid p+a}} \Lambda(m)$$

$$= \sum_{p \leqslant x} \log(p + a) + O(\log^2 x)$$

$$= x + O(x/\log x).$$

† Here we make use of Vinogradov's symbolism $C \ll D$ as an equivalent to $C = O(D)$.

Consequently,

$$\left(\sum_{m \leqslant x^{\frac{1}{2}} l^{-c}} + \sum_{x^{\frac{1}{2}} l^{-c} < m \leqslant x^{\frac{1}{2}}} + \sum_{x^{\frac{1}{2}} < m \leqslant x}\right) \Pi(x;\ m,\ -a)\, \Lambda(m) = x + O(x/\log x), \tag{1}$$

where c is some positive constant to be determined shortly.

The first sum on the left-hand side of (1) is easily evaluated by Bombieri's theorem to give

$$\begin{aligned} \sum_{m \leqslant x^{\frac{1}{2}} l^{-c}} \Pi(x;\ m,\ -a)\, \Lambda(m) &= \mathrm{Li}(x) \sum_{m \leqslant x^{\frac{1}{2}} l^{-c}} \Lambda(m)/\phi(m) + O\left(\log x \frac{x}{\log^A x}\right) \\ &= x/2 + O\left(\frac{x \log\log x}{\log x}\right) \end{aligned} \tag{2}$$

using Mertens' result [2]

$$\sum_{m \leqslant y} \Lambda(m)/\phi(m) = \log y + O(1),$$

provided A (in Bombieri's theorem) is chosen $\geqslant 2$, and $c = B(A)$.

The second sum in equation (1) can also be easily evaluated by the following theorem of Brun–Titchmarsh [3]:

THEOREM (Brun–Titchmarsh). *Let* $0 < \varepsilon < 1$; *then if* $m \leqslant x^{1-\varepsilon}$ and $(a, m) = 1$,

$$\Pi(x;\ m, a) \ll \frac{x}{\phi(m) \log x},$$

the constant implied by the $\ll$ *symbol depending at most on* ε.

We obtain, using also the above theorem of Mertens,

$$\begin{aligned} \sum_{x^{\frac{1}{2}} l^{-c} < m \leqslant x^{\frac{1}{2}}} \Pi(x;\ m,\ -a)\, \Lambda(m) &\ll \frac{x}{\log x} \sum_{x^{\frac{1}{2}} l^{-c} < m \leqslant x^{\frac{1}{2}}} \Lambda(m)/\phi(m) \\ &= O\left(\frac{x \log\log x}{\log x}\right). \end{aligned} \tag{3}$$

It therefore follows from equations (1), (2) and (3) that

$$\sum_{x^{\frac{1}{2}} < m \leqslant x} \Pi(x;\ m,\ -a)\, \Lambda(m) = x/2 + O\left(\frac{x \log\log x}{\log x}\right). \tag{4}$$

We are interested, however, in $\sum_{x^{\frac{1}{2}} < q \leqslant x} \Pi(x;\ q,\ -a) \log q$ where q runs over primes. We consequently split the left side of equation (4) into three parts

$$\left(\sum_{x^{\frac{1}{2}} < q \leqslant x} + \sum_{\substack{x^{\frac{1}{2}} < q^k \leqslant x^{\frac{3}{4}} \\ q \leqslant x^{\frac{1}{2}} \\ k > 1}} + \sum_{\substack{x^{\frac{3}{4}} < q^k \leqslant x \\ q \leqslant x^{\frac{1}{2}} \\ k > 1}}\right) \Pi(x;\ q^k,\ -a) \log q. \tag{5}$$

Let S_1, S_2 and S_3 denote the corresponding sums in (5). It follows by the Brun–Titchmarsh theorem that

$$\begin{aligned} S_2 &\ll \frac{x}{\log x} \sum_{\substack{x^{\frac{1}{2}} < q^k \leqslant x^{\frac{3}{4}} \\ q \leqslant x^{\frac{1}{2}} \\ k > 1}} \Lambda(q^k)/q^k \\ &\ll \frac{x}{\log x} \left[\sum_{m \leqslant x^{\frac{3}{4}}} \frac{\Lambda(m)}{\phi(m)} - \sum_{h \leqslant x^{\frac{1}{2}}} \frac{\log h}{h} \right] \\ &\ll \frac{x}{\log x}. \end{aligned} \tag{6}$$

Also, since there are at most $\log x \ \log 2$ powers of q less than x, and

$$\Pi(x;\, q^k, -a) \ll x^{\frac{1}{4}} \quad \text{for} \quad q^k > x^{\frac{3}{4}},$$

we see that

$$S_3 \ll x^{\frac{1}{4}} \sum_{\substack{x^{\frac{3}{4}} < q^k \leqslant x \\ q \leqslant x^{\frac{1}{2}} \\ k > 1}} \log q \ll x^{\frac{1}{4}} \sum_{q \leqslant x^{\frac{1}{2}}} \log q \log x \ll x^{\frac{3}{4}} \log^2 x. \tag{7}$$

The theorem follows immediately from equations (4), (5), (6) and (7).

We remark here that one can use the Brun–Titchmarsh theorem once more to prove that there exists a constant $c_0 > 0$ such that

$$\sum_{x^{\frac{1}{2}} < q \leqslant x^{\frac{1}{2}+c_0}} \Pi(x;\, q, -a) \ll \delta/2 \frac{x}{\log x}, \quad \text{where} \quad \delta < 1,$$

and after subtracting this from $N_a(x, x^{\frac{1}{2}})$ one can get a non-trivial lower bound for $N_a(x, x^r)$ for $\frac{1}{2} \leqslant r \leqslant \frac{1}{2} + c_0$. One sees, for example, that the constant $c_0 \approx \frac{1}{12}$ will work in this method.

3. *An Application.* We now apply the results of Theorem 1 to another related problem. Let $e_a(p)$ denote the smallest positive integer d for which $a^d \equiv 1 \bmod p$. We shall prove that for almost all primes p for which $p-1$ has a large prime factor q, then q also divides $e_a(p)$.

THEOREM 2. *Let a be any fixed positive integer other than* 1 *and let* $x > 1$. *We then have*

$$\sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid e_a(p)}} \log q = x/2 + O\left(\frac{x \log \log x}{\log x}\right),$$

where the double sum is taken over primes p and q.

Proof: Since $e_a(p) \mid p-1$, we see that

$$\sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p-1}} \log q = \sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid e_a(p)}} \log q + \sum_{p \leqslant x} \sum_{\substack{x^{\frac{1}{2}} < q \leqslant x \\ q \mid p-1 \\ q \nmid e_a(p)}} \log q. \tag{8}$$

Now, if it could be shown that the last sum on the right side of (8) is small then the theorem would follow immediately from Theorem 1. Consequently, we split this sum into two parts

$$\sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}<q\leqslant x\\ q\mid p-1\\ q\nmid e_a(p)}}\log q=\sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}<q\leqslant x^{\frac{1}{2}}l\\ q\mid p-1\\ q\nmid e_a(p)}}\log q+\sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}l<q\leqslant x\\ q\mid p-1\\ q\nmid e_a(p)}}\log q. \tag{9}$$

Let S and T denote the sums on the right side of (9). We show that each of these is small. First, by the Brun–Titchmarsh theorem

$$S \ll \sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}<q\leqslant x^{\frac{1}{2}}l\\ q\mid p-1}}\log q \ll \frac{x}{\log x}\sum_{x^{\frac{1}{2}}<q\leqslant x^{\frac{1}{2}}l}(\log q)/q$$

$$= O\left(\frac{x\log\log x}{\log x}\right).$$

Secondly, if we let

$$M_a(x)=\sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}l<q\leqslant x\\ q\mid p-1\\ q\nmid e_a(p)}}1,$$

then, for each prime counted in $M_a(x)$,

$$a^{(p-1)/q}\equiv 1 \bmod p,$$

so that

$$2^{M_a(x)} \leqslant \prod_{m\leqslant x^{\frac{1}{2}}l^{-1}}(a^m-1)$$

and therefore

$$M_a(x) \leqslant \frac{\log a}{\log 2}\sum_{m\leqslant x^{\frac{1}{2}}l^{-1}} m$$

$$= O\left(\frac{x}{\log^2 x}\right).$$

Consequently $T \leqslant M_a(x)\log x \ll x/\log x$, so that

$$S+T=O\left(\frac{x\log\log x}{\log x}\right),$$

and by equations (8) and (9) this immediately proves the theorem.

The following corollaries are easy consequences of Theorem 2.

COROLLARY 1. *Let a be any fixed positive integer other than* 1 *and let* $x>1$. *We then have*

$$\sum_{p\leqslant x}\sum_{\substack{x^{\frac{1}{2}}<q\leqslant x\\ q\mid e_a(p)}}1=\frac{1}{2}\frac{x}{\log x}+O\left(\frac{x\log\log x}{\log^2 x}\right),$$

where the double sum is taken over primes p and q.

COROLLARY 2. *For any fixed positive integer a other than* 1 *and* $x>1$

$$\sum_{p\leqslant x}e_a(p) \geqslant \frac{1}{2}\frac{x^{\frac{3}{2}}}{\log x}+O\left(\frac{x^{\frac{3}{2}}\log\log x}{\log^2 x}\right).$$

References

1. E. Bombieri, " On the large sieve ", *Mathematika*, 12 (1965), 201–225.
2. Hardy and Wright, *An introduction to the theory of numbers* (Oxford, 1965), 348–349.
3. K. Pracher, *Primzahlverteilung* (Springer, 1957), 44–45.

Columbia University,
New York 27, N.Y.

(*Received on the* 19*th of June*, 1968).