AN ELEMENTARY PROOF OF THE PRIME-NUMBER THEOREM FOR ARITHMETIC PROGRESSIONS

ATLE SELBERG

1. Introduction. In this paper we shall give an elementary proof of the theorem

$$\lim_{x \to \infty} \frac{\vartheta_{k,l}(x)}{x} = \frac{1}{\varphi(k)}, \tag{1.1}$$

where $\varphi(k)$ denotes Euler's function, and

$$\vartheta_{k,l}(x) = \sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log p, \tag{1.2}$$

where p denotes the primes, and k and l are integers with $(k,l) = 1$, k positive.

The proof proceeds essentially along the same lines as in a previous paper[1] about the case $k = 1$. However we also need in this case some of the ideas from my paper[2] on Dirichlet's theorem in order to prove that

$$\lim_{x \to \infty} \frac{\vartheta_{k,l}(x)}{x} > 0. \tag{1.3}$$

a result which we will need for our proof of (1.1).

It is possible to shorten the proof in several ways, which however would make it less elementary. For instance one could consider also the complex characters mod k, and in this way avoid Lemma 2 and most of the proof of Lemma 3. Also, by using results about the decomposition of primes in the quadratic field $K(\sqrt{D})$, we could make the proof of Lemma 1 much shorter.

As we shall see, the following proof is completely constructive, in the sense that it gives for any fixed positive ϵ, a way of finding, in a finite number of steps, an x_0—depending on ϵ and k—such that

$$\left| \frac{\vartheta_{k,l}(x)}{x} - \frac{1}{\varphi(k)} \right| < \epsilon.$$

for $x > x_0$.

Actually, it is possible in this way to prove more than (1.1). By careful estimation it is possible to show by the method below that

Received December 30, 1948.

[1]"An Elementary Proof of the Prime-number Theorem," *Ann. of Math.*, vol. 50 (1949), 305-313.

[2]"An Elementary Proof of Dirichlet's Theorem about Primes in an Arithmetic Progression," *Ann. of Math.*, vol. 50 (1949), 297-304.

$$\vartheta_{k,l}(x) = \frac{1}{\varphi(k)}\, x + O\left(\frac{x}{(\log x)^c}\right),$$

where c is a positive constant.

Throughout the paper k denotes a fixed positive integer; l denotes an integer with $(l, k) = 1$, while α, β and γ are used to designate numbers from a reduced residue system mod k; p, q and r denote primes. The letter K denotes positive constants, depending on k only; x_0 denotes a constant (not necessarily the same at each occurrence), depending on k only. In the same way x_σ denotes a number depending only on k and the positive number σ. The constants implied by the O's are generally dependent on k and in secs. **4** and **5** also on σ.

From my two previous papers mentioned above I make use of the prime-number theorem in the case $k = 1$, the formula

$$\sum_{p \leqslant x} \log^2 p + \sum_{pq \leqslant x} \log p \log q = 2x \log x + O(x), \tag{1.4}$$

and the further formula

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p + \sum_{\substack{pq \leqslant x \\ pq \equiv l(k)}} \log p \log q = \frac{1}{\varphi(k)} \Big\{ \sum_{p \leqslant x} \log^2 p + \sum_{pq \leqslant x} \log p \log q \Big\} + O(x). \tag{1.5}$$

Finally I make use of the well-known formula

$$\sum_{p \leqslant x} \frac{\log p}{p} = \log x + O(1). \tag{1.6}$$

By partial summation it is easily seen that one may also give (1.6) the forms

$$\sum_{n \leqslant x} \frac{\vartheta(n)}{n^2} = \log x + O(1). \tag{1.7}$$

$$\sum_{n \leqslant x} \vartheta\left(\frac{x}{n}\right) = x \log x + O(x), \tag{1.8}$$

where $\vartheta(x)$ denotes $\vartheta_{k,l}(x)$ for $k=1$.

Trivial estimations are often not carried out in detail, but left to the reader.

2. Fundamental formulas and inequalities. From (1.4) and (1.5) we get

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p + \sum_{\substack{pq \leqslant x \\ pq \equiv l(k)}} \log p \log q = \frac{2}{\varphi(k)}\, x \log x + O(x), \tag{2.1}$$

which may also be written[3]

$$\vartheta_l(x) \log x + \sum_{p \leqslant x} \log p\ \vartheta_{l\bar{p}}\left(\frac{x}{p}\right) = \frac{2}{\varphi(k)}\, x \log x + O(x), \tag{2.2}$$

where $\bar{p}$ denotes a solution of the congruence $p\,\bar{p} \equiv 1 \pmod{k}$.

[3]We write $\vartheta_l(x)$ instead of $\vartheta_{k,l}(x)$ where no misunderstanding can occur.

Since

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log p + \sum_{\substack{pq \leqslant x \\ pq \equiv l(k)}} \frac{\log p \log q}{\log pq} = \frac{2}{\varphi(k)} x + O\left(\frac{x}{\log x}\right), \tag{2.3}$$

which follows by partial summation using (2.1), then

$$\sum_{\substack{pq \leqslant x \\ pq \equiv l(k)}} \log p \log q = \sum_{p \leqslant x} \log p \sum_{\substack{q \leqslant x/p \\ q \equiv l\bar{p}(k)}} \log q = \frac{2}{\varphi(k)} x \sum_{p \leqslant x} \frac{\log p}{p}$$

$$- \sum_{p \leqslant x} \log p \sum_{\substack{qr \leqslant x/p \\ qr \equiv l\bar{p}(k)}} \frac{\log q \log r}{\log qr} + O\left(x \sum_{p \leqslant x} \frac{\log p}{p\,(1 + \log x/p)}\right)$$

$$= \frac{2}{\varphi(k)} x \log x - \sum_{qr \leqslant x} \frac{\log q \log r}{\log qr} \vartheta_{l\overline{qr}}\left(\frac{x}{qr}\right) + O(x \log \log x).$$

Inserting this for the second term on the left-hand side of (2.2), we get

$$\vartheta_l(x) \log x = \sum_{pq \leqslant x} \frac{\log p \log q}{\log pq} \vartheta_{l\overline{pq}}\left(\frac{x}{pq}\right) + O(x \log \log x). \tag{2.4}$$

Writing now

$$\vartheta_l(x) = \frac{1}{\varphi(k)} x + R_l(x), \tag{2.5}$$

we get from (2.2)

$$R_l(x) \log x = - \sum_{p \leqslant x} \log p\, R_{l\bar{p}}\left(\frac{x}{p}\right) + O(x). \tag{2.6}$$

In the same manner (2.4) yields

$$R_l(x) \log x = \sum_{pq \leqslant x} \frac{\log p \log q}{\log pq} R_{l\overline{qp}}\left(\frac{x}{pq}\right) + O\,(x \log \log x), \tag{2.7}$$

since

$$\sum_{pq \leqslant x} \frac{\log p \log q}{pq \log pq} = \log x + O\,(\log \log x),$$

which follows by partial summation using

$$\sum_{pq \leqslant x} \frac{\log p \log q}{pq} = \tfrac{1}{2} \log^2 x + O\,(\log x),$$

which again follows from (1.6).

Combining (2.6) and (2.7) we get

$$2\,|\,R_l(x)\,|\log x \leqslant \sum_{p \leqslant x} \log p\,|\,R_{l\bar{p}}\left(\frac{x}{p}\right)| + \sum_{pq \leqslant x} \frac{\log p \log q}{\log pq} |R_{l\overline{qp}}\left(\frac{x}{pq}\right)|$$
$$+ O(x \log \log x),$$

or if a runs over a reduced residue system mod k,

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

$$2\ |R_l(x)|\log x \leqslant \sum_a \Biggl\{ \sum_{\substack{p \leqslant x \\ p \equiv l\bar{a}(k)}} \log p \left| R_a\left(\frac{x}{p}\right)\right| + \sum_{\substack{pq \leqslant x \\ pq \equiv l\bar{a}(k)}} \frac{\log p \log q}{\log pq}\left| R_a\left(\frac{x}{pq}\right)\right| \Biggr\} + O(x \log\log x).$$

From this we get by partial summation, using (2.3),

$$2|R_l(x)|\log x \leqslant \sum_a \sum_{n \leqslant x} \Biggl\{ \sum_{\substack{p \leqslant n \\ p \equiv l\bar{a}(k)}} \log p + \sum_{\substack{pq \leqslant n \\ pq \equiv l\bar{a}(k)}} \frac{\log p \log q}{\log pq} \Biggr\}$$

$$\cdot \left\{ \left| R_a\left(\frac{x}{n}\right)\right| - \left| R_a\left(\frac{x}{n+1}\right)\right| \right\} + O(x \log\log x)$$

$$= \frac{2}{\varphi(k)} \sum_a \sum_{n \leqslant x} n \left\{ \left| R_a\left(\frac{x}{n}\right)\right| - \left| R_a\left(\frac{x}{n+1}\right)\right| \right\}$$

$$+ O\left(\sum_a \sum_{n \leqslant x} \frac{n}{1+\log n} \left| R_a\left(\frac{x}{n}\right) - R_a\left(\frac{x}{n+1}\right)\right| \right)$$

$$+ O\ (x \log\log x) = \frac{2}{\varphi(k)} \sum_a \sum_{n \leqslant x} \left| R_a\left(\frac{x}{n}\right)\right|$$

$$+ O\left(\sum_{n \leqslant x} \frac{x}{n\ (1+\log n)} \right) + O\left(\sum_{n \leqslant x} \frac{n}{1+\log n}\left(\vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right)\right)\right)$$

$$+ O\ (x \log\log x) = \frac{2}{\varphi(k)} \sum_a \sum_{n \leqslant x} \left| R_a\left(\frac{x}{n}\right)\right|$$

$$+ O\left(\sum_{n \leqslant x} \frac{1}{1+\log n}\, \vartheta\left(\frac{x}{n}\right)\right) + O\ (x \log\log x)$$

$$= \frac{2}{\varphi(k)} \sum_a \sum_{n \leqslant x} \left| R_a\left(\frac{x}{n}\right)\right| + O\ (x \log\log x)$$

or finally[4]

$$|R_l(x)| \leqslant \frac{1}{\varphi(k)\log x} \sum_a \sum_{n \leqslant x} \left| R_a\left(\frac{x}{n}\right)\right| + O\left(\frac{x \log\log x}{\log x}\right). \tag{2.8}$$

3. A lower bound for $\vartheta_l(x)$. From (2.4) in the form

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p = \sum_{\substack{pqr \leqslant x \\ pqr \equiv l(k)}} \frac{\log p \log q \log r}{\log pq} + O\ (x \log\log x)$$

[4]Instead of (2.8) we might use the somewhat sharper inequality

$$|R_l(x)| \leqslant \frac{2}{\varphi(k)\log^2 x} \sum_a \sum_{n \leqslant x} \log n \left|R_a\left(\frac{x}{n}\right)\right| + O\left(\frac{x}{\log x}\right),$$

which can be proved in a similar way.

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

we get by partial summation

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log^2 \frac{x}{p} = \sum_{\substack{pqr \leqslant x \\ pqr \equiv l(k)}} \frac{\log p \log q \log r}{\log pq} \log^2 \frac{x}{pqr} + O(x \log \log x)$$

$$\geqslant \frac{1}{\log x} \sum_{\substack{pqr \leqslant x \\ pqr \equiv l(k)}} \log p \log q \log r \ \log^2 \frac{x}{pqr} + O(x \log \log x).$$

Now it is easily seen[5] that if $pqr \leqslant x$,

$$\log^2 \frac{x}{pqr} = 2 \sum_{\substack{p \leqslant \mu,\, q \leqslant \nu \\ \mu\nu \leqslant x/r}} \frac{1}{\mu\nu} + O\left(\left(\frac{1}{p} + \frac{1}{q}\right)\left(1 + \log \frac{x}{pqr}\right)\right).$$

Inserting this expression in the preceding inequality, we get

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log^2 \frac{x}{p}$$

$$\geqslant \frac{2}{\log x} \sum_{\substack{pqr \leqslant x \\ pqr \equiv l(k)}} \log p \log q \log r \sum_{\substack{p \leqslant \mu,\, q \leqslant \nu \\ \mu\nu \leqslant x/r}} \frac{1}{\mu\nu} + O(x \log \log x).$$

$$= \frac{2}{\log x} \sum_{\alpha\beta\gamma \equiv l(k)} \sum_{\mu\nu \leqslant x} \frac{1}{\mu\nu} \vartheta_\alpha(\mu)\, \vartheta_\beta(\nu)\, \vartheta_\gamma\left(\frac{x}{\mu\nu}\right) + O(x \log \log x),$$

or

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log^2 \frac{x}{p} \geqslant \frac{2}{\log x} \sum_{\alpha\beta\gamma \equiv l(k)} \sum_{\substack{x^{\frac{1}{4}} \leqslant \mu \leqslant x^{\frac{1}{3}} \\ x^{\frac{1}{4}} \leqslant \nu \leqslant x^{\frac{1}{3}}}} \frac{1}{\mu\nu} \vartheta_\alpha(\mu)\vartheta_\beta(\nu)\vartheta_\gamma\left(\frac{x}{\mu\nu}\right) + O(x \log \log x). \tag{3.1}$$

Lemma 1. *If χ is a real non-principal character mod k, then for $x > x_0$ we have*

$$\sum_{\chi(a) = 1} \vartheta_a(x) > K_1 x, \quad \sum_{\chi(a) = -1} \vartheta_a(x) > K_1 x.$$

It is obviously sufficient if we prove that

$$\sum_{\substack{p \leqslant x \\ \chi(p) = 1}} \frac{\log p}{p} = \tfrac{1}{2} \log x + O(1), \tag{3.2}$$

because then, if $0 < \delta < 1$ is a fixed number, we get

$$\sum_{\chi(a) = 1} \vartheta_a(x) > \delta x \sum_{\substack{\delta x < p \leqslant x \\ \chi(p) = 1}} \frac{\log p}{p} = \tfrac{1}{2}\left(\delta \log \frac{1}{\delta}\right) x + O(\delta x) > K_1 x,$$

if δ is chosen small enough and $x > x_0$. The second part of the lemma follows in the same way by combining (1.6) and (3.2).

---

[5]For example, by noting that $\displaystyle\sum_{q \leqslant \nu \leqslant x/\mu r} \frac{1}{\nu} = \sum_{\mu \leqslant \nu \leqslant x/qr} \frac{1}{\nu} + O\left(\frac{1}{q} + \frac{1}{\mu}\right).$

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

To prove (3.2) we remark that[6] to each such character χ there corresponds an integer D, which is not a square, with $|D| < k^2$, and such that $\chi(p) = (D|p)$ for all primes p. Here $(D|p)$ is the usual quadratic residue symbol. Hence (3.2) is equivalent to

$$\sum_{\substack{p \leqslant x \\ (D|p)=1}} \frac{\log p}{p} = \tfrac{1}{2}\log x + O(1). \tag{3.3}$$

To prove (3.3) we consider the product

$$P = \prod_{\substack{|u| \leqslant \sqrt{x/2} \\ |v| \leqslant \sqrt{x/2|D|}}}' |u^2 - Dv^2|, \tag{3.4}$$

where the dash Π' indicates that the term $u = v = 0$ is omitted. It is easily seen that[7]

$$\log P = \frac{2x}{\sqrt{|D|}} \log x + O(x). \tag{3.5}$$

Let us estimate the highest power dividing P of a prime $p \leqslant x$.

First assume that $(D|p) = 1$. We first estimate how many solutions the congruence

$$u^2 - Dv^2 \equiv 0 \pmod{p}, \tag{3.6}$$

has in the given range for u and v. Since $(D|p) = 1$ there clearly exist solutions of (3.6) which are nontrivial i.e. with $(u, p) = (v, p) = 1$. Let now u_0, v_0 be a fixed such nontrivial solution. Then if u, v also is a solution we have

$$(uv_0)^2 - (u_0v)^2 \equiv 0 \pmod{p},$$

or one of the congruences

$$u\, v_0 \mp u_0.\, v \equiv 0 \pmod{p}, \tag{3.7}$$

must be satisfied. Conversely a solution u, v of (3.7) is a solution of (3.6). Consider (3.7) with the upper sign. Obviously the "vectors" (u, v) satisfying (3.7) form a two-dimensional lattice. Since there exist integers (u, v) with $uv_0 - u_0v = p$, the area of a "period-parallelogram" or single "cell" in the lattice is p, because it obviously could not be less than p. Also no "vector" in the lattice has a length less than $\sqrt{p/|D|}$, since for $(u, v) \neq (0,0)$ we have

$$u^2 + v^2 \geqslant |u^2 - Dv^2|/|D| \geqslant p/|D|.$$

From this it is easily seen[8] that the lattice has a basis[9] of two vectors both $<2\sqrt{|D|p}$, and hence that the rectangle $|u| \leqslant \sqrt{x/2}$, $|v| \leqslant \sqrt{x/2|D|}$ with area $2x/\sqrt{|D|}$ contains

[6]See for instance Dirichlet-Dedekind: *Vorlesungen über Zahlentheorie* (the beginning of §135).

[7]For example, by showing that the number of terms with $|u^2 - Dv^2| \leqslant T$ is $O(\sqrt{xT})$.

[8]For example, by noting that a "period-parallelogram" may always be chosen so that neither of its sides is greater than a diagonal.

[9]Or, otherwise expressed, that the lattice may be built up of "period-parallelograms" with both sides $< 2\sqrt{|D|p}$.

$$\frac{2x}{p\sqrt{|D|}} + O\left(\sqrt{\frac{x}{p}}\right)$$

such lattice points (u, v). Hence we have as many solutions of (3.7) with the upper sign.

Treating the case of the lower sign in the same way, we get altogether

$$\frac{4x}{p\sqrt{|D|}} + O\left(\sqrt{\frac{x}{p}}\right) + O\left(\frac{x}{p^2}\right),$$

solutions of (3.6) in the given range for u and v, because the two congruences (3.7) only have common solutions with $u \equiv v \equiv 0 \pmod{p}$, the number of which is $O(x/p^2)$ since we exclude the solution $u = v = 0$.

Thus

$$\frac{4x}{p\sqrt{|D|}} + O\left(\sqrt{\frac{x}{p}}\right) + O\left(\frac{x}{p^2}\right)$$

of the factors $|u^2 - Dv^2|$ of P contain p as a factor.

In the same way we find that $O(x/p^i)$ of them contain p^i as a factor for $i > 1$. Thus the highest power of p dividing P has the exponent

$$\frac{4x}{p\sqrt{D}} + O\left(\sqrt{\frac{x}{p}}\right) + O\left(\frac{x}{p^2}\right).$$

On the other hand, if $(D|p) = -1$, we see that p has to divide both u and v in order to divide $|u^2 - Dv^2|$. From this it is easily seen that P in this case contains p only to a power with exponent $O\,(x/p^2)$.

Finally if $(D|p) = 0$ or p divides D, we see that P contains p to a power with exponent $O(x/p)$. Combining these results we get

$$\log P = \frac{4x}{\sqrt{|D|}} \sum_{\substack{p \leqslant x \\ (D|p)=1}} \frac{\log p}{p} + O\left(\sqrt{x} \sum_{p \leqslant x} \frac{\log p}{\sqrt{p}}\right) + O\left(x \sum_{p \leqslant x} \frac{\log p}{p^2}\right)$$
$$+ O\left(x \sum_{p/D} \frac{\log p}{p}\right) = \frac{4x}{\sqrt{|D|}} \sum_{\substack{p \leqslant x \\ (D|p)=1}} \frac{\log p}{p} + O(x).$$

Comparing this with (3.5) we get (3.3), which proves our lemma.

LEMMA 2. *Let $h = \frac{1}{2}\varphi(k)$, and suppose that we have three sets of h different residues*[10] *mod k: $\alpha_1, \alpha_2, \ldots, \alpha_h; \beta_1, \beta_2, \ldots, \beta_h; \gamma_1, \gamma_2, \ldots, \gamma_h$. Further suppose that for each non-principal real character χ mod k there is at least one α_i with $\chi(\alpha_i) = 1$, and at least one with $\chi(\alpha_i) = -1$. Then there exists a triple α, β, γ belonging to the sets, such that $\alpha\beta\gamma \equiv l \pmod{k}$.*

Suppose that always $\alpha\beta\gamma \not\equiv l \pmod{k}$, or $\alpha\beta \not\equiv l\bar{\gamma} \pmod{k}$. This implies that there are h different values[11] the product $\alpha\beta$ cannot assume, or since $h = \frac{1}{2}\varphi(k)$, that the product $\alpha\beta$ can assume only h different values. Writing $\alpha_i = \alpha_1 \alpha'_i$ and $\beta_i = \beta_1 \beta'_i$ for $i = 1, 2, 3, \ldots, h$, this means that the products $\alpha'\beta'$ can assume only h different values. Since $\alpha'\beta'$ can assume the values 1,

[10]By residues, we understand here residues belonging to the reduced residue system.

[11]By values we mean here residues mod k.

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

$a'_2, \ldots, a'_h$ and also the values $1, \beta'_2, \ldots, \beta'_h$, the sets a' and β' are identical, and it follows easily that the set $1, a'_2, \ldots, a'_h$, forms a group with respect to multiplication. Now we define a character χ having the value 1 for all residues of the set a', and the value -1 for all remaining residues of the reduced residue system. Then we have $\chi(a_i) = \chi(a_1)\chi(a'_i) = \chi(a_1)$, which contradicts the assumption that the set a contains both a_i's with $\chi(a_i) = 1$, and such with $\chi(a_i) = -1$. This proves our lemma.

LEMMA 3. *We have for* $x > x_0$,

$$\vartheta_l(x) > K_2 x.$$

From (2.3) follows for all a and $x > x_0$,

$$\vartheta_a(x) \leqslant \frac{2}{\varphi(k)}x + O\left(\frac{x}{\log x}\right) < \frac{2}{\varphi(k)}\left(1 + \frac{1}{2\varphi(k)}\right)x. \tag{3.8}$$

Also from the prime-number theorem in the case $k = 1$, we get for $x > x_0$,

$$\sum_a \vartheta_a(x) = x + o\,(x) > \left(1 - \frac{1}{2\varphi(k)}\right)x. \tag{3.9}$$

The inequality

$$\vartheta_a(x) > \frac{x}{\varphi^2(k)}$$

must hold for at least $h = \frac{1}{2}\varphi(k)$ values of a. For if $\vartheta_a(x) \leqq x/\varphi^2(k)$ for $m > h$ values of a then, by (3.8) and (3.9),

$$\left(1 - \frac{1}{2\varphi(k)}\right)x < \sum_a \vartheta_a(x) < \left(\frac{m}{\varphi^2(k)} + \frac{2(\varphi(k) - m)}{\varphi(k)}\left(1 + \frac{1}{2\varphi(k)}\right)\right)x,$$

which leads to a contradiction.

Also, from Lemma 1, there is at least one a with $\chi(a) = 1$ and at least one with $\chi(a) = -1$ satisfying

$$\vartheta_a(x) > \frac{2K_1 x}{\varphi(k)}.$$

Hence there exists a set of $h = \frac{1}{2}\varphi(k)$ different residues $a_1, a_2, \ldots, a_h$ mod k, for each μ in the range $x^{\frac{1}{4}} \leqslant \mu \leqslant x^{\frac{1}{3}}$, $x > x_0$ such that

$$\vartheta_{a_i}(\mu) > K_3 \mu, \tag{3.10}$$

for $i = 1, 2, \ldots, h$ where

$$K_3 = \min\left(\frac{2}{\varphi(k)}K_1, \frac{1}{\varphi^2(k)}\right),$$

and such that for any real non-principal character χ mod k there is an a_i in the set with $\chi(a_i) = 1$, and another with $\chi(a_i) = -1$.

Arguing in the same way for $\vartheta_\beta(\nu)$ and $\vartheta_\gamma\,(x/\mu\nu)$ where $x^{\frac{1}{4}} \leqslant \mu \leqslant x^{\frac{1}{3}}$, $x^{\frac{1}{4}} \leqslant \nu \leqslant x^{\frac{1}{3}}$, $x > x_0$, we find from Lemma 2 that for each pair μ, ν in the ranges $x^{\frac{1}{4}} \leqslant \mu \leqslant x^{\frac{1}{3}}$, $x^{\frac{1}{4}} \leqslant \nu \leqslant x^{\frac{1}{3}}$ for $x > x_0$ there is a triple a, β, γ with

$$\vartheta_\alpha(\mu) > K_3\mu,\ \vartheta_\beta(\nu) > K_3\nu,\ \vartheta_\gamma(x/\mu\nu) > K_3x/\mu\nu,$$

and $\alpha\beta\gamma \equiv l \pmod{k}$. Hence (3.1) gives

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log^2 \frac{x}{p} > \frac{2}{\log x} K_3^3\, x \Big(\sum_{x^{\frac{1}{3}} \leqslant \mu \leqslant x^{\frac{1}{2}}} \frac{1}{\mu}\Big)^2 + O(x \log \log x)$$

$$= \frac{2}{\log x} K_3^3\, x\, (\tfrac{1}{12} \log x)^2 + O(x \log \log x) > K_4\, x \log x,$$

for $x > x_0$.

Thus if $\delta < 1$ is a fixed positive number, we get

$$\log x \log^2 1/\delta \,.\, \vartheta_l(x) \geqslant \sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log^2 x/p$$

$$- \sum_{p \leqslant \delta x} \log^2 p \log^2 x/p > K_4\, x \log x$$

$$- K_5\, \delta x \log x \log^2 1/\delta = (K_4 - K_5\, \delta \log^2 1/\delta)\, x \log x,$$

or if δ is chosen small enough,

$$\vartheta_l(x) > K_2 x,$$

for $x > x_0$, which proves Lemma 3.

From (2.2) we get, using Lemma 3,

$$\vartheta_l(x) \leqslant \frac{2}{\varphi(k)}\, x - \frac{K_2 x}{\log x} \sum_{p \leqslant x} \frac{\log p}{p} + O\Big(\frac{x}{\log x}\Big),$$

or

$$\vartheta_l(x) < \Big(\frac{2}{\varphi(k)} - K_6\Big)\, x,$$

for $x > x_0$. This combined with Lemma 3 and (2.5) gives

$$|R_l(x)| < \frac{\sigma_0}{\varphi(k)}\, x, \tag{3.11}$$

for $x > x_0$, where $\sigma_0 < 1$ is a positive number depending on k only.

4. Properties of $R_l(x)$. From (2.1) we get by partial summation

$$\sum_{\substack{p \leqslant x \\ p \equiv l(k)}} \log^2 p \log \frac{x}{p} + \sum_{\substack{pq \leqslant x \\ pq \equiv l(k)}} \log p \log q \log \frac{x}{pq} \tag{4.1}$$

$$= \frac{2}{\varphi(k)} \sum_{n \leqslant x} \log n \log \frac{x}{n} + O(x) = \frac{2}{\varphi(k)}\, x \log x + O(x).$$

Now

$$\log \frac{x}{p} = \sum_{p \leqslant n \leqslant x} \frac{1}{n} + O\Big(\frac{1}{p}\Big),$$

and

$$\log \frac{x}{pq} = \sum_{p \leqslant n \leqslant x/q} \frac{1}{n} + O\Big(\frac{1}{p}\Big).$$

Inserting this in (4.1), we get

$$\sum_{n \leqslant x} 1/n \sum_{\substack{p \leqslant n \\ p \equiv l(k)}} \log^2 p + \sum_{n \leqslant x} 1/n \sum_{p \leqslant n} \log p \sum_{\substack{q \leqslant x/n \\ q \equiv \bar{l}p(k)}} \log q$$

$$= \frac{2}{\varphi(k)} x \log x + O(x),$$

which may be written in the form

$$\sum_{n \leqslant x} \frac{\log n}{n} \vartheta_l(n) + \sum_{n \leqslant x} \frac{1}{n} \sum_{\alpha\beta \equiv l(k)} \vartheta_\alpha(n)\vartheta_\beta\left(\frac{x}{n}\right) = \frac{2}{\varphi(k)} x \log x + O(x).$$

This gives

$$\sum_{n \leqslant x} \frac{\log n}{n} R_l(n) + \frac{1}{\varphi(k)} x \log x + O(x) = \frac{1}{\varphi(k)} x \log x$$

$$- \sum_{n \leqslant x} \frac{1}{n} \sum_{\alpha\beta \equiv l(k)} R_\alpha(n) R_\beta\left(\frac{x}{n}\right) - \frac{1}{\varphi(k)} \sum_{n \leqslant x} \sum_{\alpha} R_\alpha\left(\frac{x}{n}\right)$$

$$- \frac{x}{\varphi(k)} \sum_{n \leqslant x} \frac{1}{n^2} \sum_{\alpha} R_\alpha(n) + O(x),$$

or, using (1.7) and (1.8), and noticing that

$$\sum_{\alpha} R_\alpha(y) = \vartheta(y) - y + O(1),$$

we get finally

$$\sum_{n \leqslant x} \frac{\log n}{n} R_l(n) = - \sum_{n \leqslant x} \frac{1}{n} \sum_{\alpha\beta \equiv l(k)} R_\alpha(n) R_\beta\left(\frac{x}{n}\right) + O(x). \tag{4.2}$$

Suppose now that for a positive fixed number $\sigma \leqslant \sigma_0$, we have for $x > x_\sigma$,

$$|R_l(x)| < \frac{\sigma}{\varphi(k)} x, \tag{4.3}$$

for all $(l, k) = 1$. (4.2) then yields

$$\left| \sum_{n \leqslant x} \frac{\log n}{n} R_l(n) \right| < \frac{\sigma^2}{\varphi(k)} x \sum_{n \leqslant x} \frac{1}{n} + O(x) = \frac{\sigma^2}{\varphi(k)} x \log x + O(x)$$

or if $x_1 < x$,

$$\left| \sum_{x_1 \leqslant n \leqslant x} \frac{\log n}{n} R_l(n) \right| < \frac{\sigma^2}{\varphi(k)} (x + x_1) \log x + O(x).$$

Now let

$$x_1 = (1 - \sigma)^{16} x < \frac{1 - \sigma}{1 + 15\sigma} x.$$

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

If $R_l(n)$ does not change sign in the interval $x_1 < n \leqslant x$, then the above inequality implies that there exists a y in the interval $x_1 < y \leqslant x$, such that

$$\left|\frac{R_l(y)}{y}\right| \sum_{x_1 \leqslant n \leqslant x} \log n < \frac{\sigma^2}{\varphi(k)} (x + x_1) \log x + O(x),$$

or

$$\left|\frac{R_l(y)}{y}\right| (x - x_1) \log x < \frac{\sigma^2}{\varphi(k)} (x + x_1) \log x + O(x),$$

and

$$(4.4) \qquad \left|\frac{R_l(y)}{y}\right| < \frac{\sigma^2}{\varphi(k)} \frac{x + x_1}{x - x_1} + O\left(\frac{1}{\log x}\right)$$
$$< \frac{1}{\varphi(k)} \frac{\sigma(1 + 7\sigma)}{8} + O\left(\frac{1}{\log x}\right) < \frac{1}{\varphi(k)} \frac{\sigma(1 + 3\sigma)}{4},$$

for $x > x_\sigma$. Obviously there exists such a y also in the case that $R_l(n)$ changes sign[12] in the interval $x_1 < n \leqslant x$.

For $y_1 < y_2$, it follows from (2.3) that

$$0 \leqslant \sum_{\substack{y_1 < p \leqslant y_2 \\ p \equiv l(k)}} \log p \leqslant \frac{2}{\varphi(k)} (y_2 - y_1) + O\left(\frac{y_2}{\log y_2}\right),$$

or

$$|R_l(y_2) - R_l(y_1)| < \frac{1}{\varphi(k)} (y_2 - y_1) + O\left(\frac{y_2}{\log y_2}\right),$$

so that if $\frac{1}{2} \leqslant y'/y \leqslant 2$, and $x_1 < y \leqslant x$, $x_1 < y' \leqslant x$, we get

$$|R_l(y') - R_l(y)| \leqslant \frac{1}{\varphi(k)} |y' - y| + O\left(\frac{y}{\log x}\right).$$

Thus

$$|R_l(y')| < |R_l(y)| + \frac{1}{\varphi(k)} (y' - y) + O\left(\frac{y}{\log x}\right),$$

or

$$\left|\frac{R_l(y')}{y'}\right| < \left|\frac{R_l(y)}{y}\right| \frac{y}{y'} + \frac{1}{\varphi(k)} \left|1 - \frac{y}{y'}\right| + O\left(\frac{1}{\log x}\right).$$

Now let

$$e^{-\delta} \leqslant \frac{y'}{y} \leqslant e^{\delta}$$

where

$$\delta = \frac{\sigma(1 - \sigma)}{32}.$$

The above inequality then gives, by (4.4),

[12]For there is then a y in the interval with $|R_l(y)| < \log y$.

$$\left|\frac{R_l(y')}{y'}\right| < \frac{1}{\varphi(k)} \frac{\sigma(1+3\sigma)}{4} e^{\delta} + \frac{1}{\varphi(k)} (e^{\delta} - 1) + O\left(\frac{1}{\log x}\right)$$
$$< \frac{1}{\varphi(k)} \frac{\sigma(3+5\sigma)}{8} + O\left(\frac{1}{\log x}\right) < \frac{1}{\varphi(k)} \frac{\sigma(1+\sigma)}{2},$$

for $x > x_\sigma$. Thus we have proved, assuming (4.3),

LEMMA 4. *For $x > x_\sigma$, any interval $((1-\sigma)^{16} x, x)$ contains a sub-interval $(y, e^{\delta}y)$ where $\delta = \dfrac{\sigma(1-\sigma)}{32}$, such that for all $y \leqslant z \leqslant e^{\delta}y$ we have*

$$\left|\frac{R_l(z)}{z}\right| < \frac{1}{\varphi(k)} \frac{\sigma+\sigma^2}{2}.$$

5. Proof of the prime-number theorem for the arithmetic progression. We shall now prove the

THEOREM.

$$\lim_{x \to \infty} \frac{\vartheta_l(x)}{x} = \frac{1}{\varphi(k)}.$$

Obviously this is equivalent to

$$\lim_{x \to \infty} \frac{R_l(x)}{x} = 0. \tag{5.1}$$

We have that for all $x > 1$ and $(k,l) = 1$,

$$|R_l(x)| < K_7 x, \tag{5.2}$$

and from (3.11), that for $x > x_0$,

$$|R_l(x)| < \frac{\sigma_0}{\varphi(k)} x, \tag{5.3}$$

where $\sigma_0 < 1$.

Now assume as in the preceding paragraph, that for a fixed positive number $\sigma \leqslant \sigma_0$, we have for all l,

$$|R_l(x)| < \frac{\sigma}{\varphi(k)} x, \tag{5.4}$$

for $x > x_\sigma$. Writing further $\rho = (1-\sigma)^{-16}$, we have then from Lemma 4, that for $x > x_\sigma$ and all α each interval $(\rho^{\nu-1}, \rho^{\nu})$ where $\rho \leqslant \rho^{\nu} \leqslant x/x_\sigma$ contains a sub-interval (y_ν, y'_ν) with $y'_\nu = e^{\delta} y_\nu$ and $\delta = \dfrac{\sigma(1-\sigma)}{32}$, such that for $y_\nu \leqslant n \leqslant y'_\nu$ we have

$$\left|\frac{n}{x} R_\alpha\left(\frac{x}{n}\right)\right| < \frac{1}{\varphi(k)} \frac{\sigma+\sigma^2}{2}.$$

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press

(2.8) then yields

$$
\begin{aligned}
|R_l(x)| &< \frac{1}{\varphi(k)\log x}\sum_{a}\sum_{n\leqslant x}\left|R_a\left(\frac{x}{n}\right)\right| + O\left(\frac{x}{\sqrt{\log x}}\right) \\
&< \frac{K_7 x}{\log x}\sum_{x/x_\sigma\leqslant n\leqslant x}\frac{1}{n} + \frac{\sigma x}{\varphi(k)\log x}\sum_{n\leqslant x/x_\sigma}\frac{1}{n} \\
&\quad - \frac{(\sigma-\sigma^2)x}{2\varphi^2(k)\log x}\sum_{a}\sum_{\rho^\nu\leqslant x/x_\sigma}\sum_{y_\nu\leqslant n\leqslant y'_\nu}\frac{1}{n} + O\left(\frac{x}{\sqrt{\log x}}\right) \\
&= \frac{\sigma}{\varphi(k)}x - \frac{(\sigma-\sigma^2)x}{2\varphi(k)\log x}\sum_{\rho^\nu\leqslant x/x_\sigma}\delta + O\left(\frac{x}{\sqrt{\log x}}\right) \\
&= \frac{\sigma}{\varphi(k)}x - \frac{\delta\sigma(1-\sigma)}{2\varphi(k)\log\rho}x + O\left(\frac{x}{\sqrt{\log x}}\right) \\
&= \frac{\sigma}{\varphi(k)}\left(1 - \frac{\sigma(1-\sigma)^2}{1024\log 1/1-\sigma}\right)x + O\left(\frac{x}{\sqrt{\log x}}\right) \\
&< \frac{\sigma}{\varphi(k)}\left(1 - \frac{(1-\sigma)^3}{2000}\right)x,
\end{aligned}
$$

for $x > x_\sigma$.

Since the iteration-process

$$\sigma_{n+1} = \sigma_n\left(1 - \frac{(1-\sigma_n)^3}{2000}\right),$$

starting with $0 < \sigma_0 < 1$, converges to zero (one sees easily that $\sigma_n < e^{-K_8 n}$), this proves (5.1) and hence our theorem.

The Institute for Advanced Study
and
Syracuse University

https://doi.org/10.4153/CJM-1950-007-5 Published online by Cambridge University Press