An asymptotic formula in additive number theory

by

P. Erdős (Budapest), G. Jogesh Babu and K. Ramachandra
(Bombay)

1. Introduction. In his paper [1], Erdős introduced the sequences of positive integers \( b_1 < b_2 < \ldots \), with \( (b_i, b_j) = 1 \), for \( i \neq j \), and \( \sum b_i^{-1} < \infty \). With any such arbitrary sequence of integers, he associated the sequence \( \{ d_i \} \) of all positive integers not divisible by any \( b_i \), and he showed that if \( b_1 \geq 2 \), there exists a \( \theta < 1 \) (independent of the sequence \( \{ b_i \} \)) such that \( d_{i+1} - d_i < d_i^\theta \), for \( i \geq l_\theta \). Later, Szemerédi [4] made an important progress on the problem, showing that \( \theta \) can be taken to be any number greater than \( \frac{1}{2} \).

In this paper, we study this sequence from a different point of view. We study the number \( N(n) \) of solutions of the equation \( \equiv p + d \), where \( p \) is a prime and \( d \equiv 0 \pmod{b_j} \) for any \( j \). In fact we derive an asymptotic formula for \( N(n) \), when \( b_1 \geq 3 \). We also study \( N(n) \) when the condition \( (b_i, b_j) = 1 \) is dropped.

2. In what follows, we let \( C_1, C_2, \ldots \) denote positive absolute constants and let \( C \) be a positive constant. \( p, q \) with or without subscript, always denote primes.

Theorem 1. Let \( 2 \leq b_1 < b_2 < \ldots \) be a sequence of natural numbers with the properties \( (b_i, b_j) = 1 \) whenever \( i \neq j \) and

\[
\sum_{j=1}^{\infty} b_j^{-1} < \infty.
\]

Then the number \( N(n) \) of solutions of the equation \( \equiv p + t \), where \( p \) is a prime and \( t \) is a natural number not divisible by any \( b_j \), is given by

\[
N(n) = n \left( \frac{\log n}{n} \right)^{-1} \prod_{b_j \mid n} \left( 1 - \frac{1}{(p(b_j))^{-1}} \right) + o \left( n \left( \frac{\log n}{n} \right)^{-1} \right).
\]

Remarks. If either \( b_1 \geq 3 \) or if \( n \) is even then \( N(n) \) is asymptotic to the main term in (2.2). Similar remarks apply to Theorem 2 below, which can be proved along the same lines as Theorem 1. Also it easily follows from
the prime number theorem for arithmetic progressions and the sieve of Eratosthenes, that if \((b_1, b_2) = 1\) and \(\sum_{i=1}^{\infty} \frac{1}{b_i} = \infty\) then \(N'(n) = o\left(\frac{n}{\log n}\right)\).

**Theorem 2.** Let \(l\) be any non-zero integer. Under the assumptions of Theorem 1, the number \(N_1(x)\), of primes \(p\) not exceeding \(x\) such that \(p + l\) is not divisible by any \(b_i\), satisfies

\[N_1(x) = x(\log x)^{-1} \sum_{\delta \neq 1} \left(1 - \left(\frac{\varphi(b_i)}{\delta}\right)^{-1}\right) + o\left(x(\log x)^{-1}\right).\]

**3. Proof of Theorem 1.** We denote by \(\gamma\), natural numbers not divisible by any \(b_i\), and by \(\delta\) all finite power products \(\prod b_i\delta\) where \(\delta \neq 0\) or \(1\), and we write \(h(\delta) = (-1)^{\delta_i}\). We begin with

**Lemma 1.** We have

\[\sum_{\delta} \varphi(\delta) = \zeta(1) \prod_{n=2}^{\infty} \left(1 - b_i^{-s}\right) \text{ and } \prod_{n=2}^{\infty} \left(1 - b_i^{-s}\right) = \sum_{\delta} h(\delta) \delta^{-s}\]

**Proof.** The proof follows from the fact that every natural number \(m\) can be written uniquely in the form

\[m = \left(\prod b_i\right)^{\alpha_i} \quad (\alpha_i \geq 0 \text{ are integers}).\]

This can be proved in the following way. Define \(\alpha_i\) as the greatest integer such that \(b_i\delta\) divides \(m\). This gives existence and the uniqueness is trivial.

**Lemma 2.** The two series

\[\sum_{\delta} \varphi(\delta)^{-1} \quad \text{and} \quad \sum_{\delta} h(\delta)^{-1}\]

are convergent.

**Proof.** Let \(B_1\) be the set of those \(b_i\)'s which are primes and let \(B_2\) be the set of the remaining \(b_i\)'s. Clearly, the number of \(b_i\)'s in \(B_2\) not exceeding \(x\) is less than \(\sqrt{x}\). Thus (2.1) implies convergence of the first series. Convergence of the second series follows from convergence of the first series and the identity

\[\sum \varphi(\delta)^{-1} = \prod \left(1 - \varphi(b_i)^{-1}\right).\]

**Lemma 3.** Let \(N'(n)\) be the number of solutions of

\[n = p + t', \quad t' > 0, \quad t' \equiv 0 \pmod{b_i} \text{ for every } b_i \leq \log n.\]

Then

\[N'(n) = n(\log n)^{-1} \sum_{\delta \neq 1} \left(1 - \left(\frac{\varphi(b_i)}{\delta}\right)^{-1}\right) + o\left(n(\log n)^{-1}\right).\]

**Proof.** Let \(d'\) denote a product of the form \(\prod b_i\delta\), where \(\delta \neq 0\) or \(1\) and \(b_i \leq \log n\). By Siegel–Walfisz theorem (see [3], Satz 3.3, p. 144)

and by Lemmas 1 and 2, we have

\[N'(n) = \sum_{n=p+t'} h(d') = \sum_{n=p+t'} h(d') + \sum_{n=p+t'} h(d') = \Sigma_1 + \Sigma_2.\]

Note that, if \(\delta(\delta)\) denotes the number of divisors of \(\delta\), then

\[\Sigma_2 = \sum_{p+t' \leq n} h(d') \leq \sum_{p+t' \leq n} h(d') \leq \sum_{p+t' \leq n} \delta(d') \leq \sum_{p+t' \leq n} \delta(d') \leq n^{1/2} \log n,\]

since \(h(d') \leq 1\) and \(\delta(d) \leq n^s\) for some \(s > 0\).

\[\Sigma_1 = \sum_{p+t' \leq n} \left(\frac{h(d')}{\delta(d')} \frac{n}{\log n} \left(1 + O\left(\frac{1}{\log n}\right)\right)\right)\]

\[= \frac{n}{\log n} \sum_{p+t' \leq n} \frac{h(d')}{\delta(d')} + o(n^s).\]

Thus

\[N'(n) = \Sigma_1 + \Sigma_2 = n(\log n)^{-1} \sum_{p+t' \leq n} \left(1 - \left(\frac{\varphi(b_i)}{\delta}\right)^{-1}\right) + o(n(\log n)^{-1}).\]

This completes the proof of the lemma.

**Lemma 4.** There exists a function \(\eta(x) \to 0\) as \(x \to \infty\), such that the number of primes \(p \leq n\) satisfying

\[n = p - b_i \pmod{b_i}, \quad \text{for some } b_i \in \{n^{1/2}, n\}\]

is less than

\[(\eta(x) + o(1))n(\log n)^{-1}, \quad \text{for every } x > 0.\]

**Proof.** First note that the number of composite \(b_i\)s not exceeding \(x\) is at most \(x^{1/2}\). For a fixed \(b_i \in \{n^{1/2}, n\}\), \(n - p \equiv 0 \pmod{b_i}\) has at most \(n/b_i < n^{1/2}\) solutions. Thus the contribution of the composite \(b_i\)s is less than \(n^{1/2-1}\). To complete the proof it, thus, suffices to show that the number of solutions of

\[n = p \pmod{b}, \quad n^{1/2} - q < q < n, \quad q \text{ prime},\]

is less than

\[(\eta(x) + o(1))n(\log n)^{-1}.\]

In other words we have to prove that the number of solutions of

\[n = p + q, \quad p, q \text{ primes not exceeding } n \text{ and } a < n, \text{ is less than}\]

\[(\eta(x) + o(1))n(\log n)^{-1}.\]

First note that the number of solutions of

\[n = p + q, \quad a < n', \quad (a, n) = 1 \text{ and } p, q \text{ primes not exceeding } n\]
is less than
\[ \sum_{a < n^\epsilon} \sum_{p < \log n} 1 \leq \eta(n) = o(n/(\log n))^{-1}, \]

since \( \epsilon < 1/4 \).

Now let \( a_n \) be any sequence of non-negative integers, and \( (n, a) = 1 \), the number of primes \( q < n \), for which \( n - aq \) is a prime, by Lemma 1.4 of [2]. If \( C_4 \) is a sufficiently small constant, then
\[ C_4 \sum_{a < n^\epsilon} \prod_{p < \log n \atop p \mid (n - aq)} \left( 1 - \frac{1}{p} \right) \leq C_4 \prod_{a < n^\epsilon} \left( 1 - \frac{2}{p} \right) \prod_{a < n^\epsilon} \left( \frac{1}{p} \right), \]

Thus summing for all \( a < n^\epsilon \), \( (a, n) = 1 \) and \( p \), \( q \) primes (\( \leq n \)) is less than
\[ \eta(n) \prod_{a < n^\epsilon} \left( 1 + \frac{1}{p} \right). \]

Thus summing for all \( a < n^\epsilon \), \( (a, n) = 1 \), we immediately obtain that the number of solutions of
\[ n - aq = p, \quad a < n^\epsilon, \quad (a, n) = 1 \text{ and } p, q \text{ primes (\( \leq n \))} \]
is less than
\[ \eta(n) \prod_{a < n^\epsilon} \left( 1 + \frac{1}{p} \right). \]

Now the lemma follows easily.

To complete the proof of Theorem 1, it is enough to show, in view of Lemma 3, that
\[ N(n) - N'(n) = o(n/(\log n))^{-1}. \]

To show this it will clearly be sufficient to show that the number of solutions of
\[ n = p + R, \quad R > 0, \quad R = 0 \pmod{b_i} \] for some \( b_i > \log \log n \)
is
\[ o(n/(\log n))^{-1}. \]

First observe that if \( b_i < n^{-\epsilon} \) (\( \epsilon > 0 \), small), then the number of primes \( p \leq n \) with \( n = p \pmod{b_i} \) is, by Brun-Titchmarsh Theorem (see [3], Satz 1.1, p. 44), less than \( C_3 n/(\log(b_i) \log n) \). Thus the number of primes \( p \leq n \) for which \( n = p \pmod{b_i} \) is less than \( C_3 n/(\log(b_i) \log n) \sum_{b_i > \log n} \left( \phi(b_i) \right)^{-1} = o(n/(\log n)). \]

Now the theorem follows from Lemma 4.

4. If \( (b_i, b_j) = 1 \), for \( i \neq j \), is not assumed, it is easy to give a sequence \( 2 < b_1 < b_2 < \ldots \) for which
\[ \sum_{i=1}^{\infty} \left( \phi(b_i) \right)^{-1} < \infty, \]

but there is an infinite sequence \( 0 < n_1 < n_2 < \ldots \) so that the number of solutions of
\[ n_i = p + t, \quad p \text{ prime, } t > 0 \text{ and } t \equiv 0 \pmod{b_i}, \] for all \( j \), is
\[ o(n_i/(\log n_i)) \text{ as } i \to \infty. \]

We define \( b_1 < b_2 < \ldots \) as follows. Suppose \( \{n_i\} \) be an increasing sequence of natural numbers tending to infinity sufficiently fast and \( \epsilon_i = (\log \log n_i)^{-1} \). Now take the \( b_i 's \) to be the integers of the form
\[ n_i - p, \quad p < (1 - \epsilon_i) n_i, \quad i = 1, 2, \ldots \]
Clearly the number of
\[ n_i = p + t, \quad t > 0, \quad t \equiv 0 \pmod{b_i}, \] for all \( j \), is less than
\[ \{ t_i + o(1) \} n_i/(\log n_i) = o(n_i/(\log n_i)). \]

Since
\[ (4.1) \quad \phi(m) \geq C_4 m/(\log \log m)^{-1}, \]
we have
\[ \sum_{p < (1 - \epsilon_i) n_i} \frac{1}{\phi(n_i - p)} < C_4 \frac{n_i/(\log n_i)}{\epsilon_i} = C_4 \frac{\log n_i}{\log n_i}. \]

Thus
\[ \sum_{i=1}^{\infty} \frac{1}{\phi(b_i)} < \sum_{i=1}^{\infty} \sum_{p < (1 - \epsilon_i) n_i} (\phi(n_i - p))^{-1} \leq C_4 \sum_{i=1}^{\infty} \frac{(\log \log n_i)^2}{\log n_i} < \infty, \]
if \( n_i \to \infty \) sufficiently fast.

It might be possible to construct a sequence \( 2 < b_1 < b_2 < \ldots \) of integers such that \( \sum b_i^{-1} \) is convergent and for which
\[ n_i = p + t, \quad p \text{ prime, } t > 0 \text{ and } t \equiv 0 \pmod{b_i}, \] for all \( i \), has no solution for infinitely many \( n_i \). But we are unable to find such a sequence.

On the other hand, if \( B(\sigma) \) defined by
\[ B(\sigma) = \sum_{b_i < \epsilon} 1, \]
is not too large, then the condition \( (b_i, b_j) = 1 \), for \( i \neq j \), is quite unnecessary. In this direction, we have the following

Theorem 3. Let \( 3 \leq b_1 < b_2 < \ldots \) be a sequence of integers such that
\[ (4.2) \quad B(\sigma) = o(\sigma/(\log \sigma) \log \log \sigma). \]
Then
\[ N(n) > \Theta(1/\log n)^{-1}. \]

**Proof of Theorem 3.** Let, for any \( k \geq 1 \), \( N(n, k) \) be the number of solutions of \( n = p + t \), \( p \) prime, \( t > 0 \) and \( t \equiv 0 \pmod{b_j} \), for all \( j \leq k \), and let \( A(n, k) \) be the number of solutions of \( n = p + t \), \( t > 0 \), \( t \equiv 0 \pmod{b_j} \) for some \( j > k \). We need the following lemmas.

**Lemma 5.** For every \( k \geq 1 \), there exists \( n(k) \) such that
\[ N(n, k) > C_1 n(k)/(\log n) \log k, \quad \text{for all } n \geq n(k). \]

Proof. Since each \( b_j \geq 3 \), either \( b_j = 0 \pmod{2^p} \), or there exists a prime \( p' \geq 3 \) such that \( b_j = 0 \pmod{p'} \). Let \( l(k) \) be the number of distinct primes in the set \( \{p_i\} \). Let these be denoted by \( q_i \), \( i = 1, \ldots, l(k) \).

Note that, \( N(n, k) \) is not less than the number of solutions of
\[ n = p + t, \quad t > 0, \quad t \equiv 0 \pmod{2^p} \quad \text{and} \quad t \equiv 0 \pmod{q_i} \quad \text{for all } i \leq l(k). \]

This latter number solutions, by Theorem 1, is not less than
\[
\left( 1 - \frac{1}{\varphi(4)} \right) \prod_{i=0}^{\infty} \left( 1 - \frac{1}{\varphi(q_i)} \right) \frac{n}{\log n} + o \left( \frac{n}{\log n} \right) \\
\geq \frac{1}{2} \prod_{i=0}^{\infty} \left( 1 - \frac{1}{p_i - 1} \right) \frac{n}{\log n} + o \left( \frac{n}{\log n} \right) \\
\geq C_2 \frac{n}{\log k \log n} \quad \text{for all } n \geq n(k),
\]

where \( p_i \) is the \( i \)-th odd prime number and \( n(k) \) is a sufficiently large integer. This completes the proof of Lemma 5.

**Lemma 6.** We have
\[ \sum_{i \neq 1} \varphi(b_i)^{-1} = o(\log b_i^{-1}). \]

Proof. By (4.1), (4.2) and by partial integration, we have
\[
\sum_{i \neq 1} \frac{\log \log b_i}{b_i} \leq \sum_{i \neq 1} \frac{\log \log b_i}{b_i} + \int_b^\infty \frac{\log \log t}{t} \ dt B(t) \\
= \frac{1}{t} B(t) \log \log b_i + \int_b^\infty \frac{R(t) \log t - 1}{t^2} \ dt \\
= o(\log b_i^{-1} - 1) + o \left( \int_b^\infty \frac{dt}{t^2 (1 + \log t)} \right) = o(\log b_i^{-1}) \\
= o((\log k)^{-2}).
\]

**Lemma 7.** There exists a \( b_0 \) such that, for every \( k \geq b_0 \), there exists \( n_0(k) \) satisfying
\[ A(n, k) > C_7 \frac{n}{2 \log k \log n} \quad \text{for all } n \geq n_0(k). \]

Proof. Since the number of solutions of \( n = p + t \), \( p \equiv 0 \pmod{b_j} \), is, by Brun-Titchmarsh theorem for \( b_j \leq \sqrt{n} \), less than \( C_7 n/\varphi(b_j) \log n \), thus, for any \( k \geq 1 \), the number of solutions of
\[ n = p + t, \quad p \leq n, \quad t \equiv 0 \pmod{b_j} \]

is less than
\[ C_7 \frac{n}{2 \log k \log n}. \quad \text{(4.4)} \]

By Lemma 6, there exists a \( b_0 \) such that for \( k \geq b_0 \), (4.4) is less than
\[ \frac{C_7}{10 n(k)} \quad \text{(4.5)} \]

Let, next, \( b_j > \sqrt{n} \). By Brun-Titchmarsh theorem the number of solutions of
\[ n = p + t \quad \text{is less than} \quad \left( \frac{C_7 n}{\varphi(b_j) \log n} \right) \]

is less than
\[ \left( \frac{C_7 n}{\varphi(b_j) \log n} \right). \]

So, if \( s \geq 1 \) and \( 2^s < \sqrt{n} \), then the number of solutions of
\[ n = p + t \quad \text{is less than} \quad \left( \frac{C_7 n}{\varphi(b_j) \log n} \right). \]

If we used (4.2), since, for each \( b_j \in \{n/2, n\} \), there exists at most one prime \( p \leq n \) such that \( n = p + (b_j) \), the number of solutions of
\[ n = p + t \quad \text{is less than} \quad \left( \frac{C_7 n}{\varphi(b_j) \log n} \right). \]

By summing (4.6) over \( s \) and adding (4.7) to the result, we get that the number of solutions of
\[ n = p + (b_j), \quad \text{for some } b_j > \sqrt{n}, p < n \]

is
\[ o(n \log n)^{-1}. \]
Now the lemma follows from (4.5).
To complete the proof of Theorem 3, first note that for any \( k \geq 1 \)
\[ N(n) \geq N(n, k) - A(n, k). \]
Now the theorem follows immediately from (4.8) and Lemmas 5 and 7.
Without much difficulty we could obtain an asymptotic formula for \( N(n) \) even if we only assume
\[ B(x) = o \left( \frac{x}{\log \log \log x} \right). \]
We hope to return to this problem on another occasion.

References


MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
Budapest, Hungary

SCHOOL OF MATHEMATICS
TATA INSTITUTE OF FUNDAMENTAL RESEARCH
Ooty, Bombay 4, India

Received on 20. 4. 1974 (562)

---

Some remarks on \( L \)-functions and class numbers

by

S. Chowla (University Park, Pa.)
and J. B. FRIEDLANDER (Princeton, N.J.)

§ 1. Let \( d \) denote the discriminant of the quadratic field \( K = Q(\sqrt{d}) \), and let \( \chi \) denote the associated real primitive character. \( q_0 \) will denote a positive computable constant. We simplify matters slightly by assuming \( |d| > 4 \) so that \( K \) contains no complex roots of unity. Dirichlet's formulae now give for the class number \( h(d) \),
\[ h(d) = \begin{cases} \frac{|d|^{\frac{1}{2}}}{\pi} L(1, \chi) & \text{for } d < 0, \\ \frac{|d|^{\frac{3}{2}}}{2 \log q} L(1, \chi) & \text{for } d > 0, \end{cases} \]
where \( q \) denotes the fundamental unit of \( K \).

Hecke [3] was the first to connect the magnitude of \( L(1, \chi) \) with the question of the existence of real zeros of \( L(s, \chi) \) near \( s = 1 \). For those \( d < 0 \) for which no such zero exists he was able to give a good effective lower bound for \( h(d) \).

Recently, Goldfeld [4] has given a simple proof of the celebrated theorem of Siegel [8]. His argument is easily modified to give a simple proof of Hecke's result. Furthermore, if we let \( a \) be fixed with \( \frac{1}{2} \leq a < 1 \), then an effective lower bound for \( L(1, \chi) \) (depending on \( a \)) can be given under the assumption \( L(a, \chi) \geq 0 \). In particular, we have:

(A) Let \( \frac{1}{2} < a < 1 \) and assume \( L(a, \chi) \geq 0 \). Then, there exists \( c_1(a) \) such that
\[ L(1, \chi) > c_1(a) |d|^{1-a}. \]

(B) Let \( \delta > 0 \) and assume \( L(\frac{1}{2}, \chi) \geq 0 \). Then, there exists \( c_2(\delta) \) such that
\[ L(1, \chi) > c_2(\delta) \log |d|^{1-a} |d|^{-1/2}. \]

It is to be noted that the bound gets progressively better as \( a \) increases, approaching the Siegel bound as \( a \) approaches 1.