and finally

\[ (4.43) \quad -12 \log 2\pi \sum_{d \mid n} \mu(d) + 2\pi \varphi(q) - 12A(q) + \sum_{d \mid n} \frac{\lambda(d)}{d} = \sum_{n \geq 1} a_n(q^{(0)}) c_q(n). \]

References


Received on 31.7.1987

Bilinear form of the remainder term in the Rosser–Iwaniec sieve of dimension \( x \in (1/2, 1) \)

by

JACEK POMYKAŁA (Warszawa)

1. Introduction. It is well known that the remainder term in the linear sieve can be expressed in terms of bilinear forms \( \sum_{m < M} \sum_{n < N} a_m b_n \rho(\mathcal{A}, mn) \). This result due to H. Iwaniec was established in 1977 (see [4]). This shape of the remainder term is more flexible than the conventional one and usually improves the estimates for the sifting function since the level of uniform distribution may be increased. On the other hand, it seems that an application of Rosser's weights would lead to the best sieving limit when the dimension of the sieve lies in the interval \((\frac{1}{2}, 1)\) (see [3]). In such circumstances it is natural to ask for the analogous result to that of paper [4] in the case when \( 1/2 < x < 1 \). The aim of this paper is to prove that the remainder term in the latter case can be expressed in terms of bilinear forms defined on the product \([-1, 1]^M \times [-1, 1]^N\), where \( M, N > 1 \) satisfy

\[ MN^x - 1 = \Delta. \]

Here \( \beta = \beta(x) \) is the sieving limit and \( \Delta \) reflects the level of uniform distribution.

I would also like to thank Professor Andrzej Schinzel for his critical remarks and valuable comments concerning this paper.

Notation. Let \( \mathcal{A} = \{a_1, a_2, \ldots\} \) be a finite sequence of positive integers; \( a_i \in \mathcal{A} \) means that \( a_i \) is an element of the sequence \( \mathcal{A} \). For a given set \( \mathcal{P} \) of primes and \( z \geq 2 \) we write

\[ P(z) = \prod_{p \in \mathcal{P}, p < z} p. \]

The main object in sieve theory is the sifting function \( S(\mathcal{A}, \mathcal{P}, z) \) which represents the number of elements \( a_i \in \mathcal{A} \) such that \( (a_i, P(z)) = 1 \).

For any \( d \mid P(z) \) we consider the subsequence \( \mathcal{A}_d \) which consists of those elements \( a_i \in \mathcal{A} \) for which \( a_i \equiv 0 \pmod{d} \).
We assume that the number of elements \( a_{i} \in \mathcal{X}_{d} \), which we denote by \( |\mathcal{X}_{d}| \), is approximately equal to \( \omega(d) \frac{X}{d} - 1 \) \( X \) where \( \omega(d) \) is a multiplicative function and \( X > 0 \) is a parameter (independent of \( d \)). Formally

\[
|\mathcal{X}_{d}| = \frac{\omega(d)}{d} \cdot X + r(\mathcal{X}_{d}, d)
\]

where \( r(\mathcal{X}_{d}, d) \) is to be considered as a remainder term; \( X \) is to be chosen in such a way that \( r(\mathcal{X}_{d}, d) \) should be small on average.

It is assumed that

\[
0 < \omega(p) < p \quad \text{for } p \in \mathcal{P}
\]

and that there exists a parameter \( k \in (1/2, 1) \) such that

\[
\prod_{p \in \mathcal{P}, \omega(p) = \frac{1}{p}} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \leq \frac{\left( \frac{\ln z}{\ln w} \right)^{k}}{1 + K \ln w}
\]

for all \( z > w \geq 2 \) where \( K \) is a constant \( \geq 1 \).

Every \( k \) satisfying (3) will be called the dimension of the sieve. The conditions (1)-(3) will be regarded as axioms.

For simplicity we will use the abbreviation

\[
V(z) = \prod_{p \in \mathcal{P}(z)} \left( 1 - \frac{\omega(p)}{p} \right)
\]

All constants implied in the symbols \( O(\cdot) \) and \( \ll \) may depend on \( k \) only.

2. Rosser-Iwaniec sieve. Estimation for the main term. Let \( A, \beta > 1 \). For any positive integer \( d \) we denote by \( \Omega(d) \) the number of prime factors of \( d \). Let \( d | P(z) \), \( \Omega(d) = r \). We write

\[
d = p_{1} p_{2} \ldots p_{r} \quad \text{where } p_{s+1} < \ldots < p_{r} < z.
\]

We will use the convention that the product over the empty set is equal to one. In particular, \( d = 1 \) is equivalent to \( r = 0 \) in the above notation. Now define

\[
d^{+} = \max_{0 \leq k \leq \lfloor (r-1)/2 \rfloor} p_{2k+1} p_{2k+2} \ldots p_{r},
\]

\[
d^{-} = \max_{1 \leq k \leq \lfloor r/2 \rfloor} p_{2k+1} p_{2k+2} \ldots p_{r}.
\]

Let \( \pi = \pm \). One may define Rosser's weights as follows (see [5]):

\[
\lambda^{\pi}(d) = \begin{cases} \frac{(-1)^{\pi}}{n_{p}} & \text{if } d^{\pi} < A, \\ 0 & \text{otherwise}. \end{cases}
\]

For them the following sieve inequalities are valid (see e.g. [2], p. 159):

\[
\sum_{d | P(z)} \lambda^{\pi}(d) |\mathcal{X}_{d}| \ll S(\mathcal{X}, \mathcal{P}, z) \ll \sum_{d | P(z)} \lambda^{\pi}(d) |\mathcal{X}_{d}|.
\]

Using (1) we may write

\[
P \left( S(\mathcal{X}, \mathcal{P}, z) \right) = \pi \left( X \sum_{d | P(z)} \lambda^{\pi}(d) \frac{\omega(d)}{d} + \sum_{d | P(z)} \lambda^{\pi}(d) r(\mathcal{X}_{d}, d) \right)
\]

\[
= \pi \left( XM^{\pi}(\mathcal{X}, \mathcal{P}, z) + R(\mathcal{X}_{d}) \right) \quad \text{(by definition).}
\]

The problem of evaluating the main term \( M^{\pi} \) is very difficult and has been treated in detail in [3] (cf. also [7]). We only formulate the final result:

**Lemma 1.** Let \( s = \ln A/\ln z \). Under the axioms (2), (3) we have

\[
M^{\pi}(\mathcal{X}, \mathcal{P}, z) \ll V(z) \left| f(s) + O(e^{\pi^{-1}(\ln A)^{-3/2}}) \right| \quad \text{if } z \leq A,
\]

\[
M^{\pi}(\mathcal{X}, \mathcal{P}, z) \ll V(z) \left| f(s) + O(e^{\pi^{-1}(\ln A)^{-3/2}}) \right| \quad \text{if } z \leq A^{1/2}.
\]

Here \( \beta - 1 \) is the largest real zero of

\[
y(s) = s^{2s-1} + \frac{1}{F(1-2s)} e^{-\pi s} \left\{ \exp \left( \frac{z^{1-e^{-u}}}{u} \right) - 1 \right\} z^{-2s} dz.
\]

**Remark 1.** \( f(s) \) and \( f(s) \) are the familiar functions of upper and lower bound respectively for the sifting function \( S(\mathcal{X}, \mathcal{P}, z) \). In the general case (\( \pi \geq 0 \)) they are the continuous solutions of the following system of differential-difference equations (see e.g. [6]):

\[
s^{\pi} f(s) = A_{\pi} \quad \text{if } s \leq \beta,
\]

\[
s^{\pi} f(s) = B_{\pi} \quad \text{if } s \geq \beta.
\]

The optimal \( \beta \) (sieving limit) should be equal to \( \inf \left\{ \beta \mid f(s) > 0 \text{ for all } s \geq \beta \right\} \) (or \( \beta = 1 \) if \( f(s) > 0 \) for all \( s \)). The correct choice of \( A_{\pi} \) and \( B_{\pi} \) is to be inferred from the behaviour of \( f(s) = 1 + O(e^{-\pi}) \) and \( f(s) = 1 + O(e^{-\pi}) \) as \( s \to \infty \). In the case when \( 1/2 < \pi < 1 \) we have \( B_{\pi} = 0 \) and it turns out that the sieving limit is defined by (6) (see [3]).

3. The remainder term. Main result. In view of (5) we should deal with the remainder term in the form

\[
\sum_{d \leq A} \lambda^{\pi}(d) \cdot r(\mathcal{X}_{d}, d).
\]
Transformation of this sum into a bilinear form does not proceed directly. In fact, it requires that a certain modification into the Rosser-Iwaniec sieve be introduced previously. Proceeding similarly to the proof of Theorem 1 of [4] we will prove the following

**Theorem 1.** Let \( \beta = \beta(x) \) be fixed, \( 0 < \varepsilon \leq 1/3 \), \( \Delta > 1 \). Assume that \( M, N > 1 \) satisfy the condition \( MN^{\varepsilon} = \Delta \). If the axioms (1)-(3) hold and \( z \leq \Delta^{\varepsilon/3} \) then:

\[
S(\mathcal{A}, \mathcal{P}, z) \leq XV(z) \left\{ F \left( \frac{\ln \Delta}{\ln z} \right) + E(\varepsilon, \Delta, K) \right\} + R^{+}(\mathcal{A}, M, N),
\]

(7)

\[
S(\mathcal{A}, \mathcal{P}, z) \geq XV(z) \left\{ F \left( \frac{\ln \Delta}{\ln z} \right) - E(\varepsilon, \Delta, K) \right\} + R^{-}(\mathcal{A}, M, N),
\]

(8)

where for the error term \( E(\varepsilon, \Delta, K) \) we have the estimate

\[
E(\varepsilon, \Delta, K) \leq \varepsilon + \varepsilon e\frac{1}{14} e^{K}(\ln \Delta)^{-1/3}
\]

(9)

and the remainder term \( R^{\pm}(\mathcal{A}, M, N) \) has the form

\[
R^{\pm}(\mathcal{A}, M, N) = \sum_{j \in \text{exp}(13e^{-3} \varepsilon)} \sum_{m \in \text{P}(c)} a_{n,j}(M, N, e) \sum_{n \in \text{N}} b_{n,j}(M, N, e) r(\mathcal{A}, mn).
\]

(10)

Here the coefficients \( a_{n,j}(M, N, e) \), \( b_{n,j}(M, N, e) \) are real and satisfy \( |a_{n,j}| \leq 1 \), \( |b_{n,j}| \leq 1 \).

**Remark 2.** The essential difference when compared with Theorem 1 of [4] is the more general condition \( \Delta = MN^{\varepsilon} \) which depends on the sieving limit \( \beta = \beta(x) \). We know (see [3]) that \( \beta \) is a function of \( x \) such that \( 1 < \beta(x) < 2 \) for \( 1/2 < x < 1 \), therefore one may expect a larger value of the parameter \( \Delta \) than in the traditional approach.

The proof of Theorem 1 will be based on some lemmas. Set

\[ u = \Delta^{2}. \]

The following result is known in the literature as the Fundamental Lemma (see [1]).

**Lemma 2 (see [4]).** There exist two sequences \( \{\varphi_{1}^{+}\}, \{\varphi_{1}^{-}\} \), such that

\[
\varphi_{1}^{+} = 1, \quad |\varphi_{1}^{+}| \leq 1, \quad \varphi_{1}^{-} = 0 \quad \text{if} \quad v \geq \Delta^{t},
\]

\[
\varphi_{1}^{-} \leq \frac{1}{2}, \quad |\varphi_{1}^{-}| \leq 1 \geq \varphi_{1}^{+} \leq 1,
\]

\[
\sum_{v \in \text{P}(u)} \varphi_{1}^{-}\eta(v) = V(u) \left \{ 1 + O \left ( e^{-1/2} \ln \frac{V(z)}{V(z)} \ln \frac{\Delta}{\ln z} - \frac{1}{2} \right ) \right \}.
\]

(11)

Next we quote some useful definitions. Let \( \eta = \varepsilon^{0}, \quad P(z) = P(z)/P(u), \quad \sigma_{+} = 1, \quad \sigma_{-} = 1/\beta \).

Let \( \xi = [u^{1+\eta k^{2}}; \varepsilon = 0, 1, 2, \ldots] \).

We define the following set of sequences:

\[
\{(D_{1}, \ldots, D_{r})_{r}; \ v \geq 1, \ D_{1} \in \mathcal{P}, \ i = 1, \ldots, r, \ D_{i} \leq D_{i-1} \leq \ldots \leq D_{1} < \Delta^{\varepsilon}\}.
\]

(12)

Adding to it the empty sequence \( (r = 0) \) we obtain a set of sequences which we denote by \( \mathcal{S}^{*} \). The number of sequences in \( \mathcal{S}^{*} \) is bounded by a constant depending only on \( \varepsilon \). Taking \( \varepsilon \leq 1/2 \) we can roughly estimate this number as follows:

\[
|\mathcal{S}^{*}| \leq \sum_{0 < \varepsilon < \varepsilon} (\sum_{0 < \varepsilon < \varepsilon} 1)^{\varepsilon} \leq \sum_{0 < \varepsilon < \varepsilon} \left ( \frac{\ln \varepsilon^{-2}}{\ln(1 + \eta)} \right )^{\varepsilon}
\]

\[
\leq (e^{-2} + 1)(e^{-2} / \ln(1 + \eta))^{-2} \leq (e^{-2} + 1)(2e^{-1})^{-2}
\]

\[
\leq (e^{-2} + 1)e^{-12e^{-2}} \leq 2e^{-12e^{-2}} \leq e^{-12e^{-2}} \leq e^{-12e^{-2}} \leq e^{-13e^{-2}} = \exp(-13e^{-2} \ln \varepsilon) \leq \exp(13e^{-3}).
\]

Next we define

\[
(D_{1}, \ldots, D_{r})^{+} = \max_{0 < \varepsilon < \varepsilon} D_{1}^{2}, \ D_{2}, \ldots, D_{1},
\]

\[
(D_{1}, \ldots, D_{r})^{-} = \max_{0 < \varepsilon < \varepsilon} D_{1}^{2}, D_{2}, \ldots, D_{1}.
\]

According to our convention we assign the value one to the empty sequence \( (r = 0) \). In the case when

\[
D_{1} = D_{2} = \ldots = D_{1} > D_{1+1} = \ldots > D_{1+1+\ldots+i_{k-1}+1} = \ldots = D_{1+i_{k-1}+i_{k}} = D_{r},
\]

we will write

\[
\Gamma(D_{1}, \ldots, D_{r}) = i_{1}! i_{2}! \ldots i_{k}!.
\]

Theorem 1 will be derived from the following lemma.

**Lemma 3.** Assume that the axioms (1)-(3) hold. Then the estimates (7) and (8) are valid for \( z \leq \Delta \) and \( z \leq \Delta^{\varepsilon/3} \) respectively, where the remainder term \( R^{\pm}(\mathcal{A}, M, N) \) is to be replaced by

\[
R^{\pm}(\mathcal{A}, M, N) = \sum_{(D_{1}, \ldots, D_{r}) \in \mathcal{S}^{*}} \sum_{v \in \mathcal{A}} a_{(D_{1}, \ldots, D_{r})}(v, \eta, \Delta)
\]

(12)

\[
\times \sum_{D_{1} \leq \eta^{1/2}} r(\mathcal{A}, v_{1}, \ldots, v_{r})
\]

\[
D_{1} \leq v_{1} \leq D_{2} \ldots, D_{1} \leq v_{r} \leq D_{2} \ldots
\]
where
\[ c_{(D_1, \ldots, D_r)}(v, \eta, A) = \begin{cases} 
\Gamma^{-1}(D_1, \ldots, D_r) \varphi^+ & \text{if } 2r \text{ and } (D_1, \ldots, D_r)^+ < A, \\
-\Gamma^{-1}(D_1, \ldots, D_r) \varphi^- & \text{if } 2r \text{ and } (D_1, \ldots, D_r)^- < A^{\frac{1}{1+\eta}}, \\
0 & \text{otherwise},
\end{cases} \]

\[ c_{(D_1, \ldots, D_r)}(v, \eta, A) = \begin{cases} 
\Gamma^{-1}(D_1, \ldots, D_r) \varphi^- & \text{if } 2r \text{ and } (D_1, \ldots, D_r)^- < A^{\frac{1}{1+\eta}}, \\
-\Gamma^{-1}(D_1, \ldots, D_r) \varphi^+ & \text{if } 2r \text{ and } (D_1, \ldots, D_r)^+ < A, \\
0 & \text{otherwise}.
\end{cases} \]

4. Proof of Theorem 1. In this section we will derive Theorem 1 from Lemma 3. The following assertion is of the main significance:

**Lemma 4.** Let \( M, N > 1 \), \( \beta < 2 \) and \( A = MN^{\beta-1} \). For every sequence \( (D_1, \ldots, D_r) \in \mathcal{S}^n \) such that
\[ (D_1, \ldots, D_r)^r < A \]
there exists a partition
\[ \{1, 2, \ldots, r\} = \{i_1, \ldots, i_s\} \cup \{j_1, \ldots, j_t\} \]
such that
\[ i_1 < i_2 < \ldots < i_s, \quad j_1 < j_2 < \ldots < j_t, \quad r = s + t, \]
\[ D_{i_1} D_{i_2} \ldots D_{i_s} \leq M, \quad D_{j_1} D_{j_2} \ldots D_{j_t} \leq N. \]

**Proof.** We apply induction with respect to \( r \). If \( r = 1 \), by the definition of \( \mathcal{S}^n \) we have
\[ D_1 < A^{\frac{1}{\beta}} \leq \max(M, N). \]

Now assume that the conclusion is verified for \( r-1 \) and consider the sequence \( (D_1, \ldots, D_r) \in \mathcal{S}^n \) such that
\[ (D_1, \ldots, D_r)^r < A = MN^{\beta-1}. \]

By the induction hypothesis,
\[ \{1, 2, \ldots, r-1\} = \{i_1, \ldots, i_s\} \cup \{j_1, \ldots, j_t\}, \]
where
\[ i_1 < i_2 < \ldots < i_s, \quad j_1 < j_2 < \ldots < j_t, \quad s + t = r - 1, \]
\[ D_{i_1} \ldots D_{i_s} \leq M, \quad D_{j_1} \ldots D_{j_t} \leq N. \]

We may assume that
\[ D_1 \ldots D_k D_r > M \quad \text{and} \quad D_{k+1} \ldots D_r D_r > N \]
since otherwise the conclusion is obvious. Therefore
\[ MN < D^2_r D_{r-1} \ldots D_1 < MN^{\beta-1} D_r^{2-\beta}. \]

Hence \( (\beta < 2) \) we have \( D_r > N \) and consequently
\[ M < D_1 \ldots D_k D_r \leq D_1 D_2 \ldots D_k \leq MN^{\beta-1} D_r^{2-\beta} < M. \]

This contradiction shows that Lemma 4 is valid.

Now we are in a position to derive Theorem 1 from Lemma 3. We will prove only the inequality (8) which requires more detailed considerations than the analogous inequality (7). Consider any \( M, N > 1 \) such that \( MN^{\beta-1} = A \). We have
\[ \max(M, N) = A^{\beta} \geq A^\epsilon (\epsilon \leq 1/3), \]

so the quantities
\[ M_1 = \left(\max(M, N) A^{-\epsilon}\right)^{1/(1+\eta)}, \quad N_1 = \left(\min(M, N)\right)^{1/(1+\eta)} \]
satisfy
\[ M_1 > 1, \quad N_1 > 1. \]

If \( M \geq N \) then
\[ M_1 N_1^{\beta-1} = (M A^{-\epsilon})^{1/(1+\eta)} N_1^{\beta-1}(1+\eta) = (MN^{\beta-1})^{1/(1+\eta)} A^{-\epsilon(1+\eta)} = A^{1-\epsilon(1+\eta)}. \]

If \( M < N \) then
\[ N_1 M_1^{\beta-1} = N_1^{1/(1+\eta)} (N A^{-\epsilon})^{(\beta-1)/(1+\eta)} = (MN^{\beta-1})^{1/(1+\eta)} A^{-\epsilon(\beta-1)/(1+\eta)} = A^{1-\epsilon(\beta-1)/(1+\eta)}. \]

Let
\[ z_0 = \begin{cases} 
\min(z, A^{\beta/(1+\eta)}) & \text{if } M \geq N, \\
\min(z, A^{(\beta - 1)/(1+\eta)}) & \text{if } N > M.
\end{cases} \]

and apply the Buchstab identity.

Since
\[ \beta - 1 \leq \frac{1 - \epsilon}{\beta(1+\eta)} \quad (\text{for } \epsilon \leq 1/3) \]
we have

\[ S(\mathcal{A}, \mathcal{D}, z) \geq S(\mathcal{A}, \mathcal{D}, z_0) - \sum_{p \in \mathcal{P}(t)} S(\mathcal{A}_p, \mathcal{D}, \Delta_{(\theta - 1)^{p/2}}). \]

Replacing \( \Delta \) by \( \Delta_{(\theta - 1)^{p/2}} \) or by \( \Delta_{(\theta - 1)^{(p/2) + 1}} \) and \( z \) by \( z_0 \) in Lemma 3 we have in view of the error term \( E(\epsilon, \Delta, K) \) the following inequality:

\[ S(\mathcal{A}, \mathcal{D}, z_0) \geq XV(z) \left( f \left( \frac{\ln \Delta}{\ln z} \right) - E(\epsilon, \Delta, K) \right) + R^- \]

where the remainder term \( R^- \) is equal to

\[ R^-(\mathcal{A}, \Delta^{(1 - \theta)(1 + \epsilon)}) = \sum_{(D_1, \ldots, D_h, p \in \mathcal{P}(t))} c(D_1, \ldots, D_h, p, \eta, \Delta^{(1 - \theta)(1 + \epsilon)}) \times \sum_{p_1 < \cdots < p_h, p \in \mathcal{P}(t)} r(\mathcal{A}, p_1 \ldots p_h) \text{ if } M \geq N \]

and similarly

\[ R^- = R^-(\mathcal{A}, \Delta^{(1 - \theta)(1 + \epsilon)(1 + \epsilon)}) \text{ if } N > M. \]

By Lemma 4 we find subsequences \( (D_{1, i}, \ldots, D_{h, i}) \) and \( (D_{1, j}, \ldots, D_{h, j}) \) such that

\[ D_{1, i} \cdot \ldots \cdot D_{h, i} \leq M_1, \quad D_{1, j} \cdot \ldots \cdot D_{h, j} \leq N_1. \]

Letting \( m = \max\{M_1, N_1\} \), we obtain

\[ m \leq \Delta^t M_1^{1 + \epsilon} = \max(M, N), \quad n \leq N_1^{1 + \epsilon} = \min(M, N), \]

which shows that in both cases \( (M \geq N, N > M) \) the remainder term \( R^- \) has the required form \( (10) \) (since \( a_{m, j}, b_{n, j} \) may depend on \( \Delta = MN^{1 - \epsilon} \)). By \( (14) \) it remains to estimate from above the sum

\[ \sum_{p \in \mathcal{P}(t)} S(\mathcal{A}_p, \mathcal{D}, \Delta^{(\theta - 1)^{p/2}}). \]

In view of \( (5) \) and Lemma 1 \( (s = 1) \) by standard calculations we see that the contribution of the main terms here is \( O(XV(z)E(\epsilon, \Delta, K)) \). The contribution of the remainder terms takes the form

\[ \sum_{\Delta^{(1 - \theta)(1 + \epsilon)} \in \mathcal{P}(t)} \sum_{p \in \mathcal{P}(t)} \lambda_+^\epsilon(\Delta^{(\theta - 1)^{p/2}}) r(\mathcal{A}, pd). \]

To complete the proof one has to show that the sum above has the desired form \( (10) \). Since \( \max(M, N) > \Delta^\epsilon \) we put \( m = p, \ n = d \) if \( \Delta^{(\theta - 1)^{p/2}} \)

\[ < \min(M, N) \text{ and } m = pd, n = 1 \text{ otherwise to obtain that } \]

\[ pd < \Delta^{(\theta - 1)^{p/2}}. \]

Since \( \min(M, N) < \Delta^{(\theta - 1)^{p/2}} \) in the latter case, we get

\[ \max(M, N) \geq \frac{\Delta^{(\theta - 1)^{p/2}}}{\min(M, N)^{p/2 - 1}} \geq \Delta^{(\theta - 1)^{p/2}}. \]

Hence \( pd < \max(M, N) \) and Theorem 1 follows.

5. Proof of Lemma 2. We first prove an auxiliary lemma.

**Lemma 5.** Let \( H(d) \) be any positive arithmetic function. Then

\[ \sum_{d \mid p \in \mathcal{P}(t)} \lambda_+^\epsilon(d) H(d) \leq \Sigma_1 + \Sigma_2, \]

\[ \sum_{d \mid p \in \mathcal{P}(t)} \lambda_+^\epsilon(d) H(d) \geq \Sigma_1 - \Sigma_2, \]

\[ \Sigma_1 - \Sigma_2 \leq \sum_{d \mid p \in \mathcal{P}(t)} H(d) \lambda_+^\epsilon(d) - \sum_{d \mid p \in \mathcal{P}(t)} H(d), \]

where we have set for simplicity

\[ \Sigma_1 = \sum_{(D_1, \ldots, D_h) \in \mathcal{P}(t)} \Gamma^{-1}(D_1, \ldots, D_h) \sum_{d \mid p \in \mathcal{P}(t)} H(p_1 \ldots p_d), \]

\[ \Sigma_2 = \sum_{(D_1, \ldots, D_h) \in \mathcal{P}(t)} \Gamma^{-1}(D_1, \ldots, D_h) \sum_{d \mid p \in \mathcal{P}(t)} H(p_1 \ldots p_d), \]

\[ \Sigma_1 = \sum_{(D_1, \ldots, D_h) \in \mathcal{P}(t)} \Gamma^{-1}(D_1, \ldots, D_h) \sum_{d \mid p \in \mathcal{P}(t)} H(p_1 \ldots p_d). \]

Here and in the sequel, this last sum is to be understood as being taken over those \( p_1, p_2, \ldots, p_l \) for which \( D_i \leq p_i < D_i^{1 + \epsilon} \), \( p_i \mid P(z, u) \), and \( p_i \neq p_j \) if \( i \neq j \) and \( i, j = 1, \ldots, r \).

The proof of the inequalities \( (15), (16) \) follows directly from the definitions of \( \Delta^\epsilon, \Delta^\epsilon(D_1, \ldots, D_h) \) and \( \Gamma(D_1, \ldots, D_h) \).
To show (17) we proceed as follows:

\[ \Sigma_1^+ - \Sigma_2^+ \leq \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} H(d) - \sum_{d \mid \nu(l_d, a_2) > \sqrt{d} + \eta} H(d) \]

\[ = \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} H(d) + \sum_{d \mid \nu(l_d, a_2) > \sqrt{d} + \eta} H(d) - \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} H(d) + \sum_{d \mid \nu(l_d, a_2) > \sqrt{d} + \eta} H(d) \]

\[ \leq \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} H(d) \lambda_d^+(a) + \sum_{d \mid \nu(l_d, a_2) > \sqrt{d} + \eta} H(d). \]

Similar arguments show that (18) is valid. This completes the proof of Lemma 5.

Now we are in a position to prove Lemma 3. We deal only with the lower bound (8) since the arguments for the upper bound are similar.

Denote by \( \mathcal{A} \) the subsequence of \( \mathcal{A} \) consisting of those elements \( a \in \mathcal{A} \) for which \( (a_l, P(a)) = 1 \). Applying inequality (16) for \( H(d) = S(\mathcal{A}_d, \mathcal{D}, u) \) we have by (4)

\[ S(\mathcal{A}, \mathcal{D}, u) \geq \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \lambda_d^-(a) S(\mathcal{A}_d, \mathcal{D}, u) \]

\[ = \sum_{(d_1, \ldots, d_r) \in \mathcal{A}_{\leq 2}^r} \Gamma^{-1}(D_1, \ldots, D_r) \sum_{\sum_{i=1}^r d_i < d} S(\mathcal{A}_{d_1 \ldots d_r}, \mathcal{D}, u) \]

\[ - \sum_{(d_1, \ldots, d_r) \in \mathcal{A}_{> 2}^r} \Gamma^{-1}(D_1, \ldots, D_r) \sum_{\sum_{i=1}^r d_i < d} S(\mathcal{A}_{d_1 \ldots d_r}, \mathcal{D}, u). \]

By Lemma 2 we obtain in view of (1) that

\[ \pi S(\mathcal{A}_d, \mathcal{D}, u) \leq \pi \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \varphi^+_d \mid \mathcal{A}_d \mid \]

\[ = \pi \left( \frac{\omega(d)}{d} \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \varphi_\mathcal{A}, \varphi_v \right) + \sum_{d \mid \nu(l_d, a_2) > \sqrt{d} + \eta} \varphi_v (\mathcal{A}, d, v). \]

If we insert this in (19) we find that

\[ S(\mathcal{A}, \mathcal{D}, z) \geq X \left( \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \varphi^+_d \varphi_v \right) \sum_{(d_1, \ldots, d_r) \in \mathcal{A}_{\leq 2}^r} \Gamma^{-1}(D_1, \ldots, D_r) \sum_{\sum_{i=1}^r d_i < d} \omega(p_1 \ldots p_r) \]

\[ - \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \varphi^+_d \varphi_v \sum_{(d_1, \ldots, d_r) \in \mathcal{A}_{\leq 2}^r} \Gamma^{-1}(D_1, \ldots, D_r) \sum_{\sum_{i=1}^r d_i < d} \omega(p_1 \ldots p_r) \]

\[ + R^- (\mathcal{A}, \mathcal{D}, z) \]

where \( R^- (\mathcal{A}, \mathcal{D}, z) \) has the required form (12) \( (n = -) \).

Applying the definition of \( \Gamma(D_1, \ldots, D_r) \), Lemma 2 and inequality (3), we see that the replacement in the above of \( \varphi^+_d \) by \( \varphi_v \) gives an error which does not exceed

\[ X \left\{ \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \frac{\omega(d)}{d} \left[ \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \frac{\omega(d)}{d} \right] \right\} \]

\[ \leq XV(u) e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3}) \prod_{p \mid P(u)} \left( 1 + \frac{\omega(p)}{p} \right) \]

\[ \leq XV(u) e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3}) \left( \frac{V(u)}{V(z)} \right)^2 \]

\[ \leq XV(z) e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3}) e^{-4} \left( 1 + \frac{K^2}{\ln u} \right)^2 \]

\[ \leq XV(z) e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3}) e^{-9} \left( 1 + \frac{K^2}{\ln d} \right) \]

\[ \leq XV(z) e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3}) e^{1/3} \leq XV(z) e^{\sqrt{\ln d} - 1/3 + 1} \]

\[ \leq XV(z) e^{\sqrt{\ln d} - 1/3 + 1} \text{ in view of (9)}. \]

Therefore by inequality (18) and Lemma 2 we obtain

\[ S(\mathcal{A}, \mathcal{D}, u) \geq XV(u) \left\{ 1 + O(e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3})) \right\} \]

\[ \times \left\{ \sum_{d \mid \nu(l_d, a_2) \leq \sqrt{d} + \eta} \frac{\omega(d)}{d} \lambda_d^-(a) \right\} \]

\[ + O(XV(z) E(\epsilon, A, K)) + R^- (\mathcal{A}, \mathcal{D}, z). \]

Now we are led to consider the expression

\[ XV(u) \left\{ 1 + O(e^{-1/3} (1 + e^{\sqrt{\ln d} - 1/3})) \right\} \]

\[ \times \left\{ \sum_{d \mid \nu(l_d, a_2) < d} \frac{\omega(d)}{d} \right\} \]

where \( \mathcal{A} = \{ p \in \mathcal{A}; p \geq u \} \), with the aim of showing that it is at least as large as

\[ XV(z) \left\{ f \left( \frac{\ln d}{\ln z} \right) + E(\epsilon, A, K) \right\} \]

\[ \cdot \left\{ \sum_{d \mid \nu(l_d, a_2) < d} \frac{\omega(d)}{d} \right\} \]
Since
\[ \prod_{w \leq p < z} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \leq \prod_{w \leq p < z} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \]
we may apply inequality (3) and Lemma 1 \((\not\in \rightarrow \not\in, V(z) \rightarrow V(z, u))\) to remark that
\[ M^-(A, \not\in, z) \geq \frac{V(z)}{V(u)} \left\{ f(s) + O(e^{-1/3}(\ln A)^{-1/3}) \right\}, \quad s = \frac{\ln A}{\ln z} . \]
Assuming that
\[ \sum_{d \mid P(z, u)} \frac{\omega(d)}{d} \leq e^3 + e^{-10} K^3 (\ln A)^{-1} \]
we find that the expression (20) is not less than
\[ \begin{align*}
&\leq X V(z) \left\{ f(s) + O(e^{-1/3}(\ln A)^{-1/3}) \right\} \\
&\times \left\{ f(s) + O(e^{-1/3}(\ln A)^{-1/3}) \right\} + O(e^3 + e^{-10} K^3 (\ln A)^{-1})(1 - E_1) \\
&= X V(z) \left\{ f(s) + O(e^{-1/3}(\ln A)^{-1/3}) \right\} \{ 1 + E_1 \} \\
&\quad + O \{ X V(u) (e^3 + e^{-10} K^3 (\ln A)^{-1})(1 + E_1) \}
\end{align*} \]
where
\[ E_1 \leq e^{-1/3}(1 + e^{-1/3}(\ln A)^{-1/3}) . \]
The first term on the right-hand side of (23) contributes the expected value
\[ X V(z) \left\{ f \left( \frac{\ln A}{\ln z} \right) + E(e, A, K) \right\} . \]
To handle the second one we make use of (3) to obtain
\[ V(u) = V(z) \frac{V(u)}{V(z)} = V(z) \prod_{p \mid P(z, u)} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \]
\[ \leq V(z) \frac{\ln z}{\ln u} \left( 1 + \frac{K}{\ln u} \right) \leq V(z) e^{-2} \left\{ 1 + \frac{K}{e^2 \ln A} \right\} . \]
Hence the \(O(\cdot)\) term is
\[ \leq X V(z) e^{-2} \left\{ 1 + \frac{K}{e^2 \ln A} \right\} \left\{ e^3 + e^{-10} K^3 (\ln A)^{-1} \right\} \{ 1 + E_1 \} . \]

Therefore it remains to prove (22). Every \(d \mid P(z, u)\) such that \(A^{1 + \epsilon} \leq d^\epsilon < A^{1 + \epsilon} \) can be decomposed as
\[ d = mpn \]
where
\[ m \mid P(z, u), \quad n \mid P(z, u), \quad p \in \not\in, \quad m_1 \leq p \leq m_2, \]
provided that \(m_1 = \max(u, A^{1 - \epsilon/m}), m_2 = \min(z, A^{1 + \epsilon/m}), \)
Hence
\[ \sum_{d \mid P(z, u)} \frac{\omega(d)}{d} \leq \sum_{m \mid P(z, u)} \frac{\omega(m)}{m} \sum_{m_1 \leq p \leq m_2} \frac{\omega(p)}{p} \sum_{n \mid P(z, u)} \frac{\omega(n)}{n} . \]
Since
\[ \frac{\ln m_2}{\ln m_1} \leq \frac{\ln (A^{1 - \epsilon/m})}{\ln (\max(u, A^{1 - \epsilon/m}))} = \frac{\ln (A^{1 - \epsilon/m}) + \ln A^{2\epsilon}}{\ln (\max(u, A^{1 - \epsilon/m}))} \leq 1 + \frac{2\ln A}{\epsilon^2 \ln A} = 1 + O(\epsilon^7) \]
we obtain
\[ \ln \left( \frac{\ln m_2}{\ln m_1} \right) = O(\epsilon^7) . \]
Hence in view of (3) we have
\[ \sum_{m \mid P(z, u)} \frac{\omega(p)}{p} \leq \sum_{m_1 \leq p \leq m_2} \ln \left( \prod_{p \mid \not\not\in} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \right) \]
\[ = \ln \left( \prod_{m_1 \leq p \leq m_2} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \right) \]
\[ \leq \frac{\ln m_2}{\ln m_1} + K \leq \epsilon^7 + \frac{K}{\ln u} . \]

To evaluate \(\sum \frac{\omega(m)}{m}\) we again use (3):
\[ \sum_{m \in \mathcal{P}(x)} \frac{\omega(m)}{m} = \prod_{p \in \mathcal{P}(x)} \left( 1 + \frac{\omega(p)}{p} \right) \leq \prod_{p \in \mathcal{P}(x)} \left( 1 - \frac{\omega(p)}{p} \right)^{-1} \leq \left( \frac{\ln x}{\ln u} \right)^{K} \left( 1 + \frac{K}{\ln u} \right) \leq \frac{\ln x}{\ln u} \left( 1 + \frac{K}{\ln u} \right) \leq e^{-2} \left( 1 + \frac{K}{\ln u} \right). \]

On combining the estimates above we obtain
\[ \sum_{d \mid n} \omega(d) \leq e^{-4} \left( 1 + \frac{K^2 e^{-4}}{\ln \Delta} \right) \left( e^{7} + \frac{K e^{-2}}{\ln \Delta} \right) \leq e^{-4} \left( e^{7} + e^{-6} \frac{K^3}{\ln \Delta} \right) \leq e^{3} + e^{-10} \frac{K^3}{\ln \Delta} \]
as required.

Now the proof of Theorem 1 is complete.

Acknowledgements. This paper is part of a doctoral thesis written under Professor Henryk Iwaniec. I wish to express to him my gratitude for his deep inspiration and help in my research work.

References


Elementary estimates for the Chebyshev function \( \psi(x) \) and for the Möbius function \( M(x) \)

by

N. Costa Pereira (Lisbon)

1. The general approach. We have shown in [4] that a technique first devised by Sylvester [11] to evaluate \( \lim \inf \psi(x)/x \) and \( \lim \sup \psi(x)/x \), could be transformed into an elementary method for estimating \( \psi(x) \). In this way we established several elementary bounds for \( \psi \) and for the related function \( \theta \), including Rosser's result [9]:

\[ \sup_{x > 0} \frac{\psi(x)}{x} = \frac{\psi(113)}{113} < 1.038821, \]

and also

\[ \sup_{x > 0} \frac{\theta(x)}{x} < \frac{1.01456}{69} < 68 \]

and

\[ \frac{\theta(x)}{x} > 0.985 > \frac{65}{66} \quad \text{for } x \geq 11927. \]

The present paper is devoted to a generalization and further refinement of these ideas, which allow us to obtain improved bounds for \( \psi \) and \( \theta \), as well as new estimates for the Möbius sum function.

Let \( f \) be a given function defined for all \( x > 0 \) and vanishing identically for \( 0 < x < 1 \). Assuming that the behaviour of \( f \) is sufficiently well known, we consider the problem of estimating its Möbius transform \( \varphi \) defined for \( x > 0 \) by

\[ \varphi(x) = \sum_{k \leq x} \mu(k) f \left( \frac{x}{k} \right). \]

Let \( (r_n)_{n \geq 1} \) be an increasing sequence of positive numbers which includes the positive integers. Extending \( \mu(t) \) to all \( t > 0 \) by letting \( \mu(t) = 0 \) if \( t \) is not