whence
\[ |P_f(n) - g(n)| = o(1). \]

This implies, for large \( n, P_f(n) = g(n) \). Now the result is achieved, as follows from the relation
\[ |D^k f(n)| = |D^k g(n) - P_f(n)| \sim n^{-m} \cdot o(n^{m}). \]

References


Primes in arithmetic progressions

by

E. POUVRE (Tulence) and H. IWANIEC (Warszawa)

1. Introduction. Statement of results. Let \( a \) and \( q \) be coprime integers, \( q > 1 \) and for any \( x > q \) let \( \pi(x; q, a) \) be the number of primes \( 
\leq x \) congruent to \( a \) mod \( q \). One of the basic and important problems in analytic theory of numbers is that of proving an asymptotic formula for \( \pi(x; q, a) \) that would hold, depending on \( x \), for moduli \( q \) as large as possible.

The classical prime number theorem of Siegel and Walfisz states that if \( A \) is a given positive number and \( q \leq (\log x)^A \) then
\[
\pi(x; q, a) = \frac{1}{\varphi(q)} \lfloor x / q \rfloor + O\left(\frac{x}{q^{1/2}}\right),
\]
where \( c \) and the constant implied in the symbol \( O \) depend on \( A \) alone (not effectively computable if \( A > 2 \)). A mention should be made of the two conjectures
\[
\begin{align*}
\pi(x; q, a) & = \frac{1}{\varphi(q)} \lfloor x / q \rfloor + O\left(\frac{x^{1/2}}{\log x}\right), & \text{Great Riemann Hypothesis (GRH),} \\
\pi(x; q, a) & = \frac{1}{\varphi(q)} \lfloor x / q \rfloor + O\left(\frac{x^{1/2} \log x}{q}\right), & \text{H. L. Montgomery's Hypothesis},
\end{align*}
\]
the first one giving an asymptotic formula for \( q < x^{1/2} \) and the latter for \( q < x^{1/3} \) (cf. [18]), neither of these relations is expected to be proved in a near future.

With the development of Brun's and Selberg's sieve methods it became motivated and popular to investigate statistical results which would hold for "almost all" \( q \)'s in wider ranges. After pioneering works of Yu. V. Linnik [13] and A. Renyi [18] and some others [17], [11] in 1965 E. Bombieri [2] and A. I. Vinogradov [31] proved a mean-value theorem which states, in a form given by Bombieri, that for any \( A > 0 \) the following holds
\[
\sum_{\substack{\phi(x) \leq A \leq 1}} \max_{\phi(x)} \max_{n > x} \left| \pi(x; q, a) - \frac{1}{\varphi(q)} \lfloor x / q \rfloor \right| \leq A (\log x)^{-A}
\]
with \( Q = \phi(\log q)^{-s} \) and \( B = B(A) \), the constant implied in the notation \( \lesssim \) depending on \( A \) alone. This result serves for many practical purposes as well as the GRH.

So far the limit \( Q = \phi(\log q)^{-s} \) for modulus \( q \) has not been essentially extended apart of some improvements on explicit relations between the constants \( A \) and \( B \), for instance \( B = A + 7/2 \) [19] and \( B = A + 2 \) [23]. P. X. Gallagher [7], [8] introduced major simplifications in the original Bombieri's arguments and R. C. Vaughan [20] gave an "elementary" and still simpler proof. For other variations of proofs and generalizations of results see for example [16], [12], [10].

There are essentially three principal tools that are used in all proofs of Bombieri–Vinogradov prime number theorem, namely

(i) Siegel–Walfisz theorem for small moduli,

(ii) Representation of sums over primes as bilinear forms, in case of \( \pi(x; q, a) \) the following ones

\[
B(M, N; q, a) = \sum_{M < n < 2M} a_n \sum_{N < n < 2N} \overline{b_n}, \quad mn \equiv a \pmod{q},
\]

(iii) Large sieve inequality for Dirichlet's characters

\[
\sum_{\eta \leq \chi \leq \eta'} \sum_{n \leq N} \sum_{n \leq N} a_n \chi(n) \leq (Q^2 + N) \sum_{n \leq N} |a_n|^2.
\]

It is the application of the large sieve inequality that sets the limit \( Q = \phi(\log q)^{-s} \) for modulus \( q \) in mean-value theorems for arithmetic progressions.

Recently the authors [6] succeeded to prove a mean-value theorem for

\[
\pi(x, x; q, a) = \{n \leq x; n \equiv a \pmod{q}, q \mid n - p \geq \eta}\}
\]

with \( \eta = \eta^\theta \) and \( Q = \phi(\log q)^{-s} \) where \( \eta \) is a small positive constant and \( \theta = \theta(\eta) \) \( > 1/2 \). The new key arguments are: the Linnik dispersion method and the Weil estimates for Kloosterman's sums. This method is effective for estimating forms \( B(M, N; q, a) \) with special values of \( M, N \) such for example that occur in combinatorial sieve identity for \( \pi(x, x; q, a) \) with \( s = s^\theta, \eta \) small enough but it does not work if \( \eta = 1/2 \), the case of \( \pi(x, q, a) \). In three further works [4], [5], [24] the first author has developed many other ideas getting lots of intermediate results approaching closer to prime numbers.

In this paper we enhance Fouvry's arguments with new estimates for sums of incomplete Kloosterman sums given recently by J.-M. Deshouillers and the second author [3]. Their estimates are sharper than those obtained from Weil's result or even from Hooley's R* conjecture. With all these instruments we may finally treat \( \pi(x; q, a) \) itself. Let us state the result.

An arithmetic function \( \lambda: \mathbb{N} \to \mathbb{C} \) is called to be of level \( D \) and of finite order \( \kappa \) if

\[
\lambda(d) = 0 \quad \text{if} \quad d \geq D,
\]

\[
|\lambda(d)| \leq \tau_\kappa(d) \quad \text{if} \quad d < D,
\]

where \( \tau_\kappa(d) \) is the divisor function. Next \( \lambda \) is called well-factorable if for any \( D_1, D_2, D_3, D_4, D_5, D_6 \) there exist two functions \( \mu, \nu \) of levels \( D_1 \) and \( D_2 \) and orders \( \kappa_1, \kappa_2 \) respectively such that

\[
\lambda = \mu \ast \nu.
\]

Of course \( \lambda \) is of order \( \kappa = \kappa_1 + \kappa_2 \).

**Theorem.** Let \( \alpha \neq 0, \beta > 0, \theta > 0, \varepsilon > 2 \). For any well-factorable function \( \lambda \) of level \( Q = \phi(\log q)^{-s} \) and of finite order \( \kappa \leq \kappa_1 + \kappa_2 \) we have

\[
\sum_{\alpha \leq \chi \leq \alpha'} \sum_{n \leq N} \lambda(n) \left( \pi(x; q, a) - \frac{1}{\varphi(q)} \left( \frac{x}{q} + 1 \right) \right) \ll x(\log x)^{-A}
\]

the constant implied in \( \lesssim \) depending at most on \( \varepsilon, a, \alpha, \beta \) and \( \kappa \).

This improvement of Bombieri–Vinogradov result is very small indeed and the purpose of our work is to show that a progress beyond GRH is possible by means now available. It was disappointing not to get (1.1) for arbitrary \( \lambda(q) \) relevant, e.g. to sum up the errors \( \pi(x; q, a) - \frac{1}{\varphi(q)} \left( \frac{x}{q} + 1 \right) \) with absolute values. Here the introduction of well-factorable weights builds up greater flexibility in the arguments from the four previous works. This not only makes possible to choose optimally the parameters \( M, N \) in the involved bilinear forms but also it allows us to rearrange the dispersions in a new way.

Perhaps a few words should be said to motivate our considerations of well-factorable functions. First of all they are important for the linear sieve theory. In the traditional notations and assumptions one has

\[
S(\mathcal{A}, P, x) \leq XXV(x) \left\{ F(x) + \epsilon \right\} + \sum_{i=2}^{n} B_i^+(\mathcal{A}, D),
\]

\[
S(\mathcal{A}, P, x) \geq XXV(x) \left\{ F(x) - \epsilon \right\} - \sum_{i=2}^{n} B_i^-(\mathcal{A}, D),
\]

where \( s = \log D/\log x \) and each remainder term \( B_i(\mathcal{A}, D) \) has the form

\[
\sum_{d \mid P_i} \lambda_i(d) r(\mathcal{A}, d)
\]

with a well-factorable \( \lambda_i(d) \) of level \( D \) and order 2 (for precise statement see [11]).
Let us give a typical example. Let \( D_2 \geq \ldots \geq D_f \) be such that
\[
 D_1, \ldots, D_f, D_f < D, \quad i = 1, \ldots, f
\]
and let \( \lambda_i \) be functions of level \( D_i \) and of finite orders. Then \( \lambda = \lambda_1 * \ldots * \lambda_f \) is well-factorable of level \( D \) and of finite order.

The following is obvious: if \( q \) is of level \( R \) and \( \lambda \) is well-factorable of level \( D \geq R \) then \( \varphi \cdot \lambda \) is well-factorable of level \( DR \).

It should be noted that our method does not permit the residue class \( a \) to vary freely with \( q \). Moreover, even for fixed \( a \), cannot be larger than a small power of \( \varphi \). The main reason is the lack of good estimates for Fourier coefficients \( \psi(a) \) of Maass cusp forms that are implicitly employed here by an appeal to [3]. The Petersson conjecture that \( |\psi(a)| \leq |a|^s \) would perhaps permit to take \( |a| \) as large as \( \varphi \).

As an immediate application we give the following

**Corollary.** Let \( \pi_2(a) \) denote the number of pairs of primes \( p, q + 2 \) such that \( p \leq a \). Then we have
\[
 \pi_2(a) \leq \left( \frac{34}{9} + \varepsilon \right) B x (\log x)^{-s}
\]
where \( B = \prod_{p \geq 2} \left( 1 - \frac{1}{(p-1)^s} \right) \), for any \( s > 0 \) and \( x \gg \alpha(x) \).

A similar upper bound with constant \( 4 \) in place of \( 34/9 \) can be found in [25] and with \( 3.9171 \) in [22] while the heuristically expected asymptotic formula states
\[
 \pi_2(a) \sim B \pi(a) \log a \quad \text{as} \quad a \to \infty.
\]

2. **Bilinear forms for sums over primes.** We shall begin by expressing \( \pi(a; q, a) \) or rather
\[
 \psi(a; q, a) = \sum_{\text{gcd}(a,q)=1} A(n)
\]
as bilinear forms \( B(M, N; q, a) \) of two kinds to be estimated in the next sections.

In his book [14] of 1963 Linnik gave, among other vital ideas, an obvious relation
\[
 \sum_{n=1}^{\infty} \frac{A(n)}{\log n} n^{-s} = \log \zeta(s) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \left( \frac{\log(s)}{j} \right) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \left( \sum_{n=1}^{\infty} \frac{1}{n} \right)
\]
from which it follows that
\[
 \frac{A(n)}{\log n} = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \sum_{\text{gcd}(a,q)=1} A(n)
\]
Hence it is transparent that studies of sums over primes can be reduced to sums over several integral variables (or over one variable weighted by the divisor functions \( \tau_j(n) \)). Notice that the series over \( j \) is finite because the terms with \( j \) such that \( 2^j > n \) trivially vanish. In practice large \( j \)’s cause some technical difficulties which Linnik avoids by applying (2.1) for \( n \)’s that are free of small prime factors. In his "Thèse de Doctorat d’État" [5] the first author made several iterations of the celebrated identity of Vaughan [20] obtaining a variant of (2.1) which does not involve the described inconvenience. Yet another formula (the easiest of all for practical use) was recently given by R. Heath-Brown [9], namely

**Lemma 1 (Heath-Brown).** Let \( \lambda \geq 2 \) and \( J \) be such that \( 2^J > n \). Then
\[
 A(n) = \sum_{j=1}^{J} (-1)^{j} \sum_{\text{gcd}(a,q)=1} A(n)
\]
\[
 \frac{\xi^{J}}{\xi} (\theta) \left( \frac{\log(s)}{\theta} \right) = \sum_{n=1}^{\infty} \mu(n) n^{-s} \quad \text{Proof: Letting}
\]
\[
 \sum_{n=1}^{\infty} \mu(n) n^{-s} = \frac{\xi^{J}}{\xi} (\theta) \left( \frac{\log(s)}{\theta} \right)
\]
we have on one side
\[
 \sum_{j=1}^{J} (-1)^{j} \sum_{\text{gcd}(a,q)=1} A(n)
\]
and we have on the other side
\[
 \sum_{n=1}^{\infty} \mu(n) n^{-s} = \sum_{n=1}^{\infty} \mu(n) n^{-s} = \sum_{n=1}^{\infty} \mu(n) n^{-s}
\]
which completes the proof.

For the purpose of this paper we shall arrange Heath-Brown’s identity as follows:

Let \( 0 < \epsilon \leq 1 \), \( M \) and \( N \) take values \((1 + \epsilon)^{-1}, l \) and \( l \) an integer, \( M_0 \leq \epsilon \) and let \( \Lambda_i = \{(1 + \epsilon)^{-1} M_i, N_i \} \) and \( \Lambda_0 = \{(1 + \epsilon)^{-1} M_i, M_i \} \). Then we have
\[
 A(n) = \sum_{j=1}^{J} (-1)^{j} \sum_{\text{gcd}(a,q)=1} A(n)
\]
\[
 \times \sum_{\text{gcd}(a,q)=1} \mu(m_i) \ldots \mu(m_j) \sum_{n=1}^{\infty} 1 + O \left( \theta \right) \frac{\log(s)}{\theta} \log n \]
First, by the prime number theorem and by partial summation (1.1) can be reduced to

\[ \sigma(x, Q) = \sum_{\lambda \geq 0} \lambda(q) \left( \sum_{n \leq x, a \equiv 0 \mod q} A(n) - \frac{1}{\varphi(q)} \sum_{n \leq x, a \equiv 0 \mod q} A(n) \right) \ll (\log x)^{-4}. \]

For each \( A(n) \) we introduce (2.3) with \( x \to x^{|J|}, J = 4 \). Accordingly, \( \sigma(x, Q) \) splits up into \( (A^{-1} \log x)^4 \) sums of the type

\[ \sigma(M_1, \ldots, M_J; N_1, \ldots, N_J) = \sum_{(a_0) = 1} \lambda(q) \left( \sum_{m_0 \in \mathcal{M}_1, \ldots, m_J \in \mathcal{M}_J} \mu(m_1) \ldots \mu(m_J) \right) \frac{1}{\varphi(q)} \sum_{n \leq x, \sum_{i=1}^J m_i \equiv a_0 \mod q} \mu(m_1) \ldots \mu(m_J) \]

where the variables \( m_1, \ldots, m_J, n_1, \ldots, n_J \) are additionally constrained by \( m_1 \ldots m_J n_1 \ldots n_J < x \).

Moreover, there is an error term which contributes to \( \sigma(x, Q) \) at most

\[ A \sum_{J=1}^{J-1} J \sum_{a_0 \equiv 1} \tau_{(J)}(a_0) \left( \sum_{n \leq x, a \equiv 0 \mod q} \tau_{J}(n) - \frac{1}{\varphi(q)} \sum_{n \leq x, a \equiv 0 \mod q} \tau_{J}(n) \right) \ll A x (\log x)^{-4}. \]

Analogously, the sums \( \sigma(M_1, \ldots, M_J; N_1, \ldots, N_J) \) with \( M_1 \ldots M_J \) contribute to \( \sigma(x, Q) \) at most

\[ \sum_{J=1}^{J-1} J \sum_{a_0 \equiv 1} \tau_{(J)}(a_0) \left( \sum_{n \leq x, a \equiv 0 \mod q} \tau_{J}(n) + \frac{1}{\varphi(q)} \sum_{n \leq x, a \equiv 0 \mod q} \tau_{J}(n) \right) \log x \ll A x (\log x)^{-4}. \]

Either (2.6) and (2.7) is admissible if \( A = (\log x)^{-4 - \epsilon} \), which we henceforth assume.

From the above discussion we conclude that the proof of theorem reduces to showing that for \( M_1 \ldots M_J N_1 \ldots N_J \ll x \) it holds

\[ \sigma(M_1, \ldots, M_J; N_1, \ldots, N_J) \ll x (\log x)^{-4 - \epsilon}. \]

Notice that (2.5) is redundant, therefore we suppress it to gain independent of the variables \( m_1, \ldots, m_J, n_1, \ldots, n_J \). Now, \( \sigma(M_1, \ldots, M_J; N_1, \ldots, N_J) \) can be written as bilinear form once one puts the parameters \( M_1, \ldots, M_J, N_1, \ldots, N_J \) into two disjoint sets. Our choice will depend on bounds available in the next two sections.

3. General bilinear forms. Let, for notational simplicity, \( m \sim M \) and \( n \sim N \) mean that \( M < m \leq 2M \) and \( N < n \leq 2N \) respectively. Given \( M, N, Q, R \geq 1 \) we consider

\[ \sigma(M, N; Q, R) = \sum_{a \equiv 0} \alpha_a \sum_{r \equiv 0} \beta_r \left( \sum_{m \leq M} \gamma_m \sum_{n \leq N} \delta_n - \frac{1}{\varphi(q)} \sum_{m \leq M} \gamma_m \sum_{n \leq N} \delta_n \right), \]

with coefficients \( \alpha_a, \beta_r, \gamma_m, \delta_n \) such that

(\( A_1 \)) For some constant \( \kappa \) we have

\[ \sum_{n \leq N} \gamma_n \ll \sigma(n), \quad \sum_{n \leq N} \delta_n \ll \sigma(n), \quad \sum_{n \leq N} \beta_r \ll \sigma(n). \]

(\( A_2 \)) For any \( a \neq 0 \), \( b \geq 1 \), \( q \geq 1 \), \( (ab, q) = 1 \) and \( C > 0 \) we have

\[ \sum_{n \leq N, \delta_n \leq \kappa} \delta_n = \frac{1}{\varphi(q)} \sum_{n \leq N, \delta_n \leq \kappa} \delta_n + O(N (\log x)^{2C}). \]

the constant implied in the symbol \( O \) depending on \( x \) and \( C \) only.

In this section we shall apply Linnik's dispersion method to prove the following

**Proposition 1.** Let \( MN \ll x \) and for some \( \epsilon > 0 \)

(3.1) \( x^\epsilon < N \ll x^{1/3} \),

(3.2) \( Q \ll x^{-\epsilon} N \),

(3.3) \( R \ll \min(x^{1/2-\epsilon} N^{-1/4}, x^{1/2} N^{-1+10\epsilon}) \).

Then, for any \( A > 0 \) we have

(3.4) \( \sigma(M, N; Q, R) \ll x (\log x)^{-4} \)

the constant implied in \( \ll \) depending at most on \( \epsilon, \kappa, \alpha \) and \( A \).

Before proceeding to essential transformations of \( \sigma(M, N; Q, R) \) let us make a few preliminary restrictions which do not limit the generality but are convenient for simplification of the arguments we shall use later. The result is trivial if \( MN \ll x^{1+\epsilon} \) thus assuming that \( x^{1-\epsilon} \ll MN \ll x \), it is then obvious that the worst case is

(3.5) \( MN = x \).

Next, dividing the range for modulus \( qr \) into subintervals of the type \( [Q_1, 2Q_1], [R_1, 2R_1] \) with \( Q_1 \ll Q \) and \( R_1 \ll R \) respectively one sees that there are \( (\log x)^2 \) pairs of such intervals and the worst case is \( Q_1 = Q \),
By Poisson's summation formula we deduce the following:

\[
\sum_{(m,n) = 1} \gamma(m) = \sum_{\nu | \rho} \mu(\nu) \sum_{m | \rho} \gamma(m \nu) = \sum_{\nu | \rho} \mu(\nu) \sum_{h \equiv 0 \mod \nu} \hat{\gamma}(h) / h
\]

where \( \hat{\gamma} \) stands for the Fourier transform of \( \gamma \). The zero term yields

\[
\frac{\varphi(k)}{k} \hat{\gamma}(0)
\]

whereas the terms with \( h \neq 0 \), by partial integration, yield

\[
\leq \sum_{\nu | \rho} \frac{1}{\nu} \sum_{h = 1}^{\infty} \min(M, \nu h^{-2} M^{-1}) \leq \tau(k).
\]

Hence, letting

\[
X = \hat{\gamma}(0) \sum_{(m,n) = 1} \frac{a(q)}{q} \sum_{(r_1, r_2) = 1} \frac{\beta_{r_1} \beta_{r_2} \varphi(r_1 r_2 q)}{r_1 r_2 \varphi((r_1, r_2) q)} \left( \sum_{(a, r) = 1} \delta_a \right) \left( \sum_{(a, r) = 1} \delta_a \right)
\]

we conclude that

\[
U = X + O(\omega^4 Q^{-1} N^2)
\]

because \( a(q) \varphi(q r_1 r_2) / q r_2 \varphi(q r_1) \varphi(q r_2) = a(q) / r_2 \varphi(q r_1) \varphi(q r_2) \). Here the error term \( O(\omega^4 Q^{-1} N^2) \) is admissible for (3.4) to hold.

II. Evaluation of \( V \). By definition we have

\[
V = \sum_{(a, n) = 1} \frac{\beta_{r_1 \beta_{r_2}}}{\varphi(q r_1) \varphi(q r_2)} \left( \sum_{(a, r) = 1} \delta_a \right) \left( \sum_{(a, r) = 1} \delta_a \right) \sum_{m = r_1 r_2} \gamma(m)
\]

where \( \rho_i \) stands for a solution of \( m_1 m_2 = 1 \). By the Poisson formula

\[
\sum_{m = r_1 r_2} \gamma(m) = \sum_{m = r_1 r_2} \mu(\nu) \sum_{m = m_1 m_2} \gamma(m \nu) = \sum_{m = r_1 r_2} \mu(\nu) \sum_{k / \nu \in \mathbb{Z}} e \left( - \frac{\nu k^2}{k} \right) \hat{\gamma}(h) / h
\]

where

\[
\frac{1}{k} \prod_{p | r_1, r_2} \left( 1 - \frac{1}{p} \right) \hat{\gamma}(0) + O(\tau(r)).
\]

Hence, we conclude that

\[
V = X + O(\omega^4 EN^2)
\]
because
\[
\frac{a(q)}{\varphi(gq)} \prod_{p \mid gq} \left(1 - \frac{1}{p}\right) = \frac{a(q) (r_1, r_2)}{q (r_1, r_2)} \frac{1}{\varphi((r_1, r_2), q)},
\]

Here the error term \(O(\sigma^2 N^2)\) is admissible for (3.4) to hold.

III. Main term for \(W\). By definition we have
\[
W = \sum_{(a, q) = 1} \alpha(q) \sum_{(r_1, r_2) \equiv (a, q)} \beta_{r_1} \beta_{r_2} \sum_{(n_1, q_1) = 1, \gcd(n_1, q_1) = 1} \delta_{n_1} \delta_{q_1} \sum_{(n_2, q_2) = 1, \gcd(n_2, q_2) = 1} \delta_{n_2} \delta_{q_2} 
\]
The innermost sum over \(m\) will be approximated by \(\hat{\gamma}(0) / q(r_1, r_2)\), an error \(R(q, r_1, r_2, n_1, n_2)\) to be considered later. Here we evaluate contribution of the main term, i.e. the following quantity
\[
W = \hat{\gamma}(0) \sum_{(a, q) = 1} \alpha(q) \sum_{(r_1, r_2) \equiv (a, q)} \beta_{r_1} \beta_{r_2} \sum_{(n_1, q_1) = 1, \gcd(n_1, q_1) = 1} \delta_{n_1} \delta_{q_1} \sum_{(n_2, q_2) = 1, \gcd(n_2, q_2) = 1} \delta_{n_2} \delta_{q_2} 
\]
We detect the congruence \(n_1 = n_2 \pmod{(r_1, r_2)q}\) by the well-known orthogonality of multiplicative characters \(\chi_{\pmod{(r_1, r_2)q}}\). The principal character \(\chi = \chi_0\) yields \(X\) while the non-principal characters contr\(\hat{\gamma}(0) \sum_{(r_1, r_2) \equiv (a, q)} \beta_{r_1} \beta_{r_2} \sum_{(n_1, q_1) = 1, \gcd(n_1, q_1) = 1} \delta_{n_1} \delta_{q_1} \sum_{(n_2, q_2) = 1, \gcd(n_2, q_2) = 1} \delta_{n_2} \delta_{q_2} 
\]
where
\[
T_0(r_1, r_2) := \sum_{(a, q) = 1} \alpha(q) \frac{1}{\varphi(gq)} \sum_{\substack{\chi_\pmod{q} \chi \neq \chi_0}} \left( \sum_{(n, q) = 1} \chi(n) \delta_n \right) \left( \sum_{(n, q) = 1} \chi(n) \delta_n \right)
\]
and for simplicity we denoted \(q = (r_1, r_2)\). By the Cauchy–Schwarz equality we have
\[
|T_0(r_1, r_2)|^2 \leq T_0(r_1, r_2) T_0(r_1, r_2).
\]
Each \(\chi \neq \chi_0\) is induced by exactly one primitive character \(\psi(m)\), \(e > 1, e \mid \varphi(q)\), thus letting \(ee' = e\), we have
\[
T_0(r, r) \leq Q^{-1} \sum_{1 \leq q \leq Q} \frac{1}{\varphi(f)} \sum_{1 \leq \psi \leq \varphi(f)} \sum_{\psi \sim \varphi(e)} \left| \sum_{(n, q) = 1} \psi(n) \delta_n \right|^2.
\]
By (A.4) one easily deduces that
\[
\left| \sum_{(n, q) = 1} \psi(n) \delta_n \right| \leq eN^2 \varphi^{-1}.
\]
This will be used for \(e \lesssim N^2 = E\), say, whereas if \(e > E\) we appeal to the large sieve inequality giving
\[
\sum_{E < \varphi(e) \leq E} \left| \sum_{(n, q) = 1} \psi(n) \delta_n \right|^2 \lesssim \left( \frac{Q}{f} + \frac{N}{E} \right) N^2 \varphi^2.
\]
From the three last inequalities we first infer that
\[
T_0(r, r) \leq \psi(r) \varphi^{-1} N^2 \varphi^2 + \varphi N^2 \varphi^2.
\]
and then that the non-principal characters contribute to \(W_0\) in total
\[
\leq \sum_{(r_1, r_2) \equiv (a, q)} \frac{T_r(r_1, r_2)}{|r_1, r_2|} \leq \sum_{(r_1, r_2) \equiv (a, q)} \psi(r_1, r_2)^2 T_0(r_1, r_2) \leq \psi(r_1, r_2)^2 \leq \psi(r_1, r_2)^2
\]
for \(C = 2A + 2x^2 + 5\). Concluding the above discussion let us write down what has been proved.

IV. Error term for \(W\). Now we proceed to estimate the most difficult quantity
\[
W_1 = \sum_{(a, q) = 1} \alpha(q) \sum_{(r_1, r_2) \equiv (a, q)} \beta_{r_1} \beta_{r_2} \sum_{(n_1, q_1) = 1, \gcd(n_1, q_1) = 1} \delta_{n_1} \delta_{q_1} R(q, r_1, r_2, n_1, n_2).
\]
Let \(W_1(r, \omega, q)\) be the partial sum of \(W_1\) consisting of the terms with
\[
(n_1, n_2) = (r, \omega), \quad q = \omega[r], \quad (r_1, r_2) = q.
\]
Notice that \((r, \omega) = 1\) and \((r, r) = 1\) therefore
\[
W_1 = \sum_{r} \sum_{(a, q) = 1} \sum_{(r_1, r_2) \equiv (a, q)} W_1(r, \omega, q),
\]
where
\[
W_1(r, \omega, q) = \sum_{(a, q) = 1} \alpha(q) \sum_{(r_1, r_2) \equiv (a, q)} \beta_{r_1} \beta_{r_2} \sum_{(n_1, q_1) = 1, \gcd(n_1, q_1) = 1} \delta_{n_1} \delta_{q_1} R(q, r_1, r_2, n_1, n_2).
\]
and by definition
\[
R(q, r_1, r_2, n_1, n_2) = \sum_{(m, q) = 1} \psi(m) \hat{\gamma}(0) / q'(r_1, r_2, q).
\]
Applying standard but tedious elementary arguments based mainly on the inequality \( \tau(n) \ll n^{1/2} \) it can be shown that
\[
(3.17) \quad \sum_{\gamma^2 \leq q} \sum_{\chi} \sum_{\chi(n)} |W_1(v, \omega, \chi)| \ll Q^{-1} N^{-1/2} + \omega^3, \quad v \equiv a \mod p,
\]
so it is an admissible quantity for (3.4) to hold.

For \( \nu_2 \ll \nu_1 \), we treat \( W_2(v, \omega, \chi) \) by far deeper means. The two congruences \( \nu_2 \equiv a \mod q \) and \( \nu_2 \equiv b \mod q \) from (3.16) are equivalent to the one \( n = \{(q\nu_1,\nu_2)\} \) where \( i \) is a solution of the system
\[
(3.18) \quad n_{\nu_1} = a \mod q, \quad n_{\nu_2} = b \mod q.
\]
By the Poisson sum formula we obtain
\[
\hat{\gamma}(y) = \sum_{n \equiv 0 \mod q} e \left( \frac{-ny}{q} \right) = \frac{1}{q} \sum_{n \equiv 0 \mod q} e \left( \frac{-ny}{q} \right)
\]
where \( \hat{\gamma}(y) \) is the Fourier transform of \( \gamma(m) \). For \( |h| \geq \nu_1 Q^2 M^{-1} = H \), say, by partial integration the number of times depending on \( \alpha \) one gets
\[
\hat{\gamma}(\frac{h}{\nu_1 Q}) = \nu_1^{-1} H^{-1}.
\]
This yields
\[
W_1(v, \omega, \chi) = \frac{1}{\nu_1 Q^2} \sum_{\chi} \sum_{\nu_1 \nu_2} \beta_{\nu_1 \nu_2} \sum_{\nu_3} e \left( \frac{-h\nu_3}{\nu_1 Q} \right) \sum_{\nu_4} e \left( \frac{-h\nu_4}{\nu_1 Q} \right)
\]
with the error \( O(1) \) that is much sharper than we need for (3.4) to hold.

Now our aim is to reinterpret the congruences (3.18) in order to arrive at incomplete Kloosterman sums of a certain type. Since \( n_{\nu_2} = n_{\nu_1}(\nu_1) \) there is an integer \( t \) such that
\[
n_{\nu_2} - n_{\nu_1} = \nu_1 t, \quad 1 \leq |t| \leq N/\nu_1 Q.
\]
Next (3.18) are equivalent to
\[
(3.19) \quad n_{\nu_2} - n_{\nu_1} = \nu_1 \nu_1 u_1, \quad n_{\nu_2} = a \nu_1 u_2
\]
where \( u_1, u_2, \nu_1 \) are integers. Hence \( r_{\nu_1} u_1 - r_{\nu_2} u_2 = \nu_1 t, \) so
\[
(3.20) \quad u_1 = -\nu_1 n_{\nu_1}(\nu_1) u_2.
\]
Moreover, since
\[
(3.21) \quad n_{\nu_2} - n_{\nu_1} = \nu_1 \nu_1 t.
\]
it follows that
\[
(3.22) \quad u_1 = -\nu_1 \nu_1 t, \quad u_1 = -\nu_1 \nu_1 t m(t + 1) \mod p.
\]
As a matter of fact (3.20)–(3.22) are equivalent to solvability of (3.19). By (3.20) and (3.22) we construct one congruence
\[
(3.23) \quad u_1 = -\nu_1 \nu_1 t, \quad u_1 = -\nu_1 \nu_1 t m(t + 1) \mod p,
\]
whence, by (3.19)
\[
(3.24) \quad \nu_1 \nu_1 t \equiv a \nu_1 \nu_1 t m(t + 1) \mod p.
\]
and further, since the last "non-arithmetic" term is \( \ll \nu_2^{-1} \), this enables us to write
\[
\hat{\gamma}(\frac{h}{\nu_1 Q}) = e \left( \frac{\nu_1 \nu_1 t m(t + 1)}{\nu_1 Q} \right) + O(1).
\]
The error term \( O(\nu_2^{-1}) \) contributes to \( W_1(v, \omega, \chi) \) trivially \( \ll N^2 \nu_1^2 \nu_2^{-1} \ll N^{1/2} \nu_1 \). Therefore
\[
(3.25) \quad W_1(v, \omega, \chi) = \sum_{\nu_1 \nu_2} e \left( \frac{\nu_1 \nu_1 t m(t + 1)}{\nu_1 Q} \right) \sum_{\nu_1 \nu_2} e \left( \frac{\nu_1 \nu_1 t m(t + 1)}{\nu_1 Q} \right) + O(N^{12} \nu_1 \nu_2)
\]
with the extra condition
\[
(3.26) \quad (n_{\nu_2} - n_{\nu_1}, \nu_1 \nu_1 t) = \delta t,
\]
and for given \( \sigma | a, \sigma, v = 1 \) the two resulting congruences between \( n_{\nu_1} \) and \( n_{\nu_2} \), namely \( n_{\nu_2} - n_{\nu_1} = \nu_1 \nu_1 t m(t + 1) \) and \( n_{\nu_2} - n_{\nu_1} = 0 \) \( \nu_1 \nu_1 t = 0 \) can be written, after reinterpreting \( \omega \mod p \) by \( \omega \mod \nu_1 \nu_1 t = 0 \) as one congruence
\[
(3.27) \quad n_{\nu_2} - n_{\nu_1} = \nu_1 \nu_1 t \mod \nu_1 \nu_1 t \nu_1 \nu_1 t.
\]
The reinterpretation of $\omega(\mod p^r)$ by $\omega(p^r \mod p)$ is allowed because if $\omega$ runs over the residue classes $(\mod p^r)$, prime to $p^r$, so does $\omega$.

Now our nearest aim is to make the variables of the summation $n_1$ and $n_2$ to be independent. For this purpose there are two constraints which have to be relaxed, namely the congruence (3.25) and the dependence on $q = (n_2 - n_1)/d$ of the integral

$$I(q) = \int a(q) \gamma(q, \xi) e\left(\frac{h \xi}{q \xi} + \eta q \xi\right) d\xi.$$

The first one is handled by means of additive characters $(\mod \sigma q^t)$, precisely we make use of

$$\frac{1}{\sigma q^t} \sum_{\lambda \equiv \eta (\mod q^r)} e\left(\frac{\lambda n_2 - n_1 - \eta \lambda}{q^t}\right) = \begin{cases} 1 & \text{if (3.25) holds,} \\ 0 & \text{otherwise} \end{cases}$$

and the second one by means of the Fourier integral

$$a(q) \gamma(q, \xi) = \int K(\xi, \eta) e(\eta q \xi) d\eta$$

where by the inversion formula

$$K(\xi, \eta) = \int a(q) \gamma(q, \xi) e(-\eta q \xi) dq.$$

We have $K(\xi, \eta) = 0$ unless $M/6Q \leq \xi \leq 6M/Q$, moreover $K(\xi, \eta) \ll Q$ and by partial integration two times with respect to $q$ we even get $K(\xi, \eta) \ll \eta^{-2}Q^{-1}$. All together yield

$$\int \left| K(\xi, \eta) \right| d\xi d\eta \ll MQ^{-1}$$

and

$$I(q) = \int \int K(\xi, \eta) e\left(\frac{h \xi}{q \xi} + \eta (n_2 - n_1) \right) d\xi d\eta.$$

Finally collecting (3.23), (3.26) and (3.28) we deduce that

$$\left| W_4(n, \omega, \varphi) \right| \ll \omega^2 \xi^2 R^{-1} \int \int \left| K(\xi, \eta) \right| \times

$$

\begin{align*}
& \sum_{\lambda \equiv \eta (\mod q^r)} \sum_{\eta \equiv \lambda \mod q^r} \frac{1}{\sigma q^t} \times \\
& \left| \sum_{\lambda \equiv \eta (\mod q^r)} \sum_{\eta \equiv \lambda (\mod q^r)} \delta_{n_2} e\left(\frac{\lambda + \nu \eta}{q^t} \right) e\left(ah \frac{\nu \eta}{q^t} + \frac{n_2 - n_1}{q^t}\right) \right| d\xi d\eta + N^{10/3}. 
\end{align*}

This expression is rather complicated from notational point of view. To deal with it we formulate in independent notions the following

**Lemma 2.** Denote for $C, D, H, N, T \geq 1$, $\alpha \neq 0$ and $q > 0$

$$\mathcal{S}(C, D, H, N, T)$$

$$= \sum_{(c, d) = 1} \sum_{1 < c^r < D} \sum_{1 < d^r < D} \sum_{1 < t^r < q^t} \frac{1}{\sigma q^t} \sum_{n_1 = n_2} a(h, n, t) e\left(\frac{\lambda n}{q^t} \right) e\left(\frac{ah \frac{n}{q^t} + \frac{n_2 - n_1}{q^t}}{c}\right)$$

where $a(h, n, t)$ are any numbers such that $|a(h, n, t)| \leq 1$. Then if $HT \leq N$ and $D \leq C$ we have

$$\mathcal{S}(C, D, H, N, T) \ll (CN)^{12} + (CDHN)^{14} T^{14} \left(1 + \frac{C^2}{HTN^2}\right)$$

the constant implied in $\ll$ depending at most on $a$ and $s$.

**Proof.** We shall prove a general estimate

$$\mathcal{S}(C, D, H, N, T) \ll \mathcal{S}_2(C, D, H, N, T)$$

$$\ll C^{12} \cdot D^{12} \cdot H^{12} \cdot N^{12} \cdot T^{14} \cdot C^{14} \cdot (N + HT)^{14} \times$$

$$\times \left(\frac{C^2}{HTN^2}\right) \times \times \left(\frac{C^2}{HTN^2}\right)$$

(3.31) being an easy corollary. Since the right-hand side of (3.32) is increasing in $C, D, H, N, T$ we can assume without loss of generality that $c \sim C, d \sim D, h \sim H, t \sim T$ and $n \sim N$. Then, by the Cauchy–Schwarz inequality

$$\mathcal{S}(C, D, H, N, T)$$

$$\ll \mathcal{S}_2(C, D, H, N, T)$$

$$= \sum_{h \equiv 1 (\mod H)} \sum_{t \equiv 1 (\mod T)} \sum_{n_1 = n_2} a(h, n, t) e\left(\frac{\lambda n}{q^t} \right) e\left(\frac{ah \frac{n}{q^t} + \frac{n_2 - n_1}{q^t}}{c}\right)$$

$$= \sum_{h \equiv 1 (\mod H)} \sum_{t \equiv 1 (\mod T)} \sum_{n_1 = n_2} a(h, n, t) e\left(\frac{\lambda n}{q^t} \right) e\left(\frac{ah \frac{n}{q^t} + \frac{n_2 - n_1}{q^t}}{c}\right)$$

say, where

$$b_{d^r} = \sum_{n_1 = n_2, n_1 = n_2} a(h, n, t) e\left(\frac{\lambda n}{q^t} \right) e\left(\frac{ah \frac{n}{q^t} + \frac{n_2 - n_1}{q^t}}{c}\right).$$
and $f(e, d)$ is any function which majorizes the characteristic function of the set $[0, 2C] \times [D, 2D]$. In what follows we shall require $f(e, d)$ to be the one which satisfies the assumptions of Theorem 12 of [3], i.e. of $C^4$ class with $\text{Supp} f = \frac{1}{6} C, 3C \times \frac{1}{6} D, 3D$ and

$$ \left| \frac{\partial f}{\partial e} - \frac{\partial f}{\partial d} \right| \leq C^{-1} D^{-2} $$

for any $r_1, r_2 \geq 0$, the constant implied in $\ll$ depending at most on $r_1, r_2$.

Notice that $N^2 < r < 4N^2$ and $|l| \leq 4|a| HNT = L_l$ say. The terms with $l = 0$, i.e. the diagonal $n_2 = n_2$ contribute to $\mathcal{S}(C, D, H, N, T)$ less than $4CDHTN^{1/2}$ which yields the first term on the right-hand side of (3.32). For estimating the contribution of the non-diagonal terms, i.e. those with $l \neq 0$ we first split up the interval $(0, 2L]$ into subintervals of the type $(L_j, 2L_j], L_j = 2^{-j/2}$ and then for each of $O(\log 3L)$ resulting partial sums we apply Theorem 13 of [3] giving

$$ \sum_{l \leq j} (CDHT)^{1/2} [C(N^2 + L_j)(C + DN^2 + CDNL_j^{1/2}) + D^2 NL_j^{3/2}] B^{13}_j $$

where

$$ B_j = \sum_l \sum_{r \leq L_j} |b_{nl}| \ll (HNT)^{1/2} T \sum_r \sum_{r \leq L_j} |b_{rl}|$$

$$ \ll (HNT)^{1/2} T^{1/2} \sqrt{n_2 \sigma_t} \sim L_j \ll (HNT)^{1/2} TNL_j. $$

This together with (3.33) show that the worst case is $L_j = L$ giving the remaining terms on the right-hand side of (3.32). The proof of Lemma 2 is complete.

Remark. Theorem 13 of [3] depends on the location of exceptional eigenvalues of the Laplacian for the Hecke groups $G_\nu(r)$. If the Selberg eigenvalue conjecture is true then the factor

$$ \left( 1 + \frac{C^2}{HTN^3} \right)^{1/6} $$

in (3.31) can be suppressed.

Now we wish to estimate $W_2(\nu, \omega, g)$ by an appeal to Lemma 2. To this end we have to interpret the variables $e, d, h, t, n$ from (3.30) appropriately to the situation in (3.29). Let us interpret for given values of the variables $\xi, \eta, \sigma$, the variables listed above in the following manner

$$ o \ll a/\sigma $$
$$ c \ll a/\sigma $$
$$ d \ll r_1 \sigma $$
$$ h \ll h $$
$$ t \ll t $$
$$ n \ll n $$

$$ a(h, n, t) \sim \sigma_n \sim \sigma $$

Then by (3.30) it follows that

$$ W_2(\nu, \omega, g) \ll a \int \cdots int \mathcal{S}(\xi, \eta)|R^{-2} \mathcal{S}(NR, R, H, N, \eta) \sim N^{1/2} a^2 \ll \mathcal{S}^4MQ^{-1}T^{-2} \sqrt{R^2H^2N^{1/2} + R^{1/2}H^{1/2}N^{1/2}Q^{-1} N^{1/2}} \left( 1 + \frac{a}{N^2} \right)^{1/6} \ll \mathcal{S}^4MQ^{-1}N^{1/2}a^2 + RQ^{-1/2}N^{1/2}a^{1/2}, $$

The same majorization holds for $W_4$ with possibly different but arbitrarily small $\epsilon$. This bound is admissible for (3.4) to hold.

If we collect all evaluations of $U, V, W$ and introduce them into the dispersion (3.10) we find that the main terms $X$ disappear throughout and we are left with error terms only which by (3.1)-(3.3), as we said in appropriate places, are admissible for (3.4) to hold. The proof of Proposition 1 is complete.

4. Special bilinear forms. We now consider the forms $\mathcal{S}(M, N; Q, R)$ having special coefficients $\delta_\nu$; namely we assume that

$$(A_4) \quad \delta_\nu \text{ is the characteristic function of an interval contained in } (N, 2N).$$

For such $\mathcal{S}$ (A4) is obvious. In this section we use another method to prove the following

**Proposition 2.** Let $MN \leq a$ and for some $\epsilon > 0$

$$ a^{1/4} \ll N \ll a, $$

$$ Q \ll a^{1/2}N, $$

$$ R \ll (a/N)^{1/2}. $$

Then for some $\eta = \eta(a) > 0$ we have

$$ |\mathcal{S}(M, N; Q, R)| \ll a^{-\eta}, $$

the constant implied in $\ll$ depending at most on $a$ and $\epsilon$. 

Proof. Like in the previous section we may assume without loss of
genreality that (3.5) and (3.6) hold. Moreover, for making certain Fourier
series convergent rapidly it is most convenient right now to replace the
step function $\delta_n$ by a smooth one $\beta(n)$, say, whose graph is

such that

$$\int |\delta_n - \beta(n)| \, dn \leq 2\Delta N.$$  

(4.5)

If we do this we make an error which in total is

$$\ll \sum_q \sum_r r_\sigma(q^r) r_\varphi(r) \sum_{n \equiv m(\sigma)} r_\tau(m) |\delta_n - \beta(n)| +$$

$$\ll \sigma \Delta MN \sum_n |\delta_n - \beta(n)| \ll \sigma \Delta MN = \sigma \Delta^{-1} = \sigma^{-1}$$

for $\delta = \sigma^{-2}$ which we henceforth assume.

In addition to (4.5) all we require of $\beta(n)$ is that its derivatives should satisfy

$$|\beta^{(\nu)}(n)| \ll \Delta^{-\nu}$$

for $\nu = 0, 1, \ldots$, the constant implied in $\ll$ depending on $\nu$ alone. Then, by the Poisson formula we obtain

$$\sum_{n = m(\sigma)} \delta(n) = \frac{1}{\sigma} \delta(0) + \frac{1}{\sigma} \sum_{h\neq 0} \left( -ah \frac{\overline{m}}{\sigma} \right) \delta \left( \frac{h}{\sigma} \right)$$

and, like evaluating $U$ in Section 3, we deduce that

$$\sum_{\nu = m(\sigma)} \delta(n) = \sum_{\nu = m(\sigma)} \nu^\nu \sum_n \delta(n) = \sum_{\nu = m(\sigma)} \nu^\nu \sum_h \delta \left( \frac{h}{\nu} \right)$$

$$\ll \frac{q(\nu)}{\nu} \delta(0) + O \left( \sum_{\nu = m(\sigma)} \nu^\nu \sum_{h=1}^{N} \min \left( N, \Delta^{-2} N^{-2}, \nu^2 h^{-2} \right) \right)$$

$$\ll \frac{q(\nu)}{\nu} \delta(0) + O \left( (\nu q(\nu) \Delta^{-1}) \right).$$

(4.7)

The error term $O\left( (\nu q(\nu) \Delta^{-1}) \right)$ contributes to $\sigma(M, N; Q, R) \ll \sigma \sigma^{-1} \Delta \ll \sigma^{10}$ which is admissible.

Further, for $|h| \geq \sigma^2 Q R |AN| = \sigma^2 Q R |N| = H$, say by partial integration the number of times depending on $\epsilon$ we get

$$\left| \frac{\delta \left( \frac{h}{\nu} \right)}{\nu} \right| \ll \epsilon^{-10} h^{-2}$$

which enables us to break up the series (4.6) at $|h| = H$ with an admissible total error $O(1)$.

Notice that the main terms from (4.6) and (4.7) cancel and we are left with $\sigma(M, N; Q, R, H)$, say, which stands for the contribution to $\sigma(M, N; Q, R)$ of the partial sum of (4.6) with $0 < |h| \leq H$.

$$\ll \sigma^2 \int_{\nu = \sigma Q}^{2\nu N \sigma Q} \sum_{h = M}^{H} \sum_{(h, m) = 1} r_\tau (\nu |\beta, r \tau) \epsilon \left( \frac{h}{\nu} \right) e \left( -ah \frac{\overline{m}}{\sigma} \right) dh.$$  

(4.8)

To simplify further arguments from notational point of view let us prove the following

**Lemma 3.** Let $C, D, H, S \gg 1$, $a \neq 0$, and denote

$$\mathcal{S}(C, D, H, S) = \sum_{h=0}^{CD} \sum_{a=0}^{CD} a^a \sum_{h=0}^{CD} a^h \sum_{\Delta = 1}^{CD} \Delta \sum_{b=0}^{CD} a^b$$

where $a(h, s)$ are any numbers with $|a(h, s)| \leq 1$. For any $s > 0$ we have

$$\mathcal{S}(C, D, H, S) \ll (CDHS)^{1/2} + C^{1/2} DH^{1/2} (H + S)^{1/2} + CS [DH (H + S)]^{1/2} \times$$

$$\times \left( 1 + \frac{D}{C} + \frac{D}{\sqrt{HS}} \right)^{1/2}.$$  

(4.9)

**Proof.** Since the right-hand side of (4.9) is increasing in $C, D, H, S$ we can assume without loss of generality that $a \sim C, \Delta \sim D, h \sim H$ and $s \sim S$. Then, by the Cauchy-Schwars inequality

$$|\mathcal{S}(C, D, H, S)| \ll C D \sum_{a(h, s)} a(h, s) a(h, s) \times$$

$$\times \sum_{(h, s) = 1}^{CD} f(c, d) e \left( \frac{a(h, s) - h_1 s_1}{\sigma h_2 s_2} \right)$$

$$= CD \sum_{a(h, s)} h_1 s_1 \sum_{(h, s) = 1}^{CD} f(c, d) e \left( \frac{a h_1 s_1}{\sigma h_2 s_2} \right)$$

$$\ll (CDHS)^{1/2} + C^{1/2} DH^{1/2} (H + S)^{1/2} + CS [DH (H + S)]^{1/2} \times$$

$$\times \left( 1 + \frac{D}{C} + \frac{D}{\sqrt{HS}} \right)^{1/2}.$$  

(4.9)
where \( f(c, d) \) is like that from the proof of Lemma 2 and
\[
\mathbb{b}_d = \sum_{\omega \leq \frac{D}{x}} \sum_{\omega \leq \frac{D}{x}} a(h_1, s_2) a(h_2, s_3).
\]
The terms with \( l = 0 \), i.e., such that \( h_1, s_2 = h_2, s_1 \) contribute to \( \mathcal{S}(C, D, H, S) \) at most \( O(H(S)^3 (CD)\mathcal{O}(HS)^{1/2}) \) which yields the first term of the right-hand side of (4.9). For estimating the contribution of the non-diagonal terms, i.e., those with \( l \neq 0 \), split up the interval \( (0, 2L) \), \( L = 2 \omega \mathcal{O}(HS) \), into subintervals of the type \( (L_1, 2L_1) \), \( L_1 = 2^{-3} L_2 \) and then estimate each of the \( O(\log 3L) \) resulting partial sums by applying Theorem 12 of [3] giving
\[
\sum_{x < L_2} \sum_{l = 0}^{L_2} \sum_{l = 0}^{L_2} \sum_{\omega \leq \frac{D}{x}} \sum_{l = 0}^{L_2} f(c, d) e\left(\frac{l}{x}\right) \ll (CDLS)^2 (OS^2 (S^2 + L_2) (D + CDL_2^{1/2}) + D^2 (S - 1/2)^2)^{1/2} B_2^{1/2}
\]
where
\[
B_2 = \sum_{x < L_2} \sum_{l = 0}^{L_2} b_l \ll \sum_{x < L_2} \sum_{(s_2, s_4) \in (a,b)} (H S)^{-1} (s_3, s_4) + 1)
\]
\[
\ll (HS)^2 (H + S) L_2.
\]
This shows that the worst case is \( L_2 = L \) giving the remaining terms of the right-hand side of (4.9). This completes the proof of Lemma 3.

By (4.8) and (4.9) we deduce that
\[
\mathcal{S}(M, N_1, Q, R, H) \ll \frac{a^{n_1} N_1 (Q R)^{-1}}{a^{n_1}} N Q R^{-1} \mathcal{S}(Q, M, H, R)
\]
\[
\ll a^{n_1} \frac{N}{Q R} \left[ a^{n_1} Q M (H R)^{1/2} + a^{n_1} Q M H^{1/2} (H + R)^{1/4} + Q R (M H (H + R))^{1/2} \times \left(1 + \frac{M}{Q} + \frac{M}{H R}\right)^{1/4}\right]
\]
\[
\ll a^{n_1} \left[ a^{n_1} N^{1/2} + N_1^{1/2} Q^{1/2} R^{1/2} + N_1^{1/2} M^{1/2} Q R^{1/2} N^{1/2} (Q R)^{1/2}\right] \ll a^{n_1 - 2}\]

provided (4.1)-(4.3) hold. This completes the proof of Proposition 2.

5. **Proof of theorem, conclusion.** It remains to prove (2.5). Our strategy is to arrange each \( \mathcal{S}(M_1, \ldots, M_j, N_1, \ldots, N_j) \) as a sum of the type \( \mathcal{S}(M, N; Q, R) \) which we estimated in Propositions 1 and 2. The choice of \( Q \) and \( R \) may depend on \( M \) and \( N \) because the weight function \( \lambda(q) \) is well factorable. Since the ranges (3.1) and (4.1) for \( N \) do overlap it is evident that we can deal with every sum \( \mathcal{S}(M, N; Q, R) \) that may occur. However, in order to get the maximal value for \( Q R \) we should arrange these sums as to get the optimal well location of \( N \). To this end we prove the following lemma of combinatorial nature.

**Lemma 4.** Let \( 1 \leq M_i \leq a^{n_1}, 1 \leq N_i \leq a^{n_1}, i = 1, \ldots, j \) and \( M_1 \ldots M_j N_1 \ldots N_j = y \leq a^{5^n} \). Let \( j/24 \leq 0 < 1/3 \). Then either for some \( N_i \) we have
\[
N_i \geq y^k
\]
or some partial product of \( M_1 \ldots M_j N_1 \ldots N_j \) call it \( N \), lies in the interval
\[
y^{1-\delta} \leq N \leq y^{1+\delta}.
\]

**Proof.** Suppose each \( N_i \) is \( < y^k \). Excluding all \( N_i \)'s that are \( \geq a^{n_1} \) (at most 2) the remaining \( N_i \)'s and all \( M_i \)'s yield a product, let us say \( M_1 \ldots M_j N_1 \ldots N_j \), which is \( > y^{k-\delta} \). The smallest partial product of \( M_j \ldots M_i N_i \ldots N_j \) which is \( > y^{k-\delta} \), call it \( N \), must satisfy \( y^{1-\delta} < N \leq \max(y^{1+\delta}, y^{1+\delta}) \ll y^{1+\delta} \). This completes the proof.

Proceeding to the proof of (2.5) we first observe that the result is trivial if \( y = M_1 \ldots M_j N_1 \ldots N_j < a^{n_1} \), thus we assume that \( a^{n_1} < y < a^{1} \). We shall apply Lemma 4 with \( \theta = 3/17 \). Each \( \mathcal{S}(M_1, \ldots, M_j, N_1, \ldots, N_j) \) can be written as \( \mathcal{S}(M, N; D_1, D_1, D_2) \) with arbitrary \( D_1, D_2 \geq 1 \) subject to \( D_1D_2 = D \) — the level of the weight function and with \( N \) that either
\[
N \geq a^{1-\delta }, \quad \delta s \text{ satisfies (A1) and (A2)}
\]
or
\[
a^{1-\delta} \ll N \ll a^{1+\delta}, \quad \delta s \text{ satisfies (A3) and (A4)}.
\]

In the last case \( \delta \) satisfies (A3) by the Siegel–Walfisz theorem for the Möbius \( \mu(m) \) function. According to whether (5.3) or (5.4) hold we apply Proposition 2 or 1 respectively with \( Q \) and \( R \) which equalize (4.2), (4.3) and (3.2), (3.3) giving \( QR = a^{1/2} N^{1/2} \geq a^{1/2 + 2\theta} = a^{1/2 - \delta} \) in the case of (5.3) and \( QR = a^{1/2} N^{1/2} \geq a^{1/2 - \delta} = a^{1/2 - \delta} \) in the case of (5.4). This makes possible to factorize \( D = a^{1/2 - \delta} \) as \( D_1D_2 \) with \( D_1 \leq Q, D_2 \leq R \). The proof of theorem is complete.

References

Les volumes IV et suivants sont à obtenir chez

Johnson Reprint Corporation, 111 Fifth Ave., New York, N. Y.

Les volumes I-III sont à obtenir chez

Die Bände I-III sind zu

Томы I-III можно получить через

BOOKS PUBLISHED BY THE POLISH ACADEMY OF SCIENCES
INSTITUTE OF MATHEMATICS

H. Steinhaus, Selected papers, in printing.

MONOGRAFIE MATEMATYCZNE

55. C. Bessaga and A. Pełczyński, Selected topics in infinite-dimensional topology,

BANACH CENTER PUBLICATIONS