whence

$$|P_1(n)-g(n)| = o(1).$$

This implies, for large n, $P_1(n) = g(n)$. Now the result is achieved, as follows from the relation

$$\|D^k f(n)\|_k = |D^k f(n) - g^k(n)| = |D^k f(n) - P_1^k(n)| \sim c\cdot n^{\|f\|_k}.$$

References

[1] E. Artin, *Algebraic numbers and algebraic functions*, Gordon & Breach Science Publishers, New York 1967.
[2] H. Davenport, D. J. Lewis, A. Schinzel, *Polynomials of certain special types*, Acta Arith. 9 (1964), pp. 107–116.
[3] K. R. Matthews, *Polynomials which are near to k-th powers*, Proc. Camb. Phil. Soc. 61 (1965), pp. 1–5.
[4] A. Perelli, U. Zannier, *Una proprietà aritmetica dei polinomi*, Boll. U.M.I. (5) 17-A (1980), pp. 199–202.
[5] G. Pólya, G. Szegö, *Problems and theorems in Analysis*, vol. II, Springer Verlag, Berlin 1976.
[6] P. Ribenboim, *Polynomials whose values are powers*, J. Reine Angew. Math. 268/269 (1974), pp. 34–40.
[7] J. G. Van der Corput, *Diophantische Ungleichungen — I. Zur Gleichverteilung Modulo Eins*, Acta Math. 56 (1931), pp. 373–447.

ISTITUTO DI MATEMATICA "L. TONELLI"
Via F. Buonarroti, 2
56100 Pisa, Italy
SCUOLA NORMALE SUPERIORE
Piazza dei Cavalieri, 7
56100 Pisa, Italy

Received on 16. 7. 1981
and in revised form on 4. 11. 1981 (1261)

ACTA ARITHMETICA
XLII (1983)

Primes in arithmetic progressions

by

E. Fouvry (Talence) and H. Iwaniec (Warszawa)

1. Introduction. Statement of results. Let a and q be coprime integers, $q \geqslant 1$ and for any $x \geqslant 2$ let $\pi(x; q, a)$ be the number of primes $\leqslant x$ congruent to $a \bmod q$. 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 $q \leqslant (\log x)^A$ then

$$\pi(x; q, a) = \frac{1}{\varphi(q)}\operatorname{li} x + O\big(x\exp(-c\sqrt{\log x})\big)$$

where c and the constant implied in the symbol O depend on A alone (not effectively computable if $A \geqslant 2$). A mention should be made of the two conjectures

$$\pi(x; q, a) = \frac{1}{\varphi(q)}\operatorname{li} x + O(x^{1/2+\varepsilon}), \quad \text{Great Riemann Hypothesis (GRH)},$$

$$\pi(x; q, a) = \frac{1}{\varphi(q)}\operatorname{li} x + O(q^{-1/2}x^{1/2+\varepsilon}), \quad \text{H. L. Montgomery's Hypothesis},$$

the first one giving an asymptotic formula for $q < x^{1/2-2\varepsilon}$ and the latter for $q \leqslant x^{1-3\varepsilon}$ (cf. [15]), 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], [1] in 1965 E. Bombieri [2] and A. I. Vinogradov [21] proved a mean-value theorem which states, in a form given by Bombieri, that for any $A > 0$ the following holds

$$\sum_{q\leqslant Q}\max_{(a,q)=1}\max_{y\leqslant x}\left|\pi(y; q, a) - \frac{1}{\varphi(q)}\operatorname{li} y\right| \ll x(\log x)^{-A}$$

with $Q = x^{1/2}(\log x)^{-B}$ and $B = B(A)$, the constant implied in the notation $\ll$ depending on A alone. This result serves for many practical purposes as well as the GRH.

So far the limit $Q = x^{1/2}(\log x)^{-B}$ 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

$$B(M, N; q, a) = \sum_{M<m\leqslant 2M} a_m \sum_{N<n\leqslant 2N} b_n, \quad mn \equiv a(\operatorname{mod} q),$$

(iii) Large sieve inequality for Dirichlet's characters

$$\sum_{q\leqslant Q} \sideset{}{^*}\sum_{\chi(\operatorname{mod} q)} \Big|\sum_{n\leqslant N} a_n\chi(n)\Big|^2 \ll (Q^2+N)\sum_{n\leqslant N}|a_n|^2.$$

It is the application of the large sieve inequality that sets the limit $Q = x^{1/2}$ for modulus q in mean-value theorems for arithmetic progressions.

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

$$\pi(x, z; q, a) = |\{n \leqslant x;\ n \equiv a(\operatorname{mod} q),\ p\,|\,n \Rightarrow p \geqslant z\}|$$

with $z = x^{\eta}$ and $Q = x^{\theta}$ where η 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 $\pi(x, z; q, a)$ with $z = x^{\eta}$, η—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 $\lambda\colon N \to C$ is called to be of *level* D and of finite *order* $\varkappa$ if

$$\lambda(d) = 0 \qquad \text{if } d \geqslant D,$$
$$|\lambda(d)| \leqslant \tau_{\varkappa}(d) \qquad \text{if } d < D,$$

where $\tau_{\varkappa}(d)$ is the divisor function. Next λ is called *well-factorable* if for any $D_1, D_2 \geqslant 1$, $D_1D_2 = D$ there exist two functions μ, ν of levels D_1 and D_2 and orders $\varkappa_1$, $\varkappa_2$ respectively such that

$$\lambda = \mu * \nu.$$

Of course λ is of order $\varkappa \leqslant \varkappa_1 + \varkappa_2$.

THEOREM. *Let $a \neq 0$, $\varepsilon > 0$, $A > 0$, $x \geqslant 2$. For any well-factorable function λ of level $Q = x^{9/17-\varepsilon}$ and of finite order $\varkappa$ we have*

$$\sum_{(q,a)=1} \lambda(q)\Big(\pi(x; q, a) - \frac{1}{\varphi(q)}\operatorname{li} x\Big) \ll x(\log x)^{-A} \tag{1.1}$$

the constant implied in $\ll$ depending at most on ε, a, A and $\varkappa$.

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 $\pi(x; q, a) - \frac{1}{\varphi(q)}\operatorname{li} x$ 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

$$S(\mathscr{A}, P, z) \leqslant XV(z)\{F(s)+\varepsilon\} + \sum_{j\leqslant J(\varepsilon)} R_j^+(\mathscr{A}, D), \tag{1.2}$$

$$S(\mathscr{A}, P, z) \geqslant XV(z)\{f(s)-\varepsilon\} - \sum_{j\leqslant J(\varepsilon)} R_j^-(\mathscr{A}, D), \tag{1.3}$$

where $s = \log D/\log z$ and each remainder term $R_j(\mathscr{A}, D)$ has the form

$$\sum_{d\,|\,P(z)} \lambda_j(d)\, r(\mathscr{A}, d)$$

with a well-factorable $\lambda_j(d)$ of level D and order 2 (for precise statement see [11]).

Let us give a typical example. Let $D_1 \geqslant \ldots \geqslant D_r$ be such that

$$D_1 \ldots D_i D_i < D, \quad i = 1, \ldots, r$$

and let λ_i be functions of level D_i and of finite orders. Then $\lambda = \lambda_1 * \ldots * \lambda_r$ is well-factorable of level D and of finite order.

The following is obvious; if ϱ is of level R and λ is well-factorable of level $D \geqslant R$ then $\varrho * \lambda$ 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 fixed a cannot be larger than a small power of x. The main reason is the lack of good estimates for Fourier coefficients $\varrho_j(a)$ of Maass cusp forms that are implicitly employed here by an appeal to [3]. The Petersson conjecture that $|\varrho_j(a)| \ll |a|^\varepsilon$ would perhaps permit to take $|a|$ as large as x.

As an immediate application we give the following

COROLLARY. *Let $\pi_2(x)$ denote the number of pairs of primes p, $p+2$ such that $p \leqslant x$. We then have*

$$\pi_2(x) \leqslant (34/9 + \varepsilon) Bx (\log x)^{-2}$$

where $B = 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right)$, *for any $\varepsilon > 0$ and $x \geqslant x_0(\varepsilon)$.*

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

$$\pi_2(x) \sim Bx(\log x)^{-2} \quad \text{as } x \to \infty.$$

2. Bilinear forms for sums over primes. We shall begin by expressing $\pi(x; q, a)$ or rather

$$\psi(x; q, a) = \sum_{\substack{n \leqslant x \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n)$$

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

$$\sum_{n=2}^{\infty} \frac{\Lambda(n)}{\log n} n^{-s} = \log\zeta(s) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} (\zeta(s) - 1)^j = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \Big(\sum_{n \geqslant 2} n^{-s}\Big)^j$$

from which it follows that

$$(2.1) \qquad \frac{\Lambda(n)}{\log n} = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \sum_{\substack{n_1 \ldots n_j = n \\ n_1, \ldots, n_j \geqslant 2}} 1.$$

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'Etat" [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 $z \geqslant 2$ and J be such that $2z^J > n$. Then*

$$(2.2) \qquad \Lambda(n) = \sum_{j=1}^{J} (-1)^j \binom{J}{j} \sum_{m_1, \ldots, m_j < z} \mu(m_1) \ldots \mu(m_j) \sum_{m_1 \ldots m_j n_1 \ldots n_j = n} \log n_1.$$

Proof. Letting

$$M(s) = \sum_{n<z} \mu(n) n^{-s} \quad \text{and} \quad F(s) = \frac{\zeta'}{\zeta}(s)\big(1 - \zeta(s) M(s)\big)^J$$

we have on one side

$$F(s) = \frac{\zeta'}{\zeta}(s) + \sum_{j=1}^{J} (-1)^j \binom{J}{j} M(s)^j \zeta'(s) \zeta(s)^{j-1}$$

$$= -\sum_n \Lambda(n) n^{-s} + \sum_{j=1}^{J} (-1)^j \binom{J}{j} \sum_{\substack{m_1, \ldots, m_j < z \\ m_1 \ldots m_j n_1 \ldots n_j = n}} \mu(m_1) \ldots \mu(m_j) (\log n_1) n^{-s}$$

and we have on the other side

$$F(s) = -\Big(\sum_l \frac{\Lambda(l)}{l^s}\Big)\Big(-\sum_{k \geqslant z}\Big(\sum_{m|k, m<z} \mu(m)\Big) k^{-s}\Big)^J = \sum_{n \geqslant 2z^J} a_n n^{-s},$$

which completes the proof.

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

Let $0 < \Delta \leqslant 1$, M_i and N_i take values $(1+\Delta)^l z$, l—an integer, $M_i \leqslant z$ and let $\mathscr{M}_i = [(1+\Delta)^{-1} M_i, M_i)$ and $\mathscr{N}_i = [(1+\Delta)^{-1} M_i, M_i)$. We then have

$$(2.3) \qquad \Lambda(n) = \sum_{j=1}^{J} \frac{(-1)^j}{j} \binom{J}{j} \sum_{M_1, \ldots, M_j} \sum_{N_1, \ldots, N_j} \log(N_1 \ldots N_j) \times$$

$$\times \sum_{\substack{m_i \in \mathscr{M}_i \\ m_1 \ldots m_j n_1 \ldots n_j = n}} \mu(m_1) \ldots \mu(m_j) \sum_{n_i \in \mathscr{N}_i} 1 + O\Big(\Delta \sum_{j=1}^{J} \binom{J}{j} \tau_{2j}(n)\Big),$$

the error term coming from the approximation $\log n_1 = \log N_1 + O(\Delta)$.

First, by the prime number theorem and by partial summation (1.1) can be reduced to

(2.4)
$$\mathcal{E}(x,Q):=\sum_{(q,a)=1}\lambda(q)\Big(\sum_{n<x,\,n\equiv a[q]}\Lambda(n)-\frac{1}{\varphi(q)}\sum_{n<x,\,(n,q)=1}\Lambda(n)\Big)\ll x(\log x)^{-A}.$$

For each $\Lambda(n)$ we introduce (2.3) with $z=x^{1/4}$, $J=4$. Accordingly, $\mathcal{E}(x,Q)$ splits up into $\ll(\Delta^{-1}\log x)^8$ sums of the type

$$\mathcal{E}(M_1,\dots,M_j|N_1,\dots,N_j)=\sum_{(q,a)=1}\lambda(q)\Big(\sum_{\substack{m_i\in\mathcal{M}_i,n_i\in\mathcal{N}_i\\ m_1\dots m_jn_1\dots n_j\equiv a[q]}}\mu(m_1)\dots\mu(m_j)-\frac{1}{\varphi(q)}\sum_{\substack{m_i\in\mathcal{M}_i,n_i\in\mathcal{N}_i\\ (m_1\dots m_jn_1\dots n_j,q)=1}}\mu(m_1)\dots\mu(m_j)\Big)$$

where the variables $m_1,\dots,m_j$, $n_1,\dots,n_j$ are additionally constrained by

(2.5)
$$m_1\dots m_jn_1\dots n_j<x.$$

Moreover, there is an error term which contributes to $\mathcal{E}(x,Q)$ at most

(2.6)
$$\Delta\sum_{j=1}^{J}\binom{J}{j}\sum_{\substack{q\le Q\\(a,q)=1}}\tau_k(q)\Big(\sum_{\substack{n<x\\ n\equiv a[q]}}\tau_{2j}(n)+\frac{1}{\varphi(q)}\sum_{\substack{n<x\\(n,q)=1}}\tau_{2j}(n)\Big)\ll\Delta x(\log x)^{\varkappa+8}.$$

Anologously, the sums $\mathcal{E}(M_1,\dots,M_j|N_1,\dots,N_j)$ with $M_1\dots M_jN_1\dots N_j>x$ contribute to $\mathcal{E}(x,Q)$ at most

(2.7)
$$\sum_{j=1}^{J}\binom{J}{j}\sum_{\substack{q\le Q\\(a,q)=1}}\tau_\varkappa(q)\Big(\sum_{\substack{(1+\Delta)^{-2j}x\le n<x\\ n\equiv a[q]}}\tau_{2j}(n)+\frac{1}{\varphi(q)}\sum_{\substack{(1+\Delta)^{-2j}x\le n<x\\(n,q)=1}}\tau_{2j}(n)\Big)\log x\ll\Delta x(\log x)^{\varkappa+9}.$$

Either error (2.6) and (2.7) is admissible if $\Delta=(\log x)^{-A-\varkappa-9}$, which we henceforth assume.

From the above discussion we conclude that the proof of theorem reduces to showing that for $M_1\dots M_jN_1\dots N_j\le x$ it holds

(2.8)
$$\mathcal{E}(M_1,\dots,M_j|N_1,\dots,N_j)\ll x(\log x)^{-9A-\varkappa_1}.$$

Notice that (2.5) is redundant, therefore we suppress it to gain independence of the variables $m_1,\dots,m_j,n_1,\dots,n_j$. Now, $\mathcal{E}(M_1,\dots,M_j|N_1,\dots,N_j)$ can be written as bilinear form once one puts the parameters $M_1,\dots,M_j,N_1,\dots,N_j$ 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 $M<m\le 2M$ and $N<n\le 2N$ respectively. Given $M,N,Q,R\ge 1$ we consider

$$\mathcal{E}(M,N;Q,R)=\sum_{q\le Q}\alpha_q\sum_{\substack{r\le R\\(qr,a)=1}}\beta_r\Big(\sum_{m\sim M}\gamma_m\sum_{\substack{n\sim N\\ mn\equiv a[qr]}}\delta_n-\frac{1}{\varphi(qr)}\sum_{m\sim M}\gamma_m\sum_{\substack{n\sim N\\(mn,qr)=1}}\delta_n\Big)$$

with coefficients α_q, β_r, γ_m, δ_n such that

(A_1) For some constant $\varkappa$ we have

$$|\alpha_q|\le\tau_\varkappa(q),\quad|\beta_r|\le\tau_\varkappa(r),\quad|\gamma_m|\le\tau_\varkappa(m),\quad|\delta_n|\le\tau_\varkappa(n),$$

(A_2) For any $a\ne 0$, $b\ge 1$, $q\ge 1$, $(ab,q)=1$ and $C>0$ we have

$$\sum_{\substack{n\sim N,b|n\\ n\equiv a[q]}}\delta_n=\frac{1}{\varphi(q)}\sum_{n\sim N,b|n}\delta_n+O\big(N(\log 2N)^{-C}\big),$$

the constant implied in the symbol O depending on $\varkappa$ and C only.

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

PROPOSITION 1. *Let $MN\le x$ and for some $\varepsilon>0$*

(3.1)
$$x^\varepsilon<N\le x^{1/3},$$

(3.2)
$$Q\le x^{-\varepsilon}N,$$

(3.3)
$$R\le\min(x^{1/2-\varepsilon}N^{-3/4},\,x^{5/8}N^{-11/8}).$$

Then, for any $A>0$ we have

(3.4)
$$|\mathcal{E}(M,N;Q,R)|\ll x(\log x)^{-A}$$

the constant implied in $\ll$ depending at most on ε, $\varkappa$, a and A.

Before proceeding to essential transformations of $\mathcal{E}(M,N;Q,R)$ 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 $MN<x^{1-\varepsilon}$ thus assuming that $x^{1-\varepsilon}\le MN\le x$, 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 $(Q_i,2Q_i]$, $(R_i,2R_i]$ with $Q_i\le Q$ and $R_i\le R$ respectively one sees that there are $\ll(\log x)^2$ pairs of such intervals and the worst case is $Q_i=Q$,

$R_i = R$, i.e. we may assume that

(3.6) $$q \sim Q, \quad r \sim R.$$

Finally we may assume that

(3.7) n's are squarefree.

Indeed, the n's such that $n = k^2 l$, l squarefree, $k > K$ contribute to $\mathcal{E}(M, N; Q, R)$ at most $O\big(K^{-1}x(\log x)^{\varkappa_2}\big)$ which is admissible if K is a sufficiently large power of $\log x$. As to the remaining n's we transfer the factor k^2 from n to m getting

$$|\mathcal{E}(M, N; Q, R)| \leqslant \sum_{k\leqslant K} |\mathcal{E}^*(k^2M, k^{-2}N; Q, R)| + O\big(K^{-1}x(\log x)^{\varkappa_2}\big)$$

where $\mathcal{E}^*$ means that γ_m, δ_n are replaced by γ_{k^2m} and δ_l respectively. Proposition 1 is applicable for each $\mathcal{E}^*$ giving

$$|\mathcal{E}(M, N; Q, R)| \ll Kx(\log x)^{-A_1} + K^{-1}x(\log x)^{\varkappa_2} \ll x(\log x)^{-A}$$

for $K = (\log x)^{A+\varkappa_2}$ and $A_1 = 2A + \varkappa_2$.

Having assumed (3.5)–(3.7) we begin the proof of (3.4) with an application of the Cauchy–Schwarz inequality

(3.8) $$\mathcal{E}^2(M, N; Q, R) \ll QM\mathcal{D}(M, N; Q, R)\mathcal{L}^{2\varkappa^2}$$

where $\mathcal{L} = \log x$,

(3.9) $$\mathcal{D}(M, N; Q, R) = \sum_{(q,a)=1} \alpha(q) \sum_{(m,q)=1} \gamma(m) \Bigg(\sum_{\substack{r\sim R \\ (r,a)=1}} \beta_r \sum_{\substack{n\sim A \\ mn\equiv a[qr]}} \delta_n - \frac{1}{\varphi(qr)} \sum_{\substack{r\sim R \\ (r,am)=1}} \beta_r \sum_{\substack{n\sim N \\ (n,qr)=1}} \delta_n \Bigg)^2$$

and $\alpha(q)$, $\gamma(m)$ are any functions which majorize the characteristic functions of the intervals (3.6). In what follows we shall require these functions to be of C^∞ class with supports $[\frac{1}{2}Q, 3Q]$ and $[\frac{1}{2}M, 3M]$ respectively and with derivatives satisfying

$$|\alpha^{(\nu)}(q)| \ll Q^{-\nu}, \quad |\gamma^{(\nu)}(m)| \ll M^{-\nu}, \quad \nu \geqslant 0$$

the constant implied in $\ll$ depending at most on ν.

Squaring out (3.9) we write

(3.10) $$\mathcal{D}(M, N; Q, R) = W - 2V + U$$

with the aim of evaluating each term separately.

I. Evaluation of U. Let us begin with the simplest term

$$U = \sum_{(q,a)=1} \alpha(q) \sum_{(r_1r_2,a)=1} \frac{\beta_{r_1}\beta_{r_2}}{\varphi(qr_1)\varphi(qr_2)} \Big(\sum_{(n,qr_1)=1} \delta_n \Big)\Big(\sum_{(n,qr_2)=1} \delta_n \Big) \sum_{(m,qr_1r_2)=1} \gamma(m).$$

By Poisson's summation formula we deduce the following:

$$\sum_{(m,k)=1} \gamma(m) = \sum_{\nu|k} \mu(\nu) \sum_m \gamma(\nu m) = \sum_{\nu|k} \frac{\mu(\nu)}{\nu} \sum_{h\in Z} \hat{\gamma}\Big(\frac{h}{\nu}\Big)$$

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

$$\frac{\varphi(k)}{k}\hat{\gamma}(0)$$

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

$$\ll \sum_{\nu|k} \frac{1}{\nu} \sum_{h=1}^{\infty} \min(M, \nu^2h^{-2}M^{-1}) \ll \tau(k).$$

Hence, letting

$$X = \hat{\gamma}(0) \sum_{(q,a)=1} \frac{\alpha(q)}{q} \sum_{(r_1r_2,a)=1} \frac{(r_1, r_2)}{r_1r_2} \frac{\beta_{r_1}\beta_{r_2}}{\varphi\big((r_1, r_2)q\big)} \Big(\sum_{(n,qr_1)=1} \delta_n \Big)\Big(\sum_{(n,qr_2)=1} \delta_n \Big)$$

we conclude that

(3.11) $$U = X + O(x^\varepsilon Q^{-1}N^2)$$

because $\alpha(q)\varphi(qr_1r_2)/qr_1r_2\varphi(qr_1)\varphi(qr_2) = \alpha(q)(r_1, r_2)/qr_1r_2\varphi\big((r_1, r_2)q\big)$. Here the error term $O(x^\varepsilon Q^{-1}N^2)$ is admissible for (3.4) to hold.

II. Evaluation of V. By definition we have

$$V = \sum_{(q,a)=1} \alpha(q) \sum_{(r_1r_2,a)=1} \frac{\beta_{r_1}\beta_{r_2}}{\varphi(qr_2)} \Big(\sum_{(n_1,qr_1)=1} \delta_{n_1} \Big)\Big(\sum_{(n,qr_2)=1} \delta_{n_2} \Big) \sum_{\substack{(m,r_2)=1 \\ m\equiv a\bar{n}_1[qr_1]}} \gamma(m)$$

where $\bar{n}_1$ stands for a solution of $\bar{n}_1 n_1 \equiv 1[qr_1]$. By the Poisson formula

$$\sum_{\substack{m\equiv l[k] \\ (m,r)=1}} \gamma(m) = \sum_{\substack{\nu|r \\ (\nu,k)=1}} \mu(\nu) \sum_{m\equiv l\bar{\nu}[k]} \gamma(\nu m) = \sum_{\substack{\nu|r \\ (\nu,k)=1}} \frac{\mu(\nu)}{\nu k} \sum_h e\Big(-hl\frac{\bar{\nu}}{k}\Big)\hat{\gamma}\Big(\frac{h}{\nu k}\Big)$$

$$= \frac{1}{k} \prod_{p|r, p\nmid k} \Big(1 - \frac{1}{p}\Big)\hat{\gamma}(0) + O\big(\tau(r)\big).$$

Hence, we conclude that

(3.12) $$V = X + O(x^\varepsilon RN^2)$$

because

$$\frac{a(q)}{\varphi(qr_2)}\frac{1}{qr_1}\prod_{p|r_2,\, p\nmid qr_1}\left(1-\frac{1}{p}\right) = \frac{a(q)}{q}\frac{(r_1, r_2)}{r_1 r_2}\frac{1}{\varphi\big((r_1, r_2)q\big)}.$$

Here the error term $O(x^\varepsilon RN^2)$ is admissible for (3.4) to hold.

III. Main term for W. By definition we have

$$(3.13)\quad W = \sum_{(q,a)=1} a(q) \sum_{(r_1r_2,a)=1} \beta_{r_1}\beta_{r_2} \sum_{\substack{(n_1,qr_1)=(n_2,qr_2)=1\\ n_1\equiv n_2[(r_1,r_2)q]}} \delta_{n_1}\delta_{n_2} \sum_{\substack{mn_1\equiv a[qr_1]\\ mn_2\equiv a[qr_2]}} \gamma($$

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 evaluat contribution of the main term, i.e. the following quantity

$$W_0 = \hat{\gamma}(0)\sum_{(q,a)=1}\frac{a(q)}{q}\sum_{(r_1r_2,a)=1}\frac{\beta_{r_1}\beta_{r_2}}{[r_1, r_2]}\sum_{\substack{(n_1,qr_1)=(n_2,qr_2)=1\\ n_1\equiv n_2[(r_1,r_2)q]}}\delta_{n_1}\delta_{n_2}.$$

We detect the congruence $n_1 \equiv n_2 \big(\mathrm{mod}\,(r_1, r_2)q\big)$ by the well-k orthogonality of multiplicative characters $\chi\big(\mathrm{mod}\,(r_1, r_2)q\big)$. The prin character $\chi = \chi_0$ yields X while the non-principal characters contr

$$\hat{\gamma}(0)\sum_{(r_1r_2,a)=1}\frac{\beta_{r_1}\beta_{r_2}}{[r_1, r_2]}T_\varrho(r_1, r_2)$$

where

$$T_\varrho(r_1, r_2) := \sum_{(q,a)=1}\frac{a(q)}{q}\frac{1}{\varphi(\varrho q)}\sum_{\substack{\chi(\mathrm{mod}\,\varrho q)\\ \chi\neq\chi_0}}\Big(\sum_{(n,r_1)=1}\chi(n)\delta_n\Big)\Big(\sum_{(n,r_2)=1}\overline{\chi}(n)\delta_{\cdot}$$

and for simplicity we denoted $\varrho = (r_1, r_2)$. By the Cauchy–Schwar equality we have

$$|T_\varrho(r_1, r_2)|^2 \leqslant T_\varrho(r_1, r_1)T_\varrho(r_2, r_2).$$

Each $\chi \neq \chi_0$ is induced by exactly one primitive character ψ(m $e > 1$, $e\,|\,\varrho q$, thus letting $\varrho q = ef$ we may write

$$T_\varrho(r, r) \ll Q^{-1}\sum_{f\leqslant\varrho Q}\frac{1}{\varphi(f)}\sum_{1<e\leqslant 2\varrho Q/f}\frac{1}{\varphi(e)}\sideset{}{^*}\sum_{\psi(\mathrm{mod}\,e)}\Big|\sum_{(n,rf)=1}\psi(n)\delta_n\Big|^2.$$

By (A_2) one easily deduces that

$$\Big|\sum_{(n,rf)=1}\psi(n)\delta_n\Big| \ll \tau(rf)eN\mathscr{L}^{-2C}.$$

This will be used for $e \leqslant \mathscr{L}^C = E$, say, whereas if $e > E$ we appeal to the large sieve inequality giving

$$\sum_{E<e\leqslant 2\varrho Q/f}\frac{1}{\varphi(e)}\sideset{}{^*}\sum_{\psi(\mathrm{mod}\,e)}\Big|\sum_{(n,rf)=1}\psi(n)\delta_n\Big|^2 \ll \left(\frac{\varrho Q}{f}+\frac{N}{E}\right)N\mathscr{L}^{\varkappa_3}.$$

From the three last inequalities we first infer that

$$T_\varrho(r, r) \ll \tau^2(r)Q^{-1}N^2\mathscr{L}^{\varkappa_4} + \varrho N\mathscr{L}^{\varkappa_3}$$

and then that the non-principal characters contribute to W_0 in total

$$\ll M\sum_{r_1,r_2\sim R}\frac{\tau_\varkappa(r_1)\tau_\varkappa(r_2)}{[r_1, r_2]}T_\varrho(r_1, r_1)^{1/2}T_\varrho(r_2, r_2)^{1/2}$$

$$\ll Q^{-1}Nx\mathscr{L}^{\varkappa_5} + x\mathscr{L}^{\varkappa_6} \ll Q^{-1}Nx\mathscr{L}^{-2A-2\varkappa^2}$$

for $C = 2A + 2\varkappa^2 + 5$. Concluding the above discussion let us write down what has been proved

$$(3.14)\qquad W_0 = X + O(Q^{-1}Nx\mathscr{L}^{-2A-2\varkappa^2}).$$

Here the error term is just good for (3.4) to hold.

IV. Error term for W. Now we proceed to estimate the most difficult quantity

$$(3.15)\quad W_1 = \sum_{(q,a)=1} a(q)\sum_{(r_1r_2,a)=1}\beta_{r_1}\beta_{r_2}\sum_{\substack{(n_1,qr_1)=(n_2,qr_2)=1\\ n_1\equiv n_2[(r_1,r_2)q]}}\delta_{n_1}\delta_{n_2}R(q, r_1, r_2, n_1, n_2).$$

Let $W_1(\nu, \omega, \varrho)$ be the partial sum of W_1 consisting of the terms with

$$(n_1, n_2) = \nu,\quad q \equiv \omega[\nu],\quad (r_1, r_2) = \varrho.$$

Notice that $(\varrho, a\nu) = 1$ and $(\omega, \nu) = 1$, therefore

$$W_1 = \sum_\nu \sum_{\substack{\omega[\nu]\\ (\omega,\nu)=1}} \sum_{(\varrho,a\nu)=1} W_1(\nu, \omega, \varrho)$$

where

$$W_1(\nu, \omega, \varrho) = \sum_{\substack{(q,a)=1\\ q\equiv\omega[\nu]}} a(q) \sum_{\substack{(r_1,r_2)=1\\ (r_1r_2,a)=1}} \beta_{\varrho r_1}\beta_{\varrho r_2} \sum_{\substack{(n_1,n_2)=1\\ (n_1,r_1)=(n_2,r_2)=1\\ n_1\equiv n_2[\varrho q]}} \delta_{\nu n_1}\delta_{\nu n_2}R(q, \varrho r_1, \varrho r_2, \nu n_1, \nu n_2).$$

and by definition

$$(3.16)\qquad R(q, \varrho r_1, \varrho r_2, \nu n_1, \nu n_2) = \sum_{\substack{m\nu n_1\equiv a[\varrho r_1 q]\\ m\nu n_2\equiv a[\varrho r_2 q]}} \gamma(m) - \hat{\gamma}(0)/\varrho r_1 r_2 q.$$

Applying standard but tedious elementary arguments based mainly on the inequality $\tau_l(n) \ll n^{\varepsilon_1}$ it can be shown that

$$\sum_{\nu} \sum_{\substack{\varrho \\ \nu\varrho > x^{\varepsilon}}} \sum_{\omega[\nu]} |W_1(\nu, \omega, \varrho)| \ll Q^{-1} N x^{1-\varepsilon/2} + x^{1+\varepsilon/2} \mathcal{L}^2, \tag{3.17}$$

so it is an admissible quantity for (3.4) to hold.

For $\nu\varrho \leqslant x^{\varepsilon}$ we treat $W_1(\nu, \omega, \varrho)$ by far deeper means. The two congruences $m\nu n_1 \equiv a[\varrho r_1 q]$ and $m\nu n_2 \equiv a[\varrho r_2 q]$ from (3.16) are equivalent to the one $m \equiv l[\varrho r_1 r_2 q]$ where l is a solution of the system

$$l\nu n_1 \equiv a[\varrho r_1 q], \quad l\nu n_2 \equiv a[\varrho r_2 q]. \tag{3.18}$$

By the Poisson summation formula we obtain

$$R(q, \varrho r_1, \varrho r_2, \nu n_1, \nu n_2) = \frac{1}{\varrho r_1 r_2 q} \sum_{h \neq 0} e\left(\frac{-lh}{\varrho r_1 r_2 q}\right) \hat{\gamma}\left(\frac{h}{\varrho r_1 r_2 q}\right)$$

where $\hat{\gamma}(y)$ is the Fourier transform of $\gamma(m)$. For $|h| \geqslant x^{\varepsilon} Q R^2 M^{-1} = H$, say, by partial integration the number of times depending on ε one gets

$$\hat{\gamma}\left(\frac{h}{\varrho r_1 r_2 q}\right) \ll x^{-10} h^{-2}.$$

This yields

$$W_1(\nu, \omega, \varrho) = \sum_{\substack{(q,a)=1 \\ q \equiv \omega[\nu]}} \alpha(q) \sum_{\substack{(r_1,r_2)=1 \\ (r_1 r_2, a\nu)=1}} \frac{\beta_{\varrho r_1} \beta_{\varrho r_2}}{\varrho r_1 r_2} \sum_{\substack{(n_1, \nu\varrho r_1)=1, (n_2, \nu\varrho r_2)=1 \\ (n_1, n_2)=1, n_1 \equiv n_2[\varrho q]}} \delta_{\nu n_1} \delta_{\nu n_2} \times$$
$$\times \sum_{0 < |h| \leqslant H} e(-hl/\varrho r_1 r_2 q) \int \gamma(q\xi) e(h\xi/\varrho r_1 r_2) \, d\xi + O(1)$$

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 $n_2 \equiv n_1[\varrho q]$ there is an integer t such that

$$n_2 - n_1 = \varrho q t, \quad 1 \leqslant |t| \leqslant N/\nu\varrho Q.$$

Next (3.18) are equivalent to

$$l\nu n_1 = a + \varrho q r_1 u_1, \quad l\nu n_2 = a + \varrho q r_2 u_2 \tag{3.19}$$

where u_1, u_2 are integers. Hence $r_2 u_2 n_1 - r_1 u_1 n_2 = at$, so

$$u_1 \equiv -at\overline{r_1 n_2} (\operatorname{mod} r_2 n_1). \tag{3.20}$$

Moreover, since

$$n_2 - n_1 \equiv \omega\varrho t[\nu\varrho t], \tag{3.21}$$

it follows that

$$u_1 \equiv -a\omega\overline{\varrho r_1} (\operatorname{mod} \nu). \tag{3.22}$$

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

$$u_1 \equiv -at\nu\overline{\nu r_1 n_2} - a r_2 n_1 \overline{r_2 n_1} \omega \overline{\varrho r_1} (\operatorname{mod} \nu r_2 n_1),$$

whence, by (3.19)

$$\frac{lh}{\varrho r_1 r_2 q} = \frac{ah}{\nu\varrho n_1 r_1 r_2 q} + \frac{u_1 h}{\nu n_1 r_2}$$
$$\equiv -ah \frac{\overline{\omega\varrho r_1 r_2 n_1}}{\nu} - aht \frac{\overline{\nu r_1 n_2}}{r_2 n_1} + \frac{ah}{\nu\varrho n_1 r_1 r_2 q} \ (\operatorname{mod} 1)$$

and further, since the last "non-arithmetic" term is $\ll x^{\varepsilon-1}$ this enables us to write

$$e\left(\frac{-lh}{\varrho r_1 r_2 q}\right) = e\left(ah \frac{\overline{\omega\varrho r_1 r_2 n_1}}{\nu}\right) e\left(ath \frac{\overline{\nu r_1 n_2}}{r_2 n_1}\right) + O(x^{\varepsilon-1}).$$

The error term $O(x^{\varepsilon-1})$ contributes to $W_1(\nu, \omega, \varrho)$ trivially $\ll N^2 R^2 x^{2\varepsilon-1} \ll N^{1/2} x^{\varepsilon}$. Therefore

$$W_1(\nu, \omega, \varrho) = \sum_{\substack{(r_1,r_2)=1 \\ (r_1 r_2, a\nu)=1}} \frac{\beta_{\varrho r_1} \beta_{\varrho r_2}}{\varrho r_1 r_2} \sum_{1 \leqslant |t| \leqslant T} \sideset{}{'}\sum_{\substack{(n_1, \nu\varrho r_1)=1, (n_2, \nu\varrho r_2)=1 \\ (n_1, n_2)=1, n_2 - n_1 \equiv \omega\varrho t[\nu\varrho t]}} \delta_{\nu n_1} \delta_{\nu n_2} \times \tag{3.23}$$
$$\times \sum e\left(ah \frac{\overline{\nu\varrho r_1 r_2 n_1}}{\nu}\right) e\left(aht \frac{\overline{\nu r_1 n_2}}{r_2 n_1}\right) \times$$
$$\times \int \alpha\left(\frac{n_2 - n_1}{\varrho t}\right) \gamma\left(\frac{n_2 - n_1}{\varrho t} \xi\right) e\left(\frac{h\xi}{\varrho r_1 r_2}\right) d\xi + O(N^{1/2} x^{\varepsilon}).$$

Here $\sum'$ means that the following extra condition

$$(n_2 - n_1, a\varrho t) = \varrho t \tag{3.24}$$

must be imposed. We detect (3.24) by the well-known Möbius formula

$$\sum_{\substack{(\sigma,\nu)=1, \sigma | a \\ \sigma\varrho t | (n_2 - n_1)}} \mu(\sigma) = \begin{cases} 1 & \text{if } (n_2 - n_1, a\varrho t) = \varrho t, \\ 0 & \text{otherwise.} \end{cases}$$

And for given $\sigma | a$, $(\sigma, \nu) = 1$ the two resulting congruences between n_1 and n_2, namely $n_2 - n_1 \equiv \omega\varrho t[\nu\varrho t]$ and $n_2 - n_1 \equiv 0[\sigma\varrho t]$ can be written, after reinterpreting $\omega (\operatorname{mod} \nu)$ by $\omega\sigma (\operatorname{mod} \nu)$ as one congruence

$$n_2 - n_1 \equiv \varrho\omega\sigma t (\operatorname{mod} \nu\sigma\varrho t). \tag{3.25}$$

The reinterpretation of $\omega \pmod{\nu}$ by $\omega\sigma \pmod{\nu}$ is allowed because if ω runs over the residue classes $(\bmod\, \nu)$, prime to ν, so does $\omega\sigma$.

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 $q = (n_2 - n_1)/\varrho t$ of the integral

$$I(q) = \int a(q)\gamma(q\xi)e\left(\frac{h\xi}{\varrho r_1 r_2}\right)d\xi.$$

The first one is handled by means of additive characters $(\bmod\, \nu\sigma\varrho t)$, precisely we make use of

$$\frac{1}{\nu\sigma\varrho t}\sum_{\lambda[\nu\sigma\varrho t]} e\left(\frac{\lambda(n_2-n_1-\varrho\omega\sigma t)}{\nu\sigma\varrho t}\right) = \begin{cases}1 & \text{if (3.25) holds,}\\ 0 & \text{otherwise}\end{cases} \tag{3.26}$$

and the second one by means of the Fourier integral

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

where by the inversion formula

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

We have $K(\xi,\eta) = 0$ unless $M/6Q \leqslant \xi \leqslant 6M/Q$, 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

$$\iint |K(\xi,\eta)|\,d\xi d\eta \ll MQ^{-1} \tag{3.27}$$

and

$$I(q) = \iint K(\xi,\eta)e\left(\frac{h\xi}{\varrho r_1 r_2} + \frac{\eta(n_2-n_1)}{\varrho t}\right)d\xi d\eta. \tag{3.28}$$

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

$$\begin{aligned}(3.29)\quad |W_1(\nu,\omega,\varrho)| &\ll x^\varepsilon \varrho^2 R^{-2}\iint |K(\xi,\eta)| \times \\ &\times \sum_{\sigma|a,(\sigma,\nu)=1}\ \sum_{\substack{(r_1,r_2)=1,(r_1r_2,a\nu)=1\\ r_1,r_2\sim\varrho^{-1}R}}\ \sum_{1\leqslant|h|\leqslant H}\ \sum_{1\leqslant|t|\leqslant T}\frac{1}{\nu\sigma\varrho t}\times \\ &\times \sum_{\lambda[\nu\sigma\varrho t]}\ \sum_{\substack{(n_1,\nu\varrho r_1)=1\\ n_1\sim\nu^{-1}N}}\left|\sum_{\substack{(n_2,\nu\varrho r_2 n_1)=1\\ n_2\sim\nu^{-1}N}} \delta_{\nu n_2} e\left(n_2\frac{\lambda+\nu\sigma\eta}{\nu\sigma\varrho t}\right)e\left(aht\frac{\overline{\nu r_1 n_2}}{r_2 n_1}\right)\right| d\xi d\eta + N^{1/2}x^\varepsilon.\end{aligned}$$

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 $C, D, H, N, T \geqslant 1$, $a \neq 0$ and $\varrho > 0$*

$$\begin{aligned}(3.30)\quad &\mathscr{S}(C,D,H,N,T)\\ &= \sum_{\substack{c\leqslant C\\ (c,d)=1}}\ \sum_{d\leqslant D}\ \sum_{h\leqslant H}\ \sum_{t\leqslant T}\frac{1}{\varrho^2 t}\sum_{\lambda[\varrho t]}\left|\sum_{\substack{n\leqslant N\\ (n,c)=1}} a(h,n,t)e\left(\frac{\lambda n}{\varrho t}\right)e\left(aht\frac{\overline{dn}}{c}\right)\right|\end{aligned}$$

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

$$\begin{aligned}(3.31)\quad &\mathscr{S}(C,D,H,N,T)\\ &\ll (CN)^\varepsilon\left[CDHTN^{1/2} + (CDHN^2)^{3/4}T^{1/2}\left(1+\frac{C^2}{HTN^3}\right)^{1/8}\right],\end{aligned}$$

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

Proof. We shall prove a general estimate

$$\begin{aligned}(3.32)\quad &\mathscr{S}(C,D,H,N,T)(CDHNT)^{-\varepsilon}\\ &\ll CDHTN^{1/2} + C^{1/2}DHTN^{3/4} + D^{1/2}H^{5/8}N^{3/4}T^{3/8}C^{3/4}(N+HT)^{1/4}\times\\ &\qquad\times[C^{1/4}(HT)^{1/8} + D^{1/4}N^{1/2}(HT)^{1/8} + C^{1/4}D^{1/4}N^{1/8}],\end{aligned}$$

(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

$$\begin{aligned}&\mathscr{S}^2(C,D,H,N,T)\\ &\leqslant CDH\sum_{(c,d)=1} f(c,d)\sum_{h\sim H}\ \sum_{t\sim T}\frac{1}{\varrho}\sum_{\lambda[\varrho t]}\left|\sum_{\substack{n\sim N\\ (n,c)=1}} a(h,n,t)e\left(\frac{\lambda n}{\varrho t}\right)e\left(aht\frac{\overline{dn}}{c}\right)\right|^2\\ &= CDH\sum_h\sum_t t\sum_{n_1\equiv n_2[\varrho t]} a(h,n_1,t)\overline{a(h,n_2,t)}\sum_{(c,dn_1n_2)=1} f(c,d)\times\\ &\qquad\qquad\times e\left(aht(n_2-n_1)\frac{\overline{dn_1n_2}}{c}\right)\\ &= CDH\sum_r\sum_l b_{rl}\sum_{(c,dr)=1} f(c,d)e\left(l\frac{\overline{dr}}{c}\right),\end{aligned}$$

say, where

$$b_{rl} = \sum_{\substack{n_1n_2\sim N\\ n_1n_2=r,\,n_1\equiv n_2[\varrho t]\\ aht(n_2-n_1)=l}}\ \sum_{h\sim H}\ \sum_{t\sim T} ta(h,n_1,t)\overline{a(h,n_2,t)}$$

and $f(c, d)$ is any function which majorizes the characteristic function of the set $[C, 2C]\times[D, 2D]$. In what follows we shall require $f(c, d)$ to be the one which satisfies the assumptions of Theorem 12 of [3], i.e. of C^∞ class with $\operatorname{Supp} f = [\tfrac{1}{2}C, 3C]\times[\tfrac{1}{2}D, 3D]$ and

$$\left|\frac{\partial^{\nu_1+\nu_2}}{\partial c^{\nu_1}\partial d^{\nu_2}} f(c, d)\right| \ll C^{-\nu_1}D^{-\nu_2}$$

for any $\nu_1, \nu_2 \geqslant 0$, the constant implied in $\ll$ depending at most on ν_1, ν_2.

Notice that $N^2 < r \leqslant 4N^2$ and $|l| \leqslant 4|a|HNT = L$, say. The terms with $l = 0$, i.e. the diagonal $n_2 = n_1$ contribute to $\mathscr{S}(C, D, H, N, T)$ 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 $(L_j, 2L_j]$, $L_j = 2^{-j}L$ and then for each of $O(\log 3L)$ resulting partial sums we apply Theorem 12 of [3] giving

$$(3.33)\qquad \sum_r \sum_{l\sim L_j} b_{lr} \sum_{(c,dr)=1} f(c, d)\, e\Big(l\frac{\overline{dr}}{c}\Big) \ll (CDHNT)^\varepsilon [C(N^2+L_j)(C+DN^2+CDNL_j^{-1/2})+D^2N^2L_j]^{1/2}B_j^{1/2}$$

where

$$B_j = \sum_r \sum_{l\sim L_j} |b_{rl}|^2 \ll (HNT)^\varepsilon T \sum_r \sum_{l\sim L_j} |b_{rl}|$$
$$\ll (HNT)^\varepsilon T^2 \# \{n_1, n_2, h, t;\ n_1 \equiv n_2 [t],\ |aht(n_2-n_1)| \sim L_j\}$$
$$\ll (HNT)^\varepsilon T^2 N \# \{h, t, w;\ t\sim T,\ |aht^2w| \sim L_j\} \ll (HNT)^{2\varepsilon} TNL_j.$$

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 12 of [3] depends on the location of exceptional eigenvalues of the Laplacian for the Hecke groups $\Gamma_0(r)$. If the Selberg eigenvalue conjecture is true then the factor

$$\Big(1+\frac{C^2}{HTN^3}\Big)^{1/8}$$

in (3.31) can be suppressed.

Now we wish to estimate $W_1(\nu, \omega, \varrho)$ by an appeal to Lemma 2. To this end we have to interpret the variables c, d, h, t, n from (3.30) appropriately to the situation in (3.29). Let us interpret for given values of the variables ξ, η, σ the variables listed above in the following manner

a as a/σ			
c as n_1r_2	thus	$c \leqslant 4NR$,	
d as $r_1\nu^2$	thus	$d \leqslant 2Rx^{2\varepsilon}$,	
h as h	thus	$h \leqslant H$,	
t as $\nu\sigma t$	thus	$t \leqslant \lvert a\rvert Q^{-1}N$,	
n as n_2	thus	$n \leqslant 2N$,	
$a(h, n, t)$ as $\delta_{\nu n_2} e(\eta n_2/\varrho t)$	thus	$a(h, n, t) \ll x^\varepsilon$.	

Then by (3.30) it follows that

$$W_1(\nu, \omega, \varrho) \ll x^\varepsilon \iint |K(\xi, \eta)| R^{-2} \mathscr{S}(NR, R, H, N, N/Q)\, d\xi d\eta + N^{1/2}x^\varepsilon$$
$$\ll x^\varepsilon MQ^{-1}R^{-2}\Big[R^2HQ^{-1}N^{5/2} + R^{3/2}H^{3/4}N^{11/4}Q^{-1/2}\Big(1+\frac{x}{N^3}\Big)^{1/8}\Big]$$
$$\ll R^2Q^{-1}N^{5/2}x^\varepsilon + RQ^{-3/4}N^{17/8}x^{3/8+\varepsilon}.$$

The same majorization holds for W_1 with possibly different but arbitrarily small ε. 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 $\mathscr{E}(M, N; Q, R)$ having special coefficients δ_n; namely we assume that

(A_3) δ_n is the characterictic function of an interval contained in $(N, 2N]$.

For such δ's (A_2) is obvious. In this section we use another method to prove the following

Proposition 2. *Let $MN \leqslant x$ and for some $\varepsilon > 0$*

$$(4.1)\qquad x^{1/4} < N \leqslant x,$$

$$(4.2)\qquad Q \leqslant x^{-\varepsilon}N,$$

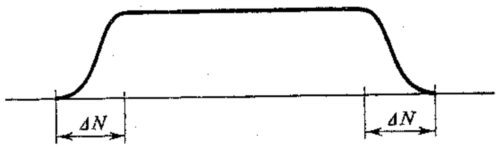
$$(4.3)\qquad R \leqslant (x/N)^{1/3}.$$

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

$$(4.4)\qquad |\mathscr{E}(M, N; Q, R)| \ll x^{1-\eta},$$

the constant implied in $\ll$ depending at most on $\varkappa$, a and ε.

Proof. Like in the previous section we may assume without loss of generality 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 δ_n by a smooth one $\delta(n)$, say, whose graph is



such that

$$\int |\delta_n - \delta(n)|\, dn \leqslant 2\Delta N. \tag{4.5}$$

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

$$\ll \sum_q \sum_r \tau_\varkappa(q)\tau_\varkappa(r) \sum_{mn\equiv a[qr]} \tau_\varkappa(m)\, |\delta_n - \delta(n)| + \\ + \sum_q \sum_r \tau_\varkappa(q)\tau_\varkappa(r)\varphi(qr)^{-1} \sum_m \sum_n \tau_\varkappa(m)\, |\delta_n - \delta(n)|$$

$$\ll x^{\varepsilon^2} M \sum_n |\delta_n - \delta(n)| \ll x^{\varepsilon^2} \Delta MN = \Delta x^{1+\varepsilon^2} = x^{1-\varepsilon^2}$$

for $\Delta = x^{-2\varepsilon^2}$ which we henceforth assume.

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

$$|\delta^{(p)}(n)| \ll (\Delta N)^{-p}$$

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

$$\sum_{n\equiv a\overline{m}[qr]} \delta(n) = \frac{1}{qr}\hat{\delta}(0) + \frac{1}{qr}\sum_{h\neq 0} e\Big(-ah\frac{\overline{m}}{qr}\Big)\hat{\delta}\Big(\frac{h}{qr}\Big) \tag{4.6}$$

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

$$\sum_{(n,qr)=1} \delta(n) = \sum_{\nu|qr} \mu(\nu) \sum_n \delta(\nu n) = \sum_{\nu|qr} \frac{\mu(\nu)}{\nu} \sum_h \hat{\delta}\Big(\frac{h}{\nu}\Big) \tag{4.7}$$
$$= \frac{\varphi(qr)}{qr}\hat{\delta}(0) + O\Big(\sum_{\nu|qr}\frac{1}{\nu}\sum_{h=1}^{\infty}\min(N, \Delta^{-2}N^{-1}\nu^2h^{-2})\Big)$$
$$= \frac{\varphi(qr)}{qr}\hat{\delta}(0) + O\big(\tau(qr)\Delta^{-1}\big).$$

The error term $O\big(\tau(qr)\Delta^{-1}\big)$ contributes to $\mathscr{E}(M, N; Q, R) \ll x^{\varepsilon}\Delta^{-1}M \ll x^{4/5}$ which is admissible.

Further, for $|h| \geqslant x^{\varepsilon^2}QR/\Delta N = x^{3\varepsilon^2}QR/N = H$, say by partial integration the number of times depending on ε we get

$$\Big|\hat{\delta}\Big(\frac{h}{qr}\Big)\Big| \ll x^{-10}h^{-2}$$

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 $\mathscr{E}(M, N; Q, R, H)$, say, which stands for the contribution to $\mathscr{E}(M, N; Q, R)$ of the partial sum of (4.6) with $0 < |h| \leqslant H$.

$$\mathscr{E}(M, N; Q, R, H) \tag{4.8}$$
$$\ll x^{\varepsilon^2}\int_{N/2Q}^{2N/Q} \sum_{q\sim Q}\sum_{\substack{m\sim M\\(q,m)=1}} \Big| \sum_{1\leqslant h\leqslant H}\sum_{\substack{r\sim R\\(r,m)=1}} r^{-1}\beta_r e\Big(\frac{h}{r}\xi\Big)e\Big(-ah\frac{\overline{m}}{qr}\Big)\Big|\, d\xi.$$

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

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

$$\mathscr{S}(C, D, H, S) = \sum_{c\leqslant C}\sum_{\substack{d\leqslant D\\(c,d)=1}} \Big| \sum_{h\leqslant H}\sum_{\substack{s\leqslant S\\(s,d)=1}} a(h,s)\, e\Big(ah\frac{\overline{d}}{cs}\Big)\Big|$$

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

$$\mathscr{S}(C, D, H, S) \tag{4.9}$$
$$\ll (CDHS)^{\varepsilon}\Big\{CD(HS)^{1/2} + C^{1/2}DH^{3/4}(H+S)^{1/4} + CS[DH(H+S)]^{1/2} \times \\ \times \Big(1 + \frac{D}{C} + \frac{D}{\sqrt{HS}}\Big)^{1/4}\Big\}.$$

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

$$|\mathscr{S}(C, D, H, S)|^2 \ll CD \sum_{h_1, h_2\sim H}\sum_{s_1, s_2\sim S} a(h_1, s_1)\overline{a(h_2, s_2)} \times \\ \times \sum\sum_{\substack{(c,d)=1\\(d, s_1s_2)=1}} f(c, d)\, e\Big(a(h_1s_2 - h_2s_1)\frac{\overline{d}}{cs_1s_2}\Big)$$
$$= CD\sum_s\sum_l b_{ls}\sum\sum_{(sc,d)=1} f(c, d)\, e\Big(l\frac{\overline{d}}{cs}\Big)$$

where $f(c, d)$ is like that from the proof of Lemma 2 and

$$b_{ls} = \sum_{\substack{s_1, s_2 \sim S \\ s_1 s_2 = s}} \sum_{\substack{h_1, h_2 \sim H \\ a(h_1 s_2 - h_2 s_1) = l}} a(h_1, s_1) \overline{a(h_2, s_2)}.$$

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 $O\big((HS)^\varepsilon CD(HS)^{1/2}\big)$ 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 $(0, 2L]$, $L = 2|a|HS$, into subintervals of the type $(L_j, 2L_j]$, $L_j = 2^{-j} L_j$ and then estimate each of the $O(\log 3L)$ resulting partial sums by applying Theorem 12 of [3] giving

$$\sum_s \sum_{l \sim L_j} b_{ls} \sum\sum_{(cs, d) = 1} f(c, d)\, e\left(l \frac{\bar{d}}{cs}\right)$$
$$\ll (CDLS)^\varepsilon [CS^2(S^2 + L_j)(C + D + CDL_j^{-1/2}) + D^2 L_j S^{-2}]^{1/2} B_j^{1/2}$$

where

$$B_j = \sum_s \sum_{l \sim L_j} |b_{ls}|^2 \ll \sum_{s_1, s_2} \sum_{|a(h_1 s_2 - h_2 s_1)| \sim L_j} \Big(HS^{-1} \sum_{\substack{s_3 s_4 = s_1 s_2 \\ (s_3, s_4) | (h_1 s_2 - h_2 s_1)}} (s_3, s_4) + 1\Big)$$
$$\ll (HS)^\varepsilon H(H + S) L_j.$$

This shows that the worst case is $L_j = L$ 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

$$\mathcal{E}(M, N; Q, R, H) \ll x^{2\varepsilon^2} N(QR)^{-1} \mathcal{S}(Q, M, H, R)$$
$$\ll x^{3\varepsilon^2} \frac{N}{QR} \Big\{ QM(HR)^{1/2} + Q^{1/2} MH^{3/4}(H + R)^{1/4} + QR\big(MH(H + R)\big)^{1/2} \times$$
$$\times \Big(1 + \frac{M}{Q} + \frac{M}{\sqrt{HR}}\Big)^{1/4} \Big\}$$
$$\ll x^{6\varepsilon^2} \{x(Q/N)^{1/2} + N^{1/2} M^{3/4} R Q^{1/4} + N^{5/8} M^{3/4} (QR^2)^{3/8}\} \ll x^{1 - 2\varepsilon^2}$$

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

5. Proof of theorem, conclusion. It remains to prove (2.8). Our strategy is to arrange each $\mathcal{E}(M_1, \ldots, M_j | N_1, \ldots, N_j)$ as a sum of the type $\mathcal{E}(M, N; Q, R)$ 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 $\mathcal{E}(M, N; Q, R)$ that may occur. However, in order to get the maximal value for QR 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* $1 \leqslant M_i \leqslant x^{1/4}$, $1 \leqslant N_i \leqslant x$, $i = 1, \ldots, j$ *and* $M_1 \ldots M_j N_1 \ldots N_j = y \leqslant x$. *Let* $7/24 \leqslant \theta \leqslant 1/3$. *Then either for some* N_i *we have*

$$N_i \geqslant y^\theta \tag{5.1}$$

or some partial product of $M_1 \ldots M_j N_1 \ldots N_j$, *call it* N, *lies in the interval*

$$y^{1-3\theta} < N \leqslant x^{1/4}. \tag{5.2}$$

Proof. Suppose each N_i is $< y^\theta$. Excluding all N_i's that are $> x^{1/4}$ (at most 3) the remaining N_i's and all M_i's yield a product, let us say $M_1 \ldots M_j N_1 \ldots N_r$, which is $> y^{1-3\theta}$. The smallest partial product of $M_1 \ldots M_j N_1 \ldots N_r$ which is $> y^{1-3\theta}$, call it N, must satisfy $y^{1-3\theta} < N \leqslant \max(x^{1/4}, y^{2(1-3\theta)}) \leqslant x^{1/4}$. This completes the proof.

Proceeding to the proof of (2.8) we first observe that the result is trivial if $y = M_1 \ldots M_j N_1 \ldots N_j < x^{1-\varepsilon}$, thus we assume that $x^{1-\varepsilon} < y < x$. We shall apply Lemma 4 with $\theta = 5/17$. Each $\mathcal{E}(M_1, \ldots, M_j | N_1, \ldots, N_j)$ can be written as $\mathcal{E}(M, N; D_1, D_2)$ with arbitrary $D_1, D_2 \geqslant 1$ subject to $D_1 D_2 = D$ — the level of the weight function and with N such that either

$$N \geqslant x^{\theta(1-\varepsilon)}, \quad \delta_n \text{ satisfies } (\mathrm{A}_1) \text{ and } (\mathrm{A}_3) \tag{5.3}$$

or

$$x^{(1-3\theta)(1-\varepsilon)} < N \leqslant x^{1/4}, \quad \delta_n \text{ satisfies } (\mathrm{A}_1) \text{ and } (\mathrm{A}_2). \tag{5.4}$$

In the last case δ_n satisfies (A_2) 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 $QR = x^{1/3-\varepsilon} N^{2/3} \geqslant x^{(1+2\theta)/3 - 2\varepsilon} = x^{9/17-\varepsilon}$ in the case of (5.3) and $QR = x^{-\varepsilon} \min(x^{1/2-\varepsilon} N^{1/4}, x^{5/8} N^{-3/8}) \geqslant x^{1/2 - (1-3\theta)/4 - 2\varepsilon} = x^{9/17 - 2\varepsilon}$ in the case of (5.4). This makes possible to factorize $D = x^{9/17-2\varepsilon}$ as $D_1 D_2$ with $D_1 \leqslant Q$, $D_2 \leqslant R$. The proof of theorem is complete.

References

[1] M. B. Barban, *New applications of the large sieve of Yu. V. Linnik* (Russian), Akad. Nauk. UzSSR Trudy Inst. Mat. 22 (1961), pp. 1–20.

[2] E. Bombieri, *On the large sieve*, Mathematika 12 (1965), pp. 201–225.

[3] J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cups forms*, Invent. Math. 70 (1982), pp. 219–288.

[4] E. Fouvry, *Répartition des suites dans les progressions arithmétiques*, Acta Arith. 41 (1982), pp. 359–382.

[5] — *Répartition des suites dans les progressions arithmétiques*, Thèse de Doctorat d'Etat à l'Université de Bordeaux I, 1981.

[6] E. Fouvry and H. Iwaniec, *On a theorem of Bombieri–Vinogradov type*, Ma matika 27 (1980), pp. 135–172.
[7] P. X. Gallagher, *The large sieve*, ibid. 14 (1967), pp. 14–20.
[8] — *Bombieri's mean value theorem*, ibid. 15 (1968), pp. 1–6.
[9] D. R. Heath-Brown, *Sieve identities and gaps between primes*, Astérisqu (1982), pp. 61–65.
[10] M. N. Huxley and H. Iwaniec, *Bombieri's theorem in short intervals*, Ma matika 22 (1975), pp. 188–194.
[11] H. Iwaniec, *A new form of the error term in the linear sieve*, Acta Arith (1980), pp. 307–320.
[12] M. Jutila, *A statistical density theorem for L-functions with applications*, i 16 (1969), pp. 207–216.
[13] Yu. V. Linnik, *The large sieve*, Dokl. Akad. Nauk SSSR 30 (1941), pp. 292–
[14] — *The Dispersion Method in Binary Additive Problems*, AMS Translation Math. Monographs, No 4, Providence, Rhode Island, 1963.
[15] H. L. Montgomery, *Problems concerning prime numbers*, Proc. Symp. I Math. 28 (1976), pp. 307–310.
[16] Y. Motohashi, *An induction principle for the generalization of Bombieri's P Number Theorem*, Proc. Japan Acad. 52 (1976), pp. 273–275.
[17] Ch.-D. Pan, *On the representation of large even integer as a sum of a prime an almost prime*, Acta Math. Sin. 12 (1962), pp. 95–106.
[18] A. Renyi, *On the representation of an even number as the sum of a single p and a single almost-prime number* (Russian), Dokl. Akad. Nauk SSSR, Ser. Ma 12 (1948), pp. 57–78.
[19] R. C. Vaughan, *Mean value theorems in prime number theory*, J. London M Soc. (2), 10 (1975), pp. 153–162.
[20] — *An elementary method in prime number theory*, Acta Arith. 37 (1980), pp. –115.
[21] A. I. Vinogradov, *The density hypothesis for Dirichlet L-series* (Russi Izv. Akad. Nauk SSSR, Ser. Math., 29 (1965), pp. 903–934.
[22] J.-R. Chen, *On the Goldbach's problem and the sieve methods*, Sci. Sinica 21 (1 pp. 701–739.
[23] F. Dress, H. Iwaniec and G. Tenenbaum, *Sur une somme liée á la fonc de Möbius* (to appear).
[24] E. Fouvry, *Autour du théorème de Bombieri–Vinogradov* (to appear).
[25] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, Lonc New York 1974.

U.E.R. DE MATHÉMATIQUES
ET D'INFORMATIQUE
UNIVERSITÉ DE BORDEAUX I
351, Cours de la Liberation
33405 Talence, France
MATHEMATICS INSTITUTE
POLISH ACADEMY OF SCIENCES
ul. Śniadeckich 8
00-950 Warszawa, Poland

Received on 9. 12. 1981 (1