ACTA ARITHMETICA
XLVII (1986)

The divisor function $d_3(n)$ in arithmetic progressions

by

D. R. HEATH-BROWN (Oxford)

1. Introduction. We define the general divisor function $d_k(n)$, for positive integral k, as the coefficient of n^{-s} in the Dirichlet series

$$\zeta(s)^k = \sum_{n=1}^{\infty} d_k(n)\, n^{-s}.$$

The primary object of this paper is to consider the case $k = 3$ of the sum

$$D_k(X, q, a) = \sum_{\substack{n \leqslant X \\ n \equiv a \pmod q}} d_k(n),$$

and, in particular, to investigate the ranges of q and X for which the sum can be estimated asymptotically.

To set matters in their proper perspective we first describe the situation for general k. We assume $(a, q) = 1$ for simplicity in this discussion. For each k there exists $\vartheta_k > 0$ such that

$$D_k(X, q, a) \sim M_k(X, q)$$

uniformly for $q \leqslant X^{\vartheta_k - \varepsilon}$ (any $\varepsilon > 0$); here

$$M_k(X, q) = \frac{X}{\varphi(q)} \operatorname{Res}\left(s^{-1} L(s, \chi_0)^k X^{s-1};\ s = 1\right),$$

where χ_0 is the principal character (mod q). When $k = 2$ the best known value is $\vartheta_2 = 2/3$. This is due to Selberg (unpublished), Hooley [15] (essentially), or see Heath-Brown [14], Corollary 1, p. 409. In each case Weil's estimate for the Kloosterman sum is a crucial ingredient in the proof. For $k \geqslant 4$, Lavrik [18] showed that one may take $\vartheta_k = 8/(3k+4)$; the case $k = 4$ is implicit in the earlier work of Linnik [21]. The method here uses estimates for character sums due to Burgess [6], together with the fourth power moment estimate for the average of $L(s, \chi)$ over characters χ (mod q). This technique also applies for $k = 3$, and yields $\vartheta_3 = 1/2$. Recently Smith

[29] (¹) has used Deligne's bound for the generalized Kloosterman sum to estimate $D_k(X, q, a)$. The argument is analogous to Selberg's treatment of the case $k=2$ and yields $\vartheta_k = 2/(k+1)$. This however is inferior to the value given by Lavrik, as soon as $k \geqslant 4$. Further improvements for $k \geqslant 5$ have been obtained by Friedlander and Iwaniec [11] who gave $\vartheta_5 = 9/20$, $\vartheta_6 = 5/12$, $\vartheta_k = 8/3k$ $(k \geqslant 7)$. We shall say more about ϑ_3 in due course.

One important application of the above estimates is to the asymptotic formula

$$\sum_{n \leqslant x} d(n) d_k(n+1) = M_k(x) + E_k(x), \tag{1.1}$$

where $M_k(x)$ is a main term of exact order $x(\log x)^k$, and $E_k(x)$ is an error term; here we have written $d(n)$ for $d_2(n)$. For $k=2$ the sum was first estimated by Ingham [17] who obtained $E_2(x) \ll x \log x$. Estermann [8] gave $E_2(x) \ll x^{11/12+\varepsilon}$, and the exponent was subsequently improved to $5/6+\varepsilon$ by Heath-Brown ([14], Theorem 2) and then to $2/3+\varepsilon$ by Deshouillers and Iwaniec [7]. The method of [14] uses $D_2(X, q, a)$ explicitly, and the exponent $5/6$ arises essentially as $3/2-\vartheta_2$. The improvement over Estermann's result is due solely to the use of sharper bounds for the Kloosterman sum. The technique of Deshouillers and Iwaniec is far more intricate and uses bounds for averages of Kloosterman sums. The case $k=3$ of (1.1) was first settled by Hooley [15], who obtained $E_3(x) \ll x(\log x \log\log x)^2$. (An earlier paper by Bellman [2] contains an error; see the remarks in [15].) Hooley's method depends on the fact that $\vartheta_2 \geqslant 2/3$. The case $k \geqslant 4$ is stated by Linnik [19]; details of the proof are given by Bredikhin [4], with $E_k(x) \ll x(\log x)^{k-1}(\log\log x)^4$. A further improvement, due to Motohashi [24], gives $E_k(x) \ll x(\log x)^{-1}(\log\log x)^c$ for some $c = c(k)$. The papers by Bredikhin and Motohashi use Linnik's dispersion method. An alternative approach due to Redmond [25], [26] is erroneous. (²)

Linnik's proof of (1.1) uses, in effect, an estimate very roughly of the form

$$\sum_{Q_1 < q \leqslant Q_2} |D_2(X, q, -1) - M_2(X, q)| \ll_A X(\log X)^{-A} \tag{1.2}$$

for any $A > 0$. In Bredikhin [4] this estimate is shown, essentially, to hold for $Q_1 = X^{5/6+\varepsilon}$, $Q_2 = X^{1-\varepsilon}$. Of course (1.2) follows from the earlier results on ϑ_2, if $Q_1 = 0$, $Q_2 = X^{2/3-\varepsilon}$. In fact Linnik claimed ([19], [20]) the estimate (1.2) with $Q_1 = X^{1/2}$ and $Q_2 = X^{1-\varepsilon}$, from which indeed it would follow that one could take the complete range $1 \leqslant q \leqslant X^{1-\varepsilon}$. This would easily yield (1.1) with

$$E_k(x) \ll_A x(\log x)^{-A}. \tag{1.3}$$

However Linnik's claim was never substantiated. According to Fouvry (oral communication, 1984) one may take $Q_1 = X^{2/3+\varepsilon}$, $Q_2 = X^{1-\varepsilon}$, by using the Deshouillers–Iwaniec theory of averages of Kloosterman sums. Thus there is an annoying gap $X^{2/3-\varepsilon} \leqslant q \leqslant X^{2/3+\varepsilon}$ for which (1.2) is not yet known.

An alternative method to obtain (1.1) was given by Motohashi [23]; this is at present the easiest route. It depends on estimates of the form

$$\sum_{q \leqslant Q} |D_k(X, q, 1) - M_k(X, q)| \ll_A X(\log X)^{-A} \tag{1.4}$$

with $Q = X^{1/2-\varepsilon}$. This estimate is implicit in Wolke [31], see also Motohashi [22]. If one could take $Q = X^{1/2+\varepsilon}$ then one could obtain (1.3). A weakened form of (1.4) (essentially having the modulus signs removed), but with $Q = X^{1/2+\varepsilon}$, may be derived by the method of Bombieri, Friedlander and Iwaniec [3], using averages of Kloosterman sums. Thus (1.3) is indeed attainable

For the case $k=3$ one can do rather more. This case is of special interest because the "easy" value for ϑ_3 is $1/2$, which is the critical value for (1.4), beyond which one can obtain (1.3). Very recently Friedlander and Iwaniec [10] have shown that one may take $\vartheta_3 = 1/2+1/230$. Their method uses multiple exponential sums, estimated via Deligne's "Riemann Hypothesis", together with ideas originating in Burgess' work [6] on character sums.

The purpose of this paper is to present an alternative method for obtaining $\vartheta_3 > 1/2$, which is simpler and, it turns out, also more powerful. We go on to improve the estimate (1.3), for the case $k=3$, saving a power of x.

Our theorems are not restricted to the case $(a, q) = 1$. Consequently, to replace $M_k(X, q)$, we shall define

$$M_k(X, q, a) = \frac{X}{\varphi(q/\delta)} \operatorname{Res}\Big\{\Big(\sum_{\substack{m=1 \\ (m,q)=\delta}}^{\infty} d_k(m) m^{-s}\Big)\frac{X^{s-1}}{s};\ s=1\Big\}$$

where $(a, q) = \delta$.

Theorem 1. *Let $q \leqslant X^{21/41}$. Then*

$$D_3(X, q, a) = M_3(X, q, a) + O(X^{86/107+\varepsilon} q^{-66/107})$$

for any $\varepsilon > 0$. Hence we may take $\vartheta_3 = 21/41 = 1/2+1/82$.

Note that the error term above is $O(X^{1+\varepsilon}q^{-1})$ precisely when $q \leqslant X^{21/41}$. We can do marginally better with a theorem of Bombieri–Vinogradov type.

(¹) There is a minor error in [29] due to the use of Smith [28]. See our comment in Section 4.

(²) In the proof of Lemma 10 in [25], $H(s+c)$ is not estimated uniformly.

THEOREM 2. *We have*

$$\sum_{q\leqslant Q}\max_{x\leqslant X}\max_{a(\mathrm{mod}\,q)}|D_3(x,q,a)-M_3(x,q,a)|\ll X^{40/51+\varepsilon}Q^{7/17},$$

for any $\varepsilon>0$.

This is more general than (1.4), and non-trivial for $Q\leqslant X^{11/21-3\varepsilon}$. We shall apply Theorem 2 to (1.1).

THEOREM 3. *There is a polynomial P of degree three such that*

$$\sum_{n\leqslant x} d(n)d_3(n+1)=xP(\log x)+O(x^{1-1/102+\varepsilon})$$

for any $\varepsilon>0$.

Deshouillers (oral communication, 1984) has suggested that considerably better error terms may, in time, be obtained through the use of estimates for averages of Kloosterman sums.

The strategy for the proof of Theorem 1 is as follows. The sum $D_3(X,q,a)$ is best thought of as a triple sum, and we shall give different estimates according to the ranges of the three variables. In Section 2 we make the preparatory transformations needed for each of these estimates. The principal bound involves a multiple Kloosterman sum which is essentially of the form $K_2(1,1,rst;q)$. This occurs is Section 4, in a triple sum where r, s, t are typically of size less than $q^{1/2}$. Thus no saving can be obtained by summing over r, s, t in the obvious way. However we shall transform the sum by combining r and s to produce a variable of size greater than $q^{1/2}$, from which a saving is available. This is the key step. In the process we encounter a complete sum (mod q), with 5 variables. This is estimated in Section 3, using for the case in which q is prime, a bound due to Birch and Bombieri ([10], appendix), which uses Deligne's "Riemann Hypothesis".

The two auxilliary estimates are dealt with in Sections 5 and 6. The first of these is straightforward, and uses Weil's bound for the ordinary Kloosterman sum. The second is more complicated, but elementary, and depends on bounds for

$$\#\{n_1,n_2,n_3,n_4;\ 1\leqslant n_i\leqslant N,\ (n_i,q)=1,\ \overline{n_1}+\overline{n_2}\equiv\overline{n_3}+\overline{n_4}\ (\mathrm{mod}\ q)\}\tag{1.5}$$

when N is small compared with q. The three estimates are then compared in Section 7 and various choices of parameters are made, producing the error term which appears in Theorem 1. In each of the estimates there is a main term which is not calculated explicitly. We merely know that it is independent of a, if $(a,q)=1$. However we show in Section 8 that this suffices to determine the leading term explicitly. Finally, all the calculations of Sections 2–7 are done on the assumption that $(a,q)=1$, and we show in Section 8 how this suffices to derive Theorem 1 in the general case.

The proof of Theorem 2 differs only in the use of an averaged bound for (1.5), in the treatment of the second auxilliary estimate in Section 6. The parameter N is independent of a and x, so that we may include the double maximum in our statement of Theorem 2. Finally, the deduction of Theorem 3 from Theorem 2, in Section 9, follows standard lines.

We have tried in this paper to give estimates as accurate as this variant of the method allows. However there are places where a slightly more complicated treatment might possibly lead to improvements. In particular, we have not verified that the division of cases in Section 7 cannot be done more efficiently.

Certain notations and conventions should be described. For the entire proofs of Theorems 1 and 2 we shall assume that $q\leqslant X$, and $Q\leqslant X$. Moreover except in the latter half of Section 8 we shall take a and q to be coprime. The proofs will use small exponents ε, which we shall change from time to time, so that we may write $(qX)^{\varepsilon}\ll X^{\varepsilon}$ without comment, for example. We will use the function $e_q(m)=\exp(2\pi i m/q)$. The symbol $\bar{n}$ will normally indicate the solution of $n\bar{n}\equiv 1\,(\mathrm{mod}\ q)$ with $0<\bar{n}<q$. We only use this notation when $(n,q)=1$. In Sections 3 and 8 the modulus will not always be q however, but will be readily understood from the context. We shall use $\sum^*$ to denote summations restricted to variables coprime to q. We will also need vector notation. We will write $\boldsymbol{u}\cdot\boldsymbol{v}$ for the usual scalar product, and ∇F for the gradient of F. The summation

$$\sum_{\boldsymbol{x}=1}^{q}\quad\text{or}\quad\sum_{\boldsymbol{x}(\mathrm{mod}\,q)}$$

means that each coordinate x_i runs from 1 to q. Finally we shall use $\|\vartheta\|$ to denote the distance from ϑ to the nearest integer or integers.

2. Preliminary transformations. We have

$$D_3(X,q,a)=\#\{(u,v,w)\in N^3;\ uvw\leqslant X,\ uvw\equiv a(\mathrm{mod}\ q)\}.$$

Our first task is to remove the condition $uvw\leqslant X$. Let $qX^{-1}\leqslant\delta\leqslant 1$ and put $\zeta=1+\delta$. We split the ranges for u, v, w into intervals

$$u\in\mathscr{I}=(U,\zeta U],\quad v\in\mathscr{J}=(V,\zeta V],\quad w\in\mathscr{K}=(W,\zeta W],$$

where U, V, W run over powers of ζ and

$$U,V,W\geqslant 1,\quad UVW\leqslant X.\tag{2.1}$$

We set

$$N(U,V,W)=\#\{(u,v,w);\ u\in\mathscr{I},\ v\in\mathscr{J},\ w\in\mathscr{K},\ uvw\equiv a(\mathrm{mod}\ q)\}.$$

Then

$$D_3(X,q,a)=\sum_{UVW\leqslant X}N(U,V,W)+O\Big(\sum_{X\zeta^{-3}<UVW\leqslant X}N(U,V,W)\Big)$$

and

$$\sum_{X\zeta^{-3} < UVW \leqslant X} N(U, V, W) \ll \sum_{\substack{X\zeta^{-3} < n \leqslant X\zeta^3 \\ n \equiv a \pmod q}} d_3(n) \ll X^{1+\varepsilon}\delta q^{-1},$$

whence

$$D_3(X, q, a) = \sum_{UVW \leqslant X} N(U, V, W) + O(X^{1+\varepsilon}\delta q^{-1}). \tag{2.2}$$

We now transform $N(U, V, W)$ in three different ways, in preparation for our three alternative estimates. Firstly we have

$$\begin{aligned} N(U, V, W) &= \sum_{\substack{\alpha,\beta,\gamma=1 \\ \alpha\beta\gamma \equiv a \pmod q}}^{q} \# \{(u, v, w);\ u \in \mathscr{I},\ v \in \mathscr{J},\ w \in \mathscr{K},\ q \mid \alpha - u,\ \beta - v,\ \gamma - w\} \\ &= \sum_{\alpha,\beta,\gamma} q^{-3} \Big(\sum_{r=1}^{q} \sum_{u \in \mathscr{I}} e_q(r\alpha - ru)\Big)\Big(\sum_{s=1}^{q} \sum_{v \in \mathscr{J}} e_q(s\beta - sv)\Big)\Big(\sum_{t=1}^{q} \sum_{w \in \mathscr{K}} e_q(t\gamma - tw)\Big) \\ &= q^{-3} \sum_{r,s,t=1}^{q} \Big(\sum_{\substack{\alpha,\beta,\gamma=1 \\ \alpha\beta\gamma \equiv a \pmod q}}^{q} e_q(r\alpha + s\beta + t\gamma)\Big) F_q(r, s, t), \end{aligned} \tag{2.3}$$

where

$$F_q(r, s, t) = \Big(\sum_u e_q(-ru)\Big)\Big(\sum_v e_q(-sv)\Big)\Big(\sum_w e_q(-tw)\Big)$$
$$\ll \min(I, \|r/q\|^{-1}) \min(J, \|s/q\|^{-1}) \min(K, \|t/q\|^{-1}),$$

with $I = \# \{u \in \mathscr{I}\}$ etc. The α, β, γ sum in (2.3) is the multiple Kloosterman sum $K_2(r, s, at; q)$, so that

$$N(U, V, W) = q^{-3} \sum_{r,s,t=1}^{q} K_2(r, s, at; q) F_q(r, s, t). \tag{2.4}$$

The other two expressions for $N(U, V, W)$ are similar to (2.4). We have

$$\begin{aligned} N(U, V, W) &= \sum_{u \in \mathscr{I}} \sum_{\substack{\beta,\gamma=1 \\ u\beta\gamma \equiv a \pmod q}}^{q} \# \{(v, w);\ v \in \mathscr{J},\ w \in \mathscr{K},\ q \mid \beta - v,\ \gamma - w\} \\ &= \sum_{u,\beta,\gamma} q^{-2} \Big(\sum_{s=1}^{q} \sum_{v \in \mathscr{J}} e_q(s\beta - sv)\Big)\Big(\sum_{t=1}^{q} \sum_{w \in \mathscr{K}} e_q(t\gamma - tw)\Big) \\ &= q^{-2} \sum_u \sum_{s,t=1}^{q} \Big(\sum_{\substack{\beta,\gamma=1 \\ u\beta\gamma \equiv a \pmod q}}^{q} e_q(s\beta + t\gamma)\Big) F_q(s, t), \end{aligned} \tag{2.5}$$

where

$$F_q(s, t) = \Big(\sum_{v \in \mathscr{J}} e_q(-sv)\Big)\Big(\sum_{w \in \mathscr{K}} e_q(-tw)\Big) \tag{2.6}$$
$$\ll \min(J, \|s/q\|^{-1}) \min(K, \|t/q\|^{-1}).$$

The β, γ sum in (2.5) is the ordinary Kloosterman sum $K_1(s, at\bar{u}; q)$, whence

$$N(U, V, W) = q^{-2} \sum_{u \in \mathscr{I}} \sum_{s,t=1}^{q} K_1(s, at\bar{u}; q) F_q(s, t). \tag{2.7}$$

Alternatively we may write

(2.8)
$$\begin{aligned} N(U, V, W) &= \sum_{u \in \mathscr{I}} \sum_{v \in \mathscr{J}} \# \{w \in \mathscr{K};\ w \equiv a\overline{uv} \pmod q\} \\ &= \sum_{u,v} q^{-1} \sum_{t=1}^{q} \sum_{w \in \mathscr{K}} e_q(at\overline{uv} - tw) = q^{-1} \sum_{u \in \mathscr{I}, v \in \mathscr{J}} \sum_{t=1}^{q} e_q(at\overline{uv}) F_q(t), \end{aligned}$$

with

$$F_q(t) = \sum_{w \in \mathscr{K}} e_q(-tw) \ll \min(K, \|t/q\|^{-1}). \tag{2.9}$$

3. Exponential sums. In this section we give estimates for an exponential sum which will occur later. We begin by defining the multiple Kloosterman sum:

$$K_n(\boldsymbol{a}; q) = \sum_{\substack{\boldsymbol{x}=1 \\ \prod x_i \equiv 1 \pmod q}}^{q}{}^{*} e_q(\boldsymbol{a} \cdot \boldsymbol{x}),$$

where $\boldsymbol{a}$, $\boldsymbol{x}$ are $(n+1)$-dimensional vectors. We note some basic properties of $K_n(\boldsymbol{a}; q)$. We only need the special case $n = 2$. By Smith ([27], Theorem 6) we have

$$|K_2(\boldsymbol{a}; q)| \leqslant q(\boldsymbol{a}, q) d_3(q), \tag{3.1}$$

where $(\boldsymbol{a}; q) = (a_1, a_2, a_3; q)$; observe that $\sigma_2(\boldsymbol{a}; p^\alpha) \leqslant \min(\alpha, r)$ in the notation of [27]. Moreover by [27], Theorems 2 and 3, we have

$$K_2(p^r \boldsymbol{a}; p^\alpha) = p^{2r} K_2(\boldsymbol{a}; p^{\alpha-r}) \quad (r < \alpha), \tag{3.2}$$

$$K_2(\boldsymbol{a}; p^\alpha) = 0 \quad (\alpha \geqslant 2,\ (\boldsymbol{a}, p) = 1,\ p \mid \textstyle\prod a_i). \tag{3.3}$$

The sum in which we are interested is

$$S(k, t_1, t_2, \varrho, \sigma; q) = \sum_{j=1}^{q}{}^{*} e_q(kj) K_2(\varrho, \sigma, jt_1; q) \overline{K_2(\varrho, \sigma, jt_2; q)}.$$

This has a product formula for $(u, v) = 1$, namely

$$S(k, t_1, t_2, \varrho, \sigma; uv) = S(v^2 k, t_1, t_2, \varrho, \sigma; u) S(u^2 k, t_1, t_2, \varrho, \sigma; v); \tag{3.4}$$

a general result of this type has been given by Hooley ([16], Lemma 3). This product formula allows us to restrict attention to prime power moduli.

We shall write

$$S = S(k,\, t_1,\, t_2,\, \varrho,\, \sigma;\, p^{\alpha}),$$

for convenience. From (3.2) and (3.3) we see that $S = 0$ unless

$$(t_1,\, p^{\alpha-1}) = (t_2,\, p^{\alpha-1}) = (\varrho,\, p^{\alpha-1}) = (\sigma,\, p^{\alpha-1}) = p^r,$$

say, in which case

$$S = p^{4r} \sum_{j=1}^{p^{\alpha}}{}^{*} e_{p^{\alpha}}(kj)\, K_2(\varrho',\, \sigma',\, jt_1';\, p^{\alpha-r})\, \overline{K_2(\varrho',\, \sigma',\, jt_2';\, p^{\alpha-r})},$$

where

$$\varrho = p^r \varrho', \qquad \sigma = p^r \sigma', \qquad t_i = p^r t_i'.$$

We now replace j by $j_1 + p^{\alpha-r} j_2$, where j_1 runs (mod $p^{\alpha-r}$) with $p \nmid j_1$, and j_2 runs (mod p^r). The sum S then has a factor

$$\sum_{j_2=1}^{p^r} e_{p^r}(kj_2) = \begin{cases} p^r, & p^r \mid k, \\ 0, & p^r \nmid k. \end{cases}$$

Hence S vanishes unless $p^r \mid k$, in which case

$$S = p^{5r} S(k',\, t_1',\, t_2',\, \varrho',\, \sigma';\, p^{\alpha-r}), \tag{3.5}$$

with $k = p^r k'$. We proceed to investigate S, firstly for $\alpha = 1$, and then for $\alpha \geqslant 2$, $(p,\, t_1 t_2 \varrho\sigma) = 1$.

To help in dealing with the case $\alpha = 1$ we note the trivial results

$$K_2(u,\, v,\, w;\, p) = 1 \qquad (p \mid u,\; p \nmid vw), \tag{3.6}$$

$$K_2(u,\, v,\, w;\, p) = 1 - p \qquad (p \mid u,\, v,\; p \nmid w), \tag{3.7}$$

$$K_2(u,\, v,\, w;\, p) = (p-1)^2 \qquad (p \mid u,\, v,\, w). \tag{3.8}$$

When $p \nmid kt_1 t_2 \varrho\sigma$ we have, by (3.1),

$$\begin{aligned} S &= \sum_{j=1}^{p} e_p(kj)\, K_2(\varrho,\, \sigma,\, jt_1;\, p)\, \overline{K_2(\varrho,\, \sigma,\, jt_2;\, p)} - |K_2(\varrho,\, \sigma,\, 0;\, p)|^2 \\ &= \sum_{j=1}^{p} \sum_{x,y,X,Y=1}^{p}{}^{*} e_p(kj + \varrho x + \sigma y + jt_1 \overline{xy} - \varrho X - \sigma Y - jt_2 \overline{XY}) + O(1) \\ &= pS' + O(1), \end{aligned} \tag{3.9}$$

where

$$S' = \sum_{\substack{x,y,X,Y=1 \\ p \mid k + t_1 \overline{xy} - t_2 \overline{XY}}}^{p}{}^{*} e_p(\varrho x + \sigma y - \varrho X - \sigma Y). \tag{3.10}$$

We replace x by $\bar{\varrho} x$, y by $\bar{\sigma} y$, X by $-\bar{\varrho} X$ and Y by $-\bar{\sigma} Y$, whence

$$S' = \sum_{\substack{x,y,X,Y=1 \\ \alpha\overline{xy} + \beta\overline{XY} \equiv 1}}^{p}{}^{*} e_p(x + y + X + Y),$$

with $\alpha = -t_1 \varrho\sigma\bar{k}$, $\beta = t_2 \varrho\sigma\bar{k}$. This sum has been considered by Birch and Bombieri ([10], appendix), who obtained $|S'| \leqslant c_1 p^{3/2}$ for $p \nmid \alpha\beta$. Here c_1, and later $c_2, c_3, \ldots$, are absolute constants. It follows that

$$|S| \leqslant c_2 p^{5/2}.$$

Since

$$\sum_{j=1}^{p}{}^{*} e_p(kj) = \begin{cases} -1, & p \nmid k, \\ p-1, & p \mid k, \end{cases}$$

we can deduce from (3.6), (3.7) and (3.8) that

$$|S| \leqslant p^{-2}(k,\, p)(t_1,\, p)(t_2,\, p)(\varrho,\, p)^2(\sigma,\, p)^2,$$

whenever $p \mid (t_1, t_2)$, $p \mid \varrho$ or $p \mid \sigma$. If $p \nmid t_1$, $p \mid t_2$, say, and $p \nmid \varrho\sigma$ then (3.6) and (3.1) yield

$$|S| \leqslant \sum_{j=1}^{p}{}^{*} |K_2(\varrho,\, \sigma,\, jt_1;\, p)| \leqslant 3p(p-1).$$

Finally, when $p \mid k$ and $p \nmid t_1 t_2 \varrho\sigma$ we may use (3.9) and (3.10). The summation condition $p \mid k + t_1 \overline{xy} - t_2 \overline{XY}$ entails $Y \equiv \overline{t_1}\, t_2 xy\bar{X}$ (mod p), whence

$$\begin{aligned} S' &= \sum_{x,y,X=1}^{p}{}^{*} e_p(\varrho x + \sigma y - \varrho X - \sigma \overline{t_1}\, t_2 xy\bar{X}) \\ &= p \sum_{\substack{x,X=1 \\ t_2 x \equiv t_1 X}}^{p}{}^{*} e_p(\varrho x - \varrho X) - \sum_{x,X=1}^{p}{}^{*} e_p(\varrho x - \varrho X) = p \sum_{x=1}^{p}{}^{*} e_p\big(\varrho x(1 - t_2 \overline{t_1})\big) - 1. \end{aligned}$$

Thus

$$S' \ll \begin{cases} p^2, & p \mid t_1 - t_2, \\ p, & p \nmid t_1 - t_2, \end{cases}$$

so that

$$|S| \leqslant c_3 p^2 (t_1 - t_2,\, p) \qquad (p \mid k,\; p \nmid t_1 t_2 \varrho\sigma).$$

Collecting our various estimates we now find:

Lemma 1. *Let* $S = S(k,\, t_1,\, t_2,\, \varrho,\, \sigma;\, p)$. *Then*

$$|S| \leqslant c_4 p^{5/2} \{(k,\, t_1 - t_2,\, p)(t_1,\, p)(t_2,\, p)(\varrho,\, p)(\sigma,\, p)\}^{1/2}.$$

The other case we have to examine is $\alpha \geqslant 2$, $(p, t_1 t_2 \varrho\sigma) = 1$. Since

$$K_2(\varrho, \sigma, jt_i; q) = K_2(1, 1, j\varrho\sigma t_i; q) \qquad ((\varrho\sigma, q) = 1)$$

we shall replace t_i by $\varrho\sigma t_i$ and consider $S = S(k, t_1, t_2, 1, 1; p^f)$ for $f \geqslant 2$ and $p \nmid t_1 t_2$. On noting that $K_2(1, 1, jt_i; p^f) = 0$ for $p|j$, by (3.3), we have

$$\begin{aligned} S &= \sum_{j=1}^{p^f}{}^{*} e_{p^f}(kj) K_2(1, 1, jt_1; p^f) \overline{K_2(1, 1, jt_2; p^f)} \\ &= \sum_{j=1}^{p^f} e_{p^f}(kj) K_2(1, 1, jt_1; p^f) \overline{K_2(1, 1, jt_2; p^f)} \\ &= \sum_{j=1}^{p^f} \sum_{\boldsymbol{m}=1}^{p^f}{}^{*} e_{p^f}(kj + m_1 + m_2 + jt_1 \overline{m_1 m_2} - m_3 - m_4 - jt_2 \overline{m_3 m_4}) \\ &= p^f \sum_{\boldsymbol{m}=1}^{p^f}{}^{(1)} e_{p^f}(m_1 + m_2 - m_3 - m_4) \end{aligned}$$

where $\sum^{(1)}$ denotes the conditions $(m_i, p) = 1$ and

$$k + t_1 \overline{m_1 m_2} \equiv t_2 \overline{m_3 m_4} \pmod{p^f},$$

or equivalently,

$$km_1 m_2 m_3 m_4 + t_1 m_3 m_4 \equiv t_2 m_1 m_2 \pmod{p^f}.$$

Hence

$$S = \sum_{j=1}^{p^f} \sum_{\boldsymbol{m}=1}^{p^f}{}^{*} e_{p^f}(F(j, \boldsymbol{m})),$$

where

(3.11) $\quad F(j, \boldsymbol{m}) = m_1 + m_2 - m_3 - m_4 + j(km_1 m_2 m_3 m_4 + t_1 m_3 m_4 - t_2 m_1 m_2).$

We now require the following result.

Lemma 2. *Let $F(\boldsymbol{x}) \in \boldsymbol{Z}[x_1, \ldots, x_n]$ and let A be a set of residue classes for $\boldsymbol{x}$* (mod p). *Write*

$$S = \sum_{\substack{\boldsymbol{x}=1 \\ \boldsymbol{x}\in A}}^{p^f} e_{p^f}(F(\boldsymbol{x}))$$

and

$$B = \{\boldsymbol{x} \pmod{p^g};\ \boldsymbol{x} \in A,\ p^g | \nabla F(\boldsymbol{x})\}.$$

Then

$$|S| \leqslant p^{nf/2} \# B \qquad (f = 2g \geqslant 2),$$

$$|S| \leqslant p^{nf/2} \sum_{\boldsymbol{x}\in B} p^{(n-r(\boldsymbol{x}))/2} \qquad (f = 2g+1 \geqslant 3),$$

where $r(\boldsymbol{x})$ is the rank (mod p) *of the quadratic form $Q_{\boldsymbol{x}}$ whose matrix is*

$$\left(\frac{1}{2}\frac{\partial^2 F}{\partial x_i \partial x_j}\right)_{i,j=1}^{n}.$$

(*If $p = 2$ we define $r(\boldsymbol{x}) = 0$.*)

To prove the lemma let $\boldsymbol{x} = \boldsymbol{u} + p^g \boldsymbol{v} + p^{f-g} \boldsymbol{w}$, where $\boldsymbol{u}$ belongs to A and runs (mod p^g), $\boldsymbol{v}$ runs (mod p^{f-2g}), and $\boldsymbol{w}$ runs (mod p^g). Then

$$F(\boldsymbol{x}) \equiv F(\boldsymbol{u}) + p^g \boldsymbol{v} \cdot \nabla F(\boldsymbol{u}) + p^{f-g} \boldsymbol{w} \cdot \nabla F(\boldsymbol{u}) + p^{2g} Q_{\boldsymbol{u}}(\boldsymbol{v}) \pmod{p^f}.$$

It follows that

$$\begin{aligned} |S| &\leqslant \sum_{\substack{\boldsymbol{u}=1 \\ \boldsymbol{u}\in A}}^{p^g} \Big| \sum_{\boldsymbol{w}=1}^{p^g} e_{p^g}(\boldsymbol{w} \cdot \nabla F(\boldsymbol{u})) \Big| \cdot \Big| \sum_{\boldsymbol{v}=1}^{p^{f-2g}} e_{p^{f-g}}(\boldsymbol{v} \cdot \nabla F(\boldsymbol{u}) + p^g Q_{\boldsymbol{u}}(\boldsymbol{v})) \Big| \\ &= p^{ng} \sum_{\boldsymbol{u}\in B} \Big| \sum_{\boldsymbol{v}=1}^{p^{f-2g}} e_{p^{f-2g}}(\boldsymbol{v} \cdot \boldsymbol{c} + Q_{\boldsymbol{u}}(\boldsymbol{v})) \Big|, \end{aligned}$$

where $\nabla F(\boldsymbol{u}) = p^g \boldsymbol{c}$. If $f = 2g$ the conclusion of the lemma is immediate. For $f = 2g+1$ we need only note that $Q_{\boldsymbol{u}}$ can be diagonalised (if $p \neq 2$) and that

$$\Big| \sum_{v=1}^{p} e_p(vc + Qv^2) \Big| \leqslant p^{1/2} \qquad (p \nmid 2Q).$$

For our application A is the set of $(j, \boldsymbol{m})$ with $p \nmid m_i$. The conditions in B for the function (3.11) are

(3.12) $\quad km_1 m_2 m_3 m_4 + t_1 m_3 m_4 - t_2 m_1 m_2 \equiv 0 \pmod{p^g},$

(3.13) $\quad 1 + jkm_2 m_3 m_4 - jt_2 m_2 \equiv 0 \pmod{p^g},$

(3.14) $\quad 1 + jkm_1 m_3 m_4 - jt_2 m_1 \equiv 0 \pmod{p^g},$

(3.15) $\quad -1 + jkm_1 m_2 m_4 + jt_1 m_4 \equiv 0 \pmod{p^g},$

(3.16) $\quad -1 + jkm_1 m_2 m_3 + jt_1 m_3 \equiv 0 \pmod{p^g}.$

From (3.13) and (3.14) we have

$$m_2 \equiv \overline{jt_2 - jkm_3 m_4} \equiv m_1 \pmod{p^g},$$

and similarly (3.15) and (3.16) yield $m_4 \equiv m_3 \pmod{p^g}$. We may therefore substitute for m_2 and m_4. Then on adding (3.13) and (3.15) and dividing by j we find

$$km_1 m_3 (m_1 + m_3) \equiv t_2 m_1 - t_1 m_3 \pmod{p^g},$$

while (3.12) becomes

(3.17) $\quad km_1^2 m_3^2 \equiv t_2 m_1^2 - t_1 m_3^2 \pmod{p^g}.$

Then, on eliminating k, we obtain $m_3^3 \equiv \bar{t}_1 t_2 m_1^3 \pmod{p^g}$. Thus for some root τ of $\tau^3 \equiv \overline{t_1}\, t_2 \pmod{p^g}$, we have $m_3 = \tau m_1 \pmod{p^g}$. Hence τ, m_1 determine j, m_2, m_3, m_4. Moreover τ can take at most 3 values (mod p^g). Since (3.17) yields

$$m_1^2 k \equiv t_1(\tau-1) \pmod{p^g},$$

we have $(k, p^g) = (\tau-1, p^g)$, whence m_1 takes at most $4(k, p^g)$ values (mod p^g) (or $2(k, p^g)$ for $p \neq 2$). We conclude that $\# B \leqslant 12(k, p^g)$.

Now if $(j, \boldsymbol{m}) \in B$ we find that, for $p \neq 2$, the matrix of $2Q_{(j,\boldsymbol{m})}$ is given (mod p) by

$$\begin{bmatrix} 0 & -m_1\tau^2 t_1 & -m_1\tau^2 t_1 & m_1\tau^2 t_1 & m_1\tau^2 t_1 \\ -m_1\tau^2 t_1 & 0 & -jt_1\tau^2 & jt_1\tau(\tau-1) & jt_1\tau(\tau-1) \\ -m_1\tau^2 t_1 & -jt_1\tau^2 & 0 & jt_1\tau(\tau-1) & jt_1\tau(\tau-1) \\ m_1\tau^2 t_1 & jt_1\tau(\tau-1) & jt_1\tau(\tau-1) & 0 & jt_1\tau \\ m_1\tau^2 t_1 & jt_1\tau(\tau-1) & jt_1\tau(\tau-1) & jt_1\tau & 0 \end{bmatrix}.$$

If we divide through by $jt_1\tau^2$ and then multiply the first row and first column by $\overline{jm_1}$ (note that $p \nmid jt_1 m_1 \tau$), the matrix becomes

$$\begin{bmatrix} 0 & -1 & -1 & 1 & 1 \\ -1 & 0 & -1 & 1-\bar\tau & 1-\bar\tau \\ -1 & -1 & 0 & 1-\bar\tau & 1-\bar\tau \\ 1 & 1-\bar\tau & 1-\bar\tau & 0 & \bar\tau \\ 1 & 1-\bar\tau & 1-\bar\tau & \bar\tau & 0 \end{bmatrix}.$$

and a simple calculation gives $r(j, \boldsymbol{m}) = 5$ unless $p \mid 3(\tau-1)$, in which case $r(j, \boldsymbol{m}) = 4$. Since $(k, p^g) = (\tau-1, p^g)$ we conclude that

$$\sum_{\boldsymbol{x} \in B} p^{(n-r(\boldsymbol{x}))/2} \leqslant c_5 (k, p^g)(k, p)^{1/2}.$$

We have now established the following lemma.

LEMMA 3. *Let $p \nmid t_1 t_2$ and let $f \geqslant 2$, and write $S = S(k, t_1, t_2, 1, 1; p^f)$. Let $g = [f/2]$. Then $S = 0$ unless $(k, p^g) \mid t_1 - t_2$, in which case*

$$|S| \leqslant 12 p^{5f/2}(k, p^g) \qquad (f \text{ even}),$$
$$|S| \leqslant c_5 p^{5f/2}(k, p^g)(k, p)^{1/2} \qquad (f \text{ odd}).$$

We may combine the results of Lemmas 1 and 3 with (3.4) and (3.5) to deduce the last estimate of this section:

LEMMA 4. *There exists a positive constant A such that*

$$S(k, t_1, t_2, \varrho, \sigma; q) \leqslant d_A(q)\, q^{5/2} \{(k^2, (t_1-t_2)^2, q)(t_1, q)(t_2, q)(\varrho, q)(\sigma, q)\}^{1/2} ((t_1-t_2)^3, q)^{1/6}.$$

4. The principal estimate for $N(U, V, W)$. For our principal treatment of $N(U, V, W)$ we think of U, V, W as being, typically, of order $X^{1/3}$. Our starting point is (2.4). Let $N_1 = N_1(U, V, W)$ denote the contribution from terms with r, s or t equal to q. Since $(a, q) = 1$ these terms will be independent of a. We shift the ranges for r, s, t so that each variable runs over $(-q/2, q/2]$, and divide these intervals into subsets $R < |r| \leqslant 2R$, $S < |s| \leqslant 2S$, $T < |t| \leqslant 2T$, where $1 \leqslant R, S, T \leqslant q$. We then write

$$F(R; q) = \min(I, qR^{-1}), \qquad F(S; q) = \min(J, qS^{-1}),$$

so that

$$N(U, V, W) - N_1 \ll (\log q)^3 F(R; q) F(S; q) q^{-3} \sum_{\substack{R < |r| \leqslant 2R \\ S < |s| \leqslant 2S}} \Big| \sum_{T < |t| \leqslant 2T} K_2(r, s, at) F_q(t) \Big|,$$

for some choice of R, S and T.

We would now like to transform $K_2(r, s, at; q)$ into sums of the form $K_2(1, 1, arst; q)$. The "generalized Kuznetsov identity" of Smith [28] would be ideal for this purpose, but unfortunately there is an error in the proof and the "identity" is invalid. ([3]) Instead we argue as follows. Let $r = \varrho r'$, where ϱ is a product of powers of those primes that divide q, and $(r', q) = 1$. If $r < 0$ we take $\varrho < 0$, $r' > 0$. Similarly set $s = \sigma s'$. Then

$$K_2(r, s, at; q) = K_2(\varrho, \sigma, ar's't; q).$$

For brevity we shall write $K(n)$ to mean $K_2(\varrho, \sigma, an; q)$. Now, on replacing r', s' by r, s, we find that

$$\text{(4.1)} \qquad N(U, V, W) - N_1 \ll (\log q)^3 F(R; q) F(S; q) q^{-3} \sum_{\varrho, \sigma} \sum_{r, s} \Big| \sum_{T < |t| \leqslant 2T} K(rst) F_q(t) \Big|,$$

where $R < |\varrho| r \leqslant 2R$ and $S < |\sigma| s \leqslant 2S$.

We shall need to know how many values ϱ and σ can take. Let $q = q_1^{e_1} \ldots q_n^{e_n}$ (q_i primes) and let $p_1 < \ldots < p_n$ be the first n primes. If $\varrho = \pm q_i^{f_i} \ldots q_n^{f_n}$ satisfies $0 < |\varrho| \leqslant q$ then $\varrho' = \pm p_1^{f_1} \ldots p_n^{f_n}$ will also satisfy $0 < |\varrho'| \leqslant q$. Consequently, if we set

$$\Psi(x, y) = \# \{n;\ 1 \leqslant n \leqslant x,\ p \mid n \text{ implies } p \leqslant y\},$$

([3]) In [28], page 318, the statement "... the inner sum in (6) is zero unless $r = 1$." is wrong. Indeed the main theorem would yield $K_2(1, 0, 0; q) = K_2(1, 1, 0; q)$, whereas $K_2(1, 0, 0; q) = \mu(q)\varphi(q)$, while $K_2(1, 1, 0; q) = \mu(q)^2$. Consequently the estimate on [28], page 320, that

$$|K_n(a; q)| \leqslant q^{n/2}(a, q)^{n/2} d_{n+2}(q),$$

is not proven. Indeed $K_3(1, 0, 0, 0; q) = \mu(q)\varphi(q)^2$, while the proposed bound would be $O(q^{3/2+\varepsilon})$. As we saw in (3.1) such a bound is true for $n = 2$.

then

$$\#\{\varrho;\ 0<|\varrho|\leqslant q\}\leqslant 2\Psi(q,p_n).$$

However, from de Bruijn [5] one has

$$\Psi(x,y)\ll x(\log y)^2\exp\{-u(\log u)-u(\log\log u)+O(u)\},\qquad u=\log x/\log y.$$

Since $p_n \ll \log q$ we conclude that

(4.2) $$\#\{\varrho;\ 0<|\varrho|\leqslant q\}\ll \exp\left\{O\left(\frac{\log q}{\log\log q}\right)\right\}.$$

The number of pairs ϱ, σ is thus $O(q^\varepsilon)$ for any $\varepsilon>0$. Hence (4.1) yields, for some pair ϱ, σ,

(4.3) $$N(U,V,W)-N_1\ll q^{-3+2\varepsilon}F(R;q)F(S;q)\sum_{r,s}\Big|\sum_{T<|t|\leqslant 2T}K(rst)F_q(t)\Big|.$$

We now replace R, S by $R|\varrho|$, $S|\sigma|$, whence (4.3) becomes

(4.4) $$N(U,V,W)-N_1\ll q^{-3+2\varepsilon}F(R|\varrho|;q)F(S|\sigma|;q)\sum\nolimits_1$$

with

$$\sum\nolimits_1=\sideset{}{^*}\sum_{R\leqslant r\leqslant 2R}\ \sideset{}{^*}\sum_{S<s\leqslant 2S}\Big|\sum_{T<|t|\leqslant 2T}K(rst)F_q(t)\Big|.$$

We have now arrived at the crux of the whole proof. The key idea goes back a surprisingly long way; in the context of $d_3(n)$, for the Piltz Divisor Problem, it appears in the work of Yüh [32], but it may indeed be even older.

We write $h=rs$, $H=RS$. Then, by Cauchy's inequality, we have

$$\sum\nolimits_1\leqslant\sideset{}{^*}\sum_{H<h\leqslant 4H}d(h)\Big|\sum_{T<|t|\leqslant 2T}K(ht)F_q(t)\Big|$$
$$\leqslant\Big\{\sum_{H<h\leqslant 4H}d(h)^2\Big\}^{1/2}\Big\{\sideset{}{^*}\sum_{H<h\leqslant 4H}\Big|\sum_{T<|t|\leqslant 2T}K(ht)F_q(t)\Big|^2\Big\}^{1/2}$$
$$\ll\{H(\log H)^3\}^{1/2}\Big\{\sum_{t_1,t_2}F_q(t_1)\overline{F_q(t_2)}\sideset{}{^*}\sum_{h}K(ht_1)\overline{K(ht_2)}\Big\}^{1/2}.$$

If we now set $F(T;q)=\min(K,qT^{-1})$, then (4.4) yields

(4.5) $$N(U,V,W)-N_1\ll H^{1/2}q^{-3}X^{3\varepsilon}F(R|\varrho|;q)F(S|\sigma|;q)F(T;q)\sum\nolimits_2^{1/2},$$

where

$$\sum\nolimits_2=\sum_{T<|t_1|,|t_2|\leqslant 2T}|\Sigma(t_1,t_2)|$$

and

(4.6) $$\Sigma(t_1,t_2)=\sideset{}{^*}\sum_{H<h\leqslant 4H}K(ht_1)\overline{K(ht_2)}.$$

The effect of this key step has been to replace a sum over two variables r, s, whose ranges are rather short, by a sum over h, with a longer range. Since H is in general larger than $q^{1/2}$, while R, S are typically smaller than $q^{1/2}$, some saving is possible from the h summation, while none could be obtained from r or s.

We proceed to transform $\Sigma(t_1,t_2)$ into a sum over a complete range (mod q). We have

$$\Sigma(t_1,t_2)=\sideset{}{^*}\sum_{j=1}^{q}\Big\{\frac{1}{q}\sum_{k=1}^{q}\sum_{H<h\leqslant 4H}e_q\big(k(j-h)\big)\Big\}K(jt_1)\overline{K(jt_2)}$$
$$=\frac{1}{q}\sum_{k=1}^{q}S(\bar{a}k,t_1,t_2,\varrho,\sigma;q)\sum_{H<h\leqslant 4H}e_q(-kh),$$

in the notation of Section 3. Thus

$$\Sigma(t_1,t_2)\ll\frac{H}{q}|S(0,t_1,t_2,\varrho,\sigma;q)|+\sum_{1\leqslant|k|\leqslant q}k^{-1}|S(\bar{a}k,t_1,t_2,\varrho,\sigma;q)|.$$

We now call on Lemma 4 to bound the sums S. We note that

$$\sum_{1\leqslant k\leqslant q}k^{-1}(k^2,D)^{1/2}\leqslant\sum_{d|D}d^{1/2}\sum_{\substack{1\leqslant k\leqslant q\\ d|k^2}}k^{-1}=\sum_{d|D}d^{1/2}\sum_{\substack{1\leqslant k\leqslant q\\ \delta|k}}k^{-1}$$
$$\ll\sum_{d|D}d^{1/2}\delta^{-1}\log q\ll d(D)\log q,$$

where δ is the smallest positive integer for which $d|\delta^2$. Thus

(4.7) $$\Sigma(t_1,t_2)\ll q^{5/2+\varepsilon}\{(t_1,q)(t_2,q)(\varrho,q)(\sigma,q)\}^{1/2}$$
$$\times\big((t_1-t_2)^3,q\big)^{1/6}\big(Hq^{-1}\big((t_1-t_2)^2,q\big)^{1/2}+1\big).$$

We next consider the contribution to $\sum_2$ arising from terms $t_1\neq t_2$ with $(t_i,q)=d_i$ and $(t_1-t_2,q)=\delta$. It follows that $d_2(d_1,\delta)(d_1,d_2,\delta)^{-1}$ divides t_2, and that t_2 determines t_1 (mod $d_1\delta(d_1,\delta)^{-1}$). We may take $d_2\geqslant d_1$ by symmetry. The number of t_i with $T<|t_i|\leqslant 2T$, $t_1\neq t_2$ is then $O\big(T^2(d_1,d_2,\delta)(d_1d_2\delta)^{-1}\big)+O\big(T(d_1,d_2,\delta)d_2^{-1}(d_1,\delta)^{-1}\big)$. Since $\delta\ll T$ the second error term is $O\big(T^2(d_1d_2)^{-1/2}\delta^{-1}\big)$. Using the boinds

$$\big((t_1-t_2)^3,q\big)^{1/6}\big((t_1-t_2)^2,q\big)^{1/2}\leqslant q^{1/6}\delta,$$
$$\big((t_1-t_2),q\big)^{1/6}\leqslant\delta^{1/2},\qquad (d_1,d_2,\delta)\leqslant(d_1d_2)^{1/2},$$

we conclude that terms $t_1\neq t_2$ contribute to $\sum_2$ a total

$$\ll q^{5/2+\varepsilon}(\varrho,q)^{1/2}(\sigma,q)^{1/2}T^2\sum_{d_1,d_2,\delta}(Hq^{-5/6}+\delta^{-1/2})$$
$$\ll q^{5/2+2\varepsilon}(\varrho,q)^{1/2}(\sigma,q)^{1/2}T^2(Hq^{-5/6}+1).$$

To deal with the terms $t_1 = t_2$ we merely use (3.1) and (4.6), whence

$$\Sigma(t, t) \ll q^{2+\varepsilon}(\varrho, \sigma, t, q)^2 H \ll q^{2+\varepsilon}(\varrho, q)(\sigma, q) H.$$

Such terms contribute to $\sum_2$ a total

$$\ll q^{2+\varepsilon}(\varrho, q)(\sigma, q) TH.$$

We now insert these bounds into (4.5), using the estimates

$$TF(T; q) \leqslant q, \quad T^{1/2} F(T; q) \leqslant (Kq)^{1/2},$$
$$R(\varrho, q)^{1/4} F(R|\varrho|; q) \leqslant R(\varrho, q)^{1/2} F(R|\varrho|; q) \leqslant q,$$
$$R^{1/2}(\varrho, q)^{1/4} F(R|\varrho|; q) \leqslant (Iq)^{1/2},$$

and the analogous results for S, σ. This leads to:

LEMMA 5. *There exists $N_1(U, V, W)$, independent of a, such that*

$$N(U, V, W) - N_1(U, V, W) \ll X^{4\varepsilon}\{q^{5/6} + q^{1/4}(IJ)^{1/2} + q^{1/2} K^{1/2}\}.$$

5. The first auxiliary bound. We shall need an estimate for $N(U, V, W)$ which is efficient when one of U, V, W (U, say) is "small". We write

$$N_1(U, V, W) = N_1 = q^{-1} \sum_{u \in \mathscr{I}} \sum_{\substack{s,t=1 \\ s \text{ or } t = q}}^{q} K_1(s, at\bar{u}; q) F_q(s, t).$$

This is independent of a, since

$$K_1(s, 0; q) = c_q(s), \quad K_1(0, at\bar{u}; q) = c_q(at\bar{u}) = c_q(t),$$

where $c_q(\cdot)$ is the Ramanujan sum. We now apply (2.7), using the bound

$$|K_1(m, n; q)| \leqslant q^{1/2}(m, n, q)^{1/2} d(q).$$

This is given by Estermann [9]; the essential ingredient, when q is prime, being due to Weil. We conclude that

$$N(U, V, W) - N_1 \ll q^{-3/2} d(q) I \sum_{s,t=1}^{q-1} (s, t, q)^{1/2} F_q(s, t).$$

By (2.6) the sum over s, t is

$$\ll \sum_{0<|s|,|t| \leqslant q/2} q^2 |s|^{-1} |t|^{-1} (s, t, q)^{1/2} \ll q^2 \sum_{s,t=1}^{q} (st)^{-1} (s, t, q)^{1/2}$$

$$\ll q^2 \sum_{d|q} d^{1/2} \Big(\sum_{\substack{s=1 \\ d|s}}^{q} s^{-1}\Big)^2 \ll q^2 \sum_{d|q} d^{-3/2} (\log q)^2 \ll q^2 (\log q)^2.$$

On combining estimates we now have the following result.

LEMMA 6. *For a suitable function $N_1(U, V, W)$, independent of a, one has*

$$N(U, V, W) - N_1(U, V, W) \ll q^{1/2+\varepsilon} I, \quad \textit{for any } \varepsilon > 0.$$

6. The second auxiliary bound. This section is devoted to the estimation of $N(U, V, W)$ when one of U, V, W, (W say) is "large" and the others are "small". We shall use the methods of Heath-Brown ([13], pp. 366–368) and Balasubramanian, Conrey and Heath-Brown ([1], proof of Lemma 7).

Our starting point is (2.8). On putting

$$N_1(U, V, W) = N_1 = q^{-1} \sum_{u \in \mathscr{I}}{}^{*} \sum_{v \in \mathscr{J}}{}^{*} F_q(0),$$

which is independent of a, we find

$$N(U, V, W) - N_1 \ll q^{-1} \sum_u \sum_{t=1}^{q-1} |F_q(t)| \cdot \Big|\sum_v e_q(at\overline{uv})\Big|.$$

Let

$$n(k) = \sum_{u,t} |F_q(t)|,$$

where the sum is for

(6.1) $\qquad u \in \mathscr{I}, \quad (u, q) = 1, \quad 1 \leqslant t \leqslant q-1, \quad at\bar{u} \equiv k \pmod{q}.$

Then, by Hölder's inequality, we obtain

$$\text{(6.2)} \quad N(U, V, W) - N_1 \ll q^{-1} \sum_{k=1}^{q} n(k) \Big|\sum_v e_q(k\bar{v})\Big|$$
$$\ll q^{-1} \Big\{\sum_{k=1}^{q} n(k)\Big\}^{1/2} \Big\{\sum_{k=1}^{q} n(k)^2\Big\}^{1/4} \Big\{\sum_{k=1}^{q} \Big|\sum_v e_q(k\bar{v})\Big|^4\Big\}^{1/4}.$$

From (2.9) we have

$$\text{(6.3)} \quad \sum_k n(k) = \sum_{u,t} |F_q(t)| \ll I \sum_{t=1}^{q-1} \|t/q\|^{-1} \ll qI(\log q).$$

To estimate $\sum n(k)^2$ we shift the range of t in the definition (6.1) so that $0 < |t| \leqslant q/2$. We then split up this new range into intervals $T < |t| \leqslant 2T$, where T runs over powers of 2, and $1/2 \leqslant T \leqslant q$. We write $n(k) = \sum n(k, T)$ accordingly, and note that

$$n(k, T) \ll \min(K, qT^{-1}) \, \# \{(u, t);\ u \in \mathscr{I},\ T < |t| \leqslant 2T,\ at\bar{u} \equiv k \pmod{q}\}.$$

Thus

$$\sum_{k=1}^{q} n(k, T)^2 \ll \min(K^2, q^2 T^{-2}) \, \# \{(u_1, u_2, t_1, t_2);\ at_1\overline{u_1} \equiv at_2\overline{u_2} \pmod{q}\}.$$

The congruence condition requires $t_1 u_2 \equiv t_2 u_1 \pmod{q}$. There are $O(TI)$ pairs (t_2, u_1). Moreover, the number of integers n for which $n \equiv t_2 u_1 \pmod{q}$ and $TU < |n| \leqslant 2T\zeta U$ is $O((1 + TUq^{-1}))$; and for each such n there are

$O(X^\varepsilon)$ solutions of $t_1 u_2 = n$. Hence

$$\#\{(u_1, u_2, t_1, t_2);\ at_1\overline{u_1} \equiv at_2\overline{u_2} \pmod q\} \ll X^\varepsilon(TI+T^2UIq^{-1}).$$

It now follows from Cauchy's inequality that

$$\sum_k n(k)^2 \ll X^\varepsilon(\log q)\sum_T \min(K^2, qT^{-2})(TI+T^2UIq^{-1}).$$

In estimating the sum over T we consider seperately the cases $T \leqslant qK^{-1}$ and $T > qK^{-1}$, whence

$$\sum n(k)^2 \ll X^\varepsilon(\log q)^2(qIK+qUI). \tag{6.4}$$

Finally we consider the third sum in (6.2), which is

$$\sum_{k=1}^{q}\Big|\sum_v e_q(k\overline{v})\Big|^4 = q\,\#\{v_i \in \mathscr{J};\ \overline{v_1}+\overline{v_2} \equiv \overline{v_3}+\overline{v_4} \pmod q\} = qH(\mathscr{J}, q), \tag{6.5}$$

say. Our first treatment of $H(\mathscr{J}, q)$ follows Heath-Brown ([13], pp. 367–368). Let

$$m(s) = \#\{v_1, v_2 \in \mathscr{J};\ \overline{v_1}+\overline{v_2} \equiv s \pmod q\}, \tag{6.6}$$

so that

$$H(\mathscr{J}, q) = \sum_{s=1}^{q} m(s)^2.$$

Each v_1 in (6.6) determines $v_2 \pmod q$, and therefore

$$m(s) \ll J(1+q^{-1}J). \tag{6.7}$$

Hence

$$H(\mathscr{J}, q) \ll J(1+q^{-1}J)\sum_{s=1}^{q} m(s) \ll J^3(1+q^{-1}J). \tag{6.8}$$

This trivial bound suffices unless $J \leqslant q$, as we now assume. If $(s, q) = d$ then (6.6) requires $d \mid v_1+v_2$. Thus, for any $d \mid q$ we have

$$\sum_{d \mid s} m(s) \ll J(1+d^{-1}J). \tag{6.9}$$

Since $J \leqslant q$, it follows from (6.7) that $m(s) \ll J$, whence

$$\sum_{d \mid q,\, d \geqslant D}\ \sum_{\substack{s=1\\ (s,q)=d}}^{q} m(s)^2 \ll d(q)J^2(1+D^{-1}J). \tag{6.10}$$

For small values of $(s, q) = d$ we shall apply the strong form of Dirichlet's approximation theorem (see Hardy and Wright [12], Theorem 36, for example). This shows that for any positive integer L there exist integers α, β

for which

$$0 < \alpha \leqslant L, \quad |s/q-\beta/\alpha| \leqslant 1/(\alpha(L+1)).$$

Hence $s\alpha \equiv \gamma \pmod q$, with $0 < \alpha \leqslant L$, $|\gamma| \leqslant q/(L+1)$. We shall choose

$$L = \min(q/d-1, [(qV)^{1/2}]).$$

Thus $qd^{-1} \nmid \alpha$, whence $q \nmid s\alpha$, so that $\gamma \neq 0$. The relation $\overline{v}_1+\overline{v}_2 \equiv s \pmod q$ now yields $\alpha(v_1+v_2) \equiv \gamma v_1 v_2 \pmod q$, whence

$$\alpha(v_1+v_2) = \gamma v_1 v_2 - qt, \tag{6.11}$$

for some integer t. The bounds for α, γ and v_i yield

$$t \ll V^2dq^{-1}+V^{3/2}q^{-1/2}.$$

To each value of t there correspond $O(X^\varepsilon)$ possible pairs v_1, v_2. To prove this we write (6.11) as

$$(\gamma v_1-\alpha)(\gamma v_2-\alpha) = \alpha^2+qt.$$

Here α, γ are fixed, so that each t leads to $O(X^\varepsilon)$ possible factors $\gamma v_i-\alpha$ of α^2+qt, except when $\alpha^2+qt = 0$. In the latter case, however, one unknown must be α/γ and the other is then determined by $\overline{v}_1+\overline{v}_2 \equiv s \pmod q$. It now follows that

$$m(s) \ll X^\varepsilon(1+V^2dq^{-1}+V^{3/2}q^{-1/2}).$$

We may combine this with (6.9) to obtain

(6.12)
$$\sum_{\substack{d \mid q\\ d \leqslant D}}\ \sum_{\substack{s=1\\ (s,q)=d}}^{q} m(s)^2 \ll \sum_d X^\varepsilon J(1+d^{-1}J)(1+V^2dq^{-1}+V^{3/2}q^{-1/2})$$
$$\ll X^\varepsilon d(q)(JV^2Dq^{-1}+J^2+J^2V^2q^{-1}+J^2V^{3/2}q^{-1/2}).$$

If we now choose $D = 1+[q^{1/2}JV^{-1}]$ we may conclude from (6.10) and (6.12) that

$$H(\mathscr{J}, q) \ll X^\varepsilon d(q)J^2(1+V^2q^{-1}+V^{3/2}q^{-1/2}). \tag{6.13}$$

In deriving this we assumed that $J \leqslant q$, but it is clear that (6.13) follows from the trivial bound (6.8) if $J > q$. When we combine the estimates (6.2), (6.3), (6.4), (6.5) and (6.13) we deduce the following

LEMMA 7. *There exists $N_1(U, V, W)$, independent of a, such that*
$$N(U, V, W)-N_1(U, V, W) \ll X^\varepsilon I^{3/4}J^{1/2}(K+U)^{1/4}(1+V^2q^{-1}+V^{3/2}q^{-1/2})^{1/4},$$
for any $\varepsilon > 0$.

We now give our second treatment of $H(\mathscr{J}, q)$, following Balasubramanian, Conrey and Heath-Brown [1], proof of Lemma 7. This leads to an average bound, in which q runs over the interval $Q < q \leqslant 2Q$. From the definition (6.5), we see that

$$H(\mathscr{J}, q) = \#\{v_i \in \mathscr{J};\ q | v^*\},$$
$$v^* = v_2 v_3 v_4 + v_1 v_3 v_4 - v_1 v_2 v_4 - v_1 v_2 v_3.$$

Thus

$$\sum_{Q<q\leqslant 2Q} H(\mathscr{J}, q) \ll \sum_{v_i \in \mathscr{J}, v^* \neq 0} d(|v^*|) + \sum_{v_i \in \mathscr{J}, v^*=0} Q \ll J^{4+\varepsilon} + Q\,\#\{v_i \in \mathscr{J};\ v^* = 0\}.$$

However

$$v^*(v_1+v_2) + v_1^2 v_2^2 = \{(v_1+v_2)v_3 - v_1 v_2\}\{(v_1+v_2)v_4 - v_1 v_2\},$$

so that if $v^* = 0$, then v_1, v_2 determine $O(V^\varepsilon)$ factors $(v_1+v_2)v_3 - v_1 v_2$, and so there correspond $O(V^\varepsilon)$ pairs v_3, v_4. It follows that

$$\sum_{Q<q\leqslant 2Q} H(\mathscr{J}, q) \ll J^{4+\varepsilon} + QJ^{2+\varepsilon}. \tag{6.14}$$

We now write $N(U, V, W; q, a)$ for $N(U, V, W)$, and $N_1(U, V, W; q)$ for $N_1(U, V, W)$. Since $H(\mathscr{J}, q)$ is independent of a, we deduce the following bound, by combining (6.2), (6.3), (6.4), (6.5) and (6.14).

Lemma 8. *There exists $N_1(U, V, W; q)$ independent of a, such that*

$$\sum_{Q<q\leqslant 2Q} \max_{(a,q)=1} |N(U, V, W; q, a) - N_1(U, V, W; q)|$$
$$\ll X^\varepsilon Q^{3/4} I^{3/4} J^{1/2} (K+U)^{1/4} (J^2+Q)^{1/4}.$$

7. The proof of Theorems 1 and 2: error terms. In this section we shall feed Lemmas 5–8 into the formula (2.2) to obtain an asymptotic expression, of the form $D_3(X, q, a) \sim M_3^*(X, q)$, with an appropriate error term. In the next section we shall consider $M_3^*(X, q)$, and remove the condition $(a, q) = 1$. We shall concentrate on the proof or Theorem 1, and merely indicate the significant differences needed in establishing Theorem 2.

We begin by recalling, from Section 2, that U, V, W run over powers of $\zeta = 1+\delta$. Let U_0, V_0, W_0 run over powers of 2 and consider the contribution to $\sum N(U, V, W)$ corresponding to the ranges $U_0 < U \leqslant 2U_0$, $V_0 < V \leqslant 2V_0$, $W_0 < W \leqslant 2W_0$. The number of triples U, V, W will be $O(\delta^{-3})$, and if we apply, say, Lemma 5, to each term, we shall be able to replace $N(U, V, W)$ by $N_1(U, V, W)$ with a total error

$$\ll X^{4\varepsilon}\{\delta^{-3} q^{5/6} + \delta^{-1} q^{1/4} \sum_{U,V} (IJ)^{1/2} + \delta^{-2} q^{1/2} \sum_W K^{1/2}\}. \tag{7.1}$$

Here, by Cauchy's inequality,

$$\sum_U I^{1/2} \leqslant \{\sum_U 1\}^{1/2} \{\sum_U I\}^{1/2} \ll \{\delta^{-1}\}^{1/2} \{U_0\}^{1/2},$$

for example, whence (7.1) is

$$\ll X^{4\varepsilon}\{\delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} (U_0 V_0)^{1/2} + \delta^{-5/2} q^{1/2} W_0^{1/2}\}.$$

We may use Lemmas 6, 7 and 8 in a similar fashion. Since the number of triples U_0, V_0, W_0 is $O((\log X)^3)$ we conclude from (2.2) that, for a suitable function $M_3^*(X, q)$ independent of a, one has

$$D_3(X, q, a) - M_3^*(X, q) \ll X^{5\varepsilon}(\delta X q^{-1} + \max_{\{U_0, V_0, W_0\}} \min\{E_1, E_2, E_3\}). \tag{7.2}$$

Here $U_0, V_0, W_0 \geqslant 1$, $U_0 V_0 W_0 \ll X$ and

$$E_1 = \delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} (U_0 V_0)^{1/2} + \delta^{-5/2} q^{1/2} W_0^{1/2},$$

from Lemma 5,

$$E_2 = \delta^{-2} q^{1/2} U_0,$$

from Lemma 6, and

$$E_3 = U_0^{3/4} V_0^{1/2} (\delta^{-3/2} W_0^{1/4} + \delta^{-7/4} U_0^{1/4})(1 + V_0^{1/2} q^{-1/4} + V_0^{3/8} q^{-1/8}),$$

from Lemma 7. Note that we may permute U_0, V_0, W_0 to our best advantage, in forming the E_i.

If we use Lemma 8 in place of Lemma 7 we find similarly that

$$\sum_{Q<q\leqslant 2Q} \max_{x\leqslant X} \max_{(a,q)=1} |D_3(x, q, a) - M_3^*(x, q)|$$
$$\ll X^{5\varepsilon} Q (\delta X Q^{-1} + \max_{(U_0, V_0, W_0)} \min\{E_1, E_2, E_3^*\}).$$

Here E_1 and E_2 have had q replaced by Q, while E_3 has become

$$E_3^* = U_0^{3/4} V_0^{1/2} (\delta^{-3/2} W_0^{1/4} + \delta^{-7/4} U_0^{1/4})(1 + V_0^{1/2} Q^{-1/4} \delta^{1/2}).$$

To simplify the task of using E_1, E_2, E_3 optimally, we shall introduce parameters A, B and estimate E_i as follows. If any variable (U_0, say) satisfies $U_0 \leqslant q^{1/3}$, then

$$E_2 \ll \delta^{-2} q^{1/2} U_0 \ll \delta^{-2} q^{5/6}.$$

If $U_0, V_0, W_0 \leqslant A$, then

$$E_1 \ll \delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} A + \delta^{-5/2} q^{1/2} A^{1/2}. \tag{7.3}$$

If any variable (W_0, say) is in the range $A \leqslant W_0 \leqslant B$, then

$$E_1 \ll \delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} (X/W_0)^{1/2} + \delta^{-5/2} q^{1/2} W_0^{1/2} \tag{7.4}$$
$$\ll \delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} X^{1/2} A^{-1/2} + \delta^{-5/2} q^{1/2} B^{1/2}.$$

Finally, if $V_0 \leqslant U_0 \leqslant W_0$ say, and $V_0 \geqslant q^{1/3}$, $W_0 \geqslant B$, then

$$(7.5)\qquad \begin{aligned} E_3 &\ll U_0^{3/4} V_0^{1/2} \delta^{-7/4} W_0^{1/4} (V_0^{1/2} q^{-1/4} + V_0^{3/8} q^{-1/8}) \\ &\ll \delta^{-7/4} W_0^{1/4} ((U_0 V_0)^{7/8} q^{-1/4} + (U_0 V_0)^{13/16} q^{-1/8}) \\ &\ll \delta^{-7/4} (q^{-1/4} X^{7/8} W_0^{-5/8} + q^{-1/8} X^{13/16} W_0^{-9/16}) \\ &\ll \delta^{-7/4} q^{-1/4} X^{7/8} B^{-5/8} + \delta^{-7/4} q^{-1/8} X^{13/16} B^{-9/16}. \end{aligned}$$

Here we have used, for example, the fact that $U_0^{3/4} V_0 \leqslant (U_0 V_0)^{7/8}$ whenever $V_0 \leqslant U_0$. Comparison of (7.3) and (7.4) now gives the optimal choice

$$(7.6)\qquad A = \min(X^{1/3},\ \delta^{1/2} q^{-1/4} X^{1/2}),$$

while (7.4) and (7.5) produce the optimal value

$$B = \max(\delta^{2/3} q^{-2/3} X^{7/9},\ \delta^{12/17} q^{-10/17} X^{13/17}).$$

Hence

$$\min\{E_1, E_2, E_3\} \ll \delta^{-3} q^{5/6} + \delta^{-2} q^{1/4} X^{1/3} + \delta^{-9/4} q^{3/8} X^{1/4} + \delta^{-13/6} q^{1/6} X^{7/18} + \delta^{-73/34} q^{7/34} X^{13/34}.$$

From (7.2) we see that the best choice of δ will be

$$\delta = \max(q^{11/24} X^{-1/4},\ q^{5/12} X^{-2/9},\ q^{11/26} X^{-3/13},\ q^{7/19} X^{-11/57},\ q^{41/107} X^{-21/107}).$$

This will satisfy $qX^{-1} \leqslant \delta \leqslant 1$, as required in Section 2, providing that $q \leqslant X^{21/41}$. We now deduce from (7.2) that

$$D_3(X, q, a) = M_3^*(X, q) + O(X^{3/4+\varepsilon} q^{-13/24}) + O(X^{7/9+\varepsilon} q^{-7/12}) + O(X^{10/13+\varepsilon} q^{-15/26}) + O(X^{46/57+\varepsilon} q^{-12/19}) + O(X^{86/107+\varepsilon} q^{-66/107}).$$

When $X^{2/9} \leqslant q \leqslant X^{21/41}$ the last error term dominates, and when $q \leqslant X^{2/9}$ the result of Smith ([29], Theorem 3) yields

$$D_3(X, q, a) = M_3^*(X, q) + O(X^{1/2+\varepsilon}) = M_3^*(X, q) + O(X^{86/107+\varepsilon} q^{-66/107}).$$

We now describe the estimation of E_1, E_2, E_3^* for the proof of Theorem 2. This will entail a different choice of A, B and δ. We shall use (7.3) and (7.4) as before. If, say $W_0 \geqslant B$, $V_0 \geqslant B\delta$, then

$$E_2 \ll \delta^{-2} Q^{1/2} (X/V_0 W_0) \ll \delta^{-3} Q^{1/2} X B^{-2}.$$

If, say, $W_0 \geqslant B$ and $Q^{1/2} \delta^{-1} \leqslant V_0 \leqslant U_0 \leqslant B\delta$, then

$$\begin{aligned} E_3^* &\ll U_0^{3/4} V_0^{1/2} \cdot \delta^{-3/2} W_0^{1/4} \cdot V_0^{1/2} Q^{-1/4} \delta^{1/2} \\ &\ll \delta^{-1} Q^{-1/4} X^{1/4} (U_0 V_0)^{5/8} \ll \delta^{-1} Q^{-1/4} X^{7/8} B^{-5/8}. \end{aligned}$$

Finally, if $W_0 \geqslant B$, $U_0, V_0 \leqslant A$, $U_0 \leqslant B\delta$ and $V_0 \leqslant Q^{1/2} \delta^{-1}$, say, then

$$E_3^* \ll U_0^{3/4} V_0^{1/2} \cdot \delta^{-3/2} W_0^{1/4} \ll \delta^{-3/2} X^{1/2} A^{1/4} B^{-1/4}.$$

This last estimate will be

$$E_3^* \ll \delta^{-1} Q^{-1/4} X^{7/8} B^{-5/8},$$

providing that

$$(7.7)\qquad QAB^{3/2} \ll \delta^2 X^{3/2}.$$

We now choose A as in (7.6), and

$$B = \max(\delta^{-1/5} X^{2/5},\ \delta^{4/3} Q^{-2/3} X^{7/9}).$$

Assuming that

$$(7.8)\qquad \delta^{69} X^{17} \geqslant Q^{30},$$

we will have $B = \delta^{4/3} Q^{-2/3} X^{7/9}$, and (7.7) will be satisfied. The optimal choice of δ is then

$$\begin{aligned} \delta &= \max(Q^{11/24} X^{-1/4},\ Q^{5/12} X^{-2/9},\ Q^{11/26} X^{-3/13},\ Q^{7/17} X^{-11/51}) \\ &= Q^{7/17} X^{-11/51}, \end{aligned}$$

if $Q \leqslant X^{11/21}$, and (7.8) does indeed hold. Moreover $QX^{-1} \leqslant \delta \leqslant 1$, for $Q \leqslant X^{11/21}$, as required. We deduce that

$$\sum_{Q < q \leqslant 2Q} \max_{x \leqslant X} \max_{(a,q)=1} |D_3(x, q, a) - M_3^*(x, q)| \ll X^{40/51+\varepsilon} Q^{7/17}.$$

8. The proof of Theorems 1 and 2: main terms. At present we have an asymptotic formula $D_3(X, q, a) \sim M_3^*(X, q)$ for $(a, q) = 1$ in which we know nothing about $M_3^*(X, q)$, save that it is independent of a. In this section we shall show firstly how to replace $M_3^*(X, q)$ by

$$(8.1)\qquad M_3(X, q) = \frac{X}{\varphi(q)} \operatorname{Res}(s^{-1} L(s, \chi_0)^3 X^{s-1};\ s = 1),$$

where χ_0 is the principal character (mod q). We shall then go on to evaluate $D_3(X, q, a)$ for $(a, q) > 1$, by using our results for $(a, q) = 1$.

We begin by observing that, by the estimates of the previous section,

$$(8.2)\qquad \sum_{a=1}^{q}{}^{*} D_3(X, q, a) = \varphi(q) M_3^*(X, q) + O(X^{86/107+\varepsilon} q^{41/107}).$$

On the other hand, by Perron's formula (see Titchmarsh [30], Lemma 3.12), the left-hand side is

$$\sum_{n \leqslant X}{}^{*} d_3(n) = \frac{1}{2\pi i} \int_{1+\varepsilon-iX}^{1+\varepsilon+iX} L(s, \chi_0)^3 \frac{X^s}{s}\, ds + O(X^{2\varepsilon}).$$

We shift the line of integration to run from $\frac{1}{2}-iX$ to $\frac{1}{2}+iX$, and use the bound

$$L(s,\chi_0) \ll q^{\varepsilon}|\zeta(s)| \ll (qX)^{\varepsilon} X^{(1-\mathrm{Re}(s))/3}.$$

The two horizontal line segments then contribute $O((qX)^{3\varepsilon})$ while the vertical line segment produces

$$\ll q^{3\varepsilon} X^{1/2} \int_{-X}^{X} |\zeta(\tfrac{1}{2}+it)|^3 \frac{dt}{1+|t|} \ll q^{3\varepsilon} X^{1/2+\varepsilon},$$

by the fourth power moment estimate (see [30], (7.6.1)). We may now conclude, on choosing a new value for ε, that

$$\sum_{a=1}^{q}{}^{*} D_3(X,q,a) = \varphi(q) M_3(X,q) + O(X^{1/2}(qX)^{\varepsilon}).$$

Comparison with (8.2) then produces

$$M_3^*(X,q) = M_3(X,q) + O(X^{68/107+\varepsilon} q^{-66/107})$$

for $q \leqslant X^{21/41}$. An analogous argument for Theorem 2 shows that

$$\sum_{Q<q\leqslant 2Q} \max_{x\leqslant X} |M_3^*(x,q) - M_3(x,q)| \ll X^{40/51+\varepsilon} Q^{7/17}$$

for $Q \leqslant X^{11/21}$. We can now deduce Theorems 1 and 2, at least for $(a,q)=1$.

For use in treating the case $(a,q)>1$ we introduce the function

(8.3) $$F(n,\delta) = \sum_{\alpha_1\alpha_2\alpha_3\beta=n} \mu(\alpha_1)\mu(\alpha_2)\mu(\alpha_3) d_3(\beta\delta).$$

This is multiplicative, in the sense that

$$F(n_1 n_2, \delta_1\delta_2) = F(n_1,\delta_1) F(n_2,\delta_2) \qquad ((n_1\delta_1, n_2\delta_2)=1).$$

An easy calculation shows that $F(p^e,1)=0$ for $e \geqslant 1$ and that $F(p^e,p^f)=0$ for $e \geqslant 3$; and consequently $F(n,\delta)=0$ unless $n|\delta^2$. We now take $(a,q)=1$ and observe that

(8.4) $$\sum_{\substack{n|\delta^2\\(n,q)=1}} F(n,\delta) D_3(Xn^{-1},q,a\bar{n}) = \sum_{\substack{m\leqslant X\\ m\equiv a(\mathrm{mod}\, q)}} \sum_{\substack{n|m\\(n,q)=1}} F(n,\delta) d_3(m/n).$$

Since $m \equiv a(\mathrm{mod}\ q)$ and $n|m$, the condition $(n,q)=1$ is redundant. We shall now prove the identity

(8.5) $$\sum_{n|m} F(n,\delta) d_3(m/n) = d_3(m\delta).$$

The right-hand side is

$$\sum_{m=\alpha_1\alpha_2\alpha_3\beta\gamma} \mu(\alpha_1)\mu(\alpha_2)\mu(\alpha_3) d_3(\beta\delta) d_3(\gamma) = \sum_{m=\alpha_1\alpha_2\alpha_3\beta\gamma_1\gamma_2\gamma_3} \mu(\alpha_1)\mu(\alpha_2)\mu(\alpha_3) d_3(\beta\delta)$$
$$= \sum_{\beta|m} d_3(\beta\delta) \sum_{m/\beta=\mu_1\mu_2\mu_3} \prod_{i\leqslant 3} \Big(\sum_{\mu_i=\alpha_i\gamma_i} \mu(\alpha_i)\Big).$$

The three innermost sums vanish unless $\mu_1=\mu_2=\mu_3=1$, and (8.5) follows. From (8.4) and (8.5) we now have

$$\sum_{\substack{n|\delta^2\\(n,q)=1}} F(n,\delta) D_3(Xn^{-1},q,a\bar{n}) = D_3(X\delta, q\delta, a\delta).$$

Taking new values for X, q and a, we set $(q,a)=\delta$, $q=\delta q_1$, $a=\delta a_1$, and deduce, in the case of Theorem 1, that

$$D_3(X,q,a) = \sum_{\substack{n|\delta^2\\(n,q_1)=1}} F(n,\delta) D_3(X(n\delta)^{-1}, q_1, a_1\bar{n})$$
$$= \sum_{n} F(n,\delta)\{M_3(X(n\delta)^{-1}, q\delta^{-1}) + O(X^{86/107+\varepsilon} q^{-66/107})\}.$$

It is clear from (8.3) that $F(n,\delta) \ll (n\delta)^{\varepsilon}$. Since $n|\delta^2$ and $\delta|q$ we therefore find that

$$D_3(X,q,a) = \sum_{n|\delta^2,(n,q_1)=1} F(n,\delta) M_3(X(\delta n)^{-1}, q\delta^{-1}) + O(X^{86/107+\varepsilon} q^{-66/107+4\varepsilon}).$$

According to (8.1) the leading term here is

$$\frac{X}{\delta\varphi(q_1)} \mathrm{Res}\,(s^{-1}(X/\delta)^{s-1} L(s,\chi_0)^3 \sum_n F(n,\delta) n^{-s}; s=1),$$

where χ_0 is the principal character (mod q_1). From (8.5) we see that this is just

$$\frac{X}{\delta\varphi(q_1)} \mathrm{Res}\,\{s^{-1}(X/\delta)^{s-1} (\sum_{\substack{m=1\\(m,q_1)=1}}^{\infty} d_3(m\delta) m^{-s}); s=1\} = M_3(X,q,a),$$

as required. This completes the proof of Theorem 1, and the final stage of the demonstration of Theorem 2 may be carried out in an analogous manner.

9. The proof of Theorem 3. For the proof of Theorem 3 we shall need to know a little about $M_3(X,q)$. We write

$$f_q(s) = \prod_{p|q}\Big(1-\frac{1}{p^s}\Big)^3 = \sum_{n|q^{\infty}} a_n n^{-s},$$

where $n|q^{\infty}$ means that $n|q^e$ for some exponent $e=e(n)$. We note that

(9.1) $$a_n \ll n^{\varepsilon}.$$

Then

$$M_3(X,q) = \frac{X}{\varphi(q)} \operatorname{Res}\{s^{-1}\zeta(s)^3 X^{s-1} f_q(s);\ s=1\}$$

$$= X \sum_{0\leqslant\alpha,\beta\leqslant 2} C_{\alpha,\beta} \sum_{n|q^\infty} \frac{a_n}{n} (\log n)^\alpha (\log X)^\beta \varphi(q)^{-1},$$

for certain constants $C_{\alpha,\beta}$. We also have

$$q/\varphi(q) = \sum_{n|q^\infty} n^{-1},$$

whence

$$M_3(X,q) = \frac{X}{q} \sum_{0\leqslant\beta\leqslant 2} (\log X)^\beta \sum_{n|q^\infty} b_{n,\beta} n^{-1}, \tag{9.2}$$

with

$$b_{n,\beta} = \sum_{0\leqslant\alpha\leqslant 2} C_{\alpha,\beta} \sum_{m|n} a_m (\log m)^\alpha \ll n^{2\varepsilon}, \tag{9.3}$$

by (9.1).

We start the proof of Theorem 3 by observing that

$$d(n) = 2\#\{q|n;\ q\leqslant x^{1/2}\} - \#\{q|n;\ nx^{-1/2}\leqslant q\leqslant x^{1/2}\}$$

whenever $n\leqslant x$. Hence

$$\sum_{n\leqslant x} d(n)d_3(n+1) = \sum_{q\leqslant x^{1/2}} \big(2\{D_3(x+1,q,1)-1\} - \{D_3(qx^{1/2}+1,q,1)-1\}\big)$$

$$= 2\sum_{q\leqslant x^{1/2}} D_3(x,q,1) - \sum_{q\leqslant x^{1/2}} D_3(qx^{1/2},q,1) + O(x^{1/2+\varepsilon}).$$

By Theorem 2 this becomes

$$2\sum_{q\leqslant x^{1/2}} M_3(x,q) - \sum_{q\leqslant x^{1/2}} M_3(qx^{1/2},q) + O(x^{101/102+\varepsilon}), \tag{9.4}$$

on taking $Q = x^{1/2}$. Thus it remains to consider the contribution from the two sums above.

Let $m = m(n)$ be the product of all the primes dividing n. Then (9.2) yields

$$\sum_{q\leqslant x^{1/2}} M_3(x,q) = x \sum_{0\leqslant\beta\leqslant 2} (\log x)^\beta \sum_{\substack{n=1\\ m(n)\leqslant x^{1/2}}}^{\infty} b_{n,\beta} n^{-1} \sum_{\substack{q\leqslant x^{1/2}\\ m|q}} q^{-1} \tag{9.5}$$

$$= x \sum_{0\leqslant\beta\leqslant 2} (\log x)^\beta \sum_{\substack{n=1\\ m\leqslant x^{1/2}}}^{\infty} b_{n,\beta} (mn)^{-1} \big(\log(x^{1/2}/m) + \gamma + O(mx^{-1/2})\big),$$

where γ is Euler's constant. The error term here contributes

$$\ll x^{1/2}(\log x)^2 \sum_{\substack{n=1\\ m\leqslant x^{1/2}}}^{\infty} n^{2\varepsilon-1}, \tag{9.6}$$

by (9.3). We now observe that

$$\#\{n\leqslant N;\ m(n)\leqslant M\} \ll \min(M,N)N^\varepsilon \tag{9.7}$$

by the argument used for (4.2). Consequently (9.6) is $O(x^{1/2+\varepsilon})$. The main terms of (9.5) are

$$x \sum_{0\leqslant\beta\leqslant 2} (\log x)^\beta \sum_{n=1}^{\infty} b_{n,\beta} (mn)^{-1} \big(\log(x^{1/2}/m) + \gamma\big) + O\Big(x(\log x)^3 \sum_{\substack{n=1\\ m\geqslant x^{1/2}}}^{\infty} m^{-1} n^{3\varepsilon-1}\Big) \tag{9.8}$$

$$= x \sum_{0\leqslant\beta\leqslant 3} C_\beta (\log x)^\beta + O(x^{1/2+3\varepsilon})$$

for suitable constants C_β, by (9.3) and (9.7). This completes our consideration of the first sum in (9.4). The only major difference in the treatment of the second sum lies in the use of the estimate

$$\sum_{\substack{q\leqslant x^{1/2}\\ m|q}} (\log q)^e = \frac{x^{1/2}}{m} \sum_{\substack{0\leqslant f,g\leqslant 2\\ f+g\leqslant 2}} (\log m)^f (\log x)^g + O(x^\varepsilon) \qquad (e\leqslant 2).$$

The outcome in this case is also of the form (9.8), and Theorem 3 follows.

Note added in proof. Fouvry's treatment of (1.2), with $Q_1 = x^{2/3+\varepsilon}$, $Q_2 = x^{1-\varepsilon}$, and the estimate $x\exp(-c(\varepsilon)(\log x)^{1/2})$ on the right, has now been published: *Sur le problème des diviseurs de Titchmarsh*, J. Reine Angew. Math. 357 (1985), pp. 51–76 (Corollary 5). Moreover Fouvry states the asymptotic formula (1.1) with $E_k(x) \ll x\exp(-c(k)(\log x)^{1/2})$. (Ibid. Corollary 4. In fact the proof is to appear as joint work between Fouvry and Tenenbaum.)

References

[1] R. Balasubramanian, J. B. Conrey, and D. R. Heath-Brown, *Asymptotic mean square of the product of the Riemann Zeta-function and a Dirichlet polynomial*, J. Reine Angew. Math. 357 (1985), pp. 161–181.
[2] R. Bellman, *On some divisor sums associated with Diophantine equations*, Quart. J. Math. Oxford Ser. (2) 1 (1950), pp. 136–146.
[3] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, to appear.
[4] B. M. Bredikhin, *Binary additive problems of indefinite type*. III. *The additive problem of divisors*, Izv. Akad. Nauk SSSR Ser. Mat. 27 (1963), pp. 777–794.
[5] N. G. de Bruijn, *On the number of positive integers $\leqslant x$ and free of prime factors $> y$*, Indag. Math. 13 (1951), pp. 2–12.

[6] D. A. Burgess, *On character sums and L-series II*, Proc. London Math. Soc. (3) 13 (1963), pp. 524–536.
[7] J.-M. Deshouillers and H. Iwaniec, *An additive divisor problem*, ibid. (2) 26 (1982), pp. 1–14.
[8] T. Estermann, *Über die Darstellungen einer Zahl als Differenz von zwei Produkten*, J. Reine Angew. Math. 164 (1931), pp. 173–182.
[9] — *On Kloosterman's sum*, Mathematika 8 (1961), pp. 83–86.
[10] J. B. Friedlander and H. Iwaniec, *Incomplete Kloosterman sums and a divisor problem*, Annals of Math. 121 (1985), pp. 319–350.
[11] — — *The divisor problem for arithmetic progressions*, Acta Arith. 45 (1985), pp. 273–277.
[12] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford 1962.
[13] D. R. Heath-Brown, *Almost-primes in arithmetic progressions and short intervals*, Math. Proc. Camb. Phil. Soc. 83 (1978), pp. 357–375.
[14] — *The fourth power moment of the Riemann Zeta-function*, Proc. London Math. Soc. (3) 38 (1979), pp. 385–422.
[15] C. Hooley, *An asymptotic formula in the theory of numbers*, Proc. London Math. Soc. (3) 7 (1957), pp. 396–413.
[16] C. Hooley, *On the representations of a number as the sum of four cubes: I*, ibid. (3) 36 (1978), pp. 117–140.
[17] A. E. Ingham, *Some asymptotic formulae in the theory of numbers*, J. London Math. Soc. 2 (1927), pp. 202–208.
[18] A. F. Lavrik, *On the problem of divisors in segments of arithmetical progressions*, Dokl. Akad. Nauk SSSR 164 (1965), pp. 1232–1234.
[19] Yu. V. Linnik, *Dispersions of divisors and quadratic forms in progressions and certain binary additive problems*, ibid. 120 (1958), pp. 960–962.
[20] — *Dispersion problems and related binary additive problems*, Proc. Internat. Congress Math. 1958, Cambridge Univ. Press, New York, 1960, pp. 313–321.
[21] — *All large numbers are sums of a prime and two squares (A problem of Hardy and Littlewood)*. II, Mat. Sb. (N. S.) 53 (95) (1961), pp. 3–38.
[22] Y. Motohashi, *An induction principle for the generalization of Bombieri's prime number theorem*, Proc. Japan Acad. 52 (1976), pp. 273–275.
[23] — *On some additive divisor problems*, J. Math. Soc. Japan 28 (1976), pp. 772–784.
[24] — *An asymptotic series for an additive divisor problem*, Math. Zeitschr. 170 (1980), pp. 43–63.
[25] D. Redmond, *An asymptotic formula in the theory of numbers*, Math. Ann. 224 (1976), pp. 247–268.
[26] — *An asymptotic formula in the theory of numbers II*, ibid. 234 (1978), pp. 221–238.
[27] R. A. Smith, *On n-dimensional Kloosterman sums*, J. Number Theory 11 (1979), pp. 324–343.
[28] — *A generalized Kuznietsov's identity for Kloosterman sums*, C. R. Math. Rep. Acad. Sci. Canada 2 (1980), pp. 315–320.
[29] — *The generalized divisor problem over arithmetic progressions*, Math. Ann. 260 (1982), pp. 255–268.
[30] E. C. Titchmarsh, *The theory of the Riemann Zeta-function*, Oxford University Press, 1951.
[31] D. Wolke, *Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen, I*, Math. Ann. 202 (1973), pp. 1–25.
[32] M.-I. Yüh, *A divisor problem*, Acta Math. Sinica 8 (1958), pp. 496–506.

Received on 18.9.1984 (1456)

ACTA ARITHMETICA
XLVII (1986)

On the distribution of multiplicative arithmetical functions

by

Janos Galambos (Philadelphia, Pa.) and Peter Szüsz (Stony Brook, N. Y.)

1. Introduction. Throughout the paper, $g(n)$ denotes a strongly multiplicative real valued arithmetical function. That is, for coprime integers m and n, $g(mn) = g(m)g(n)$, and $g(p^k) = g(p)$ for all integers $k \geqslant 1$ and primes p.

We use the notation $A_N\{\ldots\}$ for the number of positive integers $n \leqslant N$ for which the property stated in the dotted space holds.

We say that an arithmetical function $h(n)$ has an *asymptotic distribution* $F(x)$ if, as $N \to +\infty$,

$$F_N(x) = (1/N) A_N\{h(n) \leqslant x\}$$

converges to $F(x)$ at each of its continuity points, and $F(x)$ is a proper distribution function (that is, $F(x)$ is nondecreasing, continuous from the right, and its limits at plus and minus infinity are one and zero, respectively).

Note that in this definition of convergence one may always disregard a denumerable set of points x.

Our interest is the existence of the asymptotic distribution of $g(n)$. Therefore, we may assume, and we do so in the remainder of the paper, that $g(n) \neq 0$. Indeed, if P denotes the collection of those primes p for which $g(p) = 0$, then one can easily see that, as $N \to +\infty$,

$$\lim (1/N) A_N\{g(n) = 0\} = 1$$

whenever

$$\sum_{p \in P} 1/p = +\infty.$$

On the other hand, if the series above is finite then $g(n)$ and $g^*(n)$ have asymptotic distribution concurrently, where $g^*(n)$ is the strongly multiplicative function defined as $g(n)$ if $g(n) \neq 0$, and $g^*(p) = 1$ if $g(p) = 0$.

Now, the combined results of Bakshtys [1] and Galambos [3] on a strongly multiplicative function $g(n)$ reads as follows.