ACTA ARITHMETICA
LXXIX.3 (1997)

A special case of Vinogradov's mean value theorem

by

R. C. Vaughan (London) and T. D. Wooley (Ann Arbor, Mich.)

In honorem J. W. S. Cassels annos LXXV nati

1. Introduction. In analytic number theory, estimates for the number, $J_{s,k}(P)$, of solutions of the system of equations

$$\sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \quad (1 \le j \le k) \tag{1.1}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$ are of great utility. This is perhaps best illustrated by the seminal works of Vinogradov from the first half of this century (see, for example, [1, 6]). Despite modern developments, such estimates remain the primary tool in establishing the best known results concerning the zero-free region of the Riemann zeta function, and the smallest number $\widetilde{G}(k)$ of variables for which the asymptotic formula holds in Waring's problem. When $s < \frac{1}{2}k(k+1)$ and P is large compared to s, it is widely conjectured that $J_{s,k}(P) \sim s!P^s$. This is an immediate consequence of Newton's formulae on the powers of the roots of a polynomial when $1 \le s \le k$, but when $s > k+1$ the latter asymptotic formula seems far beyond the grasp of current technology. Our primary purpose in this memoir is to establish in a rather sharp form the desired asymptotic formula in the case $s = k+1$.

When s is a natural number, let $T_s(P)$ denote the number of s-tuples $\mathbf{x}$ and $\mathbf{y}$ in which $1 \le x_i, y_i \le P$ $(1 \le i \le s)$, and the x_i are a permutation of the y_j, so that in particular, $T_s(P) = s!P^s + O_s(P^{s-1})$. In Section 2 we establish the strong form below of the asymptotic formula $J_{k+1,k}(P) \sim T_{k+1}(P)$, and in connection with this we define

$$\alpha_n = \min_{\substack{1 \le r \le n \\ r \in \mathbb{N}}} (r + n/r). \tag{1.2}$$

Research of the first author supported by an EPSRC Senior Fellowship.

Research of the second author supported by NSF grant DMS-9303505 and a Fellowship from the David and Lucile Packard Foundation.

THEOREM 1. *When* $k \ge 3$,

$$J_{k+1,k}(P) - T_{k+1}(P) \ll_{\varepsilon,k} P^{\alpha_{k+1}+\varepsilon}, \tag{1.3}$$

and consequently,

$$J_{k+1,k}(P) = T_{k+1}(P) + O_k(P^{\sqrt{4k+5}}). \tag{1.4}$$

For comparison, Hua [3, Lemma 5.4] provides the upper bound $J_{k+1,k}(P) \ll_k P^{k+1}(\log 2P)^{2^k-1}$, and very recently Vaughan and Wooley [5, Theorem 1.4] have obtained the bound (1.3) with α_{k+1} replaced by $\frac{1}{2}(k+5)$. The upper bound (1.3) is non-trivial for $k \ge 4$, and is superior to those obtained hitherto for $k \ge 6$. The methods developed here are susceptible to further small improvements, but for larger k they are of no great significance. However, it is possible to obtain (1.3) with the exponent α_{k+1} replaced by 33/8 and 23/5 when $k = 4$ and $k = 5$ respectively. We briefly outline this refinement at the end of Section 2.

For the sake of completeness we remark that in the cases $k = 2, 3$, Rogovskaya [4] and Vaughan and Wooley [5, Theorem 1.5], respectively, have established the estimates

$$J_{3,2}(P) = \frac{18}{\pi^2} P^3 \log P + O(P^3),$$

and, when P is large,

$$P^2 \log P \ll J_{4,3}(P) - T_4(P) \ll P^{10/3}(\log 2P)^{35}.$$

We note that the strength of the upper bound (1.3) is sufficient for applications to quasi-diagonal behaviour in the context of Vinogradov's mean value theorem (see [7, Lemmata 2.2 and 4.2] for details).

It seems worth remarking that when P is large, the existence of one non-trivial solution, $\mathbf{x}$, $\mathbf{y}$, of the system (1.1) implies the existence of $\gg_{\mathbf{x},\mathbf{y}} P^2$ non-trivial solutions $\mathbf{x}'$, $\mathbf{y}'$ with $1 \le x_i', y_i' \le P$ $(1 \le i \le s)$. This follows by taking

$$\mathbf{x}' = q\mathbf{x} + r \quad \text{and} \quad \mathbf{y}' = q\mathbf{y} + r,$$

with $1 \le q < P/\max\{x_i, y_i\}$ and $1 \le r \le P - q\max\{x_i, y_i\}$. Thus whenever $J_{s,k}(Q) - T_s(Q) > 0$ and $P \ge Q$, one has $J_{s,k}(P) - T_s(P) \gg_k P^2$. The current state of knowledge concerning the problem of Prouhet and Tarry (see Theorem 411 and the note on page 339 of [2]) therefore suffices to demonstrate that when $1 \le k \le 9$ and P is large, one has $J_{k+1,k}(P) - T_{k+1}(P) \gg_k P^2$. Whether or not there exist non-trivial solutions of the system (1.1) when $s = k+1$ and $k > 9$ remains open to speculation.

Denote by $S_k(P)$ the number of solutions of the system

$$\sum_{i=1}^{k}(x_i^j - y_i^j) = 0 \quad (j = 1, 2, \ldots, k-2 \text{ and } k), \tag{1.5}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$ $(1 \le i \le k)$. Similarities in the underlying algebraic structure enable us in Section 3 to adapt our methods successfully in order to estimate $S_k(P) - T_k(P)$.

THEOREM 2. *When* $k \ge 3$,

$$S_k(P) - T_k(P) \ll_{\varepsilon,k} P^{\alpha_k+\varepsilon}, \tag{1.6}$$

and consequently,

$$S_k(P) = T_k(P) + O_k(P^{\sqrt{4k+1}}). \tag{1.7}$$

In this situation, Hua [3, Lemma 5.2] provides the upper bound $S_k(P) \ll_k P^k(\log 2P)^{k(2^{k-1}-1)}$, and very recently Vaughan and Wooley [5, Theorem 1.3] have obtained the bound (1.6) with α_k replaced by $\frac{1}{2}(k+3)$. When k is large the superiority of (1.6) over the latter estimates is amply illuminated by (1.7). For the sake of completeness we remark that when $k = 3$ and P is large, Vaughan and Wooley [5, Theorem 1.2] have established the estimate

$$P^2(\log P)^5 \ll S_3(P) - 6P^3 \ll P^2(\log P)^5.$$

Our proof of Theorem 1 in Section 2 is elementary, and forms a natural extension to that used in [5, Section 9]. We use polynomial identities to bound the number of solutions of the system (1.1) counted by $J_{k+1,k}(P) - T_{k+1}(P)$ in terms of the number of solutions of a linear system subject to multiplicative constraints. The latter constraints lead, via extraction of common factors, to a system amenable to linear algebra and divisor function estimates. For smaller k one may refine the estimate (1.3) somewhat by better exploiting certain of the auxiliary variables which arise in our argument. We briefly sketch at the end of Section 2 how such refinements may be established. By a fortunate coincidence, a very similar system also arises through the use of polynomial identities in the treatment of the system (1.5), and thus in Section 3 we are able to establish Theorem 2 through a similar argument.

Throughout, $\ll$ and $\gg$ denote Vinogradov's well-known notation. Implicit constants in both the notations of Vinogradov and Landau will depend at most on ε, k and r. For the sake of concision, we make frequent use of vector notation. Thus, for example, we abbreviate $(c_1, \ldots, c_t)$ to $\mathbf{c}$. Finally, we write $(a_1, \ldots, a_s)$ for the greatest common divisor of $a_1, \ldots, a_s$, and we have been careful to ensure that any possible ambiguity can be resolved by the context.

2. The proof of Theorem 1. Let $U_k(P)$ denote the number of solutions of the system

$$\sum_{i=1}^{k+1}(x_i^j - y_i^j) = 0 \quad (1 \le j \le k) \tag{2.1}$$

with $1 \leq x_i, y_i \leq P$ $(1 \leq i \leq k+1)$, and satisfying the condition that $(x_1, \ldots, x_{k+1})$ is not a permutation of $(y_1, \ldots, y_{k+1})$. In this section we establish the estimate

$$U_k(P) \ll P^{\alpha_{k+1}+\varepsilon}, \tag{2.2}$$

from which the main conclusion of Theorem 1 follows immediately. Meanwhile, (1.4) follows by taking r to be the integer closest to $\sqrt{k+1}$ in the formula for α_{k+1}, and then applying some mundane analysis.

We start by observing that the polynomial $p(\xi; \mathbf{z})$, defined by

$$p(\xi; \mathbf{z}) = \prod_{i=1}^{k+1} (z_i - \xi) - \prod_{j=1}^{k+1} z_j,$$

considered as a polynomial in ξ, has coefficients which are symmetric polynomials in $z_1, \ldots, z_{k+1}$ of degree at most k. Thus for each solution $\mathbf{x}$, $\mathbf{y}$ of the system (2.1) counted by $U_k(P)$, one has $p(\xi; \mathbf{x}) = p(\xi; \mathbf{y})$. Consequently, for each s with $1 \leq s \leq k+1$,

$$\prod_{j=1}^{k+1} (y_j - x_s) = y_1 \ldots y_{k+1} - x_1 \ldots x_{k+1}, \tag{2.3}$$

whence

$$\prod_{i=1}^{k+1} (y_i - x_s) = \prod_{j=1}^{k+1} (y_j - x_t) \quad (1 \leq s < t \leq k+1). \tag{2.4}$$

Further, if $x_i = y_j$ for any i and j, then the equation (2.3) with $s = i$ implies that $x_1 \ldots x_{k+1} = y_1 \ldots y_{k+1}$. In combination with the equations (2.1), therefore, the use of elementary properties of symmetric polynomials leads to the conclusion that $(x_1, \ldots, x_{k+1})$ is a permutation of $(y_1, \ldots, y_{k+1})$, contradicting the assumption that $\mathbf{x}$, $\mathbf{y}$ is a solution counted by $U_k(P)$. We may thus suppose that $x_i = y_j$ for no i and j.

We divide the solutions $\mathbf{x}$, $\mathbf{y}$ of (2.1) counted by $U_k(P)$ into two types according to an integer parameter r with $1 < r \leq k+1$. Let $V_{1,r}(P)$ denote the number of such solutions in which there are fewer than r distinct values amongst the x_i, and let $V_{2,r}(P)$ denote the corresponding number of solutions in which there are at least r distinct values amongst the x_i. Then

$$U_k(P) = V_{1,r}(P) + V_{2,r}(P). \tag{2.5}$$

Consider first the solutions counted by $V_{1,r}(P)$. Fix any one of the $O(P^{r-1})$ possible choices for $\mathbf{x}$, and fix also one of the $O(P)$ available choices for y_1. By interchanging the rôles of $\mathbf{x}$ and $\mathbf{y}$ in (2.4), we obtain

$$\prod_{i=1}^{k+1} (x_i - y_s) = \prod_{j=1}^{k+1} (x_j - y_1) \quad (1 \leq s \leq k+1).$$

Thus, since each of the integers $x_j - y_1$ is fixed, when $2 \le s \le k+1$ each y_s is determined by a non-trivial polynomial. Consequently, there are $O(1)$ possible choices for $y_2, \dots, y_{k+1}$, whence

$$V_{1,r}(P) \ll P^r. \tag{2.6}$$

Next consider a solution $\mathbf{x}, \mathbf{y}$ counted by $V_{2,r}(P)$. By relabelling variables we may suppose that $x_1, \dots, x_r$ are distinct. Suppose temporarily that the integers y_1 and $y_i - x_s$ $(1 \le i \le k+1,\ 1 \le s \le r)$ are determined. Then plainly x_s is determined for $1 \le s \le r$, whence y_i is determined for $1 \le i \le k+1$. Moreover, when $r < s \le k+1$, the integers x_s may be determined from the polynomial equations (2.4) with $t = 1$. Then since there are $O(P)$ possible choices for y_1, we may conclude that given $y_i - x_s$ $(1 \le i \le k+1,\ 1 \le s \le r)$, there are $O(P)$ possible choices for $\mathbf{x}, \mathbf{y}$. Substituting $u_{0j} = x_j - y_1$ and $u_{ij} = y_{i+1} - x_j$ $(1 \le i \le k,\ 1 \le j \le r)$, we deduce from (2.4)–(2.6) that

$$U_k(P) \ll PW_r(P) + P^r, \tag{2.7}$$

where $W_r(P)$ denotes the number of solutions of the system

$$\prod_{i_1=0}^{k} u_{i_1 1} = \prod_{i_2=0}^{k} u_{i_2 2} = \ldots = \prod_{i_r=0}^{k} u_{i_r r}, \tag{2.8}$$

with

$$u_{01} + u_{i1} = u_{02} + u_{i2} = \ldots = u_{0r} + u_{ir} \quad (1 \le i \le k), \tag{2.9}$$

and

$$1 \le |u_{ij}| \le P \quad (0 \le i \le k,\ 1 \le j \le r), \tag{2.10}$$

and with the u_{0j} distinct for $1 \le j \le r$.

We now use the equations (2.8) to eliminate common factors amongst the u_{ij}. In order to make our description of this process precise, we record some notational devices. Let $\mathcal{I}$ denote the set of indices $\mathbf{i} = (i_1, \dots, i_r)$ with $0 \le i_m \le k$ $(1 \le m \le r)$. Define a map $\phi : \mathcal{I} \to [0, (k+1)^r) \cap \mathbb{Z}$ by

$$\phi(\mathbf{i}) = \sum_{m=1}^{r} i_m (k+1)^{m-1}.$$

Then ϕ is bijective, and we can define the successor, $\mathbf{i}+1$, of the index $\mathbf{i}$ by

$$\mathbf{i} + 1 = \phi^{-1}(\phi(\mathbf{i}) + 1).$$

When $h \in \mathbb{N}$, we define $\mathbf{i}+h$ inductively by $\mathbf{i}+(h+1) = (\mathbf{i}+h)+1$. Further, when $\mathbf{i} \in \mathcal{I}$, we write $\mathcal{J}(\mathbf{i})$ for the set of $\mathbf{j} \in \mathcal{I}$ such that for some $h \in \mathbb{N}$ one has $\mathbf{j} + h = \mathbf{i}$. We now define the integers $\alpha_{\mathbf{i}}$, with $\mathbf{i} \in \mathcal{I}$, as follows. We put $\alpha_{\mathbf{0}} = (u_{01}, u_{02}, \dots, u_{0r})$, and suppose at stage $\mathbf{i}$ that $\alpha_{\mathbf{j}}$ has been defined for

$\mathbf{j} \in \mathcal{J}(\mathbf{i})$. We then define $\alpha_{\mathbf{i}}$ by

$$\alpha_{\mathbf{i}} = \left(\frac{u_{i_1 1}}{\beta_{\mathbf{i}}^{(1)}}, \frac{u_{i_2 2}}{\beta_{\mathbf{i}}^{(2)}}, \ldots, \frac{u_{i_r r}}{\beta_{\mathbf{i}}^{(r)}} \right), \quad \text{where} \quad \beta_{\mathbf{i}}^{(m)} = \prod_{\substack{\mathbf{j} \in \mathcal{J}(\mathbf{i}) \\ j_m = i_m}} \alpha_{\mathbf{j}},$$

and here we adopt the convention that the empty product is unity. It follows that when $0 \le l \le k$ and $1 \le m \le r$, one has

$$u_{lm} = \prod_{\substack{\mathbf{j} \in \mathcal{I} \\ j_m = l}} \alpha_{\mathbf{j}}. \tag{2.11}$$

We now consider $\alpha_{\mathbf{i}}$, with $\mathbf{i} \in \mathcal{I}$, as variables, and for the sake of transparency write

$$\widetilde{\alpha}_{lm} = \prod_{\substack{\mathbf{j} \in \mathcal{I} \\ j_m = l}} \alpha_{\mathbf{j}}. \tag{2.12}$$

Then it follows from the discussion of the preceding paragraph that $W_r(P) \le X_r(P)$, where $X_r(P)$ denotes the number of solutions of the system

$$\widetilde{\alpha}_{01} + \widetilde{\alpha}_{i1} = \widetilde{\alpha}_{02} + \widetilde{\alpha}_{i2} = \ldots = \widetilde{\alpha}_{0r} + \widetilde{\alpha}_{ir} \quad (1 \le i \le k), \tag{2.13}$$

with the $\widetilde{\alpha}_{0j}$ distinct for $1 \le j \le r$, and with

$$1 \le |\widetilde{\alpha}_{ij}| \le P \quad (0 \le i \le k,\ 1 \le j \le r). \tag{2.14}$$

Thus by (2.7),

$$U_k(P) \ll P X_r(P) + P^r. \tag{2.15}$$

Having eliminated the multiplicative conditions inherent in our system, we are left to investigate the system (2.13). When $1 \le p \le r$, we write

$$A_p = \prod_{\substack{\mathbf{i} \in \mathcal{I} \\ i_l > i_p \ (l \ne p)}} \alpha_{\mathbf{i}}. \tag{2.16}$$

It follows easily that

$$\Big| \prod_{p=1}^{r} A_p \Big| \le \prod_{\mathbf{i} \in \mathcal{I}} |\alpha_{\mathbf{i}}| \le P^{k+1},$$

and thus in any solution $\boldsymbol{\alpha}$ counted by $X_r(P)$, there exists a p with $1 \le p \le r$ such that $|A_p| \le P^{(k+1)/r}$. Moreover, given l with $1 \le l \le r$, it follows from (2.13) and (2.14) that for each solution $\boldsymbol{\alpha}$ counted by $X_r(P)$, there exist integers L_j with $0 < |L_j| \le 2P$ such that when $1 \le j \le r$ and $j \ne l$,

$$\widetilde{\alpha}_{0l} - \widetilde{\alpha}_{0j} = -L_j, \quad \widetilde{\alpha}_{il} - \widetilde{\alpha}_{ij} = L_j \quad (1 \le i \le k).$$

By relabelling variables, therefore, we deduce that $X_r(P) \ll Y_r(P)$, where $Y_r(P)$ denotes the number of solutions of the system

$$\widetilde{\alpha}_{01} - \widetilde{\alpha}_{0j} = -L_j, \quad \widetilde{\alpha}_{i1} - \widetilde{\alpha}_{ij} = L_j \quad (2 \le j \le r,\ 1 \le i \le k), \tag{2.17}$$

with

$$1 \le |L_j| \le 2P \quad (2 \le j \le r), \tag{2.18}$$

and with the $\alpha_{\mathbf{i}}$ satisfying (2.14) and the inequality

$$|A_1| \le P^{(k+1)/r}, \tag{2.19}$$

where A_1 is defined by (2.16). Further, by (2.15),

$$U_k(P) \ll PY_r(P) + P^r. \tag{2.20}$$

We claim that when the variables $L_2, \ldots, L_r$, and $\alpha_{\mathbf{i}}$ with

$$\mathbf{i} \in \mathcal{I} \quad \text{and} \quad i_l > i_1 \quad (2 \le l \le r), \tag{2.21}$$

are fixed, then there are $O(P^\varepsilon)$ possible choices for the $\alpha_{\mathbf{i}}$ satisfying (2.14) and (2.17). If such is the case, then by combining (2.18)–(2.20) with standard estimates for the divisor function, we obtain $U_k(P) \ll P^{r+(k+1)/r+\varepsilon}$, and so the main conclusion of Theorem 1 follows.

It remains to establish the latter proposition, which we prove inductively as follows. For a fixed choice of the $\alpha_{\mathbf{i}}$ with $\mathbf{i}$ satisfying (2.21), we suppose at step t that there are $O(P^{t\varepsilon})$ possible choices for those variables $\alpha_{\mathbf{i}}$ for which $\mathbf{i}$ satisfies the condition that $i_l < t$ for some l with $1 \le l \le r$. Observe first that (2.17) implies that

$$\widetilde{\alpha}_{0j} = \widetilde{\alpha}_{01} + L_j \quad (2 \le j \le r). \tag{2.22}$$

We have supposed, moreover, that $L_2, \ldots, L_r$ are fixed and non-zero, and that the variables $\alpha_{\mathbf{i}}$ for which $i_1 = 0$ and $i_l > 0$ $(2 \le l \le r)$, are also fixed. Then by using standard estimates for the divisor function, it follows from (2.22) that there are $O(P^\varepsilon)$ possible choices for the $\alpha_{\mathbf{i}}$ for which $\mathbf{i}$ satisfies the condition that $i_l = 0$ for some l with $1 \le l \le r$. Thus our hypothesis holds when $t = 1$.

Suppose next that the hypothesis is satisfied for a $t \ge 1$, and consider a fixed one of the $O(P^{t\varepsilon})$ possible choices for the $\alpha_{\mathbf{i}}$ for which $i_l < t$ for some l with $1 \le l \le r$. It follows from (2.17) that

$$\widetilde{\alpha}_{tj} = \widetilde{\alpha}_{t1} - L_j \quad (2 \le j \le r). \tag{2.23}$$

Once again, $L_2, \ldots, L_r$ are fixed and non-zero. Moreover, if

$$i_1 = t \quad \text{and} \quad i_l \ne t \quad (2 \le l \le r), \tag{2.24}$$

then either some $i_l < t$, or else $i_l > t$ $(2 \le l \le r)$, and thus the variables $\alpha_{\mathbf{i}}$ for which $\mathbf{i}$ satisfies (2.24) may also be supposed fixed. Then by using standard estimates for the divisor function, it follows from (2.23) that there are $O(P^\varepsilon)$

possible choices for the variables $\alpha_{\mathbf{i}}$ for which $\mathbf{i}$ satisfies the condition that $i_l = t$ for some l with $1 \le l \le r$. Consequently, there are $O(P^{(t+1)\varepsilon})$ possible choices for the variables $\alpha_{\mathbf{i}}$ for which $\mathbf{i}$ satisfies the condition that $i_l \le t$ for some l with $1 \le l \le r$, and so the inductive hypothesis holds with t replaced by $t+1$. This completes the induction, and the proof of the main conclusion of Theorem 1.

By better exploiting the variables $\alpha_{\mathbf{i}}$ not occurring as factors of the A_p, it is possible to improve the upper bound (1.3) a little. Although for large k these improvements are not of great significance, for smaller k they may be of some interest. We sketch below one possible approach to obtaining such refinements.

We start by making an observation concerning the solutions counted by $X_r(P)$. Let $\mathcal{I}^+$ denote the set of indices $\mathbf{i} \in \mathcal{I}$ such that $i_l > 0$ $(1 \le l \le r)$, and let $\mathcal{I}^*$ denote the corresponding set of indices subject to the additional condition that for some p with $1 \le p \le r$, one has $i_l > i_p$ whenever $l \ne p$. Thus $\operatorname{card}(\mathcal{I}^+) = k^r$, and $\operatorname{card}(\mathcal{I}^*) = r\psi(k)$, where

$$\psi(k) = \sum_{i=1}^{k-1} i^{r-1} < k^r/r.$$

Observe that by considering changes of variables corresponding to permuting the indices i_l, for each fixed l, it follows with little difficulty from the argument of the proof of Theorem 1 that $W_r(P) \ll X_r(P)$, where $X_r(P)$ is defined as before, but now one may impose the additional condition

$$\prod_{\mathbf{i}\in\mathcal{I}^*} |\alpha_{\mathbf{i}}| \le \Big(\prod_{\mathbf{i}\in\mathcal{I}^+} |\alpha_{\mathbf{i}}|\Big)^{\operatorname{card}(\mathcal{I}^*)/\operatorname{card}(\mathcal{I}^+)}.$$

It follows that

$$\begin{aligned}\Big|\prod_{p=1}^{r} A_p\Big| &\le \Big(\prod_{p=1}^{r} \prod_{\substack{\mathbf{i}\in\mathcal{I}\\ i_p=0\\ i_l>0\ (l\ne p)}} |\alpha_{\mathbf{i}}|\Big)\Big(\prod_{\mathbf{i}\in\mathcal{I}^*} |\alpha_{\mathbf{i}}|\Big)\\ &\le \Big(\prod_{p=1}^{r} \prod_{\substack{\mathbf{i}\in\mathcal{I}\\ i_p=0}} |\alpha_{\mathbf{i}}|\Big)^{1-r\psi(k)/k^r} \Big(\prod_{\mathbf{i}\in\mathcal{I}} |\alpha_{\mathbf{i}}|\Big)^{r\psi(k)/k^r}\\ &\le (P^r)^{1-r\psi(k)/k^r}(P^{k+1})^{r\psi(k)/k^r}.\end{aligned}$$

Consequently, in any solution $\boldsymbol{\alpha}$ counted by $X_r(P)$, there exists a p with $1 \le p \le r$ such that

$$|A_p| \le P^{1+(k+1-r)\psi(k)/k^r}.$$

We may now prosecute the same argument as before, but now delivering the conclusion

$$U_k(P) \ll P^{\beta_k+\varepsilon},$$

where

$$\beta_k = \min_{\substack{2\le r\le k+1\\ r\in\mathbb{N}}} \left(r+1+\frac{k+1-r}{k^r}\sum_{i=1}^{k-1} i^{r-1}\right). \tag{2.25}$$

When $r=2$, the expression on the right-hand side of (2.25) yields

$$\beta_k \le \tfrac{1}{2}(k+4+1/k).$$

Thus when $k=4$, and when $k=5$, this refined argument with $r=2$ yields the sharpest bounds available to us, namely

$$U_4(P) \ll P^{33/8+\varepsilon} \quad \text{and} \quad U_5(P) \ll P^{23/5+\varepsilon}.$$

3. The proof of Theorem 2. Having illustrated our method in Section 2 we can afford to be brief in our proof of Theorem 2. We start by recording an observation from [5, Section 8]. From [5, (8.24)], together with the equation obtained by reversing the rôles of $\mathbf{x}$ and $\mathbf{y}$ in that equation, it follows that

$$S_k(P) - T_k(P) \ll R_k(kP), \tag{3.1}$$

where $R_k(Q)$ denotes the number of solutions of the system

$$x_v \prod_{i=1}^{k}(y_i - x_u) = x_u \prod_{j=1}^{k}(y_j - x_v) \quad (1\le u<v\le k), \tag{3.2}$$

$$y_v \prod_{i=1}^{k}(x_i - y_u) = y_u \prod_{j=1}^{k}(x_j - y_v) \quad (1\le u<v\le k), \tag{3.3}$$

with $1\le x_i, y_i \le Q$ $(1\le i\le k)$, and satisfying the condition that $x_i = y_j$ for no i and j.

We divide the solutions $\mathbf{x}$, $\mathbf{y}$ of (3.2) and (3.3) counted by $R_k(Q)$ into two types according to an integer parameter r with $1<r\le k$. Let $N_{1,r}(Q)$ denote the number of such solutions in which there are fewer than r distinct values amongst the x_i, and let $N_{2,r}(Q)$ denote the corresponding number of solutions in which there are at least r distinct values amongst the x_i. Then

$$R_k(Q) = N_{1,r}(Q) + N_{2,r}(Q). \tag{3.4}$$

Consider first the solutions counted by $N_{1,r}(Q)$. Fix any one of the $O(Q^{r-1})$ possible choices for $\mathbf{x}$, and fix also any one of the $O(Q)$ possible choices for y_1. Then since each of the integers $x_j - y_1$ $(1\le j\le k)$ is fixed, when $2\le u\le k$ each y_u is determined by the non-trivial polynomial equation (3.3) with $v = 1$. Consequently, there are $O(1)$ possible choices for $y_2, \ldots, y_k$, whence

$$N_{1,r}(Q) \ll Q^r. \tag{3.5}$$

Next consider a solution $\mathbf{x}$, $\mathbf{y}$ counted by $N_{2,r}(Q)$. By relabelling variables we may suppose that $x_1, \ldots, x_r$ are distinct. Suppose temporarily that the integers x_u and $y_i - x_u$ $(1 \le i \le k,\ 1 \le u \le r)$ are determined. Then plainly x_u and y_i are determined for $1 \le i \le k$ and $1 \le u \le r$. Moreover, when $r < u \le k$, the integers x_u may be determined from the polynomial equations (3.2) with $v = 1$. Then since there are $O(Q^r)$ possible choices for $x_1, \ldots, x_r$, we may conclude that given $y_i - x_u$ $(1 \le i \le k,\ 1 \le u \le r)$, there are $O(Q^r)$ possible choices for $\mathbf{x}$, $\mathbf{y}$. Substituting $u_{ij} = y_i - x_j$ $(1 \le i \le k,\ 1 \le j \le r)$, we deduce from (3.2)–(3.5) that

$$R_k(Q) \ll Q^r \max_{\mathbf{x}} M_r(Q; \mathbf{x}) + Q^r, \tag{3.6}$$

where the maximum is taken over $x_1, \ldots, x_r$ with

$$1 \le x_i \le Q \quad (1 \le i \le r),$$

and with the x_i distinct, and where $M_r(Q; \mathbf{x})$ denotes the number of solutions of the system (2.8) with

$$\begin{gathered} x_1 + u_{i1} = x_2 + u_{i2} = \ldots = x_r + u_{ir} \quad (1 \le i \le k), \\ 1 \le |u_{ij}| \le Q \quad (1 \le i \le k,\ 1 \le j \le r), \end{gathered} \tag{3.7}$$

and

$$u_{0i} = x_i^{-1} \prod_{j=1}^{r} x_j \quad (1 \le i \le r). \tag{3.8}$$

We may now extract common factors between the variables u_{ij} precisely as in Section 2. Thus, on recalling the notation of Section 2, we deduce that there are integers $\alpha_{\mathbf{i}}$ $(\mathbf{i} \in \mathcal{I})$ such that when $0 \le l \le k$ and $1 \le m \le r$, one has (2.11). We note that in view of (3.8), the u_{0i} are fixed. Thus, by making use of standard estimates for the divisor function, we deduce that there are $O(Q^\varepsilon)$ possible choices for the $\alpha_{\mathbf{j}}$ for which $j_m = 0$ for some m with $1 \le m \le r$. Treating the $\alpha_{\mathbf{i}}$ now as variables, and recalling the notation (2.12), we conclude that $M_r(Q; \mathbf{x}) \ll Q^\varepsilon K_r(Q; \mathbf{x})$, where $K_r(Q; \mathbf{x})$ denotes the number of solutions of the system

$$x_1 + \widetilde{\alpha}_{i1} = x_2 + \widetilde{\alpha}_{i2} = \ldots = x_r + \widetilde{\alpha}_{ir} \quad (1 \le i \le k), \tag{3.9}$$

with

$$1 \le |\widetilde{\alpha}_{ij}| \le Q \quad (1 \le i \le k,\ 1 \le j \le r), \tag{3.10}$$

and with the variables $\alpha_{\mathbf{i}}$, for which $i_m = 0$ for some m with $1 \le m \le r$, fixed.

We investigate the system (3.9) following the trail laid down in Section 2. When $1 \le p \le r$, we write $B_p = \prod_{\mathbf{i}}^* \alpha_{\mathbf{i}}$, where the product is over $\mathbf{i} \in \mathcal{I}$ for which $i_l > i_p$ $(l \ne p)$, and $i_l > 0$ $(1 \le l \le r)$. It follows that

$$\Big|\prod_{p=1}^{r} B_p\Big| \le \prod_{\substack{\mathbf{i}\in\mathcal{I} \\ i_l>0\ (1\le l\le r)}} |\alpha_{\mathbf{i}}| \le Q^k,$$

and thus in any solution $\boldsymbol{\alpha}$ counted by $K_r(Q;\mathbf{x})$, there exists a p with $1 \le p \le r$ such that $|B_p| \le Q^{k/r}$. By relabelling variables, we therefore deduce that

$$K_r(Q;\mathbf{x}) \ll I_r(Q;\mathbf{x}),$$

where $I_r(Q;\mathbf{x})$ denotes the number of solutions of the system

$$\widetilde{\alpha}_{i1} - \widetilde{\alpha}_{ij} = L_j \quad (2 \le j \le r,\ 1 \le i \le k), \tag{3.11}$$

with $L_j = x_j - x_1$ $(2 \le j \le r)$, and with the $\alpha_{\mathbf{i}}$ satisfying (3.10) and the inequality

$$|B_1| \le Q^{k/r}. \tag{3.12}$$

We claim that when the variables $\alpha_{\mathbf{i}}$, with $\mathbf{i}$ satisfying (2.21), are fixed, then there are $O(Q^{\varepsilon})$ possible choices for the $\alpha_{\mathbf{i}}$ satisfying (3.10) and (3.11). If such is the case, then by combining (3.12) with standard estimates for the divisor function, we obtain $I_r(Q;\mathbf{x}) \ll Q^{k/r+\varepsilon}$, whence by (3.6) we have $R_k(Q) \ll Q^{r+k/r+\varepsilon}$. The main conclusion of Theorem 2 follows immediately.

But the claimed conclusion may be established precisely as in the argument of the final paragraphs of Section 2, noting only that the $\alpha_{\mathbf{i}}$, for which $i_m = 0$ for some m with $1 \le m \le r$, are in this instance already fixed. This completes the proof of the main conclusion of Theorem 2, the estimate (1.7) following directly.

References

[1] J. W. S. Cassels and R. C. Vaughan, *Obituary: Ivan Matveevich Vinogradov*, Bull. London Math. Soc. 17 (1985), 584–600; see Biogr. Mem. Fellows Royal Society 31 (1985), 613–631.

[2] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., 4th reprint, Clarendon Press, Oxford, 1989.

[3] L.-K. Hua, *Additive Theory of Prime Numbers*, Amer. Math. Soc., Providence, 1965.

[4] N. N. Rogovskaya, *An asymptotic formula for the number of solutions of a system of equations*, in: Diophantine Approximations, Part II, Moskov. Gos. Univ., Moscow, 1986, 78–84 (in Russian).

[5] R. C. Vaughan and T. D. Wooley, *On a certain nonary cubic form and related equations*, Duke Math. J. 80 (1995), 669–735.

[6] I. M. Vinogradov, *Selected Works*, Springer, Berlin, 1985.

[7] T. D. Wooley, *Quasi-diagonal behaviour in certain mean value theorems of additive number theory*, J. Amer. Math. Soc. 7 (1994), 221–245.

Mathematics Department
Huxley Building
Imperial College
180 Queen's Gate
London, SW7 2BZ, U.K.
E-mail: rvaughan@ma.ic.ac.uk

Mathematics Department
University of Michigan
Ann Arbor, Michigan 48109-1003
U.S.A.
E-mail: wooley@math.lsa.umich.edu

Received on 11.2.1996 (2933)