Vinogradov systems with a slice off

Mathematika 63 (2017) 797–817
doi:10.1112/S0025579317000134

VINOGRADOV SYSTEMS WITH A SLICE OFF

JULIA BRANDES AND TREVOR D. WOOLEY

In memoriam Klaus Friedrich Roth

Abstract. Let $I_{s,k,r}(X)$ denote the number of integral solutions of the modified Vinogradov system of equations

$$x_1^j + \cdots + x_s^j = y_1^j + \cdots + y_s^j \quad (1 \leqslant j \leqslant k,\ j \neq r),$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$. By exploiting sharp estimates for an auxiliary mean value, we obtain bounds for $I_{s,k,r}(X)$ for $1 \leqslant r \leqslant k-1$. In particular, when $s, k \in \mathbb{N}$ satisfy $k \geqslant 3$ and $1 \leqslant s \leqslant (k^2-1)/2$, we establish the essentially diagonal behaviour $I_{s,k,1}(X) \ll X^{s+\varepsilon}$.

§1. *Introduction.* Systems of symmetric diagonal equations are, by orthogonality, intimately connected with mean values of exponential sums, and consequently find numerous applications in the analytic theory of numbers. In this paper we consider the number $I_{s,k,r}(X)$ of integral solutions of the system of equations

$$x_1^j + \cdots + x_s^j = y_1^j + \cdots + y_s^j \quad (1 \leqslant j \leqslant k,\ j \neq r), \tag{1.1}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$. This system is related to that of Vinogradov in which the equations (1.1) are augmented with the additional slice

$$x_1^r + \cdots + x_s^r = y_1^r + \cdots + y_s^r,$$

and may be viewed as a testing ground for progress on systems not of Vinogradov type. Relatives of such systems have been employed in work on the existence of rational points on systems of diagonal hypersurfaces as well as cognate paucity problems (see for example [**2**–**4**]). The main conjecture for the system (1.1) asserts that whenever $r, s, k \in \mathbb{N}$, $r < k$ and $\varepsilon > 0$, then

$$I_{s,k,r}(X) \ll X^{s+\varepsilon} + X^{2s-(k^2+k-2r)/2}. \tag{1.2}$$

Here and throughout, the constants implicit in Vinogradov's notation may depend on s, k, and ε. It is an easy exercise to establish a lower bound for $I_{s,k,r}(X)$ that shows the estimate (1.2) to be best possible, save that when $k > 2$

Received 23 April 2017.
MSC (2010): 11L15, 11D45, 11L07, 11P55 (primary).

one may expect to be able to take ε to be zero. Our focus in this memoir is the diagonal regime $I_{s,k,r}(X) \ll X^{s+\varepsilon}$, and this we address with some level of success in the case $r = 1$.

THEOREM 1.1. *Let $s, k \in \mathbb{N}$ satisfy $k \geqslant 3$ and $1 \leqslant s \leqslant (k^2-1)/2$. Then for each $\varepsilon > 0$, one has $I_{s,k,1}(X) \ll X^{s+\varepsilon}$.*

In view of the main conjecture (1.2), one would expect the conclusion of Theorem 1.1 to hold in the extended range $1 \leqslant s \leqslant (k^2+k-2)/2$. Previous work already in the literature falls far short of such ambitious assertions. Work of the second author from the early 1990s shows that $I_{s,k,r}(X) \ll X^{s+\varepsilon}$ only for $1 \leqslant s \leqslant k$ (see [**7**, Theorem 1]). Meanwhile, as a consequence of the second author's resolution of the main conjecture in the cubic case of Vinogradov's mean value theorem [**9**, Theorem 1.1], one has the bound $I_{s,3,1}(X) \ll X^{s+\varepsilon}$ for $1 \leqslant s \leqslant 4$ (see [**8**, Theorem 1.3]). This conclusion is matched by that of Theorem 1.1 above in the special case $k = 3$. The ideas underlying recent progress on Vinogradov's mean value theorem can, however, be brought to bear on the problem of estimating $I_{s,k,r}(X)$. Thus, it is a consequence of the second author's work on nested efficient congruencing [**10**, Corollary 1.2] that one has $I_{s,k,r}(X) \ll X^{s+\varepsilon}$ for $1 \leqslant s \leqslant k(k-1)/2$. Such a conclusion could also be established through methods related to those of Bourgain, Demeter and Guth [**1**], though the necessary details have yet to be elucidated in the published literature. Both the aforementioned estimate $I_{4,3,1}(X) \ll X^{4+\varepsilon}$, and the new bound reported in Theorem 1.1 go well beyond this work based on efficient congruencing and l^2-decoupling. Indeed, when $r = 1$ we achieve an estimate tantamount to square-root cancellation in a range of $2s$-th moments extending the interval $1 \leqslant s \leqslant k(k-1)/2$ roughly half way to the full conjectured range $1 \leqslant s \leqslant (k^2+k-2)/2$.

Our strategy for proving Theorem 1.1 is based on the proof of the estimate $I_{4,3,1}(X) \ll X^{4+\varepsilon}$ in [**8**, Theorem 1.3], though it is flexible enough to deliver estimates for the mean value $I_{s,k,r}(X)$ with $r \geqslant 1$, as we now outline. For each integral solution $\mathbf{x}, \mathbf{y}$ of the system (1.1) with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$, one has the additional equation

$$\sum_{i=1}^{s} (x_i^r - y_i^r) = h, \tag{1.3}$$

for some integer h with $|h| \leqslant sX^r$. We seek to count all such solutions with h thus constrained. For each integer z with $1 \leqslant z \leqslant X$, we find that whenever $\mathbf{x}, \mathbf{y}, h$ satisfy (1.1) and (1.3), then one has

$$\sum_{i=1}^{s} (u_i^j - v_i^j) = \omega_j h z^{j-r} \quad (1 \leqslant j \leqslant k), \tag{1.4}$$

where ω_j is 0 for $1 \leqslant j < r$ and $\binom{j}{r}$ for $r \leqslant j \leqslant k$, and in which we write $u_i = x_i + z$ and $v_i = y_i + z$ $(1 \leqslant i \leqslant s)$. If we are able to obtain significant cancellation

in the number of solutions of the system (1.4), now with $\mathbf{u}$, $\mathbf{v}$ constrained only by the conditions $1 \leqslant u_i, v_i \leqslant 2X$ $(1 \leqslant i \leqslant s)$, then the overcounting by z may be reversed to show that there is significant cancellation in the system (1.1) underpinning the mean value $I_{s,k,r}(X)$. This brings us to consider the number of solutions of the system

$$\sum_{i=1}^{2t} h_i z_i^{j-r} = 0 \quad (r \leqslant j \leqslant k), \tag{1.5}$$

with $|h_i| \leqslant sX^r$ and $1 \leqslant z_i \leqslant X$ $(1 \leqslant i \leqslant 2t)$. This auxiliary mean value may be analysed through the use of multiplicative polynomial identities engineered using ideas related to those employed in [7].

The reader may be interested to learn the consequences of this strategy when r is permitted to exceed 1. The conclusion of Theorem 1.1 is in fact a special case of a more general result which, for $r \geqslant 2$, unfortunately fails to deliver diagonal behaviour.

THEOREM 1.2. *Let $r, s, k \in \mathbb{N}$ satisfy $k > r \geqslant 1$ and*

$$1 \leqslant s \leqslant \frac{k(k+1)}{2} - \frac{k(k+1) - r(r-1)}{4\kappa},$$

where κ is an integer satisfying $1 \leqslant \kappa \leqslant (k-r+2)/2$. Then for each $\varepsilon > 0$, one has

$$I_{s,k,r}(X) \ll X^{s+(r-1)(1-1/(2\kappa))+\varepsilon}.$$

When $r > 1$, although we do not achieve diagonal behaviour, we do improve on the estimate $I_{s,k,r}(X) \ll X^{s+r+\varepsilon}$ that follows for $1 \leqslant s \leqslant k(k+1)/2$ from the main conjecture in Vinogradov's mean value theorem via the triangle inequality. When $r > 2$, the bound for $I_{s,k,r}(X)$ obtained in the conclusion of Theorem 1.2 remains weaker than what could be obtained by interpolating between the aforementioned bounds $I_{s,k,r}(X) \ll X^{s+\varepsilon}$ $(1 \leqslant s \leqslant k(k-1)/2)$ and $I_{s,k,r}(X) \ll X^{s+r+\varepsilon}$ $(1 \leqslant s \leqslant k(k+1)/2)$. The former bound is, however, yet to enter the published literature.

In §2 we speculate concerning what bounds might hold for a class of mean values associated with the system (1.5). In particular, should a suitable analogue of the main conjecture hold for this auxiliary mean value, then the conclusion of Theorem 1.2 would be valid with a value of κ now permitted to be as large as

$$\kappa = \left\lfloor \frac{(k-r)(k+r+1)+2}{4} \right\rfloor.$$

We refer the reader to Conjecture 2.2 below for precise details, and we note in particular the constraint (2.4). When $r = 1$ and $k \equiv 0$ or 3 modulo 4, this would conditionally establish the estimate $I_{s,k,1}(X) \ll X^{s+\varepsilon}$ in the range $1 \leqslant s \leqslant (k^2+k-2)/2$, and hence the main conjecture (1.2) in full for these cases.

When $r > 1$, this conditional result establishes a bound slightly stronger than $I_{s,k,r}(X) \ll X^{s+r-1}$ when $1 \leqslant s \leqslant (k^2+k-4)/2$, which seems quite respectable.

We begin in §2 by announcing an auxiliary mean value estimate generalizing that associated with the system (1.5). This we establish in §§3–6, obtaining a polynomial identity in §3 of appropriate multiplicative type, establishing a lemma to count integral points on auxiliary equations in §4, and classifying solutions according to the vanishing of certain sets of coefficients in §5. In §6 we combine these ideas with a divisor estimate to complete the proof of this auxiliary estimate. Finally, in §7, we provide the details of the argument sketched above which establishes Theorems 1.1 and 1.2.

Throughout, the letters r, s and k will denote positive integers with $r < k$, and ε will denote a sufficiently small positive number. We take X to be a large positive number depending at most on s, k and ε. The implicit constants in the notations of Landau and Vinogradov will depend at most on s, k, ε, and the coefficients of fixed polynomials that we introduce. We adopt the following convention concerning the number ε. Whenever ε appears in a statement, we assert that the statement holds for each $\varepsilon > 0$. Finally, we employ the non-standard convention that whenever $G : [0,1)^k \to \mathbb{C}$ is integrable, then

$$\oint G(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \int_{[0,1)^k} G(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}.$$

Here and elsewhere, we use vector notation liberally in a manner that is easily discerned from the context.

§2. *An auxiliary mean value.* Our focus in this section and those following lies on the system of equations (1.5), since this is intimately connected with the Vinogradov system missing the slice of degree r. Since little additional effort is required to proceed in wider generality, we establish a conclusion in which the monomials z^{j-r} $(r \leqslant j \leqslant k)$ in (1.5) are replaced by independent polynomials $f_j(z)$. We begin in this section by introducing the notation required to state our main auxiliary result.

Let t be a natural number. When $1 \leqslant j \leqslant t$, consider a non-zero polynomial $f_j \in \mathbb{Z}[x]$ of degree k_j. We say that $\mathbf{f} = (f_1, \dots, f_t)$ is *well-conditioned* when the degrees of the polynomials f_j satisfy the condition

$$0 \leqslant k_t < k_{t-1} < \cdots < k_1, \tag{2.1}$$

and there is no positive integer z for which $f_1(z) = \cdots = f_t(z) = 0$.

Let X be a positive number sufficiently large in terms of t, $\mathbf{k}$ and the coefficients of f. We define the exponential sum $\mathfrak{g}(\boldsymbol{\alpha}; X)$ by putting

$$\mathfrak{g}(\boldsymbol{\alpha}; X) = \sum_{|h| \leqslant X^r} \sum_{1 \leqslant z \leqslant X} e(h(f_1(z)\alpha_1 + \cdots + f_t(z)\alpha_t)).$$

Finally, we define the mean value

$$A_{s,r}(X; \mathbf{f}) = \oint |\mathfrak{g}(\boldsymbol{\alpha}; X)|^{2s}\, d\boldsymbol{\alpha}. \tag{2.2}$$

By orthogonality, the mean value $A_{s,r}(X;\mathbf{f})$ counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{2s} h_i f_j(z_i) = 0 \quad (1 \leqslant j \leqslant t), \tag{2.3}$$

with $|h_i| \leqslant X^r$ and $1 \leqslant z_i \leqslant X$ $(1 \leqslant i \leqslant 2s)$. The system (2.3) plainly generalizes (1.5). Our immediate goal is to establish the mean value estimate recorded in the following theorem.

THEOREM 2.1. *Let r, s and t be natural numbers with $t \geqslant 2s-1$. Then whenever $\mathbf{f}$ is a well-conditioned t-tuple of polynomials having integral coefficients, one has $A_{s,r}(X;\mathbf{f}) \ll X^{r(2s-1)+1+\varepsilon}$.*

Note that when $r = 1$, the conclusion of this theorem is tantamount to exhibiting square-root cancellation in the mean value (2.2), so is essentially best possible. Indeed, even in situations wherein $r > 1$, the solutions of (2.3) in which $z_1 = z_2 = \cdots = z_{2s}$ make a contribution to $A_{s,r}(X;\mathbf{f})$ of order $X \cdot (X^r)^{2s-1}$, and so the conclusion of Theorem 2.1 is again essentially best possible. Henceforth, we restrict our attention to the situation described by the hypotheses of Theorem 2.1. Thus, we may suppose that $t \geqslant 2s-1$, and that $\mathbf{f}$ is a well-conditioned t-tuple of polynomials $f_j \in \mathbb{Z}[x]$ with $\deg(f_j) = k_j \geqslant 0$.

It seems not unreasonable to speculate that the estimate claimed in the statement of Theorem 2.1 should remain valid when s is significantly larger than $(t+1)/2$. The total number of choices for the $2s$ pairs of variables h_i, z_i occurring in the system (2.3) is of order $(X^{r+1})^{2s}$. Meanwhile, the t equations comprising (2.3) involve monomials having typical size of asymptotic order X^{r+k_j} $(1 \leqslant j \leqslant t)$. Thus, for large s, one should expect that

$$A_{s,r}(X;\mathbf{f}) \ll (X^r)^{2s-t} X^{2s-k_1-\cdots-k_t}.$$

Keeping in mind the diagonal solutions discussed above, one is led to the following conjecture.

CONJECTURE 2.2. *Let r, s and t be natural numbers, and suppose that $\mathbf{f}$ is a well-conditioned t-tuple of polynomials having integral coefficients, with $\deg(f_j) = k_j$ $(1 \leqslant j \leqslant t)$. Then one has*

$$A_{s,r}(X;\mathbf{f}) \ll X^{\varepsilon}(X^{r(2s-1)+1} + X^{2s(r+1)-tr-k_1-\cdots-k_t}).$$

In the special case in which $t = k-r+1$ and $k_j = j-1$ $(1 \leqslant j \leqslant t)$ relevant to the system (1.5), this conjectural bound reads

$$A_{s,r}(X;\mathbf{f}) \ll X^{\varepsilon}(X^{r(2s-1)+1} + X^{2s(r+1)-(k+r)(k-r+1)/2}).$$

In such circumstances, one finds that

$$A_{s,r}(X;\mathbf{f}) \ll X^{r(2s-1)+1+\varepsilon},$$

provided that s is an integer satisfying

$$4s \leqslant (k-r)(k+r+1)+2. \tag{2.4}$$

We finish this section by remarking that the estimate $A_{s,r}(X;\mathbf{f}) \ll X^{2rs}$ is fairly easily established when $t \geqslant 2s$, a stronger condition than that imposed in Theorem 2.1, as we now sketch. We may suppose that $t = 2s$ without loss, and in such circumstances the equations (2.3) may be interpreted as a system of $2s$ linear equations in $2s$ variables h_i. There are $O(X^{2s})$ choices for the variables z_i, contributing $O(X^{2s})$ to $A_{s,r}(X;\mathbf{f})$ from those solutions with $\mathbf{h} = \mathbf{0}$. Meanwhile, if $\mathbf{h} \neq \mathbf{0}$ one must have

$$\det(f_j(z_i))_{1\leqslant i,j\leqslant 2s} = 0. \tag{2.5}$$

By applying the theory of Schur functions (see Macdonald [5, Ch. I]) as in the proof of [6, Lemma 1], one finds that

$$\det(f_j(z_i))_{1\leqslant i,j\leqslant 2s} = \Theta(\mathbf{z};\mathbf{f}) \prod_{1\leqslant i<j\leqslant 2s} (z_i - z_j),$$

where the polynomial $\Theta(\mathbf{z};\mathbf{f})$ is asymptotically definite, meaning that whenever z_i is sufficiently large for $1 \leqslant i \leqslant 2s$, then $|\Theta(\mathbf{z};\mathbf{f})| \geqslant 1$.

The contribution to $A_{s,r}(X;\mathbf{f})$ arising from the solutions of (2.3) with $z_i = O(1)$, for some index i, is $O((X^r)^{2s})$. For if $z_i = O(1)$, then we may fix h_i, and interpret the system as a mean value of exponential sums, applying the triangle inequality. An application of Hölder's inequality reveals that if such solutions dominate, then

$$A_{s,r}(X;\mathbf{f}) \ll X^r \oint |\mathfrak{g}(\boldsymbol{\alpha};X)|^{2s-1}\, d\boldsymbol{\alpha} \ll X^r A_{s,r}(X;\mathbf{f})^{1-1/(2s)},$$

and the desired conclusion follows. Meanwhile, if z_i is sufficiently large for each index i, then $|\Theta(\mathbf{z};\mathbf{f})|$ is strictly positive and hence (2.5) can hold only when $z_i = z_j$ for some indices i and j with $1 \leqslant i < j \leqslant 2s$. By symmetry we may suppose that $i = 2s-1$ and $j = 2s$, and then we obtain from (2.3) the new system of equations

$$\sum_{i=1}^{2s-1} h_i' f_j(z_i) = 0 \quad (1 \leqslant j \leqslant 2s),$$

with $h_i' = h_i$ $(1 \leqslant i \leqslant 2s-2)$ and $h_{2s-1}' = h_{2s-1} + h_{2s}$. This new system is of similar shape to (2.3), and we may apply an obvious inductive argument to bound the number of its solutions. Here, we keep in mind that given h_{2s-1}', there are $O(X^r)$ possible choices for h_{2s-1} and h_{2s}. Thus we conclude that if this second class of solutions dominates, then one has

$$A_{s,r}(X;\mathbf{f}) \ll X^r \cdot X^{r(2s-1)} \ll X^{2rs}.$$

This completes our sketch of the proof that when $t = 2s$, the total number of solutions counted by $A_{s,r}(X;\mathbf{f})$ is $O(X^{2rs})$. The reader will likely have no difficulty in refining this argument to deliver the conclusion of Theorem 2.1 when $t = 2s$.

§3. *A polynomial identity.* The structure of the polynomials $hf_j(z)$ underlying the mean value $A_{s,r}(X;\mathbf{f})$ permits polynomial identities to be constructed of utility in constraining solutions of the underlying system of equations (2.3). In this section we construct such identities.

For the sake of concision, when n is a natural number and $1 \leqslant j \leqslant t$, we define the polynomial $\sigma_{j,n} = \sigma_{j,n}(\mathbf{z};\mathbf{h})$ by putting

$$\sigma_{j,n}(\mathbf{z};\mathbf{h}) = h_1 f_j(z_1) + \cdots + h_n f_j(z_n).$$

LEMMA 3.1. *Suppose that $n \geqslant 1$ and that $\mathbf{f} = (f_1, \dots, f_{2n+1})$ is a well-conditioned $(2n+1)$-tuple of polynomials having integral coefficients. Then there exists a polynomial $\Psi_n(\mathbf{w}) \in \mathbb{Z}[w_1, \dots, w_{2n+1}]$ whose total degree and coefficients depend at most on n, $\mathbf{k}$ and the coefficients of $\mathbf{f}$, having the property that*

$$\Psi_n(\sigma_{1,n}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n}(\mathbf{z};\mathbf{h})) = 0 \tag{3.1}$$

identically in $\mathbf{z}$ and $\mathbf{h}$, and yet

$$\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h})) \neq 0. \tag{3.2}$$

Proof. We apply an argument similar to that of [7, Lemma 1] based on a consideration of transcendence degrees. Let $K = \mathbb{Q}(\sigma_{1,n}, \dots, \sigma_{2n+1,n})$. Then $K \subseteq \mathbb{Q}(z_1, \dots, z_n, h_1, \dots, h_n)$, so that K has transcendence degree at most $2n$ over $\mathbb{Q}$. It follows that the $2n+1$ polynomials $\sigma_{1,n}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n}(\mathbf{z};\mathbf{h})$ cannot be algebraically independent over $\mathbb{Q}$. Consequently, there exists a non-zero polynomial $\Psi_n \in \mathbb{Z}[w_1, \dots, w_{2n+1}]$ satisfying the property (3.1).

It remains now only to confirm that a choice may be made for this non-trivial polynomial Ψ_n in such a manner that property (3.2) also holds. In order to establish this claim, we begin by considering any non-zero polynomial Ψ_n of smallest total degree satisfying (3.1). Suppose, if possible, that $\Psi_n(\sigma_{1,n+1}, \dots, \sigma_{2n+1,n+1})$ is also identically zero. Then the polynomials

$$\frac{\partial}{\partial z_i}\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h})) \tag{3.3}$$

and

$$\frac{\partial}{\partial h_i}\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h})) \tag{3.4}$$

must also be identically zero for $1 \leqslant i \leqslant n+1$. Write

$$u_j = \frac{\partial}{\partial w_j}\Psi_n(w_1, \dots, w_{2n+1}) \quad (1 \leqslant j \leqslant 2n+1),$$

in which we evaluate the right-hand side at $w_i = \sigma_{i,n+1}(\mathbf{z};\mathbf{h})$ $(1 \leqslant i \leqslant 2n+1)$. Then it follows from an application of the chain rule that the vanishing of the polynomials (3.3) and (3.4) implies the relations

$$\sum_{j=1}^{2n+1} h_i f_j'(z_i) u_j = 0 \quad (1 \leqslant i \leqslant n) \tag{3.5}$$

and

$$\sum_{j=1}^{2n+1} f_j(z_i)u_j = 0 \quad (1 \leqslant i \leqslant n+1). \tag{3.6}$$

Notice here that we have deliberately omitted the index $i = n + 1$ from the relations (3.5), since this is superfluous to our needs.

In order to encode the coefficient matrix associated with the system of linear equations in $\mathbf{u}$ described by the relations (3.5) and (3.6), we introduce a block matrix as follows. We define the $n \times (2n + 1)$ matrix

$$A_n = (h_i f_j'(z_i))_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant 2n+1}}$$

and the $(n + 1) \times (2n + 1)$ matrix

$$B_n = (f_j(z_i))_{\substack{1 \leqslant i \leqslant n+1 \\ 1 \leqslant j \leqslant 2n+1}},$$

and then define the $(2n + 1) \times (2n + 1)$ matrix D_n via the block decomposition

$$D_n = \begin{pmatrix} A_n \\ B_n \end{pmatrix}.$$

We claim that $\det(D_n)$ is not identically zero as a polynomial. The confirmation of this fact we defer to the end of this proof.

With the assumption $\det(D_n) \neq 0$ in hand, one sees that the system of equations (3.5) and (3.6) has only the trivial solution $\mathbf{u} = \mathbf{0}$ over K. However, since $\Psi_n(\mathbf{w})$ is a non-constant polynomial, at least one of the derivatives

$$\frac{\partial}{\partial w_j}\Psi_n(w_1, \dots, w_{2n+1}) \quad (1 \leqslant j \leqslant 2n+1)$$

must be non-zero. Suppose that the partial derivative with respect to w_J is non-zero. Then there exists a non-constant polynomial

$$\Psi_n^*(\mathbf{w}) = \frac{\partial}{\partial w_J}\Psi_n(w_1, \dots, w_{2n+1})$$

having the property that, since $u_J = 0$, one has

$$\Psi_n^*(\sigma_{1,n}(\mathbf{z}; \mathbf{h}), \dots, \sigma_{2n+1,n}(\mathbf{z}; \mathbf{h})) = 0.$$

But the total degree of Ψ_n^* is strictly smaller than that of Ψ_n, contradicting our hypothesis that Ψ_n has minimal total degree. We are therefore forced to conclude that the relation (3.2) does indeed hold.

We now turn to the problem of justifying our assumption that $\det(D_n) \neq 0$. We prove this assertion for any well-conditioned $(2n + 1)$-tuple of polynomials $\mathbf{f}$ by induction on n. Observe first that when $n = 0$, one has $\det(D_0) = f_1(z_1)$. Since $f_1(z)$ is not identically zero, it follows that $\det(D_0) \neq 0$, confirming the

base case of our inductive hypothesis. We suppose next that $n \geqslant 1$ and that $\det(D_{n-1}) \neq 0$ for all well-conditioned $(2n-1)$-tuples of polynomials $\mathbf{f}$, and we seek to show that $\det(D_n) \neq 0$.

Denote by $\mathcal{I}$ the set of all 2-element subsets $\mathfrak{a} = \{a_1, a_2\}$ contained in $\mathcal{N} = \{1, 2, \dots, 2n+1\}$. When $\mathfrak{a} = \{a_1, a_2\} \in \mathcal{I}$, we define the matrices

$$A(\mathfrak{a}) = (h_i f_j'(z_i))_{\substack{2 \leqslant i \leqslant n \\ j \in \mathcal{N} \setminus \mathfrak{a}}} \quad \text{and} \quad B(\mathfrak{a}) = (f_j(z_i))_{\substack{2 \leqslant i \leqslant n+1 \\ j \in \mathcal{N} \setminus \mathfrak{a}}}.$$

Equipped with this notation, we define the minors

$$U(\mathfrak{a}) = \det \begin{pmatrix} h_1 f_{a_1}'(z_1) & h_1 f_{a_2}'(z_1) \\ f_{a_1}(z_1) & f_{a_2}(z_1) \end{pmatrix} \quad \text{and} \quad V(\mathfrak{a}) = \det \begin{pmatrix} A(\mathfrak{a}) \\ B(\mathfrak{a}) \end{pmatrix}.$$

In this way, we discern that for appropriate choices of $\sigma(\mathfrak{a}) \in \{1, -1\}$, the precise nature of which need not detain us, one has

$$\det(D_n) = \sum_{\mathfrak{a} \in \mathcal{I}} \sigma(\mathfrak{a}) U(\mathfrak{a}) V(\mathfrak{a}).$$

By relabelling indices and then applying the inductive hypothesis for the $(2n-1)$-tuple $(f_3, \dots, f_{2n+1})$, it is apparent that $V(\{1,2\})$ is not identically zero. In view of (2.1), moreover, if the leading coefficients of f_1 and f_2 are c_1 and c_2, respectively, then the leading monomial in $U(\{1,2\})$ is

$$(k_1 - k_2) c_1 c_2 h_1 z_1^{k_1+k_2-1} \neq 0.$$

It follows that $U(\{1,2\})$ is also not identically zero. Also, since no other minor of the shape $U(\mathfrak{a})$, with $\mathfrak{a} \in \mathcal{I}$ and $\mathfrak{a} \neq \{1,2\}$, has degree $k_1 + k_2 - 1$ or greater with respect to z_1, we deduce that $\det(D_n)$ is not identically zero. This confirms the inductive hypothesis for the index n and completes the proof of our claim for all n. □

Henceforth, when $n \geqslant 1$, we consider a fixed choice for the polynomials $\Psi_n(\mathbf{w}) \in \mathbb{Z}[w_1, \dots, w_{2n+1}]$, of minimal total degree, satisfying the conditions (3.1) and (3.2). It is useful to extend this definition by taking $\Psi_0(w) = w$. We may now establish our fundamental polynomial identity.

LEMMA 3.2. *Suppose that $n \geqslant 0$ and the $(2n+1)$-tuple $\mathbf{f} = (f_1, \dots, f_{2n+1})$ of polynomials in $\mathbb{Z}[x]$ is well-conditioned. Then there exists a non-zero polynomial $\Phi_n(\mathbf{z}; \mathbf{h}) \in \mathbb{Z}[\mathbf{z}, \mathbf{h}]$ with the property that*

$$\begin{aligned} &\Psi_n(\sigma_{1,n+1}(\mathbf{z}; \mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z}; \mathbf{h})) \\ &\quad = \Phi_n(\mathbf{z}; \mathbf{h}) h_1 \cdots h_{n+1} \prod_{1 \leqslant i < j \leqslant n+1} (z_i - z_j). \end{aligned} \tag{3.7}$$

Proof. In the case $n = 0$, the product over i and j on the right-hand side of (3.7) is empty, and by convention we take this empty product to be 1. In such

circumstances, we see that $\Psi_0(\sigma_{1,1}(z_1;h_1)) = h_1 f_1(z_1)$, and the conclusion of the lemma is immediate.

Suppose next that $n \geqslant 1$. Then, when $h_{n+1} = 0$, we have

$$\sigma_{i,n+1}(\mathbf{z};\mathbf{h}) = \sigma_{i,n}(\mathbf{z};\mathbf{h}) \quad (1 \leqslant i \leqslant 2n+1),$$

and thus we deduce from property (3.1) of Lemma 3.1 that in this situation, one has

$$\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h})) = 0. \tag{3.8}$$

It follows that h_{n+1} divides $\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h}))$, and by symmetry the same holds for $h_1, \dots, h_n$. Meanwhile, when $z_n = z_{n+1}$, we have

$$\sigma_{i,n+1}(\mathbf{z};\mathbf{h}) = \sigma_{i,n}(\mathbf{z};h_1,\dots,h_{n-1},h_n+h_{n+1}),$$

and again we find from property (3.1) of Lemma 3.1 that in this special situation one has (3.8). We thus conclude that $z_n - z_{n+1}$ divides the polynomial $\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h}))$, and by symmetry the same holds for $z_i - z_j$ whenever $1 \leqslant i < j \leqslant n+1$.

In light of these observations, it is apparent that

$$\Psi_n(\sigma_{1,n+1}(\mathbf{z};\mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z};\mathbf{h}))$$

is divisible by

$$h_1 \cdots h_{n+1} \prod_{1 \leqslant i < j \leqslant n+1} (z_i - z_j).$$

The quotient of the former polynomial by the latter cannot be zero, since this former polynomial is non-zero, by virtue of property (3.2) of Lemma 3.1. We therefore conclude that a non-zero polynomial $\Phi_n(\mathbf{z};\mathbf{h}) \in \mathbb{Z}[\mathbf{z},\mathbf{h}]$ does indeed exist satisfying (3.7). This completes the proof of the lemma. □

It seems quite likely that additional potentially useful structure might be extracted from the polynomial identities provided by Lemma 3.2. For example, the relation

$$(h_1+h_2)(h_1z_1^2+h_2z_2^2) - (h_1z_1+h_2z_2)^2 = h_1h_2(z_1-z_2)^2$$

plays a prominent role in the proof of [**8**, Lemma 2.1]. Meanwhile, writing

$$s_j = h_1z_1^j + h_2z_2^j + h_3z_3^j \quad (0 \leqslant j \leqslant 4),$$

one may verify that

$$\begin{aligned}(s_1s_4 - s_2s_3)^2(s_0s_2 - s_1^2) - (s_0s_4 - s_2^2)(s_1s_3 - s_2^2)^2 \\ = h_1h_2h_3(z_1-z_2)^2(z_2-z_3)^2(z_3-z_1)^2F_{6,3}(\mathbf{z};\mathbf{h}),\end{aligned}$$

for a suitable bihomogeneous polynomial $F_{6,3}(\mathbf{z};\mathbf{h}) \in \mathbb{Z}[\mathbf{z},\mathbf{h}]$, of degree 6 with respect to $\mathbf{z}$ and degree 3 with respect to $\mathbf{h}$.

§4. *Counting integral solutions pairwise.* The polynomial identity furnished by Lemma 3.2 is of multiplicative type, and particularly powerful when $\Psi_n(\sigma_{1,n+1}, \dots, \sigma_{2n+1,n+1})$ is non-zero for a fixed integral choice of $\mathbf{z}$ and $\mathbf{h}$, for then we may exploit elementary estimates for the divisor function. However, it is possible that the latter quantity vanishes. This brings us into the domain of the classification of solutions according to the vanishing or non-vanishing of various intermediate coefficients. We begin with an elementary lemma concerning polynomials in two variables similar to [7, Lemma 2], the proof of which we include for the sake of completeness.

LEMMA 4.1. *Let $\psi \in \mathbb{Z}[z, h]$ be a non-trivial polynomial of total degree d. Then the number of integral solutions of the equation $\psi(z, h) = 0$ with $|z| \leqslant X$ and $|h| \leqslant X^r$ is at most $2d(2X^r + 1)$.*

Proof. We may write $\psi(z, h) = a_d(z)h^d + \cdots + a_1(z)h + a_0(z)$, with $a_i \in \mathbb{Z}[z]$ of degree at most d for $0 \leqslant i \leqslant d$. The solutions to be counted are of two types. Firstly, one has solutions (z, h) with $|z| \leqslant X$ for which $a_i(z) \neq 0$ for some index i, and secondly one has solutions for which $a_i(z) = 0$ $(0 \leqslant i \leqslant d)$. Given any fixed one of the (at most) $2X + 1$ possible choices of z in a solution of the first type, one finds that h satisfies a non-trivial polynomial equation of degree at most d, to which there are at most d integral solutions. There are consequently at most $d(2X + 1)$ solutions of this first type. On the other hand, whenever (z, h) is a solution of the second type, then z satisfies some non-trivial polynomial equation $a_i(z) = 0$ of degree at most d. Since this equation has at most d integral solutions and there are at most $2X^r + 1$ possible choices for h, one has at most $d(2X^r + 1)$ solutions of the second type. The conclusion of the lemma now follows. □

We now announce an initial classification of intermediate coefficients. We define sets $\mathcal{T}_{n,m} \subseteq \mathbb{Z}[z_1, \dots, z_m, h_1, \dots, h_m]$ for $0 \leqslant m \leqslant n + 1$ inductively as follows. First, let $\mathcal{T}_{n,n+1}$ denote the singleton set containing the polynomial

$$\Psi_n(\sigma_{1,n+1}(\mathbf{z}; \mathbf{h}), \dots, \sigma_{2n+1,n+1}(\mathbf{z}; \mathbf{h})). \tag{4.1}$$

Next, suppose that we have already defined the set $\mathcal{T}_{n,m+1}$, and consider an element $\psi \in \mathcal{T}_{n,m+1}$. We may interpret ψ as a polynomial in z_{m+1} and h_{m+1} with coefficients $\phi(z_1, \dots, z_m; h_1, \dots, h_m)$. We now define $\mathcal{T}_{n,m}$ to be the set of all non-zero polynomials $\phi \in \mathbb{Z}[z_1, \dots, z_m, h_1, \dots, h_m]$ occurring as coefficients of elements $\psi \in \mathcal{T}_{n,m+1}$ in this way. Note in particular that since the polynomial (4.1) is not identically zero, it is evident that each set $\mathcal{T}_{n,m}$ is non-empty.

This classification of coefficients yields a consequence of Lemma 4.1 of utility to us in §6.

LEMMA 4.2. *Let m and n be natural numbers with $1 \leqslant m \leqslant n \leqslant t$. Suppose that z_i and h_i are fixed integers for $1 \leqslant i \leqslant m$ with $1 \leqslant z_i \leqslant X$ and $|h_i| \leqslant X^r$.*

Suppose also that there exists $\phi \in \mathcal{T}_{n,m}$ having the property that

$$\phi(z_1, \dots, z_m; h_1, \dots, h_m) \neq 0.$$

Then the number $N_m(X)$ of integral solutions of the system of equations

$$\psi(z_1, \dots, z_{m+1}; h_1, \dots, h_{m+1}) = 0 \quad (\psi \in \mathcal{T}_{n,m+1}),$$

with $1 \leqslant z_{m+1} \leqslant X$ and $|h_{m+1}| \leqslant X^r$, satisfies $N_m(X) \ll X^r$.

Proof. It follows from the iterative definition of the sets $\mathcal{T}_{n,m}$ that any element $\phi \in \mathcal{T}_{n,m}$ occurs as a coefficient polynomial of an element $\psi \in \mathcal{T}_{n,m+1}$, when viewed as a polynomial in h_{m+1} and z_{m+1}. Fixing any one such polynomial ψ, we find that for the fixed choice of $z_1, \dots, z_m, h_1, \dots, h_m$ presented by the hypotheses of the lemma, the polynomial $\psi(\mathbf{z}; \mathbf{h})$ is a non-trivial polynomial in z_{m+1}, h_{m+1}. We therefore conclude from Lemma 4.1 that $N_m(X) \ll X^r$. This completes the proof of the lemma. □

§5. *Classification of solutions.* We now address the classification of the set $\mathcal{S}$ of all solutions of the system of equations

$$\sigma_{j,2s}(\mathbf{z}; \mathbf{h}) = 0 \quad (1 \leqslant j \leqslant 2s - 1), \tag{5.1}$$

with $1 \leqslant \mathbf{z} \leqslant X$ and $|\mathbf{h}| \leqslant X^r$. This we execute in two stages. Our discussion is eased by the use of some non-standard notation. When $(i_1, \dots, i_m)$ is an m-tuple of positive integers with $1 \leqslant i_1 < \cdots < i_m \leqslant 2s$, we abbreviate $(z_{i_1}, \dots, z_{i_m})$ to $\mathbf{z_i}$ and $(h_{i_1}, \dots, h_{i_m})$ to $\mathbf{h_i}$.

In the first stage of our classification, when $0 \leqslant n < s$, we say that $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}$ is of type S_n when:

(i) for all $(n+1)$-tuples $(i_1, \dots, i_{n+1})$ with $1 \leqslant i_1 < \cdots < i_{n+1} \leqslant 2s$, one has

$$\Psi_n(\sigma_{1,n+1}(\mathbf{z_i}; \mathbf{h_i}), \dots, \sigma_{2n+1,n+1}(\mathbf{z_i}; \mathbf{h_i})) = 0;$$

and

(ii) for some n-tuple $(j_1, \dots, j_n)$ with $1 \leqslant j_1 < \cdots < j_n \leqslant 2s$, one has

$$\Psi_{n-1}(\sigma_{1,n}(\mathbf{z_j}; \mathbf{h_j}), \dots, \sigma_{2n-1,n}(\mathbf{z_j}; \mathbf{h_j})) \neq 0.$$

Here, we interpret the condition (ii) to be void when $n = 0$. Finally, we say that $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}$ is of type S_s when the condition (ii) holds with $n = s$. It follows that every solution $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}$ is of type S_n for some index n with $0 \leqslant n \leqslant s$. We denote the set of all solutions of type S_n by $\mathcal{S}_n$.

In the second stage of our classification, when $1 \leqslant n < s$ we subdivide the solutions $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_n$ as follows. When $0 \leqslant m \leqslant n$, we say that a solution $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_n$ is of type $T_{n,m}$ when condition (ii) holds for the n-tuple $\mathbf{j}$, and:

(iii) for all $(m+1)$-tuples $(i_1, \dots, i_{m+1})$ with $1 \leqslant i_1 < \cdots < i_{m+1} \leqslant 2s$ and $i_l \notin \{j_1, \dots, j_n\}$ $(1 \leqslant l \leqslant m+1)$, and for all $\psi \in \mathcal{T}_{n,m+1}$, one has $\psi(\mathbf{z_i}; \mathbf{h_i}) = 0$; and

(iv) for some m-tuple $(\iota_1, \dots, \iota_m)$ with $1 \leqslant \iota_1 < \cdots < \iota_m \leqslant 2s$ and $\iota_l \notin \{j_1, \dots, j_n\}$ $(1 \leqslant l \leqslant m)$, and for some $\phi \in \mathcal{T}_{n,m}$, one has $\phi(\mathbf{z}_\iota; \mathbf{h}_\iota) \neq 0$.

Here, we interpret the condition (iv) to be void when $m = 0$. It follows that whenever $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_n$ with $1 \leqslant n < s$, then it is of type $T_{n,m}$ for some index m with $0 \leqslant m \leqslant n$. As before, we introduce the notation $\mathcal{S}_{n,m}$ to denote the set of all solutions of type $T_{n,m}$. We thus have the decomposition

$$\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_s \cup \bigcup_{n=1}^{s-1} \bigcup_{m=0}^{n} \mathcal{S}_{n,m}. \tag{5.2}$$

§6. *A divisor estimate.* Having enunciated our classification of solutions in the previous section, we are equipped to estimate the number of solutions of the system (5.1) with $1 \leqslant \mathbf{z} \leqslant X$ and $|\mathbf{h}| \leqslant X^r$. This will establish Theorem 2.1, since by discarding superfluous equations if necessary, we may always suppose that $t = 2s - 1$. Before embarking on the main argument, we establish a simple auxiliary result.

LEMMA 6.1. *Suppose that $f \in \mathbb{Z}[x]$ is a polynomial of degree $k \geqslant 1$. Let u be an integer with $1 \leqslant u \leqslant k$, and let h_i and a_i be fixed integers for $1 \leqslant i \leqslant u$ with $\mathbf{h} \neq \mathbf{0}$ and $a_i \neq a_j$ $(1 \leqslant i < j \leqslant u)$. Then for any integer n, the equation*

$$\sum_{i=1}^{u} h_i f(z + a_i) = n \tag{6.1}$$

has at most k solutions in z.

Proof. It suffices to show that the polynomial in z on the left-hand side of (6.1) has positive degree. We therefore assume the opposite and seek a contradiction. Suppose that f is given by

$$f(z) = c_k z^k + c_{k-1} z^{k-1} + \cdots + c_1 z + c_0,$$

where $c_k \neq 0$. The polynomial on the left-hand side of (6.1) takes the shape

$$F(z) = d_k z^k + d_{k-1} z^{k-1} + \cdots + d_1 z + d_0,$$

with

$$d_i = \sum_{j=i}^{k} c_j \binom{j}{i} (h_1 a_1^{j-i} + \cdots + h_u a_u^{j-i}) \quad (0 \leqslant i \leqslant k).$$

In particular, we see directly that d_k can vanish only if $h_1 + \cdots + h_u = 0$. Let i be a positive integer with $i < k$, and suppose that one has

$$h_1 a_1^{k-j} + \cdots + h_u a_u^{k-j} = 0 \tag{6.2}$$

for all integers j with $i < j \leqslant k$. Then the vanishing of d_i implies that (6.2) holds also for $j = i$. Proceeding inductively in this way, we deduce that (6.2) is

satisfied for the entire range $1 \leqslant j \leqslant k$. Restricting attention to the system of equations with indices $k-u+1 \leqslant j \leqslant k$, we find that this system of equations can hold simultaneously only when either $\mathbf{h} = \mathbf{0}$, or else

$$0 = \det(a_i^{j-1})_{1 \leqslant i,j \leqslant u} = \prod_{1 \leqslant i < j \leqslant u} (a_i - a_j).$$

In the latter case, one has $a_i = a_j$ for some indices i and j with $1 \leqslant i < j \leqslant u$. Both these cases are excluded by the hypotheses of the statement of the lemma, so the system of equations (6.2) cannot hold for all $1 \leqslant j \leqslant k$, and hence the polynomial F is non-trivial of positive degree. Consequently, the equation (6.1) has at most $\deg(F) \leqslant k$ solutions in z. □

The proof of Theorem 2.1. We begin by examining the solutions of (5.1) of type S_0, recalling that $1 \leqslant \mathbf{z} \leqslant X$ and $|\mathbf{h}| \leqslant X^r$. When $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_0$, one has $h_i f_1(z_i) = 0$ for $1 \leqslant i \leqslant 2s$. Suppose that the indices i for which $h_i = 0$ are $i_1, \dots, i_a$, and the indices j for which $h_j \neq 0$ are $j_1, \dots, j_b$. In particular, one has $a+b = 2s$. By relabelling variables, if necessary, there is no loss of generality in supposing that $\mathbf{j} = (1, \dots, b)$ and $\mathbf{i} = (b+1, \dots, 2s)$. There are $O(X^{2s-b})$ possible choices for h_i and z_i with $b+1 \leqslant i \leqslant 2s$, since $h_i = 0$ for these indices i. Meanwhile, for $1 \leqslant j \leqslant b$, one has $f_1(z_j) = 0$, and so there are at most k_1 possible choices for z_j. For each fixed such choice, since the polynomials $f_1, \dots, f_t$ are well-conditioned, we find that $f_l(z_j) \neq 0$ for some index l with $2 \leqslant l \leqslant t$. Thus, the variables $h_1, \dots, h_b$ satisfy a system of t linear equations in which there are non-vanishing coefficients. We deduce that when $b \geqslant 1$, there are $O((X^r)^{b-1})$ possible choices for h_j and z_j with $1 \leqslant j \leqslant b$. Finally, combining these estimates for all possible choices of $\mathbf{i}$ and $\mathbf{j}$, we discern that

$$\operatorname{card} \mathcal{S}_0 \ll X^{2s} + \sum_{b=1}^{2s} X^{2s-b} \cdot X^{r(b-1)} \ll X^{(2s-1)r+1}. \tag{6.3}$$

Next we consider the solutions of (5.1) of type S_s. When $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_s$, there is an s-tuple $\mathbf{i}$ with $1 \leqslant i_1 < \cdots < i_s \leqslant 2s$ for which one has

$$\Psi_{s-1}(\sigma_{1,s}(\mathbf{z_i}; \mathbf{h_i}), \dots, \sigma_{2s-1,s}(\mathbf{z_i}; \mathbf{h_i})) \neq 0.$$

Write $\mathbf{i}'$ for the s-tuple $(i'_1, \dots, i'_s)$ with $1 \leqslant i'_1 < \cdots < i'_s \leqslant 2s$ for which

$$\{i_1, \dots, i_s\} \cup \{i'_1, \dots, i'_s\} = \{1, 2, \dots, 2s\}.$$

It follows from (5.1) that $\sigma_{j,s}(\mathbf{z_i}; \mathbf{h_i}) = \sigma_{j,s}(\mathbf{z_{i'}}; -\mathbf{h_{i'}})$ $(1 \leqslant j \leqslant 2s-1)$, and hence there is a non-zero integer $N = N(\mathbf{z_{i'}}; \mathbf{h_{i'}})$ for which

$$\Psi_{s-1}(\sigma_{1,s}(\mathbf{z_{i'}}; -\mathbf{h_{i'}}), \dots, \sigma_{2s-1,s}(\mathbf{z_{i'}}; -\mathbf{h_{i'}})) = N \tag{6.4}$$

and

$$\Psi_{s-1}(\sigma_{1,s}(\mathbf{z_i}; \mathbf{h_i}), \dots, \sigma_{2s-1,s}(\mathbf{z_i}; \mathbf{h_i})) = N.$$

By relabelling variables, if necessary, there is no loss of generality in supposing that $\mathbf{i} = (1, 2, \dots, s)$ and $\mathbf{i}' = (s+1, s+2, \dots, 2s)$.

Fix any one of the $O(X^{(r+1)s})$ possible choices for $\mathbf{z}_{\mathbf{i}'}$, $\mathbf{h}_{\mathbf{i}'}$ with $1 \leqslant \mathbf{z}_{\mathbf{i}'} \leqslant X$, $|\mathbf{h}_{\mathbf{i}'}| \leqslant X^r$, and satisfying (6.4). Then we infer from Lemma 3.2 that

$$h_1 \cdots h_s \prod_{1 \leqslant i < j \leqslant s} (z_i - z_j) \quad \text{divides } N(\mathbf{z}_{\mathbf{i}'}; \mathbf{h}_{\mathbf{i}'}). \tag{6.5}$$

Moreover, one has $N(\mathbf{z}_{\mathbf{i}'}; \mathbf{h}_{\mathbf{i}'}) \neq 0$. Since the latter integer is fixed, we see by means of an elementary divisor function estimate that there are $O(X^{\varepsilon})$ possible choices for $h_1, \dots, h_s$ and integers $a_2, \dots, a_s$ equipped with the property that $z_i = z_1 + a_i$ $(2 \leqslant i \leqslant s)$. With the exception of the undetermined variable z_1, it follows that there are at most $O(X^{(r+1)s+\varepsilon})$ possible choices for all the variables in question. However, the integer z_1 satisfies the system of equations

$$h_1 f_j(z_1) + \sum_{i=2}^{s} h_i f_j(z_1 + a_i) = n_j \quad (1 \leqslant j \leqslant 2s-1), \tag{6.6}$$

in which h_i, a_i and n_j are all fixed for all indices i and j. Consider the polynomial with index $j = 1$ of largest degree $k_1 \geqslant 2s-2$. If a_i is zero for any index i, then we have $z_1 = z_i$. Meanwhile, if $a_i = a_j$ for any indices i and j with $2 \leqslant i < j \leqslant s$, one sees that $z_i = z_j$. Consequently, in either of these scenarios, and also in the situation with $\mathbf{h} = \mathbf{0}$, one finds via (6.5) that $N(\mathbf{z}_{\mathbf{i}'}; \mathbf{h}_{\mathbf{i}'}) = 0$, contradicting our assumption that $N(\mathbf{z}_{\mathbf{i}'}; \mathbf{h}_{\mathbf{i}'}) \neq 0$. We may thus safely assume that the conditions of Lemma 6.1 are satisfied for the polynomial f_1 with $a_1 = 0$. By the conclusion of the lemma, it follows that there are at most k_1 choices for z_1 satisfying (6.6), and hence

$$\operatorname{card} \mathcal{S}_s \ll X^{(r+1)s+\varepsilon}. \tag{6.7}$$

Next we consider the set $\mathcal{S}_{n,m}$ for a given pair of indices n and m with $1 \leqslant n < s$ and $0 \leqslant m \leqslant n$. For any $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_{n,m}$, condition (ii) holds for some n-tuple $\mathbf{j}$. By relabelling variables, if necessary, we may suppose that $\mathbf{j} = (1, \dots, n)$. Write $\mathbf{j}'$ for the $(2s-n)$-tuple $(n+1, \dots, 2s)$. Then given any one fixed choice of the variables $\mathbf{z}_{\mathbf{j}'}$, $\mathbf{h}_{\mathbf{j}'}$, we have

$$\begin{aligned} &\Psi_{n-1}(\sigma_{1,n}(\mathbf{z}_{\mathbf{j}}; \mathbf{h}_{\mathbf{j}}), \dots, \sigma_{2n-1,n}(\mathbf{z}_{\mathbf{j}}; \mathbf{h}_{\mathbf{j}})) \\ &\quad = \Psi_{n-1}(\sigma_{1,2s-n}(\mathbf{z}_{\mathbf{j}'}; -\mathbf{h}_{\mathbf{j}'}), \dots, \sigma_{2n-1,2s-n}(\mathbf{z}_{\mathbf{j}'}; -\mathbf{h}_{\mathbf{j}'})) \neq 0. \end{aligned}$$

Thus, there is a fixed non-zero integer N with the property that

$$\Psi_{n-1}(\sigma_{1,n}(\mathbf{z}_{\mathbf{j}}; \mathbf{h}_{\mathbf{j}}), \dots, \sigma_{2n-1,n}(\mathbf{z}_{\mathbf{j}}; \mathbf{h}_{\mathbf{j}})) = N,$$

and we deduce from Lemma 3.2 that

$$h_1 \cdots h_n \prod_{1 \leqslant i < j \leqslant n} (z_i - z_j) \quad \text{divides } N.$$

From here, the argument applied above in the case $n = s$ may be employed *mutatis mutandis* to conclude that there are $O(X^{\varepsilon})$ possible choices for $h_1, \dots, h_n, z_1 - z_2, \dots, z_1 - z_n$. If we put $a_i = z_i - z_1$ $(2 \leqslant i \leqslant n)$ and $a_1 = 0$, then we find just as in our earlier analysis that z_1 satisfies a non-trivial polynomial equation of degree at most k_1, whence there are at most k_1 choices for z_1. We therefore conclude that, given any one fixed choice of $\mathbf{z}_{\mathbf{j}'}, \mathbf{h}_{\mathbf{j}'}$, the number of choices for $\mathbf{z}_{\mathbf{j}}, \mathbf{h}_{\mathbf{j}}$ is $O(X^{\varepsilon})$.

It thus remains to count the number of choices for $\mathbf{z}_{\mathbf{j}'}$ and $\mathbf{h}_{\mathbf{j}'}$. Note in particular that, since $(\mathbf{z}, \mathbf{h}) \in \mathcal{S}_{n,m}$, we have the additional information that conditions (iii) and (iv) are satisfied. We may therefore suppose that there exists some $\phi \in \mathcal{T}_{n,m}$, and some m-tuple $(\iota_1, \dots, \iota_m)$ with $n+1 \leqslant \iota_1 < \cdots < \iota_m \leqslant 2s$, for which

$$\phi(\mathbf{z}_{\boldsymbol{\iota}}; \mathbf{h}_{\boldsymbol{\iota}}) \neq 0. \tag{6.8}$$

With a fixed choice of $\boldsymbol{\iota}$, we may suppose further that for all i satisfying $n+1 \leqslant i \leqslant 2s$ and $i \notin \{\iota_1, \dots, \iota_m\}$, and for all $\psi \in \mathcal{T}_{n,m+1}$, one has

$$\psi(z_{\iota_1}, \dots, z_{\iota_m}, z_i; h_{\iota_1}, \dots, h_{\iota_m}, h_i) = 0. \tag{6.9}$$

Given any such $\boldsymbol{\iota}$ and ϕ, there are $O(X^{(r+1)m})$ possible choices for $\mathbf{z}_{\boldsymbol{\iota}}, \mathbf{h}_{\boldsymbol{\iota}}$, with $1 \leqslant \mathbf{z}_{\boldsymbol{\iota}} \leqslant X$ and $|\mathbf{h}_{\boldsymbol{\iota}}| \leqslant X^r$, satisfying (6.8). We claim that for any fixed such choice, the number of possible choices for the integers z_i and h_i with $n+1 \leqslant i \leqslant 2s$ and $i \notin \{\iota_1, \dots, \iota_m\}$ is $O((X^r)^{2s-n-m})$. In order to confirm this claim, observe that there is a polynomial $\psi \in \mathcal{T}_{n,m+1}$ having the property that some coefficient of $\psi(z_1, \dots, z_{m+1}; h_1, \dots, h_{m+1})$, considered as a polynomial in z_{m+1} and h_{m+1}, is equal to $\phi(z_1, \dots, z_m; h_1, \dots, h_m)$. It then follows from (6.8) that the equation (6.9) is a non-trivial polynomial equation in z_i and h_i. We therefore deduce from Lemma 4.2 that for each fixed choice of $\mathbf{z}_{\boldsymbol{\iota}}$ and $\mathbf{h}_{\boldsymbol{\iota}}$ under consideration, and for each i with $n+1 \leqslant i \leqslant 2s$ and $i \notin \{\iota_1, \dots, \iota_m\}$, there are $O(X^r)$ possible choices for z_i and h_i satisfying (6.9). Thus we infer that there are $O(X^{r(2s-n-m)})$ possible choices for z_i and h_i with $n+1 \leqslant i \leqslant 2s$ for each fixed choice of $\mathbf{z}_{\boldsymbol{\iota}}, \mathbf{h}_{\boldsymbol{\iota}}$. Since the number of choices for $\boldsymbol{\iota}$ and $\phi \in \mathcal{T}_{n,m}$ is $O(1)$, the total number of choices for $\mathbf{z}_{\mathbf{j}'}$ and $\mathbf{h}_{\mathbf{j}'}$ available to us is $O(X^{(r+1)m} \cdot X^{r(2s-n-m)})$. Furthermore, our discussion above showed that for each fixed such choice of $\mathbf{z}_{\mathbf{j}'}, \mathbf{h}_{\mathbf{j}'}$, the number of possible choices for $\mathbf{z}_{\mathbf{j}}, \mathbf{h}_{\mathbf{j}}$ is $O(X^{\varepsilon})$. Thus altogether we conclude that

$$\operatorname{card} \mathcal{S}_{n,m} \ll X^{r(2s-n)+m+\varepsilon}. \tag{6.10}$$

By combining our estimates (6.3), (6.7) and (6.10) via (5.2), we discern that

$$\operatorname{card} \mathcal{S} \ll X^{(2s-1)r+1} + X^{(r+1)s+\varepsilon} + \sum_{n=1}^{s-1} \sum_{m=0}^{n} X^{r(2s-n)+m+\varepsilon} \ll X^{r(2s-1)+1+\varepsilon},$$

and the conclusion of Theorem 2.1 follows. □

§7. *The proof of Theorems 1.1 and 1.2.* Our preparations now complete, we establish the mean value estimates recorded in Theorems 1.1 and 1.2. Let X be a large positive number, and suppose that s and k are natural numbers with $k \geqslant 2$ and $1 \leqslant s \leqslant (k^2-1)/2$. We define the exponential sum $\mathfrak{g}_r(\boldsymbol{\alpha}; X)$ by putting

$$\mathfrak{g}_r(\boldsymbol{\alpha}; X) = \sum_{|h| \leqslant sX^r} \sum_{1 \leqslant z \leqslant X} e\bigg(\binom{r}{r}h\alpha_r + \binom{r+1}{r}hz\alpha_{r+1} + \cdots + \binom{k}{r}hz^{k-r}\alpha_k\bigg). \tag{7.1}$$

Also, when $1 \leqslant d \leqslant k$, we put

$$\mathfrak{h}_d(\boldsymbol{\alpha}; X) = \sum_{1 \leqslant x \leqslant X} e(\alpha_1 x + \cdots + \alpha_d x^d).$$

Then, with the standard notation associated with Vinogradov's mean value theorem in mind, we put

$$J_{\sigma,d}(X) = \oint |\mathfrak{h}_d(\boldsymbol{\alpha}; X)|^{2\sigma}\, d\boldsymbol{\alpha}. \tag{7.2}$$

We note that the main conjecture in Vinogradov's mean value theorem is now known to hold for all degrees. This is a consequence of work of the second author for degree 3, and for degrees exceeding 3 it follows from the work of Bourgain, Demeter and Guth (see [**9**, Theorem 1.1] and [**1**, Theorem 1.1]). Thus, one has

$$J_{\sigma,d}(X) \ll X^{\sigma+\varepsilon} \quad (1 \leqslant \sigma \leqslant d(d+1)/2). \tag{7.3}$$

In addition, one finds via orthogonality that for each integer κ, one has

$$\oint |\mathfrak{g}_r(\boldsymbol{\alpha}; X)|^{2\kappa}\, d\boldsymbol{\alpha} \leqslant A_{\kappa,r}(sX; \mathbf{f}),$$

where $f_j(z) = z^{k-r+1-j}$ $(1 \leqslant j \leqslant k-r+1)$.

LEMMA 7.1. *When s is a natural number, one has*

$$I_{s,k,r}(X) \ll X^{-1} \oint |\mathfrak{h}_k(\boldsymbol{\alpha}; 2X)|^{2s} \mathfrak{g}_r(-\boldsymbol{\alpha}; X)\, d\boldsymbol{\alpha}.$$

Proof. Define δ_j to be 1 when $j = r$, and 0 otherwise. We start by noting that the mean value $I_{s,k,r}(X)$ counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{s} (x_i^j - y_i^j) = \delta_j h \quad (1 \leqslant j \leqslant k), \tag{7.4}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$ and $|h| \leqslant sX^r$. We remark that the constraint on

$$\sum_{i=1}^{s} (x_i^r - y_i^r) \tag{7.5}$$

imposed by the equation of degree r in (7.4) is void, since the range for h automatically accommodates all possible values of the expression (7.5) within (7.4).

We next consider the effect of shifting every variable by an integer z with $1 \leqslant z \leqslant X$. By the binomial theorem, for any shift z, one finds that $(\mathbf{x}, \mathbf{y})$ is a solution of (7.4) if and only if it is also a solution of the system

$$\sum_{i=1}^{s}((x_i+z)^j-(y_i+z)^j)=\omega_j h z^{j-r} \quad (1 \leqslant j \leqslant k),$$

where ω_j is 0 for $1 \leqslant j < r$ and $\binom{j}{r}$ for $r \leqslant j \leqslant k$. Thus, for each fixed integer z with $1 \leqslant z \leqslant X$, the mean value $I_{s,k,r}(X)$ is bounded above by the number of integral solutions of the system

$$\sum_{i=1}^{s}(u_i^j-v_i^j)=\omega_j h z^{j-r} \quad (1 \leqslant j \leqslant k),$$

with $1 \leqslant \mathbf{u}, \mathbf{v} \leqslant 2X$ and $|h| \leqslant sX^r$. On applying orthogonality, we therefore infer that

$$I_{s,k,r}(X) \ll X^{-1} \sum_{1 \leqslant z \leqslant X} \oint |\mathfrak{h}_k(\boldsymbol{\alpha}; 2X)|^{2s} \mathfrak{f}(-\boldsymbol{\alpha}; z)\, d\boldsymbol{\alpha},$$

where

$$\mathfrak{f}(\boldsymbol{\alpha}; z) = \sum_{|h| \leqslant sX^r} e(\omega_r h \alpha_r + \omega_{r+1} h z \alpha_{r+1} + \cdots + \omega_k h z^{k-r} \alpha_k).$$

The proof of the lemma is completed by reference to (7.1). □

The proof of Theorem 1.2. Let s, k and r be integers with $k > r \geqslant 1$. Also, let κ be a positive integer with $\kappa \leqslant (k-r+2)/2$. Observe that it suffices to restrict attention to the special case

$$s = \left\lfloor \frac{k(k+1)}{2} - \frac{k(k+1)-r(r-1)}{4\kappa} \right\rfloor,$$

since one may interpolate via Hölder's inequality to recover the conclusion of the theorem for smaller values of s. Put

$$v = \frac{r(r-1)}{4\kappa} \quad \text{and} \quad u = s - v.$$

Furthermore, set

$$w = \left(1 - \frac{1}{2\kappa}\right)\frac{k(k+1)}{2},$$

so that $s = \lfloor v + w \rfloor$. In particular, we have $w \geqslant u$.

On applying Hölder's inequality in combination with Lemma 7.1, we find that

$$I_{s,k,r}(X) \ll X^{-1}U_1^{1-1/(2\kappa)}U_2^{1/(2\kappa)}, \tag{7.6}$$

where

$$U_1 = \oint |\mathfrak{h}_k(\boldsymbol{\alpha}; 2X)|^{(u/w)k(k+1)}\, d\boldsymbol{\alpha} \tag{7.7}$$

and

$$U_2 = \oint |\mathfrak{h}_k(\boldsymbol{\alpha}; 2X)^{r(r-1)}\mathfrak{g}_r(\boldsymbol{\alpha}; X)^{2\kappa}|\, d\boldsymbol{\alpha}. \tag{7.8}$$

A comparison of (7.7) with (7.2) leads us via (7.3) to the estimate

$$U_1 \ll X^{(u/w)k(k+1)/2+\varepsilon}. \tag{7.9}$$

Meanwhile, by orthogonality, we discern from (7.8) that U_2 counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{r(r-1)/2}(x_i^j - y_i^j) = \binom{j}{r}\sum_{l=1}^{2\kappa} h_l z_l^{j-r} \quad (r \leqslant j \leqslant k) \tag{7.10}$$

$$\sum_{i=1}^{r(r-1)/2}(x_i^j - y_i^j) = 0 \quad (1 \leqslant j < r), \tag{7.11}$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant 2X$, $1 \leqslant \mathbf{z} \leqslant X$ and $|\mathbf{h}| \leqslant sX^r$. By interpreting (7.11) through the prism of orthogonality, it follows from (7.2) that the number of available choices for $\mathbf{x}$ and $\mathbf{y}$ is bounded above by $J_{r(r-1)/2,r-1}(2X)$. For each fixed such choice of $\mathbf{x}$ and $\mathbf{y}$, it follows from (7.10) via orthogonality and the triangle inequality that the number of available choices for $\mathbf{z}$ and $\mathbf{h}$ is at most $A_{\kappa,r}(sX; \mathbf{f})$. Thus we deduce from (7.3) and Theorem 2.1 that

$$U_2 \leqslant J_{r(r-1)/2,r-1}(2X)A_{\kappa,r}(sX; \mathbf{f}) \ll X^{r(r-1)/2+r(2\kappa-1)+1+\varepsilon}. \tag{7.12}$$

On substituting (7.9) and (7.12) into (7.6), we infer that

$$I_{s,k,r}(X) \ll X^{\varepsilon-1}(X^{(u/w)k(k+1)/2})^{1-1/(2\kappa)}(X^{2r\kappa+1+r(r-3)/2})^{1/(2\kappa)} \ll X^{s+\Delta+\varepsilon},$$

where

$$\Delta = (r-1) - \frac{r-1}{2\kappa}.$$

This completes the proof of Theorem 1.2. □

The proof of Theorem 1.1. The conclusion of Theorem 1.1 is an immediate consequence of Theorem 1.2 in the special case $r = 1$. Making use of the notation of the statement of the latter theorem, we note that when $k = 2l + 1$ is odd, one may take $\kappa = \lfloor (k+1)/2 \rfloor = l+1$, and we deduce that $I_{s,k,1}(X) \ll X^{s+\varepsilon}$ provided that s is a natural number not exceeding

$$\frac{k(k+1)}{2} - \frac{k(k+1)}{4(l+1)} = \frac{k(k+1)}{2} - \frac{k}{2}.$$

Meanwhile, when $k = 2l$ is even, one may instead take $\kappa = l$, and the same conclusion holds provided that s is a natural number not exceeding

$$\frac{k(k+1)}{2} - \frac{k(k+1)}{4l} = \frac{k(k+1)}{2} - \frac{k+1}{2}.$$

The desired conclusion therefore follows in both cases, and the proof of Theorem 1.1 is complete. □

Acknowledgements. Both authors thank the Fields Institute in Toronto for excellent working conditions and support that made this work possible during the Thematic Program on Unlikely Intersections, Heights, and Efficient Congruencing. The work of the first author was supported by the National Science Foundation under Grant No. DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. The second author's work was supported by a European Research Council Advanced Grant under the European Union's Horizon 2020 research and innovation programme via grant agreement No. 695223.

References

1. J. Bourgain, C. Demeter and L. Guth, Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math.* (2) **184**(2) (2016), 633–682.
2. J. Brandes and S. T. Parsell, Simultaneous additive equations: repeated and differing degrees. *Canad. J. Math.* **69**(2) (2017), 258–283.
3. J. Brüdern and O. Robert, A paucity estimate related to Newton sums of odd degree. *Mathematika* **58**(2) (2012), 225–235.
4. J. Brüdern and O. Robert, Rational points on linear slices of diagonal hypersurfaces. *Nagoya Math. J.* **218** (2015), 51–100.
5. I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford Mathematical Monographs (Oxford, 1979).
6. S. T. Parsell and T. D. Wooley, A quasi-paucity problem. *Michigan Math. J.* **50** (2002), 461–469.
7. T. D. Wooley, A note on symmetric diagonal equations. In *Number Theory with an Emphasis on the Markoff Spectrum (Provo, UT, 1991)* (eds A. D. Pollington and W. Moran), Dekker (New York, 1993), 317–321.
8. T. D. Wooley, Rational solutions of pairs of diagonal equations, one cubic and one quadratic. *Proc. Lond. Math. Soc.* (3) **110**(2) (2015), 325–356.
9. T. D. Wooley, The cubic case of the main conjecture in Vinogradov's mean value theorem. *Adv. Math.* **294** (2016), 532–561.
10. T. D. Wooley, Nested efficient congruencing and relatives of Vinogradov's mean value theorem. *Preprint*, 2017, arXiv:1708.01220.

Julia Brandes,
Mathematical Sciences Research Institute,
17 Gauss Way,
Berkeley,
CA 94720-5070,
U.S.A.
and
Mathematical Sciences,
Chalmers Institute of Technology and
University of Gothenburg,
412 96 Göteborg,
Sweden
E-mail: brjulia@chalmers.se
Current address:
Pure Mathematics,
University of Waterloo,
200 University Avenue West,
Waterloo, ON, N2L 3G1,
Canada
E-mail: jbrandes@uwaterloo.ca

Trevor D. Wooley,
School of Mathematics,
University of Bristol,
University Walk,
Clifton,
Bristol BS8 1TW,
U.K.
E-mail: matdw@bristol.ac.uk