THE PAUCITY PROBLEM FOR CERTAIN SYMMETRIC DIOPHANTINE EQUATIONS

TREVOR D. WOOLEY

Abstract. Let $\varphi_1, \ldots, \varphi_r \in \mathbb{Z}[z_1, \ldots z_k]$ be integral linear combinations of elementary symmetric polynomials with $\deg(\varphi_j) = k_j$ $(1 \leqslant j \leqslant r)$, where $1 \leqslant k_1 < k_2 < \ldots < k_r = k$. Subject to the condition $k_1 + \ldots + k_r \geqslant \frac{1}{2}k(k-1) + 2$, we show that there is a paucity of non-diagonal solutions to the Diophantine system $\varphi_j(\mathbf{x}) = \varphi_j(\mathbf{y})$ $(1 \leqslant j \leqslant r)$.

1. Introduction

Our focus in this memoir lies on systems of simultaneous Diophantine equations defined by symmetric polynomials. Let $\varphi_1, \ldots, \varphi_r \in \mathbb{Z}[z_1, \ldots, z_k]$ be symmetric polynomials with $\deg(\varphi_j) = k_j$ $(1 \leqslant j \leqslant r)$, so that for any permutation π of $\{1, 2, \ldots, k\}$, one has

$$\varphi_j(x_{\pi 1}, \ldots, x_{\pi k}) = \varphi_j(x_1, \ldots, x_k) \quad (1 \leqslant j \leqslant r).$$

Given any such permutation π, the system of Diophantine equations

$$\varphi_j(x_1, \ldots, x_k) = \varphi_j(y_1, \ldots, y_k) \quad (1 \leqslant j \leqslant r) \tag{1.1}$$

plainly has the trivial solutions obtained by putting $y_i = x_{\pi i}$ $(1 \leqslant i \leqslant k)$. Denoting by $T_k(X)$ the number of these trivial solutions with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant k)$, one is led to the question of whether there is a paucity of non-diagonal solutions. Thus, fixing k and $\boldsymbol{\varphi}$, and writing $N_k(X; \boldsymbol{\varphi})$ for the number of solutions of the system (1.1) in this range, one may ask whether as $X \to \infty$, one has

$$N_k(X; \boldsymbol{\varphi}) = T_k(X) + o(T_k(X)),$$

or equivalently, whether $N_k(X; \boldsymbol{\varphi}) = k!X^k + o(X^k)$. This question has been examined extensively in the diagonal case where the polynomials under consideration take the shape $\varphi_j(\mathbf{x}) = x_1^{k_j} + \ldots + x_k^{k_j}$, as can be surmised from the references to this memoir. Relatively little consideration has been afforded to more general symmetric polynomials. It transpires that by adapting a strategy applied previously in a special case of Vinogradov's mean value theorem (see [17]), we are able to settle this paucity problem for numerous systems of the type (1.1) in a particularly strong form.

Further notation is required to describe our conclusions. Define the elementary symmetric polynomials $\sigma_j(\mathbf{z}) \in \mathbb{Z}[z_1, \ldots, z_k]$ for $j \geqslant 0$ by means of the

2010 *Mathematics Subject Classification.* 11D45, 11D72.

Key words and phrases. Paucity, symmetric polynomials, Diophantine equations.

generating function identity

$$\sum_{j=0}^{k}\sigma_j(\mathbf{z})t^{k-j}=\prod_{i=1}^{k}(t+z_i). \tag{1.2}$$

We restrict attention primarily to symmetric polynomials of the shape

$$\varphi_j(\mathbf{z})=\sum_{l=1}^{k}a_{jl}\sigma_l(\mathbf{z})\quad (1\leqslant j\leqslant r), \tag{1.3}$$

with fixed coefficients $a_{jl}\in\mathbb{Z}$ for $1\leqslant j\leqslant r$ and $1\leqslant l\leqslant k$. By taking appropriate integral linear combinations of the polynomials $\varphi_1,\ldots,\varphi_r$, it is evident that in our investigations concerning $N_k(X;\boldsymbol{\varphi})$, we may suppose that $\deg(\varphi_j)=k_j$ $(1\leqslant j\leqslant r)$, where the exponents k_j satisfy the condition

$$1\leqslant k_1<k_2<\ldots<k_r=k. \tag{1.4}$$

This can be seen by applying elementary row operations on the reversed $r\times k$ matrix of coefficients $A=(a_{j,k-l+1})$, reducing to an equivalent system having the same number of solutions, in which the new coefficient matrix A' has upper triangular form and $r'\leqslant r$ non-vanishing rows.

A simple paucity result is provided by our first theorem. It is useful here and elsewhere to introduce the auxiliary quantity

$$w(\boldsymbol{\varphi})=\tfrac{1}{2}k(k+1)-k_1-k_2-\ldots-k_r. \tag{1.5}$$

Here and throughout this paper, implicit constants in the notations of Landau and Vinogradov may depend on ε, k, and the coefficients of $\boldsymbol{\varphi}$.

Theorem 1.1. *Let $\varphi_1,\ldots,\varphi_r$ be symmetric polynomials of the shape* (1.3) *having respective degrees $k_1,\ldots,k_r$ satisfying* (1.4)*. Then for each $\varepsilon>0$,*

$$N_k(X;\boldsymbol{\varphi})=T_k(X)+O(X^{w(\boldsymbol{\varphi})+1+\varepsilon}).$$

In particular, when $k_1+\ldots+k_r\geqslant \frac{1}{2}k(k-1)+2$, one has

$$N_k(X;\boldsymbol{\varphi})=k!X^k+O(X^{k-1+\varepsilon}).$$

A specialisation of the system (1.1) illustrates the kind of results made available by Theorem 1.1. Fix a choice of coefficients $a_{jl}\in\mathbb{Z}$ for $1\leqslant l\leqslant k-r$ and $k-r+1\leqslant j\leqslant k$, and denote by $M_{k,r}(X;\mathbf{a})$ the number of integral solutions of the simultaneous equations

$$\sigma_j(\mathbf{x})+\sum_{l=1}^{k-r}a_{jl}\sigma_l(\mathbf{x})=\sigma_j(\mathbf{y})+\sum_{l=1}^{k-r}a_{jl}\sigma_l(\mathbf{y})\quad (k-r+1\leqslant j\leqslant k), \tag{1.6}$$

in variables $\mathbf{x}=(x_1,\ldots,x_k)$ and $\mathbf{y}=(y_1,\ldots,y_k)$ with $1\leqslant x_i,y_i\leqslant X$.

Corollary 1.2. *Suppose that $k,r\in\mathbb{N}$ and $(k-r)(k-r+1)<2k-2$. Then there is a paucity of non-diagonal solutions in the system of equations* (1.6)*. In particular, for each $\varepsilon>0$ one has*

$$M_{k,r}(X;\mathbf{a})=T_k(X)+O(X^{\frac{1}{2}(k-r)(k-r+1)+1+\varepsilon}).$$

Results analogous to those of Theorem 1.1 and Corollary 1.2 for diagonal Diophantine systems involving r equations are almost always limited to systems possessing only $r+1$ pairs of variables, and these we now describe. In the case $r=1$, Hooley applied sieve methods to investigate the equation

$$x_1^k + x_2^k = y_1^k + y_2^k,$$

with $k \geqslant 3$, and ultimately established the paucity of non-diagonal solutions with a power saving (see [9, 10]). Strong conclusions have been derived when $k=3$ by Heath-Brown [8] using ideas based on quadratic forms, enhancing earlier work of the author that extends from cubes to general cubic polynomials [20]. When $r=2$, $k_1 < k_2$ and $k_2 \geqslant 3$, the paucity of non-diagonal solutions has been established for the pair of simultaneous equations

$$\left.\begin{aligned} x_1^{k_1} + x_2^{k_1} + x_3^{k_1} &= y_1^{k_1} + y_2^{k_1} + y_3^{k_1} \\ x_1^{k_2} + x_2^{k_2} + x_3^{k_2} &= y_1^{k_2} + y_2^{k_2} + y_3^{k_2} \end{aligned}\right\}.$$

Sharp results are available in the case $(k_1,k_2)=(1,3)$ (see [16]), with non-trivial conclusions available for the exponent pair $(1,k)$ when $k \geqslant 3$ (see [5, 14]). The cases $(2,3)$ and $(2,4)$ were tackled successfully via affine slicing methods (see [15, 19]), with the remaining cases of this type covered by Salberger [12] using variants of the determinant method. Most other examples in which it is known that there is a paucity of non-diagonal solutions are closely related to the Vinogradov system of equations

$$x_1^j + \ldots + x_{r+1}^j = y_1^j + \ldots + y_{r+1}^j \quad (1 \leqslant j \leqslant r).$$

The paucity problem has been solved here by Vaughan and the author [17], with similar conclusions when the equation of degree r is replaced by one of degree $r+1$. Recent work [21] shows that the missing equation of degree r in this last result can be replaced by one of degree $r-d$, provided that d is not too large. Meanwhile, when the exponents k_j satisfy (1.4), results falling just short of paucity have been obtained in general for systems of the shape

$$x_1^{k_j} + \ldots + x_{r+1}^{k_j} = y_1^{k_j} + \ldots + y_{r+1}^{k_j} \quad (1 \leqslant j \leqslant r),$$

(see [11, 18]). In the special case $\mathbf{k}=(1,3,\ldots,2r-1)$, Brüdern and Robert [2] have even established the desired paucity result. We should note also that the existence of non-diagonal solutions, often exhibited by remarkable parametric formulae, has long been the subject of investigation, as recorded in Gloden's book [4], with notable recent contributions by Choudhry [3].

Taking k large and $r = k - \lfloor \sqrt{2k} \rfloor + 1$, we see that (1.6) constitutes a Diophantine system of r equations in $k = r + \sqrt{2r} + O(1)$ pairs of variables having a paucity of non-diagonal solutions. Corollary 1.2 therefore exhibits a large class of Diophantine systems in which the barrier described in the previous paragraph is emphatically surmounted. There are two exceptions to the rule noted in that paragraph. First, Salberger and the author [13, Corollary 1.4 and Theorem 5.2] have established the paucity of non-diagonal solutions in situations where the underlying equations have very large degree in terms of the number of variables. Second, work of Bourgain *et al.* [1, Theorem 26] and Heap *et al.* [7, Theorem 1.2] examines systems of equations determined by relations of divisor type having the shape

$$(x_1+\theta)(x_2+\theta)\cdots(x_k+\theta)=(y_1+\theta)(y_2+\theta)\cdots(y_k+\theta), \tag{1.7}$$

in which $\theta\in\mathbb{C}$ is algebraic of degree d over $\mathbb{Q}$. When $d\leqslant k$, this relation generates d independent symmetric Diophantine equations, and it is shown that the number of integral solutions of (1.7) with $1\leqslant x_i,y_i\leqslant X$ is asymptotically $T_k(X)+O(X^{k-d+1+\varepsilon})$. In particular, when $d=2$, we obtain a pair of simultaneous Diophantine equations in k variables having a paucity of non-diagonal solutions. For example, when $k=4$ and $\theta=\sqrt{-1}$, we obtain the system

$$\left.\begin{aligned} x_1x_2x_3x_4-&x_1x_2-x_2x_3-x_3x_4-x_4x_1-x_2x_4-x_1x_3\\ &=y_1y_2y_3y_4-y_1y_2-y_2y_3-y_3y_4-y_4y_1-y_2y_4-y_1y_3\\ x_1x_2x_3+&x_2x_3x_4+x_3x_4x_1+x_4x_1x_2-x_1-x_2-x_3-x_4\\ &=y_1y_2y_3+y_2y_3y_4+y_3y_4y_1+y_4y_1y_2-y_1-y_2-y_3-y_4 \end{aligned}\right\},$$

having $4!X^4+O(X^{3+\varepsilon})$ integral solutions with $1\leqslant x_i,y_i\leqslant X$.

We have avoided discussion of systems (1.1) containing the equation

$$x_1x_2\cdots x_k=y_1y_2\cdots y_k. \tag{1.8}$$

Here, when $r\geqslant 2$, one may parametrise the solutions of (1.8) in the shape

$$\begin{aligned} x_1&=\alpha_1\alpha_2\cdots\alpha_k, & x_2&=\beta_1\beta_2\cdots\beta_k, & \ldots,\quad x_k&=\omega_1\omega_2\cdots\omega_k,\\ y_1&=\alpha_1\beta_1\cdots\omega_1, & y_2&=\alpha_2\beta_2\cdots\omega_2, & \ldots,\quad y_k&=\alpha_k\beta_k\cdots\omega_k, \end{aligned}$$

to provide a paucity result by simple elimination. The reader will find all of the ideas necessary to complete this elementary exercise in [16].

This memoir is organised as follows. In §2 we derive a multiplicative relation amongst the variables $\mathbf{x},\mathbf{y}$ of the system (1.1). This may be applied to obtain Theorem 1.1 and Corollary 1.2. The polynomials $\varphi_j(\mathbf{z})$ that are the subject of Theorem 1.1 are integral linear combinations of the elementary symmetric polynomials $\sigma_1(\mathbf{z}),\ldots,\sigma_k(\mathbf{z})$. In §3 we examine the extent to which our methods are applicable when the polynomials $\varphi_j(\mathbf{z})$ are permitted to depend non-linearly on $\sigma_1(\mathbf{z}),\ldots,\sigma_k(\mathbf{z})$.

Our basic parameter is X, a sufficiently large positive number. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon>0$. We make frequent use of vector notation in the form $\mathbf{x}=(x_1,\ldots,x_k)$. Here, the dimension k will be evident to the reader from the ambient context.

Acknowledgements: The author's work is supported by NSF grants DMS-2001549 and DMS-1854398. We thank the referee for useful comments.

2. Multiplicative relations from symmetric polynomials

Our initial objective in this section is to obtain a multiplicative relation between the variables underlying the system (1.1). Let $\varphi_1,\ldots,\varphi_r$ be polynomials of the shape (1.3) having respective degrees $k_1,\ldots,k_r$ satisfying (1.4).

We define the complementary set of exponents $\mathcal{R} = \mathcal{R}(\boldsymbol{\varphi})$ by putting

$$\mathcal{R}(\boldsymbol{\varphi}) = \{1, \ldots, k\} \setminus \{k_1, \ldots, k_r\}. \tag{2.1}$$

The counting function $N_k(X; \boldsymbol{\varphi})$ remains unchanged when we replace the polynomials φ_j $(1 \leqslant j \leqslant r)$ by any collection of linearly independent integral linear combinations, and thus we may suppose that these polynomials take the shape

$$\varphi_j(\mathbf{z}) = a_j \sigma_{k_j}(\mathbf{z}) - \sum_{\substack{1 \leqslant l < k_j \\ l \in \mathcal{R}}} b_{jl} \sigma_l(\mathbf{z}) \quad (1 \leqslant j \leqslant r), \tag{2.2}$$

where $a_j \in \mathbb{Z} \setminus \{0\}$ and $b_{jl} \in \mathbb{Z}$.

Given a solution $\mathbf{x}, \mathbf{y}$ of the system (1.1) counted by $N_k(X; \boldsymbol{\varphi})$, we define the integers $h_l = h_l(\mathbf{x}, \mathbf{y})$ for $l \in \mathcal{R}$ by putting

$$h_l(\mathbf{x}, \mathbf{y}) = \sigma_l(\mathbf{x}) - \sigma_l(\mathbf{y}). \tag{2.3}$$

Thus, since $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant k)$, one has $|h_l(\mathbf{x}, \mathbf{y})| \leqslant 2^k X^l$ $(l \in \mathcal{R})$. By making use of the relations (2.2) and (2.3), the system (1.1) becomes

$$a_j \left(\sigma_{k_j}(\mathbf{x}) - \sigma_{k_j}(\mathbf{y}) \right) = \sum_{\substack{1 \leqslant l < k_j \\ l \in \mathcal{R}}} b_{jl} h_l \quad (1 \leqslant j \leqslant r). \tag{2.4}$$

Then, by wielding (1.2) in combination with (2.3) and (2.4), we obtain

$$\begin{aligned} \prod_{i=1}^k (t + x_i) - \prod_{i=1}^k (t + y_i) &= \sum_{m=0}^k t^{k-m} \left(\sigma_m(\mathbf{x}) - \sigma_m(\mathbf{y}) \right) \\ &= \sum_{m \in \mathcal{R}} h_m t^{k-m} + \sum_{j=1}^r a_j^{-1} t^{k-k_j} \sum_{\substack{1 \leqslant l < k_j \\ l \in \mathcal{R}}} b_{jl} h_l. \end{aligned} \tag{2.5}$$

Put $A = a_1 a_2 \cdots a_r$ and $c_j = A / a_j$ $(1 \leqslant j \leqslant r)$. Also, define

$$\psi_l(t) = A t^{k-l} + \sum_{\substack{1 \leqslant j \leqslant r \\ k_j > l}} c_j b_{jl} t^{k-k_j} \quad (l \in \mathcal{R}),$$

and then set

$$\Psi(t; \mathbf{h}) = \sum_{l \in \mathcal{R}} h_l \psi_l(t). \tag{2.6}$$

Notice here that the polynomial $\Psi(t; \mathbf{h})$ has integral coefficients and degree at most $k - 1$ with respect to t. Moreover, it follows from (2.5) that

$$A \Bigg(\prod_{i=1}^k (t + x_i) - \prod_{i=1}^k (t + y_i) \Bigg) = \Psi(t; \mathbf{h}). \tag{2.7}$$

We may now record the multiplicative relation employed in our proof of Theorem 1.1. Suppose that $\mathbf{x}, \mathbf{y}$ is an integral solution of the system (1.1), where $\varphi_1, \dots, \varphi_r$ are symmetric polynomials as described above. Then, with $h_l = \sigma_l(\mathbf{x}) - \sigma_l(\mathbf{y})$ $(l \in \mathcal{R})$, we deduce by substituting $t = -y_j$ into (2.7) that

$$A \prod_{i=1}^{k} (x_i - y_j) = \Psi(-y_j; \mathbf{h}) \quad (1 \leqslant j \leqslant k). \tag{2.8}$$

The proof of Theorem 1.1. We divide the solutions of the system (1.1) with $1 \leqslant x_i, y_i \leqslant X$ into two types. A solution $\mathbf{x}, \mathbf{y}$ will be called *potentially diagonal* when $\{x_1, \dots, x_k\} = \{y_1, \dots, y_k\}$, and *non-diagonal* when, for some index j with $1 \leqslant j \leqslant k$, one has either $y_j \notin \{x_1, \dots, x_k\}$ or $x_j \notin \{y_1, \dots, y_k\}$.

We first consider potentially diagonal solutions $\mathbf{x}, \mathbf{y}$. Suppose, if possible, that $\mathbf{x}, \mathbf{y}$ satisfies the condition that the polynomial $\Psi(t; \mathbf{h})$ defined in (2.6) is identically zero as a polynomial in t. Since $A \neq 0$, it follows from (2.7) that

$$\prod_{i=1}^{k} (t + x_i) = \prod_{i=1}^{k} (t + y_i).$$

The roots of the polynomials on the left and right hand sides here must be identical, and so too must be their respective multiplicities. Thus $(x_1, \dots, x_k)$ is a permutation of $(y_1, \dots, y_k)$, and it follows that the number of solutions $\mathbf{x}, \mathbf{y}$ of this type counted by $N_k(X; \boldsymbol{\varphi})$ is precisely $T_k(X)$.

Suppose next that $\mathbf{x}, \mathbf{y}$ is a potentially diagonal solution with $\Psi(t; \mathbf{h})$ not identically zero, where $\mathbf{h}$ is defined via (2.3). In this situation, there are distinct integers $w_1, \dots, w_s$, for some integer s with $1 \leqslant s \leqslant k$, such that

$$\{x_1, \dots, x_k\} = \{y_1, \dots, y_k\} = \{w_1, \dots, w_s\}. \tag{2.9}$$

Consider any tuple of integers $\mathbf{h}$ with $|h_l| \leqslant 2^k X^l$ $(l \in \mathcal{R})$ for which the polynomial $\Psi(t; \mathbf{h})$ is not identically zero. This polynomial has degree at most $k - 1$, and hence there is an integer ξ with $1 \leqslant \xi \leqslant k$ for which the integer $\Theta = \Psi(\xi; \mathbf{h})$ is non-zero. We see from (2.7) that $(\xi + w_1) \cdots (\xi + w_s)$ divides Θ. Thus, an elementary divisor function estimate (see for example [6, Theorem 317]) shows the number of possible choices for $\xi + w_1, \dots, \xi + w_s$ to be at most

$$\sum_{\substack{d_1, \dots, d_s \in \mathbb{N} \\ d_1 \cdots d_s | \Theta}} 1 \leqslant \Bigl(\sum_{d | \Theta} 1 \Bigr)^s = d(|\Theta|)^s \ll |\Theta|^{\varepsilon}.$$

Since $\Theta = O(X^k)$, we see that there are at most $O(X^{\varepsilon})$ possible choices for $\xi + w_1, \dots, \xi + w_s$, and hence also for $w_1, \dots, w_s$. From here, the relation (2.9) implies that for these fixed choices of $\mathbf{h}$, the number of possible choices for $\mathbf{x}$ and $\mathbf{y}$ is also $O(X^{\varepsilon})$. On recalling (1.5), we discern that the total number of choices for the tuple $\mathbf{h}$ with $|h_l| \leqslant 2^k X^l$ $(l \in \mathcal{R})$ is $O(X^{w(\boldsymbol{\varphi})})$. Let $T_k^*(X; \boldsymbol{\varphi})$ denote the number of potentially diagonal solutions $\mathbf{x}, \mathbf{y}$ not counted by $T_k(X)$. For each of the $O(X^{w(\boldsymbol{\varphi})})$ choices for $\mathbf{h}$ associated with these solutions, we have shown that there are $O(X^{\varepsilon})$ solutions $\mathbf{x}, \mathbf{y}$, whence

$$T_k^*(X; \boldsymbol{\varphi}) \ll X^{w(\boldsymbol{\varphi}) + \varepsilon}. \tag{2.10}$$

Finally, suppose that $\mathbf{x}, \mathbf{y}$ is a non-diagonal solution of the system (1.1). By invoking symmetry (twice), we may suppose that $y_k \notin \{x_1, \dots, x_k\}$, whence (2.8) shows the integer $\Theta = \Psi(-y_k; \mathbf{h})$ to be non-zero. There are $O(X^{w(\boldsymbol{\varphi})})$ possible choices for $\mathbf{h}$ and $O(X)$ possible choices for y_k corresponding to this situation. Fixing any one such, and noting that $\Theta = O(X^k)$, an elementary divisor function estimate again shows that there are $O(X^\varepsilon)$ possible choices for $x_1 - y_k, \dots, x_k - y_k$ satisfying (2.8). Fix any one such choice for these divisors. Since y_k is already fixed, it follows that $x_1, \dots, x_k$ are likewise fixed.

It remains to determine $y_1, \dots, y_{k-1}$. Since the tuple $\mathbf{h}$ has already been fixed and $\Psi(t; \mathbf{h})$ has degree at most $k-1$ with respect to t, the polynomial

$$A \prod_{i=1}^{k} (x_i - t) - \Psi(-t; \mathbf{h})$$

has degree k with respect to t, and has all of its coefficients already fixed. It therefore follows from (2.8) that there are at most k choices for each of the variables $y_1, \dots, y_{k-1}$. Let $T_k^\dagger(X; \boldsymbol{\varphi})$ denote the number of non-diagonal solutions $\mathbf{x}, \mathbf{y}$ counted by $N_k(X; \boldsymbol{\varphi})$. Then we may conclude that

$$T_k^\dagger(X; \boldsymbol{\varphi}) \ll X^{w(\boldsymbol{\varphi})+1+\varepsilon}. \tag{2.11}$$

On recalling our opening discussion together with the estimates (2.10) and (2.11), we arrive at the upper bound

$$N_k(X; \boldsymbol{\varphi}) - T_k(X) = T_k^*(X; \boldsymbol{\varphi}) + T_k^\dagger(X; \boldsymbol{\varphi}) \ll X^{w(\boldsymbol{\varphi})+1+\varepsilon}. \tag{2.12}$$

This delivers the first conclusion of Theorem 1.1. On recalling the definition (1.5) of $w(\boldsymbol{\varphi})$, it follows that when $k_1 + \ldots + k_r \geqslant \frac{1}{2}k(k-1) + 2$, one has $w(\boldsymbol{\varphi}) \leqslant k - 2$, and thus the second conclusion of Theorem 1.1 is immediate from (2.12) and the asymptotic formula $T_k(X) = k!X^k + O(X^{k-1})$. $\square$

The proof of Corollary 1.2. Write

$$\varphi_j(\mathbf{z}) = \sigma_j(z_1, \dots, z_k) + \sum_{l=1}^{k-r} a_{jl}\sigma_l(z_1, \dots, z_k) \quad (k - r + 1 \leqslant j \leqslant k).$$

Then we see that the system (1.6) is comprised of symmetric polynomials having degrees $k - r + 1, \dots, k$. For this system, it follows from (1.5) that

$$w(\boldsymbol{\varphi}) = \tfrac{1}{2}k(k+1) - \sum_{j=k-r+1}^{k} j = \tfrac{1}{2}(k-r)(k-r+1).$$

We therefore deduce from Theorem 1.1 that

$$M_{k,r}(X; \mathbf{a}) = N_k(X; \boldsymbol{\varphi}) = T_k(X) + O(X^{\frac{1}{2}(k-r)(k-r+1)+1+\varepsilon}),$$

and so there is a paucity of non-diagonal solutions in the system (1.6) provided that $(k-r)(k-r+1) < 2k-2$. This completes the proof of the corollary. $\square$

3. Non-linear variants

The system of Diophantine equations (1.1) underlying Theorem 1.1 and Corollary 1.2 possesses features inherently linear in nature. Indeed, as is evident from the discussion initiating §2, we are able to restrict attention to integral linear combinations of elementary symmetric polynomials of the shape (2.2). This is a convenient but not essential simplification, as we now explain.

We now describe the symmetric polynomials presently in our field of view. Recall the exponents $k_1,\ldots,k_r$ satisfying (1.4), the complementary set of exponents $\mathcal{R}=\mathcal{R}(\boldsymbol{\varphi})$ defined in (2.1), and the definition (1.5) of $w(\boldsymbol{\varphi})$. Putting $R=\mathrm{card}(\mathcal{R})$, we label the elements of $\mathcal{R}$ so that $\mathcal{R}=\{l_1,\ldots,l_R\}$. We consider symmetric polynomials in the variables $\mathbf{z}=(z_1,\ldots,z_k)$ of the shape

$$\varphi_j(\mathbf{z})=a_j\sigma_{k_j}(\mathbf{z})-\Upsilon_j(\sigma_{l_1}(\mathbf{z}),\ldots,\sigma_{l_R}(\mathbf{z}))\quad (1\leqslant j\leqslant r),\tag{3.1}$$

where $a_j\in\mathbb{Z}\setminus\{0\}$ and $\Upsilon_j\in\mathbb{Z}[s_1,\ldots,s_R]$.

Theorem 3.1. *Let $\varphi_1,\ldots,\varphi_r$ be symmetric polynomials of the shape* (3.1) *with respective degrees $k_1,\ldots,k_r$ satisfying* (1.4). *Then*

$$N_k(X;\boldsymbol{\varphi})=T_k(X)+O(X^{2w(\boldsymbol{\varphi})+1+\varepsilon}).$$

In particular, when $k_1+\ldots+k_r\geqslant \frac{1}{2}k^2+1$, one has

$$N_k(X;\boldsymbol{\varphi})=k!X^k+O(X^{k-1+\varepsilon}).$$

A specialisation again makes for accessible conclusions. Fix the polynomials $\Upsilon_j\in\mathbb{Z}[s_1,\ldots,s_R]$ for $k-r+1\leqslant j\leqslant k$, and denote by $L_{k,r}(X;\boldsymbol{\Upsilon})$ the number of solutions of the simultaneous equations

$$\sigma_j(\mathbf{x})+\Upsilon_j(\sigma_1(\mathbf{x}),\ldots,\sigma_{k-r}(\mathbf{x}))=\sigma_j(\mathbf{y})+\Upsilon_j(\sigma_1(\mathbf{y}),\ldots,\sigma_{k-r}(\mathbf{y}))\quad (k-r+1\leqslant j\leqslant k),\tag{3.2}$$

in variables $\mathbf{x}=(x_1,\ldots,x_k)$ and $\mathbf{y}=(y_1,\ldots,y_k)$ with $1\leqslant x_i,y_i\leqslant X$.

Corollary 3.2. *Suppose that $k,r\in\mathbb{N}$ satisfy $(k-r)(k-r+1)<k-1$. Then there is a paucity of non-diagonal solutions in the system* (3.2). *In particular,*

$$L_{k,r}(X;\boldsymbol{\Upsilon})=T_k(X)+O(X^{(k-r)(k-r+1)+1+\varepsilon}).$$

The proof of Theorem 3.1. Our present strategy is very similar to that wrought against Theorem 1.1. Given a solution $\mathbf{x},\mathbf{y}$ of the system (1.1) counted by $N_k(X;\boldsymbol{\varphi})$, define $h_m(\mathbf{z})=\sigma_{l_m}(\mathbf{z})$ for $\mathbf{z}\in\{\mathbf{x},\mathbf{y}\}$ and $1\leqslant m\leqslant R$. In view of (3.1), the system (1.1) now becomes

$$a_j(\sigma_{k_j}(\mathbf{x})-\sigma_{k_j}(\mathbf{y}))=\Upsilon_j(\mathbf{h}(\mathbf{x}))-\Upsilon_j(\mathbf{h}(\mathbf{y}))\quad (1\leqslant j\leqslant r).\tag{3.3}$$

In the present non-linear scenario there may be some index j for which the integer on the right hand side of (3.3) is not equal to $\Upsilon_j(\mathbf{h}(\mathbf{x})-\mathbf{h}(\mathbf{y}))$. The argument here therefore contains extra complications, with weaker quantitative conclusions than Theorem 1.1.

Put $A = a_1 a_2 \cdots a_r$ and $c_j = A/a_j$ $(1 \leqslant j \leqslant r)$. Then, by applying the identity (1.2) together with (3.3), we obtain for $1 \leqslant j \leqslant r$ the relation

$$A\Big(\prod_{i=1}^{k}(t+x_i) - \prod_{i=1}^{k}(t+y_i)\Big) = \Psi(t;\mathbf{h}) - \Psi(t;\mathbf{g}), \tag{3.4}$$

where $\mathbf{h} = (h_1(\mathbf{x}), \ldots, h_R(\mathbf{x}))$, $\mathbf{g} = (h_1(\mathbf{y}), \ldots, h_R(\mathbf{y}))$, and we write

$$\Psi(t;\mathbf{e}) = A \sum_{m \in \mathcal{R}} e_m t^{k-m} + \sum_{j=1}^{r} c_j t^{k-k_j} \Upsilon_j(\mathbf{e}).$$

If $\mathbf{x}, \mathbf{y}$ is a solution of (1.1) counted by $N_k(X;\boldsymbol{\varphi})$, then with $h_m = \sigma_{l_m}(\mathbf{x})$ and $g_m = \sigma_{l_m}(\mathbf{y})$ for $1 \leqslant m \leqslant R$, we deduce by setting $t = -y_j$ in (3.4) that

$$A\prod_{i=1}^{k}(x_i - y_j) = \Psi(-y_j;\mathbf{h}) - \Psi(-y_j;\mathbf{g}) \quad (1 \leqslant j \leqslant k). \tag{3.5}$$

We now follow the path already trodden in the proof of Theorem 1.1 presented in §2. Suppose first that $\mathbf{x}, \mathbf{y}$ is a potentially diagonal solution. When the polynomial $\Psi(-t;\mathbf{h}) - \Psi(-t;\mathbf{g})$ is identically zero, we find that $\mathbf{x}, \mathbf{y}$ is counted by $T_k(X)$. Meanwhile, when instead this polynomial is not identically zero, a divisor function argument shows that for each fixed choice of $\mathbf{h}$ and $\mathbf{g}$, there are $O(X^\varepsilon)$ possible choices for $\mathbf{x}$ and $\mathbf{y}$. We have $|h_m(\mathbf{x})| \leqslant 2^k X^{l_m}$ and $|h_m(\mathbf{y})| \leqslant 2^k X^{l_m}$, so the total number of choices for $(h_1, \ldots, h_R)$ and $(g_1, \ldots, g_R)$ is $O(X^{2w(\boldsymbol{\varphi})})$, where $w(\boldsymbol{\varphi})$ is defined by (1.5). Thus, the total number of potentially diagonal solutions is equal to $T_k(X) + O(X^{2w(\boldsymbol{\varphi})+\varepsilon})$.

Suppose next that $\mathbf{x}, \mathbf{y}$ is a non-diagonal solution of the system (1.1). By symmetry, we may again suppose that $y_k \notin \{x_1, \ldots, x_k\}$, and (3.5) shows that the integer $\Theta = \Psi(-y_j;\mathbf{h}) - \Psi(-y_j;\mathbf{g})$ is non-zero. There are $O(X^{2w(\boldsymbol{\varphi})})$ possible choices for $\mathbf{h}$ and $\mathbf{g}$, and $O(X)$ possible choices for y_k in this scenario. Fix any one such choice, and note from (3.5) that $\Theta = O(X^k)$. An elementary divisor function estimate shows there to be $O(X^\varepsilon)$ choices for $x_1 - y_k, \ldots, x_k - y_k$ satisfying (3.5). Fixing any one such choice fixes the integers $x_1, \ldots, x_k$. An argument essentially identical to that applied in the proof of Theorem 1.1 shows from here that there are $O(1)$ possible choices for $y_1, \ldots, y_{k-1}$. Hence, the number of non-diagonal solutions of the system (1.1) is $O(X^{2w(\boldsymbol{\varphi})+1+\varepsilon})$.

Combining the two contributions to $N_k(X;\boldsymbol{\varphi})$ that we have obtained yields

$$N_k(X;\boldsymbol{\varphi}) - T_k(X) \ll X^{2w(\boldsymbol{\varphi})+1+\varepsilon},$$

confirming the first conclusion of Theorem 3.1. The definition (1.5) of $w(\boldsymbol{\varphi})$, moreover, implies that when $k_1 + \ldots + k_r \geqslant \frac{1}{2}k^2 + 1$, one has $2w(\boldsymbol{\varphi}) \leqslant k-2$. The second conclusion of Theorem 3.1 therefore follows directly from the first. □

The proof of Corollary 3.2. When $k - r + 1 \leqslant j \leqslant k$, write

$$\varphi_j(\mathbf{z}) = \sigma_j(z_1, \ldots, z_k) + \Upsilon_j(\sigma_1(z_1, \ldots, z_k), \ldots, \sigma_{k-r}(z_1, \ldots, z_k)).$$

Then we see from (1.5) that the system (3.2) is comprised of symmetric polynomials with $w(\boldsymbol{\varphi}) = \frac{1}{2}(k-r)(k-r+1)$. Theorem 3.1 therefore shows that

$$L_{k,r}(X;\boldsymbol{\Upsilon}) = N_k(X;\boldsymbol{\varphi}) = T_k(X) + O(X^{(k-r)(k-r+1)+1+\varepsilon}),$$

and provided that $(k-r)(k-r+1) < k-1$, there is a paucity of non-diagonal solutions in the system (3.2). This completes the proof of the corollary. □

References

[1] J. Bourgain, M. Z. Garaev, S. V. Konyagin and I. E. Shparlinski, *Multiplicative congruences with variables from short intervals*, J. Anal. Math. **124** (2014), 117–147.

[2] J. Brüdern and O. Robert, *Rational points on linear slices of diagonal hypersurfaces*, Nagoya Math. J. **218** (2015), 51–100.

[3] A. Choudhry, *Symmetric Diophantine systems*, Acta Arith. **59** (1991), no. 3, 291–307.

[4] A. Gloden, *Mehrgradige Gleichungen*, P. Noordhoff, Groningen, 1944.

[5] G. Greaves, *Some Diophantine equations with almost all solutions trivial*, Mathematika **44** (1997), no. 1, 14–36.

[6] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6th edition, Oxford University Press, Oxford, 2008.

[7] W. Heap, A. Sahay and T. D. Wooley, *A paucity problem associated with a shifted integer analogue of the divisor function*, J. Number Theory, in press, arxiv:2108.00287; doi:10.1016/j.jnt.2022.05.006.

[8] D. R. Heath-Brown, *The density of rational points on cubic surfaces*, Acta Arith. **79** (1997), no. 1, 17–30.

[9] C. Hooley, *On another sieve method and the numbers that are a sum of two hth powers*, Proc. London Math. Soc. (3) **43** (1981), no. 1, 73–109.

[10] C. Hooley, *On another sieve method and the numbers that are a sum of two hth powers: II*, J. Reine Angew. Math. **475** (1996), 55–75.

[11] S. T. Parsell and T. D. Wooley, *A quasi-paucity problem*, Michigan Math. J. **50** (2002), no. 3, 461–469.

[12] P. Salberger, *Rational points of bounded height on threefolds*, Analytic Number Theory, Clay Math. Proc. **7** (2007), pp. 207–216, Amer. Math. Soc., Providence, RI.

[13] P. Salberger and T. D. Wooley, *Rational points on complete intersections of higher degree, and mean values of Weyl sums*, J. London Math. Soc. (2) **82** (2010), no. 2, 317–342.

[14] C. M. Skinner and T. D. Wooley, *On the paucity of non-diagonal solutions in certain diagonal Diophantine systems*, Quart. J. Math. Oxford (2) **48** (1997), no. 2, 255–277.

[15] W. Y. Tsui and T. D. Wooley, *The paucity problem for simultaneous quadratic and biquadratic equations*, Math. Proc. Cambridge Philos. Soc. **126** (1999), no. 2, 209–221.

[16] R. C. Vaughan and T. D. Wooley, *On a certain nonary cubic form and related equations*, Duke Math. J. **80** (1995), no. 3, 669–735.

[17] R. C. Vaughan and T. D. Wooley, *A special case of Vinogradov's mean value theorem*, Acta Arith. **79** (1997), no. 3, 193–204.

[18] T. D. Wooley, *A note on symmetric diagonal equations*, Number theory with an emphasis on the Markoff spectrum (Provo, UT, 1991), Editors: A. D. Pollington and W. Moran, Dekker, New York, 1993, pp. 317–321.

[19] T. D. Wooley, *An affine slicing approach to certain paucity problems*, Analytic Number Theory: Proceedings of a Conference in Honor of Heini Halberstam (B.C. Berndt, H. G. Diamond and A. J. Hildebrand, eds.), vol. 2, 1996, pp. 803–815, Prog. Math. **139**, Birkhäuser, Boston.

[20] T. D. Wooley, *Sums and differences of two cubic polynomials*, Monatsh. Math. **129** (2000), no. 2, 159–169.

[21] T. D. Wooley, *Paucity problems and some relatives of Vinogradov's mean value theorem*, submitted, arXiv:2107.12238.

Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA

E-mail address: twooley@purdue.edu