WARING'S PROBLEM WITH POLYNOMIAL SUMMANDS

KEVIN FORD

1. *Introduction*

Let $f(x)$ be an integer-valued polynomial with no fixed integer divisor $\geqslant 2$, that is, for no integer $d \geqslant 2$ does $d | f(x)$ for all integers x. One generalization of the famous Waring problem is to determine whether, for large enough s, the equation

$$f(x_1)+f(x_2)+\ldots+f(x_s) = n \tag{1.1}$$

is solvable in positive integers $x_1, \ldots, x_s$ for sufficiently large integers n. The existence of such s for every f was established by Kamke [**5**] in 1921. Subsequent authors (Pillai, Hua [**2–4**], Vinogradov, Načaev [**7**], and others) have studied the problem of bounding $G(f)$, the least s for which (1.1) is solvable for all large n.

Questions of local solubility of (1.1), that is, solubility of the congruence

$$f(x_1)+f(x_2)+\ldots+f(x_s) \equiv n \pmod{q}, \tag{1.2}$$

play a more important and complicated role in this problem than in the classical Waring problem. Let $\Gamma_0(f)$ denote the least number s so that (1.2) is solvable for every pair n, q. It is well known that $\Gamma_0(x^k) \leqslant 4k$ for every k, but Hua [**4**] found that for every k, the polynomial

$$f_k(x) = \sum_{j=1}^{k} (-1)^{k-j} 2^{j-1} \binom{x}{j}$$

satisfies $\Gamma_0(f_k) \geqslant 2^k - 1$ (take $s = 2^k - 2$, $q = 2^k$ and $n = (-1)^k$ in (1.2)). Clearly $G(f) \geqslant \Gamma_0(f)$, but one can say more by restricting the values of n under consideration, as has been done by several authors in the case $f(x) = x^4$ (for example [**1, 6**]).

The singular series

$$\mathfrak{S}_{s,f}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left(\frac{1}{q} \sum_{r=1}^{q} e(af(r)/q) \right)^s e(-an/q),$$

where $e(z) = e^{2\pi i z}$, encapsulates the local solubility information. In particular, $\mathfrak{S}_{s,f}(n) \geqslant 0$ for every n and $\mathfrak{S}_{s,f}(n) > 0$ if and only if (1.2) is soluble for every q. Define $\bar{G}(f)$ to be the least number s so that for every $\delta > 0$ and every $n > n_0(\delta)$ with $\mathfrak{S}_{s,f}(n) \geqslant \delta$, (1.1) is soluble. The reason for taking $\mathfrak{S}_{s,f}(n) \geqslant \delta$ instead of $\mathfrak{S}_{s,f}(n) > 0$ is that we wish to exclude from consideration certain n lying in sparse sequences for which (1.1) is insoluble but $\mathfrak{S}_{s,f}(n) > 0$. For example, taking $f(x) = x^4$, $s = 15$ and $n_j = 79 \cdot 16^j$ $(j = 0, 1, \ldots)$, it can be shown that (1.1) is not soluble for $n = n_j$, that $\mathfrak{S}_{s,f}(n_j) > 0$ for all j, and that $\mathfrak{S}_{s,f}(n_j) \to 0$ as $j \to \infty$. It is known that $G(x^4) = 16$ (see [**1**]) and that $\bar{G}(x^4) \leqslant 11$ almost holds (see [**6**]).

Received 1 February 1999; revised 17 May 1999.

2000 *Mathematics Subject Classification* 11A26, 11N64 (primary).

J. London Math. Soc. (2) 61 (2000) 671–680

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

It has been known for several decades that a standard diminishing ranges argument combined with Vinogradov's bounds for exponential sums yields

$$\bar{G}(f) \leqslant k(4\log k + O(\log\log k)) \tag{1.3}$$

uniformly for all f of degree k. In the classical Waring problem, Wooley [**8**] has shown that

$$G(x^k) \leqslant k(\log k + O(\log\log k)).$$

The new methods which have been successful in reducing the upper bounds in the classical case, however, do not apply directly in the more general case. The only improvement to (1.3) has been by Wooley [**10**], who has reduced the upper bound for $\bar{G}(f)$ for certain lacunary polynomials f (polynomials with few non-zero coefficients). For the precise statement, let

$$f(x) = \sum_{i=1}^{t} a_i x^{k_i}, \quad k = k_1 > k_2 > \ldots > k_t, \tag{1.4}$$

where the numbers a_i are integers, $a_1 > 0$ and $k \geqslant 2$. We do not lose any generality in assuming that the a_i are integers, for (1.1) is soluble for a given function $f(x) = g(x)$ and given $n = m$ if and only if the corresponding equation with $f(x) = dg(x)$ is soluble with $n = dm$.

THEOREM 1.1 [**10**, Theorem 9]. *We have*

(i) $\bar{G}(f) \leqslant 2k(\log k + \log t + \log\log k + O(1))$;

(ii) $\bar{G}(f) \leqslant (1+o(1))k\log k$ *as* $k \to \infty$ *provided that either*

(a) $t = o(\sqrt{\log k})$ *and* $\log(k_2 k_3 \ldots k_t) = o(\log k)$; *or*

(b) $k_2 = o(\log k/\log\log k)$.

This gives an upper bound for $\bar{G}(f)$ of the same strength as in the classical setting when both t and k_2 are very small. The purpose of this paper is to show that $\bar{G}(f) \leqslant (1+o(1))k\log k$ for a wider class of polynomials, in particular those where $k - k_t$ is small.

THEOREM 1.2. *With f defined by* (1.4), *we have*

$$\bar{G}(f) \leqslant k(\log k + 2\log t + \log(k-k_t) + \log\log k + O(1)).$$

The main idea is to modify the efficient differencing method of [**8**] to handle general polynomial summands, and use these bounds together with the exponential sum bounds in [**10**] and the diminishing ranges method.

2. *A mean value theorem*

Let

$$f(x) = \sum_{i=1}^{t} a_i x^{k_i},$$

where

$$k = k_1 > k_2 > \ldots > k_t = k_1 - r \geqslant 1$$

and each a_i is a non-zero integer. We also assume that $a_1 > 0$, $k \geqslant 4$, $r \leqslant k/2$ and $t \geqslant 2$. With f fixed, define for positive integers w

$$g(x;w) = w^{-k_t} f(xw) = \sum_{i=1}^{t} w^{k_i - k_t} a_i x^{k_i}.$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

Let $\mathscr{A}(P,R)$ denote the set of positive integers $n \leqslant P$ which have no prime factors exceeding R. Let $I_s(P,R;w)$ denote the number of solutions of

$$\sum_{i=1}^{s} (g(x_i;w)-g(y_i;w)) = 0, \quad x_i, y_i \in \mathscr{A}(P,R)\ (1 \leqslant i \leqslant s).$$

We say a number λ_s is admissible if, for every $\varepsilon > 0$, there is a positive $\eta = \eta(\varepsilon, s)$ such that if $R \leqslant P^{\eta}$ then

$$I_s(P,R;w) \ll P^{\lambda_s+\varepsilon} \qquad \text{uniformly in } w \geqslant 1.$$

The implied constant may depend on ε, s and f only.

In the following, vectors are denoted by boldface symbols with subscripts indicating their dimension, for example $\boldsymbol{m}_j$ stands for $(m_1, m_2, \dots, m_j)$. If $\Psi(z) = \Psi(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)$ is a polynomial in $2j+2$ variables with non-negative coefficients, let

$$S_s(P,Q,R) = S_s(P,Q,R;\Psi;\boldsymbol{m}_j;w)$$

be the number of solutions of

$$\Psi(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)+\sum_{i=1}^{s} g(x_i;wm_1\dots m_j) = \Psi(z';\boldsymbol{h}'_j,\boldsymbol{m}_j,w)+\sum_{i=1}^{s} g(y_i;wm_1\dots m_j) \quad (2.1)$$

with

$$P/2 \leqslant z, z' \leqslant P;\ \forall\, i{:}\, 1 \leqslant h_i, h'_i \leqslant P/m_i^{k_t};\ x_i, y_i \in \mathscr{A}(Q,R).$$

Also, let

$$T_s(P,Q,R,M) = T_s(P,Q,R,M;\Psi;\boldsymbol{m}_{j+1};w)$$

denote the number of solutions of

$$\Psi(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)-\Psi(z';\boldsymbol{h}_j,\boldsymbol{m}_j,w) + m_{j+1}^{k_t}\sum_{i=1}^{s}(g(x_i;wm_1\dots m_{j+1})-g(y_i;wm_1\dots m_{j+1})) = 0 \quad (2.2)$$

with

$$P/2 \leqslant z,\ z' \leqslant P;\ z \equiv z' (\mathrm{mod}\, m_{j+1}^{k_t});\ \forall\ i{:}\, 1 \leqslant h_i \leqslant P/m_i^{k_t};\ x_i, y_i \in \mathscr{A}(Q/M,R).$$

From now on it will be assumed that $\Psi(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)$, interpreted as a polynomial in z with coefficients depending on $\boldsymbol{h}_j,\boldsymbol{m}_j,w$, has degree between 2 and k and the leading coefficient is positive whenever the variables h_i, m_i, w are all positive. The notation $o(1)$ will mean a function which tends to 0 as $P \to \infty$, the rate of which depends only on f and s. We also write $\Psi'(z)$ or $\Psi'(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)$ for the partial derivative of Ψ with respect to z, and assume that it is non-zero whenever $P/2 \leqslant z \leqslant P$ and each variable h_i, m_i, w is positive. We shall also suppose that $R = P^{\eta}$ for some fixed $\eta > 0$ depending on f, s, ε. Implied constants depend only on f, s, η, ε unless otherwise specified.

Our first lemma is a direct generalization of [**8**, Lemma 2.2].

LEMMA 2.1. *If* $1 < M < Q < P$ *then*

$$S_s(P,Q,R;\Psi,\boldsymbol{m}_j;w) \ll \left(\prod_{i=1}^{j} P/m_i^{k_t}\right)(MR)^{2s}P^{o(1)} \max_{M \leqslant m_{j+1} \leqslant MR} T_s(P,Q,R,M;\Psi;\boldsymbol{m}_{j+1};w).$$

Proof. As in [**8**], the notation $s_0(n)$ stands for the square-free kernel of n, that is, $s_0(n) = \prod_{p|a} p$, and $x\mathscr{D}(M)\,y$ means some $d \mid x$, $d \leqslant M$ satisfies $s_0(x/d) \mid y$.

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

We partition the solutions $z, z', \boldsymbol{h}_j, \boldsymbol{h}'_j, \boldsymbol{x}_s, \boldsymbol{y}_s$ of (2.1) into three classes: S_1 is the number of solutions with some $x_i \leqslant M$ or some $y_i \leqslant M$, S_2 is the number of remaining solutions with $x_i \mathscr{D}(M)\Psi'(z)$ or $y_i \mathscr{D}(M)\Psi'(z')$ for some i, and S_3 is the number of solutions not counted in S_1 or S_2. Clearly $S_s(P,Q,R) \leqslant 3\max(S_1, S_2, S_3)$.

Case 1: S_1 is maximal. Let

$$u(\alpha; L) = \sum_{x \in \mathscr{A}(L,R)} e(\alpha g(x; wm_1 \dots m_j)),$$
$$v(\alpha) = \sum_{z, \boldsymbol{h}_j} e(\alpha \Psi(z; \boldsymbol{h}_j, \boldsymbol{m}_j, w)).$$

Then

$$S_s(P,Q,R) \ll \int_0^1 |v(\alpha)^2 u(\alpha; M)\, u(\alpha; Q)^{2s-1}|\, d\alpha$$

and an application of Hölder's inequality gives

$$S_s(P,Q,R) \ll (S_s(P,M,R))^{1/(2s)} (S_s(P,Q,R))^{1-1/(2s)},$$

whence

$$S_s(P,Q,R) \ll S_s(P,M,R) \leqslant M^2 S_{s-1}(P,Q,R).$$

Case 2: S_2 is maximal. Since the variable w plays no role in the estimation of $S_s(P,Q,R)$ in this case, the analysis of [**8**] may be used without modification, giving

$$S_s(P,Q,R) \ll QMP^{o(1)} S_{s-1}(P,Q,R).$$

Case 3: S_2 is maximal. As in [**8**], for each solution of (2.1) counted by S_3, there are numbers $c_i \mid x_i$ and $d_i \mid y_i$ $(1 \leqslant i \leqslant s)$ with $M < c_i$, $d_i \leqslant MR$ and $(c_i, \Psi'(z)) = (d_i, \Psi'(z')) = 1$ for each i. Define the generating functions

$$u_d(\alpha) = \sum_{x \in \mathscr{A}(Q/d, R)} e(\alpha g(dx; wm_1 \dots m_j)),$$
$$v_d(\alpha) = \sum_{\substack{z, \boldsymbol{h}_j \\ (d, \Psi'(z))=1}} e(\alpha \Psi(z; \boldsymbol{h}_j, \boldsymbol{m}_j, w)).$$

Writing $C = c_1 \dots c_s$ and $D = d_1 \dots d_s$ and applying Hölder's inequality twice gives

$$\begin{aligned} S_s(P,Q,R) &\ll \sum_{\boldsymbol{c}, \boldsymbol{d}} \int_0^1 v_C(\alpha)\, v_D(\alpha) \prod_{i=1}^s u_{c_i}(\alpha)\, u_{d_i}(\alpha)\, d\alpha \\ &\leqslant \sum_{\boldsymbol{c}, \boldsymbol{d}} \prod_{i=1}^s \left(\int_0^1 |v_C^2\, u_{c_i}^{2s}| \right)^{1/2s} \left(\int_0^1 |v_D^2\, u_{d_i}^{2s}| \right)^{1/2s} \\ &\leqslant \sum_{\boldsymbol{c}, \boldsymbol{d}} \prod_{i=1}^s \left(\int_0^1 |v_{c_i}^2\, u_{c_i}^{2s}| \right)^{1/2s} \left(\int_0^1 |v_{d_i}^2\, u_{d_i}^{2s}| \right)^{1/2s} \\ &\leqslant \left(\sum_{M < d \leqslant MR} \left(\int_0^1 |v_d^2\, u_d^{2s}| \right)^{1/2s} \right)^{2s} \\ &\leqslant (MR)^{2s-1}\, V, \end{aligned}$$

where V is the number of solutions of

$$\Psi(z; \boldsymbol{h}_j, \boldsymbol{m}_j, w) - \Psi(z'; \boldsymbol{h}'_j, \boldsymbol{m}_j, w) + \sum_{i=1}^s (g(dx_i; m_1 \dots m_j w) - g(dy_i; m_1 \dots m_j w)) = 0$$

with

$$P/2 \leqslant z,\, z' \leqslant P;\ M < d \leqslant MR;\ (d, \Psi'(z)) = (d, \Psi'(z')) = 1$$
$$\forall\ i: 1 \leqslant h_i,\, h'_i \leqslant P/m_i^{k_i},\ x_i, y_i \in \mathscr{A}(Q/M, R).$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

We now make the critical observation that

$$g(dx; m_1 \dots m_j w) = d^{k_t} g(x; m_1 \dots m_j wd),$$

which implies that

$$\Psi(z) \equiv \Psi(z') \,(\mathrm{mod}\, d^{k_t}).$$

For some constant K, depending only on k, the number of solutions of $\Psi(z) \equiv u \,(\mathrm{mod}\, d^{k_t})$ with $(d, \Psi'(z)) = 1$ is $\ll d^{K/\log\log d} \ll P^{o(1)}$. As in [**8**], an application of the Cauchy–Schwarz inequality reduces the problem to the case where $z \equiv z' \,(\mathrm{mod}\, d^{k_t})$ and $\boldsymbol{h}_j = \boldsymbol{h}'_j$, introducing a factor $HMRP^{o(1)}$, where $H = \prod_{i=1}^{j} (P/m_i^{k_t})$. Writing $m_{j+1} = d$, and recalling (2.2), we obtain

$$S_s(P, Q, R) \ll H(MR)^{2s} P^{o(1)} \max_{M < m_{j+1} \leqslant MR} T_s(P, Q, R, M; \Psi; \boldsymbol{m}_{j+1}; w).$$

Combining the three cases yields

$$S_s(P, Q, R) \ll QMP^{o(1)} S_{s-1}(P, Q, R) + H(MR)^{2s} P^{o(1)} \max_{m_{j+1}} T_s(P, Q, R, M).$$

If the second term on the right is the larger, we are done. If the first term dominates, by Hölder's inequality (as in case 1)

$$\begin{aligned} S_{s-1}(P, Q, R) &\ll \left(\int_0^1 |v(\alpha)^2 u(\alpha; Q)^{2s}| \, d\alpha \right)^{1-1/s} \left(\int_0^1 |v(\alpha)|^2 \, d\alpha \right)^{1/s} \\ &\ll (S_s(P, Q, R))^{1-1/s} (PH^2)^{1/s}, \end{aligned}$$

thus

$$S_s(P, Q, R) \ll PH^2 (QMP^{o(1)})^s. \tag{2.3}$$

On the other hand, counting only the trivial solutions of (2.2), that is, solutions with $x_i = y_i$ for each i, $z = z'$ and $\boldsymbol{h}_j = \boldsymbol{h}'_j$ gives the lower bound

$$T_s(P, Q, R, M) \geqslant PH |\mathscr{A}(Q/M, R)|^s \gg PH(Q/M)^s.$$

Combined with (2.3) this yields

$$S_s(P, Q, R) \ll HM^{2s} \max_{m_{j+1}} T_s(P, Q, R, M; \Psi; \boldsymbol{m}_{j+1}; w)$$

and the lemma follows in this case. □

Suppose now that

$$0 < \phi_j \leqslant \frac{1}{k_t} \quad (1 \leqslant j \leqslant k_t - 1)$$

and for each j set

$$M_j = P^{\phi_j}, \quad H_j = PM_j^{-k_t}, \quad \theta_j = \phi_1 + \dots + \phi_j, \quad Q_j = P^{1-\theta_j}.$$

As in [**8**], define the modified forward difference operator Δ_j by

$$\Delta_0(p(z)) = p(z), \quad \Delta_1(p(z); h_1, m_1) = \frac{1}{m_1^{k_t}} (p(z + h_1 m_1^{k_t}) - p(z))$$

and

$$\Delta_j(p(z); \boldsymbol{m}_j; \boldsymbol{h}_j) = \frac{1}{(m_1 \dots m_j)^{k_t}} \sum_{\substack{\varepsilon_1, \dots, \varepsilon_j \\ \varepsilon_i \in \{0, 1\}}} (-1)^{j + \varepsilon_1 + \dots + \varepsilon_j} p(z + \varepsilon_1 h_1 m_1^{k_t} + \dots + \varepsilon_j h_j m_j^{k_t})$$

for $j \geqslant 1$. Now define polynomials Ψ_j, in $2j+2$ variables, by

$$\begin{aligned} \Psi_0(z, w) &= g(z; w), \\ \Psi_j(z; \boldsymbol{h}_j, \boldsymbol{m}_j, w) &= \Delta_j(g(z; w); \boldsymbol{m}_j; \boldsymbol{h}_j) \quad (j \geqslant 1). \end{aligned}$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

It is straightforward to show that

$$\Psi_j(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)=(h_1\dots h_j)\sum_{i=1}^{t} w^{k_i-k_t}a_i\sum_{\boldsymbol{l}}\binom{k_i}{l_0,\dots,l_j}z^{l_0}(h_1m_1^{k_t})^{l_1-1}\dots(h_jm_j^{k_t})^{l_j-1},$$

where the inner sum is over $\boldsymbol{l}=(l_0,l_1,\dots,l_j)$ satisfying $l_0+\dots+l_j=k_i$, $l_0\geqslant 0$ and $l_i\geqslant 1$ $(i=1,2,\dots,j)$. As a polynomial in z the leading term is $a_1w^{k-k_t}h_1\dots h_jz^{k-j}$. Also, if each $0<h_i\leqslant P/m_i^{k_t}$, $m_i>0$, $w>0$, $P/2\leqslant z\leqslant P$ and P is sufficiently large depending only on f, we have $\Psi_j'(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w)>0$.

Further define

$$S_{s,j}(\boldsymbol{m}_j)=S_s(P,Q_j,R;\Psi_j;\boldsymbol{m}_j;w)$$

and

$$T_{s,j}(\boldsymbol{m}_{j+1})=T_s(P,Q_j,R,M_{j+1};\Psi_j;\boldsymbol{m}_{j+1};w).$$

In the following we assume that w is fixed and that $M_j\leqslant m_j\leqslant M_jR$ for each j.

By Lemma 2.1 we obtain

$$S_{s,j}(\boldsymbol{m}_j)\ll\left(\prod_{i=1}^{j}H_i\right)(M_{j+1}R)^{2s}P^{o(1)}\max_{M_{j+1}\leqslant m_{j+1}\leqslant M_{j+1}R}T_{s,j}(\boldsymbol{m}_{j+1}).$$

LEMMA 2.2. *We have*

$$T_{s,j}(\boldsymbol{m}_{j+1})\ll P\left(\prod_{i=1}^{j}H_i\right)I_s(Q_{j+1},R;wm_1\dots m_{j+1})+\{I_s(Q_{j+1},R;wm_1\dots m_{j+1})\,S_{s,j+1}(\boldsymbol{m}_{j+1})\}^{1/2}.$$

Proof. We have $T_{s,j}(\boldsymbol{m}_{j+1})\leqslant U_0+2U_1$, where U_0 is the number of solutions of (2.2) with $z=z'$ and U_1 is the number of solutions of (2.2) with $z<z'$. Clearly

$$U_0\leqslant P\left(\prod_{i=1}^{j}H_i\right)I_s(Q_{j+1},R;wm_1\dots m_{j+1}).$$

To bound U_1, we set

$$z'=z+h_{j+1}m_{j+1}^{k_t}$$

so that $1\leqslant h_{j+1}\leqslant H_{j+1}$ and

$$m_{j+1}^{-k_t}(\Psi_j(z';\boldsymbol{h}_j,\boldsymbol{m}_j,w)-\Psi_j(z;\boldsymbol{h}_j,\boldsymbol{m}_j,w))=\Psi_{j+1}(z;\boldsymbol{h}_{j+1},\boldsymbol{m}_{j+1},w).$$

Thus (2.2) becomes

$$\Psi_{j+1}(z;\boldsymbol{h}_{j+1},\boldsymbol{m}_{j+1},w)+\sum_{i=1}^{s}(g(x_i;m_1\dots m_{j+1}w)-g(y_i;m_1\dots m_{j+1}w))=0.$$

Write

$$v(\alpha)=\sum_{z,\boldsymbol{h}_{j+1}}e(\alpha\Psi_{j+1}(z;\boldsymbol{h}_{j+1},\boldsymbol{m}_{j+1},w)),$$

$$u(\alpha)=\sum_{x\in\mathscr{A}(Q_{j+1},R)}e(\alpha g(x;m_1\dots m_{j+1}w)).$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned}U_1&=\int_0^1 v(\alpha)\,|u(\alpha)|^{2s}\,d\alpha\\&\leqslant\left(\int_0^1|u(\alpha)|^{2s}\right)^{1/2}\left(\int_0^1|v(\alpha)|^2\,|u(\alpha)|^{2s}\right)^{1/2}\\&=\{I_s(P,Q_{j+1};wm_1\dots m_{j+1})\,S_{s,j+1}(\boldsymbol{m}_{j+1})\}^{1/2}.\end{aligned}$$

□

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

LEMMA 2.3. *Suppose that* $\lambda_s = 2s - k_t + \Delta$ *is admissible and* $\Delta > 0$. *Set*

$$\theta = \frac{1}{k_t+\Delta} + \left(\frac{1}{k_t} - \frac{1}{k_t+\Delta}\right)\left(\frac{k_t-\Delta}{2k_t}\right)^{k_t-1},$$
$$\Delta' = \Delta(1-\theta) + k_t\theta - 1,$$
$$\lambda_{s+1} = 2(s+1) - k_t + \Delta'.$$

Then λ_{s+1} *is admissible*.

Proof. When $\Delta = 0$ then $\Delta' = 0$ and the conclusion follows by expressing $I_s(P, R; w)$ as an integral of exponential sums and making a trivial estimate. When $\Delta > 0$ the lemma will follow by iterating Lemmas 2.1 and 2.2 as in [**8**, Lemma 3.1] using the choice

$$\phi_j = \frac{1}{k_t+\Delta} + \left(\frac{1}{k_t} - \frac{1}{k_t+\Delta}\right)\left(\frac{k_t-\Delta}{2k_t}\right)^{k_t-j} \quad (1 \leqslant j \leqslant k_t). \tag{2.4}$$

We note that (2.4) and the bound $\Delta \leqslant k_t$ imply that

$$\frac{1}{2k_t} \leqslant \phi_j \leqslant \frac{1}{k_t} \quad \forall\, j.$$

Also, since $k \geqslant 3$ and $\Delta > 0$,

$$\theta_{k_t} = 1 - \frac{(k_t-2)\Delta + \Delta^2 + 2\Delta(\frac{1}{2} - \Delta/2k_t)^{k_t-1}}{(k_t+\Delta)^2} < 1.$$

Fix $\varepsilon > 0$. Since λ_s is admissible, there is an $\eta > 0$ so that whenever $R \leqslant Q^\eta$ we have $I_s(Q, R; w) \ll Q^{\lambda_s+\varepsilon}$ uniformly in $w \geqslant 1$. Let η' be a positive real number satisfying

$$\eta' < (1-\theta_{k_t})\eta, \quad \eta' < \frac{\varepsilon}{4sk_t} \tag{2.5}$$

and put $R = P^{\eta'}$. Then, for each j, $I_s(Q_j, R; w) \ll Q_j^{\lambda_s+\varepsilon}$ uniformly in $w \geqslant 1$. We next show by induction on j that

$$T_{s,j}(\boldsymbol{m}_{j+1}) \ll P\left(\prod_{i=1}^{j} H_i\right) Q_{j+1}^{\lambda_s+\varepsilon}, \tag{2.6}$$

where the implied constant may depend on j as well as f, s, ε. First, when $j = k_t - 1$ we have $H_{j+1} = 1$. Representing $S_{s,j+1}(\boldsymbol{m})$ as an integral of exponential sums and taking a trivial bound gives

$$S_{s,j+1}(\boldsymbol{m}_{j+1}) \leqslant P^2\left(\prod_{i=1}^{j} H_i\right)^2 I_s(Q_{j+1}, R; wm_1 \dots m_{j+1})$$
$$\ll P^2\left(\prod_{i=1}^{j} H_i\right)^2 Q_{j+1}^{\lambda_s+\varepsilon}.$$

By Lemma 2.2, (2.6) holds for $j = k_t - 1$. Now suppose that $j \leqslant k_t - 2$ and (2.6) holds for all larger j. By Lemma 2.1 and the induction hypothesis,

$$S_{s,j+1}(\boldsymbol{m}_{j+1}) \ll P^{1+o(1)}\left(\prod_{i=1}^{j+1} H_i\right)^2 (M_{j+2}R)^{2s} Q_{j+2}^{\lambda_s+\varepsilon}.$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

Hence, by Lemma 2.2,

$$T_{s,j}(\boldsymbol{m}_{j+1}) \ll P\left(\prod_{i=1}^{j} H_i\right) Q_{j+1}^{\lambda_s+\varepsilon}\left\{1+R^s P^{-1/2+o(1)} H_{j+1} M_{j+2}^s \left(\frac{Q_{j+2}}{Q_{j+1}}\right)^{(1/2)(\lambda_s+\varepsilon)}\right\}.$$

By (2.4) and (2.5) the term in the braces is $1+P^{\xi+o(1)}$, where

$$\begin{aligned}\xi &= -\tfrac{1}{2}+1-k_t\phi_{j+1}+s\phi_{j+2}-\tfrac{1}{2}\phi_{j+2}(\lambda_s+\varepsilon)+\eta' s\\ &= -\tfrac{1}{2}\varepsilon\phi_{j+2}+\eta' s<0.\end{aligned}$$

Therefore (2.6) holds for $0\leqslant j\leqslant k_t-1$. Finally, by (2.5), (2.6) with $j=0$ and Lemma 2.1,

$$I_{s+1}(P,R;w)\leqslant S_{s,0}(-)\ll (M_1R)^{2s}P^{1+o(1)}Q_1^{\lambda_s+\varepsilon}=P^{L+o(1)},$$

where, since $\theta=\phi_1$,

$$L=2s\theta+2s\eta'+1+\theta(\lambda_s+\varepsilon)=\lambda_{s+1}+2s\eta'<\lambda_{s+1}+\varepsilon.$$

Since ε is arbitrary, the lemma follows. □

The admissible values of λ_s arising from Lemma 2.3 are precisely the values of λ_s arising in [**8**] for the monomial $f(x)=x^{k_t}$. Thus, by the proof of [**9**, Theorem 2.1], it follows that these values satisfy $\lambda_s<2s-k_t+k_te^{1-2s/k_t}$. Hence we have the following.

COROLLARY 2.4. *For each $s\geqslant 1$, there is an $\eta=\eta(s)>0$ such that, for $R\leqslant P^{\eta}$, uniformly in $w\geqslant 1$,*

$$I_s(P,R;w)\ll P^{2s-k_t+k_te^{1-2s/k_t}}.$$

REMARK 2.5. In Corollary 2.4, the exponent of P tends to $2s-k_t=2s-k+r$ as $s\to\infty$, which is short by r of the 'ideal' exponent $2s-k$ needed in the application to the generalized Waring problem. Fortunately, if r is small, we can overcome this deficiency with a diminishing ranges argument.

REMARK 2.6. We in fact can do a bit better than Corollary 2.4, by taking advantage of the form of $g(x;w)$ when w becomes large. For example, if $w\geqslant c(f,s)Q^k$, the only solutions counted in $I_s(Q,R;w)$ are the 'diagonal' solutions corresponding to $x_i=y_i$ for each i and thus $I_s(Q,R;w)=|\mathscr{A}(Q,R)|^s\ll Q^s$. Using such ideas, however, does not appear to break the $2s-k_t$ barrier mentioned in Remark 2.5, or increase significantly the rate at which the exponent approaches $2s-k_t$ as s increases. Thus, in the application to Waring's problem, only the second order terms will be improved.

3. *Application to Waring's problem*

Proof of Theorem 1.2. Suppose that k is large and $r\leqslant k/3$ (else Theorem 1.2 is no stronger than [**10**, Theorem 9]). Let n be a large integer, $P=[(n/a_1)^{1/k}]$ and s,u be fixed positive integers. For $i=1,\dots,u$ set

$$P_i=2^{-i}P^{(1-1/k)^{i-1}}$$

and let

$$Q=\frac{1}{s2^{u+1}}P^{(1-1/k)^u}.$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

Define the generating functions

$$D(\alpha)=\sum_{x=1}^{P} e(\alpha f(x)),\quad F(\alpha)=\sum_{x\in\mathscr{A}(P,R)} e(\alpha f(x)),\quad H(\alpha)=\sum_{x\in\mathscr{A}(Q,R)} e(\alpha f(x))$$

and

$$G(\alpha)=E_1(\alpha)\dots E_u(\alpha),\quad E_i(\alpha)=\sum_{P_i<x\leqslant 2P_i} e(\alpha f(x))\quad (i=1,\dots,u).$$

Let $R(n)$ denote the number of solutions of

$$n=f(w_1)+f(w_2)+\sum_{i=1}^{6k} f(x_i)+\sum_{i=1}^{2s} f(y_i)+\sum_{i=1}^{u}(f(z_i)+f(z_i')) \tag{3.1}$$

with

$$1\leqslant w_1,\ w_2\leqslant P,\ \forall\ i: x_i\in\mathscr{A}(P,R),\ y_i\in\mathscr{A}(Q,R),\ P_i<z_i,\ z_i'\leqslant 2P_i.$$

Then

$$R(n)=\int_0^1 D(\alpha)^2 F(\alpha)^{6k} G(\alpha)^2 H(\alpha)^{2s} e(-n\alpha)\,d\alpha.$$

We now define the major and minor arcs. For $2\leqslant a\leqslant q\leqslant a_1 P^{1/2}R$, $(a,q)=1$ define

$$\mathfrak{M}(q,a)=\left\{\alpha\in[0,1]: |\alpha-a/q|\leqslant\frac{P^{1/2}}{qP^k}\right\}.$$

Also let

$$\mathfrak{M}(1,1)=[0,P^{1/2-k}]\cup[1-P^{1/2-k},1],\quad \mathfrak{M}=\bigcup_{a,q}\mathfrak{M}(q,a),\quad \mathfrak{m}=[0,1]\backslash\mathfrak{M}.$$

The major arcs are handled as in Wooley [**10**, §9], where the function $\mathscr{F}_{1,s}(\alpha)$ there is replaced by $G(\alpha)^2H(\alpha)^{2s}$. This gives

$$\int_{\mathfrak{M}} D(\alpha)^2 F(\alpha)^{6k} G(\alpha)^2 H(\alpha)^{2s} e(-n\alpha)\,d\alpha \gg D(0)^2 F(0)^{6k} G(0)^2 H(0)^{2s} P^{-k}(\mathfrak{S}(n)+o(1)),$$

where $\mathfrak{S}(n)=\mathfrak{S}_{2s+2u+6k+2,f}(n)$. For the minor arcs, we have

$$\left|\int_{\mathfrak{m}} D(\alpha)^2 F(\alpha)^{6k} G(\alpha)^2 H(\alpha)^{2s} e(-n\alpha)\,d\alpha\right|\leqslant P^2\sup_{\alpha\in\mathfrak{m}}|F(\alpha)|^{6k}\int_0^1 |H(\alpha)^s G(\alpha)|^2\,d\alpha. \tag{3.2}$$

By [**10**, Theorem 7], for large k we have

$$\sup_{\alpha\in\mathfrak{m}}|F(\alpha)|\ll P^{1-1/(6tk\log k)}.$$

The integral on the right-hand side of (3.2) counts the number of solutions of

$$\sum_{i=1}^{u}(f(z_i)-f(z_i'))+\sum_{i=1}^{s}(f(y_i)-f(y_i'))=0$$

with

$$P_i<z_i,\ z_i'\leqslant 2P_i,\quad y_i,y_i'\in\mathscr{A}(Q,R).$$

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press

It is straightforward to show that $z_i = z_i'$ for $i = 1, \dots, u$, which is a consequence of our choice of the numbers P_i. By Corollary 2.4 we obtain

$$\begin{aligned}\int_0^1 |H(\alpha)^s G(\alpha)|^2 \, d\alpha &\ll P_1 \dots P_u \int_0^1 |H(\alpha)|^{2s} \, d\alpha \\ &\ll P_1 \dots P_u Q^{2s-(k-r)+\Delta} \\ &\ll H(0)^{2s} G(0)^2 P^{-K_1}, \quad K_1 = k - (\Delta + r)(1 - 1/k)^u\end{aligned}$$

where $\Delta = (k-r) e^{1-2s/(k-r)}$. Therefore, the left-hand side of (3.2) is

$$\ll D(0)^2 F(0)^{6k} G(0)^2 H(0)^{2s} P^{-K_2}, \quad K_2 = \frac{1}{t \log k} + k - (\Delta + r) e^{-u/k}.$$

We now choose

$$s = [\tfrac{1}{2}(k-r)\log(k/r - 1)] + 1, \quad u = [k \log(4rt \log k)] + 1,$$

which implies that $\Delta \leqslant er$ and $K_2 \geqslant k + 0.07/t \log k$. Therefore, for P large,

$$R(n) \gg D(0)^2 F(0)^{6k} G(0)^2 H(0)^{2s} P^{-k}(\mathfrak{S}(n) + o(1))$$

and hence

$$\bar{G}(f) \leqslant 6k + 2s + 2u + 2 \leqslant k(\log(krt^2) + 2 \log\log k + 10). \qquad \square$$

References

1. H. Davenport, 'On Waring's problem for fourth powers', *Ann. of Math.* 40 (1939) 731–747.
2. L.-K. Hua, 'On Waring's problem with polynomial summands', *J. Chinese Math. Soc.* 1 (1936) 21–61.
3. L.-K. Hua, 'On a generalized Waring's problem', *Proc. London Math. Soc.* (2) 43 (1937) 161–182.
4. L.-K. Hua, 'On a generalized Waring problem', *J. Chinese Math. Soc.* 2 (1940) 175–191.
5. E. Kamke, 'Verallgemeinerungen des Waring–Hilbertschen Satzes', *Math. Ann.* 83 (1921) 85–112.
6. K. Kawada and T. D. Wooley, 'Sums of fourth powers and related topics', *J. Reine Angew. Math.* 512 (1999) 173–223.
7. V. I. Načaev, 'Waring's problem for polynomials', *Trudy Mat. Inst. Steklov* 38 (1951) 190–243 (Russian); *Amer. Math. Soc. Transl. Ser.* 2 3 (1956) 39–89 (English).
8. T. D. Wooley, 'Large improvements in Waring's problem', *Ann. of Math.* 135 (1992) 131–164.
9. T. D. Wooley, 'The application of a new mean value theorem to the fractional parts of polynomials', *Acta Arith.* 65 (1993) 163–179.
10. T. D. Wooley, 'On exponential sums over smooth numbers', *J. Reine Angew. Math.* 488 (1997) 79–140.

Department of Mathematics
University of South Carolina
Columbia
SC 29208
USA

ford@math.sc.edu

https://doi.org/10.1112/S002461070000884X Published online by Cambridge University Press