JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 9, Number 4, October 1996

THE REPRESENTATION OF NUMBERS AS SUMS OF UNLIKE POWERS. II

KEVIN B. FORD

1. Introduction

In a previous paper ([Fo]), the author proved that every sufficiently large integer is representable in the form

$$n = \sum_{i=1}^{15} x_i^{i+1}, \tag{1.1}$$

where the numbers x_i are nonnegative integers. In an addendum to that paper, the author announced an improvement, for which we now supply a detailed proof. Our main result is

Theorem 1. *Every sufficiently large natural number n is representable in the form*

$$n = \sum_{i=1}^{14} x_i^{i+1}.$$

The principal tool in the proof is the Hardy-Littlewood circle method, incorporating results and techniques of a powerful new iterative method developed by Vaughan and Wooley ([Va3], [Va4], [VW1], [Wo1], [VW2]).

In section 2, an algorithm developed in [Fo] for optimizing the parameters in mixed power mean value theorems is generalized and analyzed. Section 3 details a more sophisticated method of generating mixed power mean value theorems, by a limited adaptation of the iterative method itself. The form of these estimates offers many advantages over those of section 2, and provides the key to the elimination of the 16th power from (1.1). These mean value theorems are then applied to the proof of Theorem 1 in section 4. The tools developed here are applicable to a wide range of mixed power representation problems, and we briefly illustrate in section 5 the application to the problem of determining the number of terms required to represent all large n, when the lowest power used is a kth power instead of a square.

Throughout, n is a large natural number whose representation as a sum of mixed powers is at issue, and ε is an arbitrarily small positive real number. Constants implied by the Landau O- and Vinogradov $\ll$-symbols may depend on ε or k. For a real number x, write $e(x)$ for $e^{2\pi ix}$, $[x]$ for the greatest integer not exceeding x and $||x||$ for the distance from x to the nearest integer. Unless otherwise specified, lowercase Latin letters denote natural numbers and Greek letters denote real numbers. Let $\mathcal{A}(P, R)$ denote the set of natural numbers not exceeding P with no

Received by the editors August 18, 1994 and, in revised form, March 31, 1995.
1991 *Mathematics Subject Classification.* Primary 11P05.

prime factor exceeding R. For a given representation problem, we take $R = n^\eta$ for some $\eta > 0$. Many assertions, especially (2.1) below, are valid for η less than some bound, which may depend on ε. We will take η to be as small as desired, and will frequently state $R \ll n^\varepsilon$ without comment. Since the number of such assertions is finite, there is no danger of losing control of implicit constants.

Let

$$P_k = \frac{1}{2} n^{1/k}$$

for each exponent k appearing in the representation, and define the generating functions

$$f_k(\alpha) = \sum_{m \in \mathcal{B}_k} e(\alpha m^k),$$

where $\mathcal{B}_k$ is a subset of the integers between 1 and $2P_k$. Our choices for sets $\mathcal{B}_k$ will be either $(P_k, 2P_k]$ or $\mathcal{A}(P_k, R)$, both of which are "full size", meaning that $|\mathcal{B}_k| \gg P_k$. The fact that $|\mathcal{A}(P_k, R)| \gg P_k$ is a classical sieve result, and a proof may be found in [DB]. Let $F(\alpha) = f_{k_1}(\alpha) f_{k_2}(\alpha) \cdots f_{k_r}(\alpha)$. Then

$$R(n) = \int_0^1 F(\alpha) e(-n\alpha) d\alpha \tag{1.2}$$

is the number of representations of n in the form

$$n = x_1^{k_1} + x_2^{k_2} + \cdots + x_r^{k_r} \qquad (x_i \in \mathcal{B}_{k_i}). \tag{1.3}$$

Showing $R(n) > 0$ for large n is accomplished by partitioning $[0, 1]$ into the "major arcs" $\mathfrak{M}$ and "minor arcs" $\mathfrak{m}$. The definition of $\mathfrak{M}$ varies from problem to problem, but always adheres to a special form. For $Y \geqslant 1$, let

$$\mathfrak{M}(Y) = \bigcup_{q \leqslant Y} \bigcup_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} \mathfrak{M}(Y; q, a), \tag{1.4}$$

where

$$\mathfrak{M}(Y; q, a) = \left\{ \alpha \in [0,1] : \left|\left| \alpha - \frac{a}{q} \right|\right| \leqslant \frac{Y}{nq} \right\}. \tag{1.5}$$

When $Y < \frac{1}{2} n^{1/2}$, the intervals $\mathfrak{M}(Y; q, a)$ are pairwise disjoint, hence any $\alpha \in \mathfrak{M}(Y)$ uniquely determines a, q and $\beta = ||\alpha - a/q||$. Throughout, whenever $\alpha \in \mathfrak{M}(Y)$ is given, we assume these definitions of a, q and β. Observe also that $Y < X$ implies $\mathfrak{M}(Y) \subset \mathfrak{M}(X)$.

To handle the minor arcs, we write

$$\int_{\mathfrak{m}} |F(\alpha)| d\alpha \leqslant \sup_{\alpha \in \mathfrak{m}} |F_1(\alpha)| \int_0^1 |F_2(\alpha)| d\alpha, \tag{1.6}$$

where $F(\alpha) = F_1(\alpha) F_2(\alpha)$. The supremum of F_1 is estimated by means of either Weyl's inequality for small k or an estimate such as ([Wo1], Theorem 1.4) for large

k. The Cauchy-Schwarz inequality may be used to break the integral of $|F_2|$ into two "mean-square" integrals which may be estimated in an elementary way by consideration of the underlying diophantine equations (cf. Theorem 3 of [Th] and Theorem 6.2 of [Va1]). The techniques developed in the next two sections provide an alternative method of estimating these integrals.

For the circle method to be successful, one must show that the left side of (1.6) is smaller in order of magnitude than $F(0)n^{-1}$, the expected order of $R(n)$. Incidentally, the form of (1.6) necessitates the condition

$$\sum_{i=1}^{r} \frac{1}{k_i} > 2$$

in order to obtain such a minor arc bound, even assuming a best possible upper bound for $\int_0^1 |F_2|$ (see [HL], Hypothesis K). This condition implies that proving all large n have a representation

$$n = \sum_{i=1}^{10} x_i^{i+1}$$

is the theoretical limit of the circle method.

2. Smooth mean value theorems for mixed powers

Throughout this section,

$$f_k(\alpha) = \sum_{m \in \mathcal{A}(P_k, R)} e(\alpha m^k).$$

The works of Vaughan and Wooley ([Va3], [Va4], [Wo1], [VW2]) provide, for each pair of positive integers (k, s), an estimate

$$\int_0^1 |f_k(\alpha)|^{2s} d\alpha \ll P_k^{\lambda(k,s)} \tag{2.1}$$

valid for $0 < \eta \leqslant \eta(k, s)$. Note that we may take $\lambda(k, 1) = 1$ and $\lambda(k, 2) = 2 + \varepsilon$. The first is a consequence of Parseval's identity, and the second is deduced by an elementary consideration of the underlying diophantine equation (see §6.1 of [Va1]). Table 1 lists all of the values of $\lambda(k, s)$ that will be required in the proof of Theorem 1. The values were calculated with 16 digit precision by computer, the final significant digit being rounded up. The values for $s = 3$ and $s = 4$ are given by Theorem 1.4 of [Va4]. The values for $5 \leqslant k \leqslant 9$ are taken from the appendix of [VW2]. The remaining values are the result of an iteration procedure (on s), applying at each step one of Lemma 2.2 of [Va4] (for $s = 5$ and some k), Lemma 3.2 of [Wo1] (for intermediate s), or inequality $(k - 2)$ of [Va3, §4] (for large s).

We can extend these mean value theorems to nonintegral s by a simple interpolation via Hölder's inequality. If h is an integer and $0 < \theta < 1$, we have

$$\int_0^1 |f_k(\alpha)|^{2(h+\theta)} d\alpha \leqslant \left(\int_0^1 |f_k(\alpha)|^{2h} d\alpha \right)^{1-\theta} \left(\int_0^1 |f_k(\alpha)|^{2h+2} d\alpha \right)^{\theta}. \tag{2.2}$$

TABLE 1

k	s	$\lambda(k,s)$	k	s	$\lambda(k,s)$	k	s	$\lambda(k,s)$
4	3	3.1861407	7	7	8.5410894	12	9	10.5917109
	4	4.5627210		12	17.2932208		10	12.0382168
	5	6.1876809	8	4	4.2289285		11	13.5346434
	6	8.0000001		5	5.5116307		12	15.0795792
5	3	3.1362571		7	8.3284883	13	5	5.2784087
	4	4.4386563		8	9.8579814		9	10.4462093
	5	5.9250797		14	20.3659701		10	11.8557131
	6	7.5417546	9	4	4.1822894		12	14.8136933
	7	9.2727289		5	5.4201075		13	16.3600526
	8	11.0773627		7	8.1447208		14	17.9492906
6	3	3.0909091		8	9.6154494	14	5	5.2589353
	4	4.3333334		16	23.4293887		6	6.4627737
	5	5.7246965	10	4	4.1636826		7	7.7092805
	6	7.2315633		5	5.4010244		10	11.7095544
	7	8.8505716		6	6.6996396		13	16.0997159
	8	10.5604127		9	10.9660666		14	17.6467017
	9	12.3536709	11	4	4.1372319	15	6	6.4217891
	10	14.2030055		5	5.3449419		11	12.9676813
	11	16.0860412		6	6.6133232		12	14.3984547
7	3	3.0639191		9	10.7541737		14	17.3790325
	4	4.2641175		21	31.4828795		15	18.9270975
	5	5.5891167	12	5	5.3159121		16	20.5121267
	6	7.0143820		8	9.1960407			

Thus, defining

$$\lambda(k,h+\theta)=(1-\theta)\lambda(k,h)+\theta\lambda(k,h+1) \tag{2.3}$$

extends (2.1) to all positive real s.

Now suppose μ_i are positive real numbers. If $x_1,\dots,x_r$ are positive numbers satisfying

$$x_1+\cdots+x_r=1, \tag{2.4}$$

then by Hölder's inequality, we have

$$S:=\int_0^1 |f_{k_1}^{\mu_1}\cdots f_{k_r}^{\mu_r}|^2 \leqslant \prod_{i=1}^r \left(\int_0^1 |f_{k_i}|^{2\mu_i/x_i}\right)^{x_i}. \tag{2.5}$$

It follows from (2.1), (2.3) and (2.5) that

$$S\ll n^{\phi}, \qquad \phi=\sum_{i=1}^r \frac{x_i\lambda(k_i,\mu_i/x_i)}{k_i}. \tag{2.6}$$

For a given set of exponents $\{k_i\}$, our goal is to minimize ϕ subject to (2.4). To this end, define for $x\in(0,1]$ the functions

$$g_i(x)=\frac{x\lambda(k_i,\mu_i/x)}{k_i} \qquad (1\leqslant i\leqslant r). \tag{2.7}$$

When $\mu_1 = \cdots = \mu_r = 1$, the algorithm of section 7 of [Fo] will find the optimum values x_i. The algorithm is identical in the general case, and is reproduced below. The Cauchy-Schwarz inequality gives

$$\int_0^1 |f_k|^{2h} \leqslant \left(\int_0^1 |f_k|^{2h-2}\right)^{1/2} \left(\int_0^1 |f_k|^{2h+2}\right)^{1/2}. \tag{2.8}$$

Combined with (2.3), this shows that each $\lambda(k,s)$ is convex as a function of s for $s > 0$. By Lemma 7.1 of [Fo], it follows that each function $g_i(x)$ is convex, and thus Lemma 7.2 of [Fo] implies that $\sum g_i(x_i)$ is minimized whenever

$$\min_i D^+ g_i(x_i) \geqslant \max_i D^- g_i(x_i), \tag{2.9}$$

where D^- and D^+ are the left and right differential operators, respectively.

The algorithm for finding such x_i is to iterate the following operation:

> Find i and j (with $i \neq j$) such that
>
> $$D^- g_j(x_j) - D^+ g_i(x_i)$$
>
> is maximal. If the maximal difference is zero, then by (2.9) the optimal values of x_i have been found. Otherwise, set $x_i = x_i + \delta$ and $x_j = x_j - \delta$, where δ is the least positive number for which (with the new values of x_i and x_j)
>
> $$D^- g_j(x_j) \leqslant D^+ g_i(x_i).$$

Since each $\lambda(k_i, s)$ is linear in s on each interval $[h, h+1]$ for natural numbers h, it follows from (2.7) that the functions $g_i(x)$ are piecewise linear with bends at the points $x = 1/h$ for positive integers h. In practice, $g_i'(x_i) \neq g_j'(x_j)$ for all i, j and all x_i, x_j within the range of interest. Consequently, if $x_1, \ldots, x_r$ minimize the sum in (2.6), then by (2.9), at most one of the numbers $1/x_i$ will be nonintegral. For the same reason, each step in the above algorithm will leave either $1/x_i$ or $1/x_j$ integral, so the algorithm will find the optimal set $x_1, \ldots, x_r$ in a finite number of steps.

The starting values of the algorithm are taken so that

$$\frac{k_i x_i}{\mu_i} = \frac{k_j x_j}{\mu_j} \tag{2.10}$$

for all i, j. This usually gives a value for ϕ in (2.6) that is close to optimal. Heuristically, the iteration procedure described above produces values of $\lambda(k,s)$ of the form

$$\lambda(k,s) \approx 2s - k + cke^{-2s/k} \tag{2.11}$$

for some constant c (see Theorem 2.1 of [Wo2] for a proof of an upper bound of this form). Considering the functions g_i to be continuously differentiable, the condition (2.9) becomes "$g_i'(x_i) = g_j'(x_j)$ for all i, j". By (2.7),

$$g_i(x) \approx \frac{2\mu_i}{k_i} - x + cxe^{-2\mu_i/(k_i x)},$$

so that

$$g_i'(x) \approx -1 + ce^{-2\mu_i/(k_i x)} \left(\frac{2\mu_i}{k_i x} + 1\right),$$

which is a function of $k_i x/\mu_i$. This justifies (2.10).

3. Adaptation of the new iterative method

In this section we show how the new iterative method of Vaughan and Wooley may be adapted in a limited manner to handle mixed powers. Throughout this section, $h = 3$ or $h = 4$, and $k \geqslant h$. The method presented below will apply in principle to all h, but it is most effective for smaller h.

Let $S_{k,s}^{(h)}(P)$ denote the number of solutions of

$$z_1^h + x_1^k + \cdots + x_s^k = z_2^h + y_1^k + \cdots + y_s^k$$

with

$$\begin{gathered} 1 \leqslant z_1, z_2 \leqslant P, \\ x_i, y_i \in \mathcal{A}(P^{h/k}, R) \qquad (1 \leqslant i \leqslant s). \end{gathered} \tag{3.1}$$

Let $0 < \theta < 1/k$, $M = P^\theta$, $H = P^{1-k\theta}$ and $Q = P^{h/k-\theta}$. Let $T_{k,s}^{(h)}(P;\theta)$ denote the number of solutions of the equation

$$z_1^h + m^k(u_1^k + \cdots + u_s^k) = z_2^h + m^k(v_1^k + \cdots + v_s^k)$$

with

$$\begin{gathered} u_i, v_i \in \mathcal{A}(Q, R) \qquad (1 \leqslant i \leqslant s), \\ M \leqslant m \leqslant MR, \\ 1 \leqslant z_1, z_2 \leqslant P, \quad z_1 \equiv z_2 \pmod{m^k}. \end{gathered}$$

By Lemma 2.2 of [Wo1] (the Fundamental Lemma),

$$S_{k,s}^{(h)}(P) \ll P^{h/k+\theta+\varepsilon} S_{k,s-1}^{(h)}(P) + P^{(2s-1)\theta+\varepsilon} T_{k,s}^{(h)}(P;\theta). \tag{3.2}$$

In practice the second term on the right side of (3.2) will dominate the first. Our estimation of $T_{k,s}^{(h)}(P;\theta)$ follows that in [Va4, §2]. Clearly

$$T_{k,s}^{(h)} = \sum_{|d| \leqslant H} U_d,$$

where U_d is the number of solutions with $z_1 = z_2 + dm^k$. Considering separately solutions counted in U_0, and letting $z = z_1 + z_2$, we obtain

$$T_{k,s}^{(h)}(P;\theta) \ll PMRQ^{\lambda(k,s)} + \int_0^1 F_h(\alpha) |f_k(2^k\alpha; Q)|^{2s} d\alpha, \tag{3.3}$$

where

$$f_k(\alpha; Q) = \sum_{m \in \mathcal{A}(Q,R)} e(\alpha m^k),$$

$$F_h(\alpha) = \sum_{M < m \leqslant MR} \sum_{d \leqslant H} \sum_{z \leqslant 2P} e(\alpha \Psi_h), \tag{3.4}$$

$$\Psi_h = \Psi_h(m, d, z) = m^{-k} \left((z + dm^k)^h - (z - dm^k)^h \right). \tag{3.5}$$

If $2^{1-h} \leqslant a \leqslant 1/2$, then by Hölder's inequality and (2.1), we have

$$\begin{aligned}\int_0^1 & F_h(\alpha)|f_k(2^k\alpha;Q)|^{2s}d\alpha \\ &\ll \left(\int_0^1 |f_k(2^k\alpha;Q)|^{2s/(1-a)}d\alpha\right)^{1-a}\left(\int_0^1 |F_h(\alpha)|^{1/a}d\alpha\right)^a \\ &\ll Q^{(1-a)\lambda(k,s/(1-a))}\left(\int_0^1 |F_h(\alpha)|^{1/a}d\alpha\right)^a.\end{aligned} \tag{3.6}$$

Combining (3.2), (3.3) and (3.6) yields

$$\begin{aligned}S_{k,s}^{(h)}(P) \ll{}& P^{h/k+\theta+\varepsilon}S_{k,s-1}^{(h)}(P) + P^{(2s-1)\theta+\varepsilon} \\ &\times\left\{PMRQ^{\lambda(k,s)} + Q^{(1-a)\lambda(k,s/(1-a))}\left(\int_0^1 |F_h|^{1/a}\right)^a\right\}.\end{aligned} \tag{3.7}$$

We now require estimates for the power moments of F_h appearing in (3.7).

Lemma 3.1. *If $1 \leqslant j \leqslant h-1$, then*

$$\int_0^1 |F_h(\alpha)|^{2^j}d\alpha \ll P^{2^j-j+\varepsilon}(MRH)^{2^j-1}(MR)^{e_j},$$

where

$$e_j = \begin{cases}0, & 1 \leqslant j \leqslant h-2,\\ 1, & j = h-1.\end{cases}$$

Proof. This follows from (2.13)–(2.15) of [Va4] and is similar to the proof of Hua's inequality ([Va1], Lemma 2.5). Only the case $k = h$ is treated in [Va4], but the exponent of m in (3.5) plays no role in the argument.

For the mean cube of $|F_3|$, we can do better than using Cauchy-Schwarz combined with the square and fourth power moment estimates of Lemma 3.1.

Lemma 3.2. *If $0 < \theta < \frac{3}{5k+6}$, then*

$$\int_0^1 |F_3(\alpha)|^3 d\alpha \ll P^{7/2+\varepsilon}M^{-(2k-3/2)}.$$

Proof. The case $k = 3$ is proven in §3 of [Va3], and the general case follows in a very similar manner. From (3.4), (3.5) and the Cauchy-Schwarz inequality, we have

$$|F_3(\alpha)|^2 \leqslant D(\alpha)E(\alpha), \tag{3.8}$$

where

$$D(\alpha) = \sum_{d\leqslant H}\left|\sum_{z\leqslant 2P} e(6\alpha d z^2)\right|^2$$

and

$$E(\alpha) = \sum_{d\leqslant H}\left|\sum_{M<m\leqslant MR} e(2\alpha d^3 m^{2k})\right|^2.$$

Now suppose $(a,q)=1$ and $\beta = |\alpha - a/q| \leqslant q^{-2}$. By Lemma 3.1 of [Va3],

$$D(\alpha) \ll P^{\varepsilon}\left(\frac{P^2H}{q(1+P^2H\beta)} + PH + q + P^2Hq\beta\right). \tag{3.9}$$

Further, if $M^k \leqslant X \leqslant M^kH^3$, $q \leqslant X$ and $\beta \leqslant (qX)^{-1}$, then by a slight modification of the proof of Lemma 3.4 of [Va3],

$$E(\alpha) \ll P^{\varepsilon}\left(\frac{HM^2}{q^{1/k}(1+M^{2k}H^3\beta)^{1/3}} + HM\right). \tag{3.10}$$

Incidentally, the condition $\theta < \frac{3}{5k+6}$ comes from the estimation of $E(\alpha)$. This is not a serious restriction in applications, as the optimal choice for θ is usually smaller. Let $\mathfrak{m}$ denote the set of points α in $[0,1]$ with the property that whenever there are a,q with $(a,q)=1$ and $|\alpha - a/q| \leqslant (PHq)^{-1}$ we have $q > P$, and let $\mathfrak{M} = [0,1]\backslash\mathfrak{m}$. If $\alpha \in [0,1]$, Dirichlet's Theorem (Lemma 2.1 of [Va1]) implies that there are a,q such that $(a,q)=1$, $|\alpha - a/q| \leqslant (PHq)^{-1}$ and $q \leqslant PH$. If $\alpha \in \mathfrak{m}$, then $q > P$ and thus (3.8), (3.9) and (3.10) imply

$$F_3(\alpha) \ll P^{\varepsilon}H(PM)^{1/2}.$$

Combined with the $j=1$ case of Lemma 3.1, we obtain

$$\int_{\mathfrak{m}} |F_3(\alpha)|^3 d\alpha \ll P^{\varepsilon}H^2(PM)^{3/2}. \tag{3.11}$$

If $\alpha \in \mathfrak{M}$, then α is in some interval

$$\mathfrak{M}(q,a) = \{\alpha : |\alpha - a/q| \leqslant (PHq)^{-1}\}$$

with $1 \leqslant a \leqslant q \leqslant P$ and $(a,q)=1$. Let $W = P^2H = M^{2k}H^3$. By (3.9) and (3.10),

$$F_3^2(\alpha) \ll P^{2+\varepsilon}H^2M\left(\frac{1}{q(1+W\beta)}\right)\left(\frac{M}{q^{1/k}(1+W\beta)^{1/3}} + 1\right).$$

Hence

$$\begin{aligned}\int_{\mathfrak{M}(q,a)} |F_3(\alpha)|^3 d\alpha &\ll P^{3+\varepsilon}H^3M^{3/2} \\ &\qquad \times \int_0^{\infty}\left(\frac{M^{3/2}}{q^{\frac{3}{2}(1+1/k)}(1+W\beta)^2} + \frac{1}{q^{3/2}(1+W\beta)^{3/2}}\right)d\beta \\ &\ll P^{3+\varepsilon}H^3M^{3/2}W^{-1}\left(M^{3/2}q^{-\frac{3}{2}(1+1/k)} + q^{-3/2}\right).\end{aligned}$$

Thus, summing on a and q,

$$\int_{\mathfrak{M}} |F_3(\alpha)|^3 d\alpha \ll P^{7/2+\varepsilon}H^3M^{3/2}W^{-1}. \tag{3.12}$$

The lemma now follows from (3.11) and (3.12).

Power moments of F_h for general a may be obtained by combining the estimates in the preceding two lemmas with Hölder's inequality. For a given value of a, the optimal value of θ is obtained by equating the exponents of P in the last two terms on the right side of (3.7). The best choice for a is always among the numbers

TABLE 2

h	k	s	a	θ	$\nu(h,k,s)$	$\alpha(h,k,s)$	$\alpha_1(h,k,s)$	$\alpha_2(h,k,s)$
3	4	2	1/2	0.0758695	2.6517390	0.7827536	0.7777777	0.7745321
	4	3	1/4	0.1319339	3.7608491	0.9130503	0.8839482	0.9096887
	4	4	1/4	0.1836074	5.0531506	0.9822831	0.9495585	0.9765398
	5	2	1/2	0.0408771	2.2817543	0.7060819	0.7142857	0.6894677
	5	3	1/3	0.0861487	3.1284620	0.8238460	0.8188939	0.8146398
	5	4	1/4	0.1191524	4.0875366	0.9041544	0.8879727	0.8999882
	5	5	1/4	0.1398833	5.1250612	0.9583129	0.9320412	0.9491339
	5	6	1/4	0.1571564	6.2256947	0.9914351	0.9618142	0.9883955
	7	3	1/2	0.0427554	2.4386414	0.7109290	0.7273291	0.6947423
	8	4	3/7	0.0482460	2.7677873	0.7440709	0.7604903	0.7321722
	8	5	1/3	0.0584464	3.3291908	0.8069364	0.8125386	0.7977859
	9	5	1/3	0.0476170	3.0247832	0.7695167	0.7830072	0.7588825
	10	6	1/3	0.0449987	3.2484013	0.7838662	0.7952274	0.7744844
	11	6	1/3	0.0379449	3.0080307	0.7548988	0.7716306	0.7446965
4	5	2	1/2	0.0545028	2.7090057	0.6227485	0.6029411	0.6163206
	5	3	1/4	0.0964020	3.7850763	0.7537309	0.7309310	0.7490315
	5	4	1/4	0.1309629	5.0173291	0.8456677	0.8256065	0.8452342
	5	5	1/6	0.1528162	6.3627777	0.9093055	0.8908813	0.9146962
	5	6	1/8	0.1658191	7.7726659	0.9568335	0.9373557	0.9626086
	6	2	1/2	0.0303030	2.3939394	0.5681818	0.5500000	0.5577206
	6	3	1/4	0.0725309	3.2716050	0.6820987	0.6673913	0.6757088
	6	4	1/4	0.0932991	4.2309857	0.7755869	0.7576923	0.7697900
	6	5	1/4	0.1120070	5.2953284	0.8428345	0.8257906	0.8426338
	7	2	1/2	0.0182626	2.1793824	0.5265829	0.5108695	0.5168487
	7	3	2/5	0.0542908	2.9102132	0.6295895	0.6179901	0.6224295
	7	4	1/4	0.0712835	3.7029455	0.7171207	0.7030926	0.7093090
	7	5	1/4	0.0840072	4.5643271	0.7874896	0.7707473	0.7828099
	12	8	1/4	0.0451245	4.3723723	0.7402402	0.7268689	0.7334236
	12	9	1/4	0.0490897	4.8942408	0.7764398	0.7619959	0.7723562
	13	9	1/4	0.0426035	4.5360362	0.7506063	0.7367782	0.7448091
	13	10	1/4	0.0462219	5.0243560	0.7823725	0.7682570	0.7788800
	14	10	1/4	0.0404149	4.6806444	0.7584103	0.7445539	0.7532918
	15	11	1/4	0.0384232	4.8050993	0.7653918	0.7514953	0.7609428
	15	12	1/4	0.0410742	5.2339635	0.7915091	0.7775425	0.7892236

$\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{8}$ or values for which $\frac{s}{1-a}$ is an integer (see the remarks at the conclusion of section 2). From (3.7) we thus obtain estimates of the form

$$S_{k,s}^{(h)}(P) \ll P^{\nu(h,k,s)}.$$

Values of $\nu(h,k,s)$ along with the corresponding choices for a and θ for various triples (h,k,s) are listed in Table 2. As with $\lambda(k,s)$, these values were calculated with 16 digit precision, the last significant figure being rounded up.

To illustrate the calculations, let us estimate $S_{4,3}^{(3)}(P)$. Taking $a = \frac{1}{4}$, it follows from (3.7) that

$$S_{4,3}^{(3)}(P) \ll P^{3/4+\theta+\nu(3,4,2)+\varepsilon} + P^{5\theta+\varepsilon}\left\{P^{1+\theta}Q^{\lambda(4,3)} + Q^{(3/4)\lambda(4,4)}\left(\int_0^1 |F_3|^4\right)^{1/4}\right\}.$$

By Lemma 3.1, Table 1 and the top row of Table 2,

$$S_{4,3}^{(3)}(P) \ll P^{3.4017390+\theta} + P^{3.3896056+2.81385930\theta} + P^{3.8165306-0.42204070\theta}.$$

Equating the last two terms on the right gives $\theta \approx 0.1319339$ and $S_{4,3}^{(3)}(P) \ll P^{3.7608492}$. The values in the table were computed with 16 digit precision, and this accounts for the slight discrepancy between this exponent and the value given in Table 2.

By considering the underlying diophantine equations (see (3.1)), for fixed h and k we have

$$\int_0^1 |f_h(\alpha)f_k^s(\alpha)|^2 d\alpha \leqslant S_{k,s}^{(h)}(2P_h) \ll P_h^{\nu(h,k,s)}, \tag{3.13}$$

where

$$\begin{aligned} f_h(\alpha) &= \sum_{P_h < m \leqslant 2P_h} e(\alpha m^h), \\ f_k(\alpha) &= \sum_{m \in \mathcal{A}(P_k,R)} e(\alpha m^k). \end{aligned} \tag{3.14}$$

The above bounds may be used to generate mixed-power mean value theorems of the sort described in the preceding section. By the same interpolation argument (cf. (2.2) and (2.3)), the definition of $\nu(h,k,s)$ can be extended to all positive real s. If x_i are nonnegative real numbers satisfying $x_1 + \cdots + x_r = 1$, then by Hölder's inequality,

$$\begin{aligned} \int_0^1 |f_h f_{k_1} \cdots f_{k_r}|^2 &\leqslant \prod_{i=1}^r \left(\int_0^1 \left|f_h f_{k_i}^{1/x_i}\right|^2\right)^{x_i} \\ &\ll \prod_{i=1}^r n^{(x_i/k_i)\nu(h,k_i,1/x_i)}. \end{aligned} \tag{3.15}$$

Since the functions $\nu(h,k,s)$ are convex as functions of s by the analog of (2.8), an algorithm identical to that described in the preceding section will find the optimum values of x_i in (3.15).

The inequality (3.13) may also be written as

$$\int_0^1 |f_h f_k^s|^2 \ll (f_h(0)f_k^s(0))^2\, n^{-\alpha(h,k,s)}, \tag{3.16}$$

where

$$\alpha(h,k,s) = \frac{2}{h} + \frac{2s}{k} - \frac{\nu(h,k,s)}{h}.$$

The number $\alpha(h,k,s)$ can be regarded as a measure of how much is saved over the trivial bound for the left side of (3.16). It also has an arithmetical interpretation. Let $N_{h,k,s}(X)$ denote the number of positive integers less than X which can be written as the sum of an hth power and s kth powers. A standard application of

the Cauchy-Schwarz inequality (cf. the introduction to Chapter 6 of [Va1]) shows that

$$N_{h,k,s}(X) \gg X^{\alpha(h,k,s)}. \tag{3.17}$$

The values of $\alpha(h,k,s)$ are also included in Table 2. The final two columns in the table are exponents for lower bounds for $N_{h,k,s}(X)$ obtained by two different methods. Davenport's diminishing ranges method ([Va1], Theorem 6.2), seeded with a mean value estimate of type (2.1), yields the bound $N_{h,k,s}(X) \gg X^{\alpha_1(h,k,s)}$. Combining (2.5) with the homogeneous mean value theorems (2.1) produces the bound $N_{h,k,s}(X) \gg X^{\alpha_2(h,k,s)}$.

When $h = 3$, Davenport's method gives the best result for small s, and the method of this section takes over for the larger s. When $h = 4$, inequality (3.16) is superior for both smaller s and some larger s, while applying Hölder's inequality to the homogeneous mean value theorems gives the best results for intermediate values of s.

Even when $\alpha(h,k,s)$ is smaller than $\alpha_1(h,k,s)$ or $\alpha_2(h,k,s)$, the structure of the generating function for $S_{k,s}^{(h)}(P)$ provides several advantages in applications to mixed power problems such as Theorem 1. These advantages are discussed in the next section.

4. The proof of Theorem 1

In the proof of the theorem in [Fo], the two most critical estimates are the minor arc bound and the estimate of the error resulting from the replacement of f_3 by its approximation W_3. They correspond to estimates (4.10) and (4.24) below. Together these bounds determine the optimal choice for the major arcs. As noted earlier, the key to the elimination of the 16th power is the use of the mean value theorems developed in the previous section. The chief advantage of these estimates is not the improvement in the estimates themselves as measured by the numbers $\alpha(h,k,s)$, but the form of the generating functions involved. In fact, in the application below some of the mean value theorems used are worse than those attainable by other methods (cf. (4.3) and Table 2). Using (3.15), the generating functions f_3 and f_4 are ordinary Weyl sums (see (4.1) below), which means that (4.19) may be used instead of the much weaker (4.21) on the major arcs. Also, in contrast to Davenport's method, none of the generating functions is "diminished", i.e. we have $f_k(0) \gg P_k$ for every k. The importance of this comes into play in replacing f_3 by W_3 in (4.24). The function F_3 appearing there may be taken to be "smaller" than in [Fo], and this strengthens the estimate for the mean value theorem of its "complement" F_4.

For $2 \leqslant k \leqslant 4$ define

$$f_k(\alpha) = \sum_{P_k < m \leqslant 2P_k} e(\alpha m^k), \tag{4.1}$$

and for $5 \leqslant k \leqslant 15$ define

$$f_k(\alpha) = \sum_{m \in \mathcal{A}(P_k, R)} e(\alpha m^k). \tag{4.2}$$

Let $\mathfrak{M} = \mathfrak{M}(n^\mu)$ where $\mu = 0.461039$. Let $F(\alpha) = f_2(\alpha)f_3(\alpha)\cdots f_{15}(\alpha)$ and define $F_1(\alpha)$ through $F_7(\alpha)$ as follows:

$$\begin{aligned} F_1 &= f_3f_7f_8f_9f_{10}f_{11} & F_5 &= f_{10}f_{12}f_{13}f_{14}f_{15}\\ F_2 &= f_4f_5f_6f_{12}f_{13}f_{14}f_{15} & F_6 &= f_6f_7f_{11}\\ F_3 &= f_8f_9f_{10}f_{11}f_{14} & F_7 &= f_5f_8f_9\\ F_4 &= f_4f_5f_6f_7f_{12}f_{13}f_{15} \end{aligned}$$

We first establish a number of mean value theorems of the type discussed in sections 2 and 3. The values of $\lambda(k,s)$ required are listed in Table 1, and the values of $\nu(h,k,s)$ required are listed in Table 2.

Lemma 4.1. *We have*

$$\int_0^1 |F_1(\alpha)|^2 d\alpha \ll F_1^2(0)n^{-0.777561}, \tag{4.3}$$

$$\int_0^1 |F_2(\alpha)|^2 d\alpha \ll F_2^2(0)n^{-0.761827}, \tag{4.4}$$

$$\int_0^1 |F_4(\alpha)|^2 d\alpha \ll F_4^2(0)n^{-0.795935}. \tag{4.5}$$

Proof. We utilize the mean value theorems (3.13), starting with Hölder's inequality in the general form

$$\int_0^1 \left| f_h \prod_{k\in\mathcal{K}} f_k \right|^2 \leqslant \prod_{k\in\mathcal{K}} \left(\int_0^1 |f_h f_k^{a_k}|^2 \right)^{1/a_k},$$

where $\sum 1/a_k = 1$. The optimal values of a_k are obtained by the algorithm described in section 2. For (4.3), with $h = 3$, the optimal values are $a_7 = 4, a_8 = \frac{60}{13}, a_9 = 5$ and $a_{10} = a_{11} = 6$. For (4.4), with $h = 4$, the optimal values are $a_5 = 3, a_6 = 4, a_{12} = \frac{1980}{227}, a_{13} = 9, a_{14} = 10$ and $a_{15} = 11$. Lastly, for (4.5), with $h = 4$, we take $a_5 = a_6 = 4, a_7 = 5, a_{12} = 9, a_{13} = 10$ and $a_{15} = \frac{45}{4}$.

Lemma 4.2. *We have*

$$\int_0^1 |F_3(\alpha)|^2 d\alpha \ll F_3^2(0)n^{-0.461039}, \tag{4.6}$$

$$\int_0^1 |F_5(\alpha)|^2 d\alpha \ll F_5^2(0)n^{-0.375420}, \tag{4.7}$$

$$\int_0^1 |F_6(\alpha)|^2 d\alpha \ll F_6^2(0)n^{-0.391744}, \tag{4.8}$$

$$\int_0^1 |F_6(\alpha)F_7^3(\alpha)|^2 d\alpha \ll F_6^2(0)F_7^6(0)n^{-0.968255}. \tag{4.9}$$

Proof. Here we use the homogeneous mean value theorems (2.1) along with Hölder's inequality in the general form

$$\int_0^1 \left| \prod_{k\in\mathcal{K}} f_k^{\mu_k} \right|^2 \leqslant \prod_{k\in\mathcal{K}} \left(\int_0^1 |f_k|^{2a_k} \right)^{\mu_k/a_k},$$

where $\sum \mu_k/a_k = 1$. As in Lemma 4.1 , the optimal values for a_k are obtained from the algorithm described in section 2. For (4.6), we take $a_8 = 4, a_9 = \frac{140}{29}, a_{10} = a_{11} = 5$ and $a_{14} = 7$. For (4.7), we take $a_{10} = 4, a_{12} = a_{13} = 5, a_{14} = \frac{60}{11}$ and

$a_{15} = 6$. For (4.8), we take $a_6 = \frac{12}{5}, a_7 = 3$ and $a_{11} = 4$. Finally, for (4.9), we take $a_5 = 8, a_6 = \frac{336}{31}, a_7 = 12, a_8 = 14, a_9 = 16$ and $a_{11} = 21$.

The final tool we need for the minor arcs is Weyl's inequality (Lemma 2.4 of [Va1]).

Lemma 4.3 (Weyl's Inequality). *Suppose that* $(a, q) = 1$, $|\alpha - a/q| < q^{-2}$, $\phi(x) = \alpha x^k + \alpha_1 x^{k-1} + \cdots + \alpha_{k-1}x + \alpha_k$ *and* $K = 2^{k-1}$. *Then*

$$\sum_{x=1}^{Q} e(\phi(x)) \ll Q^{1+\varepsilon}\left(\frac{1}{q} + \frac{1}{Q} + \frac{q}{Q^k}\right)^{1/K}.$$

Using the Cauchy-Schwarz inequality, (4.3), (4.4) and Lemma 4.3 to bound $f_2(\alpha)$, we obtain

$$\begin{aligned}\int_{\mathfrak{m}} |F(\alpha)|d\alpha &\leqslant \sup_{\alpha\in\mathfrak{m}} |f_2(\alpha)| \left(\int_0^1 |F_1(\alpha)|^2 d\alpha\right)^{1/2} \left(\int_0^1 |F_2(\alpha)|^2 d\alpha\right)^{1/2} \\ &\ll F(0)n^{-(\mu+.777561+.761827)/2+\varepsilon} \ll F(0)n^{-1.000213}.\end{aligned} \tag{4.10}$$

We now introduce several auxiliary functions that come into play on the major arcs. First set

$$S_k(q, a) = \sum_{m=1}^{q} e\left(\frac{am^k}{q}\right). \tag{4.11}$$

Next let

$$w_k(\theta) = \begin{cases} \displaystyle\sum_{n/2^k < m \leqslant n} \frac{1}{k} m^{1/k-1} e(\theta m), & k \leqslant 4, \\ \displaystyle\sum_{R^k < m \leqslant P_k^k} \frac{1}{k} m^{1/k-1} \varrho\left(\frac{\log m}{k \log R}\right) e(\theta m), & k \geqslant 5, \end{cases} \tag{4.12}$$

where ϱ is Dickman's function (see [DB]). The only property of ϱ that we require is that $\varrho(x) > 0$ for all positive x. For $2 \leqslant k \leqslant 15$, let

$$W_k(\alpha, q, a) = \frac{1}{q} S_k(q, a) w_k\left(\alpha - \frac{a}{q}\right). \tag{4.13}$$

For brevity, write

$$W_k(\alpha) = \begin{cases} W_k(\alpha, q, a), & \text{for } \alpha \in \mathfrak{M}, \\ 0, & \text{for } \alpha \in \mathfrak{m}, \end{cases} \tag{4.14}$$

and

$$\Delta_k(\alpha) = f_k(\alpha) - W_k(\alpha). \tag{4.15}$$

Upper bounds for W_k and Δ_k are required. For $k \leqslant 4$, Lemma 6.3 of [Va1] states

$$W_k(\alpha, q, a) \ll \left(\frac{n}{q}\right)^{1/k} (1 + n||\alpha - a/q||)^{-1} \tag{4.16}$$

and when $k \geqslant 5$, Lemma 5.4 of [Va3] states

$$W_k(\alpha, q, a) \ll \left(\frac{n}{q}\right)^{1/k} (1 + n||\alpha - a/q||)^{-1/k}. \tag{4.17}$$

When $\alpha \in \mathfrak{M}$ and $k \leqslant 4$, (4.16) gives

$$W_k(\alpha) \ll \left(\frac{n}{q}\right)^{1/k} (1+n\beta)^{-1} \tag{4.18}$$

and Theorem 2 of [Va2] states

$$\Delta_k(\alpha) \ll q^{1/2+\varepsilon}(1+n\beta)^{1/2}. \tag{4.19}$$

For $k \geqslant 5$, (4.17) gives

$$W_k(\alpha) \ll \left(\frac{n}{q}\right)^{1/k} (1+n\beta)^{-1/k} \tag{4.20}$$

and Lemma 5.4 of [Va3] gives

$$\Delta_k(\alpha) \ll \frac{qn^{1/k}}{\log n}(1+n\beta). \tag{4.21}$$

Observe that (4.21) is nontrivial only if $q \leqslant \log n$. Utilization of this inequality will necessitate a prior pruning of the major arcs to $\mathfrak{M}(Y)$, where Y is a small power of $\log n$. The following lemma contains an estimate for f_k valid on major arcs of small order.

Lemma 4.4 (Vaughan-Wooley). *If f_k is defined as in (4.2), and $\alpha \in \mathfrak{M}(\log^A n)$ for some A, then*

$$f_k(\alpha) \ll n^{1/k} q^{\varepsilon - 1/k}(1+n\beta)^{-1/k},$$

where the implied constant may depend on A.

Proof. This is a special case of Lemma 8.5 of [VW1].

The next lemma, essentially Lemma 2 of [Br], greatly aids the estimation of error terms produced by replacing the generating functions f_k with their approximations W_k.

Lemma 4.5 (Brüdern). *Let $Q \leqslant N$. For $1 \leqslant a \leqslant q \leqslant Q$, $(a,q)=1$, let $M(q,a)$ denote an arbitrary interval contained in $[(a/q)-1/2;(a/q)+1/2]$ and assume that the $M(q,a)$ are pairwise disjoint. Write M for the union of all $M(q,a)$. Let $G : M \to \mathbb{C}$ be a function satisfying*

$$G(\alpha) \ll \frac{N}{q}\left(1+N\left|\alpha-\frac{a}{q}\right|\right)^{-1}$$

for $\alpha \in M(q,a)$. Furthermore, let $\Psi : \mathbb{R} \to [0,\infty)$ be a function with a Fourier expansion

$$\Psi(\alpha) = \sum_{|h| \leqslant H} \psi_h e(\alpha h)$$

such that $\log H \ll \log N$. Then

$$\int_M G(\alpha)\Psi(\alpha)d\alpha \ll Q\psi_0 \log N + (\log N)^2 \sum_{h \neq 0} |\psi_h| d(|h|),$$

where $d(n)$ denotes the number of divisors of n.

A typical application of this lemma will be with $G(\alpha)$ equal to a product of powers of some of the $W_k(\alpha)$ (using (4.18)) and $\Psi(\alpha) = |f_{k_1} \cdots f_{k_r}|^2$, so that ψ_h is the number of solutions of

$$h = \sum_{i=1}^{r} (x_i^{k_i} - y_i^{k_i}), \qquad x_i, y_i \in \mathcal{B}_i.$$

Note that $\sum_h \psi_h = \Psi(0)$ and $\psi_0 = \int_0^1 \Psi(\alpha) d\alpha$. The hypotheses of the lemma imply $d(|h|) \ll N^{\varepsilon}$; thus we have

Corollary 4.6. *Under the hypotheses of Lemma 4.5 , suppose further that $\psi_h \geqslant 0$ and*

$$\int_0^1 \Psi(\alpha) d\alpha \ll N^{\varepsilon} Q^{-1} \Psi(0).$$

Then

$$\int_M G(\alpha)\Psi(\alpha) d\alpha \ll N^{\varepsilon} \Psi(0).$$

Returning to the proof of Theorem 1, we now replace f_2 by W_2 and f_3 by W_3 on the major arcs. When $2 \leqslant k \leqslant 4$, (4.19) and (1.5) imply

$$\sup_{\alpha \in \mathfrak{M}} |\Delta_k(\alpha)| \ll n^{\mu/2+\varepsilon}. \tag{4.22}$$

Thus, by (4.3), (4.4) and the Cauchy-Schwarz inequality,

$$\begin{aligned} \int_{\mathfrak{M}} |\Delta_2 F_1 F_2| &\ll n^{\mu/2+\varepsilon} \left(\int_0^1 |F_1|^2 \right)^{1/2} \left(\int_0^1 |F_2|^2 \right)^{1/2} \\ &\ll F(0) n^{-1.039174}. \end{aligned} \tag{4.23}$$

Now write $G(\alpha) = |W_2(\alpha)|^2$ and $\Psi(\alpha) = |F_3(\alpha)|^2$. By (4.18), (4.6) and Corollary 4.6 (with $Q = n^{\mu}$), we have

$$\int_{\mathfrak{M}} |W_2 F_3|^2 \ll n^{\varepsilon} F_3^2(0).$$

Combined with (4.5) and (4.22), it follows that

$$\begin{aligned} \int_{\mathfrak{M}} |W_2 \Delta_3 F_3 F_4| &\leqslant \sup_{\mathfrak{M}} |\Delta_3| \left(\int_{\mathfrak{M}} |W_2 F_3|^2 \right)^{1/2} \left(\int_0^1 |F_4|^2 \right)^{1/2} \\ &\ll F(0) n^{-1.000781}. \end{aligned} \tag{4.24}$$

Before replacing f_4 with W_4, a pruning of the major arcs is required. Let $\mathfrak{M}_1 = \mathfrak{M}(n^{3/8})$ and $\mathfrak{N}_1 = \mathfrak{M} \backslash \mathfrak{M}_1$. On $\mathfrak{N}_1$, whenever $q \leqslant n^{3/8}$, (1.5) gives $\beta \geqslant \frac{n^{3/8}}{qn}$. Let $\theta = \frac{3}{44}$. Then on $\mathfrak{N}_1$, either $q \geqslant n^{3/8-\theta}$ or $\beta \geqslant n^{\theta-1}$. By (4.18), we have in the first case

$$|W_2 W_3|^2 \ll \frac{n}{q}(1+n\beta)^{-1} n^{(2/3)(1-3/8+\theta)} = \frac{n}{q}(1+n\beta)^{-1} n^{61/132}$$

and in the second case

$$|W_2 W_3|^2 \ll \frac{n^{5/3}}{q}(1+n\beta)^{-1} n^{-3\theta} = \frac{n}{q}(1+n\beta)^{-1} n^{61/132}.$$

It now follows from (4.5), (4.6), Corollary 4.6 and the Cauchy-Schwarz inequality that

$$\int_{\mathfrak{N}_1} |W_2W_3F_3F_4| \leqslant \left(\int_{\mathfrak{N}_1} |W_2W_3F_3|^2\right)^{1/2} \left(\int_0^1 |F_4|^2\right)^{1/2} \ll F(0)n^{-1.000240}. \tag{4.25}$$

This is more than enough for replacing f_4. Indeed, by combining (4.7), (4.8), (4.9), (4.19), Corollary 4.6 and Hölder's inequality we have

$$\begin{aligned}\int_{\mathfrak{M}_1} |W_2W_3\Delta_4F_5F_6F_7| &\leqslant \sup_{\mathfrak{M}_1}|\Delta_4| \left(\int_{\mathfrak{M}_1} |W_2F_5|^2\right)^{1/2} \left(\int_{\mathfrak{M}_1} |W_3^3F_6^2|\right)^{1/3} \\ &\qquad \times \left(\int_0^1 |F_6^2F_7^6|\right)^{1/6} \\ &\ll F(0)n^{-1.057209}.\end{aligned} \tag{4.26}$$

We now prune the major arcs further, to a set of α where Lemma 4.4 is applicable. Let $\mathfrak{N}_2 = \mathfrak{M}_1\backslash\mathfrak{M}(X)$, where $X = (\log n)^A$ for some large constant A. Arguing as in the first pruning, on $\mathfrak{N}_2$ we have either $q \geqslant X^{1/2}$ or $\beta \geqslant n^{-1}X^{1/2}$. By (4.18), we have in the first case

$$|W_4|^6 \ll \frac{n^{3/2}}{q}(1+n\beta)^{-1}X^{-1/4}$$

and in the second case

$$|W_4|^6 \ll \frac{n^{3/2}}{q}(1+n\beta)^{-1}X^{-5/2}.$$

In either case, $|W_2|^2 \ll (n/q)(1+n\beta)^{-1}$ and $|W_3|^3 \ll (n/q)(1+n\beta)^{-1}$.

To successfully apply Lemma 4.5, we must estimate more precisely the sum $\sum |\psi_h|d(|h|)$, employing a result due to McDonagh [McD].

Lemma 4.7 (McDonagh). *For each positive integer k, there is a constant C_k such that*

$$\sum_{t<N^{1/k}} d(N-t^k) \ll N^{1/k}(\log N)^{C_k}.$$

Although not required here, it follows from the proof of Lemma 4.7 that $C_k = k2^{2k+1}$ is admissible. Suppose $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_r$ and $\Psi(\alpha) = |f_{k_1}(\alpha)\cdots f_{k_r}(\alpha)|^2$. Then in the notation of Lemma 4.5, ψ_h is the number of solutions of

$$h = \sum_{i=1}^r (x_i^{k_i} - y_i^{k_i}), \qquad x_i, y_i \in \mathcal{A}(P_{k_i}, R).$$

For brevity write $N = x_1^{k_1} + \sum_{i=2}^{r}(x_i^{k_i} - y_i^{k_i})$. By Lemma 4.7,

$$\begin{aligned}\sum_{h\neq 0}\psi_h d(|h|) &= 2\sum_{h>0}\psi_h d(h)\\ &= 2\sum_{\substack{\mathbf{x},\mathbf{y}\\ N>y_1^{k_1}}} d(N-y_1^{k_1})\\ &\leqslant 2P_{k_1}(P_{k_2}\cdots P_{k_r})^2 \max_{N\leqslant rn}\sum_{y_1<N^{1/k_1}} d(N-y_1^{k_1})\\ &\ll (P_{k_1}\cdots P_{k_r})^2(\log n)^{C_{k_1}}.\end{aligned}$$

If $G(\alpha)$ satisfies the hypothesis of Lemma 4.5 and $\int_0^1 \Psi(\alpha)d\alpha \ll \Psi(0)n^{-3/8}$, we have

$$\int_{\mathfrak{N}_2} G\Psi \ll \Psi(0)(\log n)^{C_{k_1}+2}. \tag{4.27}$$

By (4.7) and (4.8), the bound (4.27) holds with $\Psi = |F_5|^2$ and $\Psi = |F_6|^2$. Therefore, by Hölder's inequality,

$$\begin{aligned}&\int_{\mathfrak{N}_2}|W_2W_3W_4F_5F_6F_7|\\ &\quad\leqslant F_7(0)\left(\int_{\mathfrak{N}_2}|W_2F_5|^2\right)^{1/2}\left(\int_{\mathfrak{N}_2}|W_3^3F_6^2|\right)^{1/3}\left(\int_{\mathfrak{N}_2}|W_4^6F_6^2|\right)^{1/6}\\ &\quad\ll F(0)n^{-1}(\log n)^{(C_6+C_{10})/2+2-A/24}\ll F(0)n^{-1}(\log n)^{-1}\end{aligned} \tag{4.28}$$

if A is sufficiently large.

The final pruning stage takes us down to a set where (4.21) is useful. Let $Y = (\log n)^{1/4}$ and $\mathfrak{N}_3 = \mathfrak{M}(X)\backslash\mathfrak{M}(Y)$. By Lemma 4.4 and (4.18), when $\alpha\in\mathfrak{N}_3$ we have

$$|W_2W_3W_4F_5F_6F_7(\alpha)| \ll F(0)q^{-H}(1+n\beta)^{-Q},$$

where $H = \frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{15}-\varepsilon > 2.318$ and $Q = 3+\frac{1}{5}+\cdots+\frac{1}{15}$. We have

$$\begin{aligned}\int_{\mathfrak{N}_3} q^{-H}(1+n\beta)^{-Q}d\alpha &\leqslant \sum_{q\leqslant Y} q^{1-H}\int_{Y/nq}^{\infty}\frac{d\beta}{(n\beta)^Q} + \sum_{Y<q\leqslant X} q^{1-H}\int_0^{X/nq}\frac{d\beta}{(1+n\beta)^Q}\\ &\ll n^{-1}Y^{1-Q}\sum_{q\leqslant Y}q^{Q-H} + n^{-1}\sum_{Y<q\leqslant X}q^{1-H}\\ &\ll n^{-1}Y^{2-H}.\end{aligned}$$

It follows that

$$\int_{\mathfrak{N}_3}|W_2W_3W_4F_5F_6F_7| \ll F(0)n^{-1}(\log n)^{-1/13}. \tag{4.29}$$

The major arcs are now small and few enough for the replacement of each f_k with the corresponding W_k for $k\geqslant 5$. When $\alpha\in\mathfrak{M}(Y)$, (4.21) gives

$$\Delta_k(\alpha) \ll n^{1/k}(\log n)^{-3/4}.$$

If μ denotes Lebesgue measure, then by (1.4) and (1.5), $\mu(\mathfrak{M}(Y)) \ll Y^2 n^{-1}$ and hence

$$\int_{\mathfrak{M}(Y)} |\Delta_k| \ll n^{1/k-1} (\log n)^{-1/4}. \tag{4.30}$$

Combining (4.30) with the trivial bounds $f_k \ll n^{1/k}$, $W_k \ll n^{1/k}$ yields

$$\int_{\mathfrak{M}(Y)} W_2 W_3 W_4 F_5 F_6 F_7(\alpha) e(-n\alpha) d\alpha = \int_{\mathfrak{M}(Y)} W_2 W_3 \cdots W_{15}(\alpha) e(-n\alpha) d\alpha + O(F(0) n^{-1} (\log n)^{-1/4}). \tag{4.31}$$

Gathering together (4.10), (4.23)–(4.26), (4.28), (4.29) and (4.31) yields

$$\int_0^1 F(\alpha) e(-n\alpha) d\alpha = \int_{\mathfrak{M}(Y)} W_2 \cdots W_{15}(\alpha) e(-n\alpha) d\alpha + O(F(0) n^{-1} (\log n)^{-1/13}). \tag{4.32}$$

In [Fo], only the generating functions for $k = 2, 3, 4$ and 6 were replaced by the corresponding W_k. Because all of the f_k have been replaced by the corresponding W_k here, the remainder of the argument is greatly simplified.

Our next step is to modify the integral on the right side of (4.32), replacing the integral over each major arc $\mathfrak{M}(Y; q, a)$ with an integral over [0, 1]. The idea is that $W_k(\alpha, q, a)$ is small outside $\mathfrak{M}(Y; q, a)$. By (4.16) and (4.17),

$$\sum_{q \leqslant Y} \sum_{(a,q)=1} \int_{[0,1] \setminus \mathfrak{M}(Y;q,a)} |W_2 \cdots W_{15}(\alpha, q, a)| \ll F(0) \sum_{q \leqslant Y} q^{1-H} \int_{Y/qn}^{\infty} \frac{d\beta}{(n\beta)^Q} \ll F(0) n^{-1} (\log n)^{-1/13}. \tag{4.33}$$

From the definition of W_k (see (4.11)–(4.13)), we have

$$\sum_{q \leqslant Y} \sum_{(a,q)=1} \int_0^1 W_2 W_3 \cdots W_{15}(\alpha, q, a) e(-n\alpha) d\alpha = \mathfrak{S}(n, Y) I(n), \tag{4.34}$$

where

$$\mathfrak{S}(n, Y) = \sum_{q \leqslant Y} A(n, q), \tag{4.35}$$

$$A(n, q) = \frac{1}{q^{14}} \sum_{(a,q)=1} S_2 S_3 \cdots S_{15}(q, a) e(-an/q), \tag{4.36}$$

$$I(n) = \frac{1}{15!} \sum_{(4.37)} m_2^{-1/2} m_3^{-2/3} m_4^{-3/4} \prod_{k=5}^{15} \varrho\left(\frac{\log m_k}{k \log R}\right) m_k^{1/k-1},$$

$$\frac{n}{2^k} < m_k \leqslant n \quad (k = 2, 3, 4), \qquad R^k < m_k \leqslant P_k^k \quad (5 \leqslant k \leqslant 15), \qquad n = m_2 + \cdots + m_{15}. \tag{4.37}$$

Since the numbers $\left(\dfrac{\log m_k}{k\log R}\right)$ are uniformly bounded above, it is readily shown (see [Va1], Theorem 2.3) that

$$I(n) \gg F(0)n^{-1}.$$

In light of (4.32)–(4.34), it remains to show $\mathfrak{S}(n,Y) \gg 1$. First, since $A(n,q)$ is multiplicative in q (see Lemma 2.11 of [Va1]), we have

$$\mathfrak{S}(n) := \sum_{q=1}^{\infty} A(n,q) = \prod_p \chi_p, \qquad \chi_p = \sum_{h=0}^{\infty} A(n,p^h).$$

The estimate $S_k(q,a) \ll q^{1-1/k}$ (Theorem 4.2 of [Va1]) and (4.36) give

$$|\mathfrak{S}(n) - \mathfrak{S}(n,Y)| \leqslant \sum_{q>Y} |A(n,q)| \ll \sum_{q>Y} q^{1-H} \ll (\log n)^{-1/13}$$

and

$$|\chi_p - 1| \leqslant \sum_{h=1}^{\infty} |A(n,p^h)| \ll p^{1-H}. \tag{4.38}$$

We also have $\chi_p \geqslant p^{-65}$ by Lemma 6.4 of [Fo], and together with (4.38), this shows that $\mathfrak{S}(n) \gg 1$, completing the proof of Theorem 1.

5. Further applications of the method

Define $H(k)$ to be the smallest number s such that all large n admit a representation

$$n = \sum_{i=1}^{s} x_i^{i+k-1}.$$

The existence of $H(k)$ for every k was proved independently by Freiman [Fr] and Scourfield [Sc]. They also established the bound

$$H(k) \ll k^5 \log^2 k.$$

Although no explicit bound for $H(3)$ has appeared in the literature, techniques existing prior to the new iterative method (see, for example, [Br]) produce an upper bound for $H(3)$ above 200. We sketch proofs of the following two estimates.

Theorem 2. *We have* $H(3) \leqslant 72$.

Theorem 3. *We have* $H(k) \ll k^2 \log k$.

The most critical estimate for the proof of Theorem 2 is the minor arc bound. For $5 \leqslant k \leqslant 74$, let

$$f_k(\alpha) = \sum_{m \in \mathcal{A}(P_k,R)} e(\alpha m^k).$$

For $3 \leqslant k \leqslant 4$, let

$$f_k(\alpha) = \sum_{P_k < m \leqslant 2P_k} e(\alpha m^k).$$

Set $\mathfrak{M} = \mathfrak{M}(n^{1/3})$ and $\mathfrak{m} = [0,1]\backslash\mathfrak{M}$. Let $F(\alpha) = f_3(\alpha)f_4(\alpha)\cdots f_{74}(\alpha)$ and let

$$F_1 = f_5 f_{19} f_{20} \cdots f_{74}, \qquad F_2 = f_4 f_6 f_7 \cdots f_{18}.$$

By arguments similar to the proofs of Lemmas 4.1 and 4.2, we obtain

$$\int_0^1 |F_1(\alpha)|^2 d\alpha \ll F_1^2(0) n^{-0.914282},$$
$$\int_0^1 |F_2(\alpha)|^2 d\alpha \ll F_2^2(0) n^{-0.919198}.$$

Combined with Weyl's inequality and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\int_{\mathfrak{m}} |F(\alpha)| d\alpha &\leqslant \sup_{\alpha\in\mathfrak{m}} |f_3(\alpha)| \left(\int_0^1 |F_1(\alpha)|^2 d\alpha\right)^{1/2} \left(\int_0^1 |F_2(\alpha)|^2 d\alpha\right)^{1/2} \\ &\ll F(0) n^{-1.000073}.\end{aligned}$$

As in the proof of Theorem 1, handling the major arcs requires a multi-stage pruning process. As the methods are similar to those used in the proof of Theorem 1, we suppress the details.

The proof of Theorem 3 depends on the strength of the new mean value theorems (2.1) combined with the analysis of section 2, plus a bound for smooth Weyl sums on minor arcs which is vastly superior to known minor arc bounds for classical Weyl sums. Let

$$f_h(\alpha) = \sum_{P_h < m \leqslant 2P_h} e(\alpha m^h) \qquad (k \leqslant h \leqslant 9k),$$
$$f_h(\alpha) = \sum_{m\in\mathcal{A}(P_h,R)} e(\alpha m^h) \qquad (20k \leqslant h \leqslant r),$$

where r is defined below. Let $\mathfrak{M} = \mathfrak{M}(n^{1/(25k)})$, $\mathfrak{m} = [0,1]\backslash\mathfrak{M}$, and set

$$F_1 = f_k \cdots f_{9k} \qquad F_3 = f_{40k} \cdots f_r,$$
$$F_2 = f_{20k} \cdots f_{40k-1} \qquad F = F_1 F_2 F_3.$$

For technical reasons relating to the range of validity of Theorem 1.4 of [Wo1] and the strength of (4.19), the generating functions for $9k+1, \dots, 20k-1$ are not used.

When k is large and $h \geqslant 20k$, Theorem 1.4 of [Wo1] implies

$$\sup_{\alpha\in\mathfrak{m}} |f_h(\alpha)| \ll P_h^{1-1/(3h\log h)},$$

whence

$$\sup_{\alpha\in\mathfrak{m}} |F_2(\alpha)| \ll F_2(0) n^{-1/(500k\log k)}. \tag{5.1}$$

By Theorem 2.1 of [Wo2], $\lambda(k,s) \leqslant 2s - k + ke^{1-2s/k}$ when $k \geqslant 4$ and s is a positive integer. Combined with (2.3), it is easily shown that

$$\lambda(k,s) \leqslant 2s - k + 3ke^{-2s/k} \tag{5.2}$$

for all real $s \geqslant 2$. Now let $h \geqslant 40k$. By (5.2), there is a number M such that if $s = s(h) = \frac{h}{2}(\log k + \log\log k + M)$ (cf. (2.10)), then $\lambda(h,s) \leqslant 2s - h + \frac{h}{1000k\log k}$. Let r be the largest positive integer with

$$\sum_{h=40k}^{r-1} \frac{1}{2s(h)} < 1,$$

and increase M so that

$$\sum_{h=40k}^{r} \frac{1}{2s(h)} = 1.$$

It readily follows that $r \ll k^2 \log k$. Applying Hölder's inequality, we obtain

$$\int_0^1 |F_3| \leqslant \prod_{h=40k}^{r} \left(\int_0^1 |f_h|^{2s(h)} \right)^{1/(2s(h))} \ll F_3(0) n^{-K}, \tag{5.3}$$

where

$$K = \sum_{h=40k}^{r} \frac{1}{2s(h)} \left(1 - \frac{\lambda(h, s(h)) - 2s(h) + h}{h} \right) \geqslant 1 - \frac{1}{1000k \log k}. \tag{5.4}$$

By (5.1), (5.3) and (5.4), it follows that

$$\begin{aligned} \int_{\mathfrak{m}} |F(\alpha)| d\alpha &\leqslant F_1(0) \sup_{\alpha \in \mathfrak{m}} |F_2(\alpha)| \int_0^1 |F_3(\alpha)| d\alpha \\ &\ll F(0) n^{-1-1/(1000k \log k)}. \end{aligned}$$

Because we have a large number of classical Weyl sums comprising $F_1(\alpha)$, the contribution from the major arcs is easily handled by the methods of [Va1, Ch. 4] without the necessity for any pruning.

Acknowledgement

The author wishes to thank Professor Heini Halberstam for his constant support and encouragement. This work forms part of the author's Ph.D. thesis for the University of Illinois at Urbana-Champaign.

References

[Br] Brüdern, J., *A problem in additive number theory*, Math. Proc. Cambridge Philos. Soc. **103** (1) (1988), 27–33. MR **89c:**11150

[DB] De Bruijn, N. G., *On the number of of positive integers $\leqslant x$ and free of prime factors $> y$*, Nederl. Acad. Wetensch. Proc. Ser. A. **54** (1951), 50–60. MR **13:**724e

[Fo] Ford, K. B., *The representation of numbers as sums of unlike powers*, J. London Math. Soc. (2) **51** (1995), 14–26. MR **96c:**11115

[Fr] Freiman, G. A., *Solution to Waring's problem in a new form*, Uspekhi Mat. Nauk, no. 1, **4** (29) (1949), 193. MR **11:**162a

[HL] Hardy, G. H. and Littlewood, J. E., *Some problems of "Partitio Numerorum" : VI. Further researches in Waring's problem*, Math. Z. **23** (1925), 1–37.

[McD] McDonagh, S., *On the sum $\sum_{t<N^{1/k}} d(N-t^k)$*, Proc. Edinburgh Math. Soc. (2) **15** (1967), 215–219. MR **35:**4179

[Sc] Scourfield, E. J., *A generalization of Waring's problem*, J. London Math. Soc. **35** (1960), 98–116. MR **22:**2591

[Th] Thanigasalam, K., *On sums of powers and a related problem*, Acta Arith. **36** (1980), 125–141; *Addendum and corrigendum*, ibid **42** (1983), 425. MR **81i:**11061; MR **85e:**11075

[Va1] Vaughan, R. C., *The Hardy-Littlewood method*, University Press, Cambridge, 1981. MR **84b:**10002

[Va2] ———, *Some remarks on Weyl sums*, Topics in Classical Number Theory, Colloq. Math. Soc. Janos Bolyai, vol. 34, Elsevier North-Holland, 1984, pp. 1585–1602. MR **86e:**11067

[Va3] ———, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71. MR **90c:**11072

[Va4] ———, *A new iterative method in Waring's problem, II*, J. London Math. Soc. (2) **39** (1989), 219–230. MR **90c:**11074

[VW1] Vaughan, R. C., and Wooley, T. D., *On Waring's problem: some refinements*, Proc. London Math. Soc. (3) **63** (1991), 35–68. MR **92f:**11133

[VW2] ———, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240. CMP 96:01

[Wo1] Wooley, T. D., *Large improvements in Waring's problem*, Annals of Math. **135** (1992), 131–164. MR **93b:**11129

[Wo2] ———, *The application of a new mean value theorem to the fractional parts of polynomials*, Acta Arith. **65** (1993), 163–179. MR **94h:**11064

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Current address: Department of Mathematics, University of Texas at Austin, Austin, Texas 78712

E-mail address: ford@math.utexas.edu