THE REPRESENTATION OF NUMBERS AS SUMS OF UNLIKE POWERS

Kevin B. Ford

1. Introduction

The famous Waring's problem is to determine the minimum s such that every (or every sufficiently large) natural number may be expressed as the sum of s kth powers of nonnegative integers. In this connection it is natural to raise corresponding questions for certain sets of mixed powers (see [4,9]), and an elegant special case is whether there exists a number s such that every natural number is the sum of s successive powers, starting with a square. This latter question was answered affirmatively by Roth [7] who proved that all sufficiently large natural numbers n admit a representation

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

with $s = 50$. Subsequently, Thanigasalam [10,11,12], Vaughan [13,14] and Brüdern [1,2]) have improved upon the value of s. The current record is held by Brüdern [2], who showed that (1.1) is solvable for large n and $s = 17$. The aim of this paper is a further improvement in the value of s; we have the following theorem.

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

$$n = \sum_{i=1}^{15} x_i^{i+1}. \tag{1.2}$$

The proof relies on the Hardy-Littlewood circle method, and most closely follows [2]. The main new ingredient is the incorporation of mean value theorems of Vaughan [18,19] and Wooley [20].

2. Notation

Throughout, n is a large natural number and ε is an arbitrarily small positive real number. Constants implies by the Landau and Vinogradov symbols may depend on ε or k. For a real number z, let $e(z) = \exp(2\pi i z)$. Let $\mathscr{A}(P,R)$ denote the set of natural numbers not exceeding P with no prime factor exceeding R. For $2 \leqslant k \leqslant 16$, let

$$\lambda_k = \begin{cases} \frac{8}{9} & \text{for } k = 4, 6, 12, \\ 1 & \text{otherwise} \end{cases}$$

1991 *Mathematics Subject Classification.* Primary 11P05.

and set

$$P_k = \begin{cases} \frac{1}{4} n^{\lambda_k/k} & \text{for } k = 12, \\ \frac{1}{2} n^{\lambda_k/k} & \text{otherwise.} \end{cases} \tag{2.1}$$

Let $K_1 = \{3, 4, 6, 12\}, K_2 = \{5, 7, 8, 9, 10, 11, 13, 14, 15, 16\}$, and define the generating functions

$$f_k(\alpha) = \sum_{P_k < m \leqslant 2P_k} e(\alpha m^k) \qquad \text{for } k = 2, 3, 4, 6, 12, \tag{2.2}$$

$$f_k(\alpha) = \sum_{m \in \mathscr{A}(P_k, P_k^\eta)} e(\alpha m^k) \qquad \text{for } k \in K_2, \tag{2.3}$$

where η is a sufficiently small positive number depending on ε (see Lemma 3.5 below). Let

$$F_i(\alpha) = \prod_{k \in K_i} f_k(\alpha) \qquad \text{for } i = 1, 2, \tag{2.4}$$

$$F(\alpha) = \prod_{k=2}^{16} f_k(\alpha). \tag{2.5}$$

To prove the theorem it suffices to show that $R(n) > 0$, where

$$R(n) = \int_0^1 F(\alpha) e(-n\alpha) d\alpha. \tag{2.6}$$

For $1 \leqslant a \leqslant q \leqslant n^{4/9}$, $(a, q) = 1$, define the major arcs by

$$\mathfrak{M}(q, a) = \left\{ \alpha : \left| \alpha - \frac{a}{q} \right| \leqslant \frac{n^{4/9}}{qn} \right\}, \tag{2.7}$$

with the usual convention when $q = 1$. It is easily shown that the major arcs are pairwise disjoint. Denote by $\mathfrak{M}$ the union of all major arcs, and define the minor arcs $\mathfrak{m}$ by $\mathfrak{m} = [0, 1] \backslash \mathfrak{M}$. Any $\alpha \in \mathfrak{M}$ uniquely determines a, q and $\beta = \|\alpha - \frac{a}{q}\|$, where $\|x\| = \min_{y \in \mathbb{Z}} |x - y|$. Throughout, whenever $\alpha \in \mathfrak{M}$ is given, we assume these definitions of a, q and β.

3. Minor Arcs

By the Cauchy-Schwarz inequality, we have

$$\int_{\mathfrak{m}} |F(\alpha)| d\alpha \leqslant \sup_{\alpha \in \mathfrak{m}} |f_2(\alpha)| \left(\int_0^1 |F_1(\alpha)|^2 d\alpha \right)^{\frac{1}{2}} \left(\int_0^1 |F_2(\alpha)|^2 d\alpha \right)^{\frac{1}{2}}. \tag{3.1}$$

The first expression on the right is estimated by Lemma 3.1 below (Weyl's inequality). This is [15, Lemma 2.4].

Lemma 3.1. *Suppose that* $(a,q)=1$, $|\alpha-a/q|<q^{-2}$, $\phi(x)=\alpha x^k+\alpha_1x^{k-1}+\cdots+\alpha_{k-1}x+\alpha_k$, *and* $K=2^{k-1}$. *Then*

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

The next lemma is a generalization of Davenport's iteration method due to Vaughan [17, Lemma 4]).

Lemma 3.2. *Suppose* $k\geqslant 3$, $1-1/k\leqslant\lambda\leqslant 1$, $\nu=k\lambda-k+1$, $P\geqslant P_0(k,\varepsilon), C=C(k,\varepsilon)$. *Let* $R(m)$ *denote a non-negative arithmetic function with support on* $[1,CP^{k\lambda}]$. *Let*

$$R_1(m)=\sum_{P<x\leqslant 2P} R(m-x^k)$$

and define

$$S=\sum_m R(m)^2, \qquad T=\sum_m R_1(m)^2.$$

Then for $1\leqslant j\leqslant k-2$ *we have*

$$T\ll PS+P^{1+\nu-2^{1-j}}S+P^{1+\varepsilon+\nu(1-2^{-j})-(j+1)2^{-j}}S^{1-2^{-j}}\left(\sum_m R(m)\right)^{2^{1-j}}.$$

Lemma 3.3.

$$\int_0^1 |F_1(\alpha)|^2d\alpha \ll F_1(0)^2n^{-7/9+\varepsilon}.$$

Proof. Let $R(m)$ denote the number of solutions of $x_4^4+x_6^6+x_{12}^{12}=m$ with $P_k<x_k\leqslant 2P_k$. Then in the notation of Lemma 3.2 with $k=3$ and $\lambda=8/9$, S is the number of solutions of

$$x_4^4+x_6^6+x_{12}^{12}=y_4^4+y_6^6+y_{12}^{12} \tag{3.2}$$

with $P_k<x_k,y_k\leqslant 2P_k$, and from (2.4) if follows that

$$T=\int_0^1 |F_1(\alpha)|^2d\alpha.$$

The number of solutions of (3.2) with $x_4\neq y_4$ is $O(P_6P_{12})^{2+\varepsilon}=O(n^{4/9+\varepsilon})$, since for any $k\geqslant 2$, the number of solutions of $x^k-y^k=m$ with $x\neq y$ is $O(|m|^\varepsilon)$. Likewise, the number of solutions of (3.2) with $x_4=y_4$ and $x_6\neq y_6$ is $O(P_4P_{12}^{2+\varepsilon})=O(n^{5/12+\varepsilon})$. Finally, the number of solutions with $x_4=y_4, x_6=y_6, x_{12}=y_{12}$ is $O(P_4P_6P_{12})=O(n^{4/9})$. Thus applying Lemma 3.2 with $j=1$ yields

$$T\ll n^{7/9+\varepsilon}\ll F_1(0)^2n^{-7/9+\varepsilon}.$$

The second integral in (3.1) is estimated with the aid of the following three lemmas. The first is a classical sieve result, and is discussed in [5, Chapter 1].

Lemma 3.4. *If* $\eta > 0$ *then* $\operatorname{card} \mathscr{A}(P, P^\eta) \gg P$.

Lemma 3.5. *If* $\eta > 0$ *is sufficiently small, then for each triple* $(k, s, \lambda(k,s))$ *given in Table 3.1 we have*

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

Proof. These follow from a combination of [20,Lemma 3.2], [19, Theorem 1.4 and Lemma 2.2] and inequality $(k-2)$ of [18, §4] (cf. [20, §5, Tables 5.1 and 5.2; 19, Table 1.1]). Each number in Table 3.1 was calculated with 16 digit precision, and has been rounded up in the last decimal place.

k	s	$\lambda(k,s)$	k	s	$\lambda(k,s)$
5	5	5.987384	11	3	3.022614
5	6	7.660993	11	12	15.385063
7	3	3.063920	13	5	5.278409
7	7	8.624130	13	14	17.949291
8	6	6.924136	14	5	5.258936
8	8	9.929058	14	15	19.233401
9	3	3.035806	15	5	5.232864
9	9	11.213914	15	16	20.512127
9	10	12.819123	16	5	5.218977
10	5	5.401025	16	18	23.411517
10	11	14.107494			

Table 3.1

Lemma 3.6.

$$\int_0^1 |F_2(\alpha)|^2 d\alpha \ll F_2(0)^2 n^{-7/9-\nu},$$

where $\nu > .00279$.

Proof. Let $a_5 = 5, a_7 = 7, a_8 = 8, a_9 = 9 + \frac{4671}{5641} = 9 + \theta, a_k = k+1$ for $10 \leqslant k \leqslant 15$, $a_{16} = 18$. Since $\sum 1/a_k = 1$, we have by Hölder's inequality

$$\int_0^1 |F_2(\alpha)|^2 d\alpha \leqslant \prod_{k \in K_2} \left(\int_0^1 |f_k(\alpha)|^{2a_k} d\alpha \right)^{1/a_k}.$$

By another application of Hölder's inequality,

$$\int_0^1 |f_9|^{2a_9} \leqslant \left(\int_0^1 |f_9|^{18} \right)^{1-\theta} \left(\int_0^1 |f_9|^{20} \right)^{\theta}.$$

The result now follows from Lemmas 3.4 and 3.5.

Remark. The values a_k given above are the optimum values and were found by computer. The algorithm that was used to find them is discussed in section 7 below.

Completing the minor arc estimate, it follows from (3.1) and Lemmas 3.1, 3.3 and 3.6 that

$$\begin{aligned}\int_{\mathfrak{m}} |F(\alpha)| d\alpha &\ll P_2 n^{-2/9+\varepsilon} F_1(0) n^{-7/18+\varepsilon} F_2(0) n^{-7/18-\nu/2} \\ &\ll F(0) n^{-1-\nu/2+\varepsilon}.\end{aligned} \tag{3.3}$$

4. The Major Arcs

We first introduce several auxillary functions. In this section, the variable k will take on only the values $2, 3, 4, 6$. Let

$$S_k(q,a) = \sum_{m=1}^{q} e\left(\frac{am^k}{q}\right), \tag{4.1}$$

$$w_k(\alpha) = \sum_{P_k^k < m \leqslant (2P_k)^k} \frac{1}{k} m^{(1/k)-1} e(\alpha m), \tag{4.2}$$

$$W_k(\alpha, q, a) = \frac{1}{q} S_k(q,a) w_k\left(\alpha - \frac{a}{q}\right). \tag{4.3}$$

For brevity, write

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

and

$$\Delta_k(\alpha) = f_k(\alpha) - W_k(\alpha).$$

Our first goal is to replace each f_k by the corresponding W_k without introducing too large an error. The next three lemmas comprise the chief tools used to handle the error terms. The first is [15,Lemma 6.3] and the second is [16, Theorem 2]. The third, a special case of ([2, Lemma 2]), greatly simplifies the estimation of certain 'error' terms arising from the Hardy-Littlewood method, and allows one to take the major arcs much larger and more numerous than traditional methods (see, for example, [15, Chapter 4]). Our treatment of the major arcs closely follows that of Brüdern [2].

Lemma 4.1. *Suppose that* $(q,a) = 1$ *and* $|\beta| < 1/2$. *Then*

$$W_k(a/q + \beta, q, a) \ll P_k q^{-1/k} (1 + P_k^k |\beta|)^{-1}.$$

Lemma 4.2. *Suppose that* $(a,q) = 1$ *and* $\alpha = a/q + \beta$. *Then*

$$\Delta_k(\alpha) \ll q^{1/2+\varepsilon} (1 + P_k^k |\beta|)^{1/2}.$$

Lemma 4.3. *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*

$$G(\alpha) \ll \frac{N}{q}\left(1+N\left|\alpha-\frac{a}{q}\right|\right)^{-1} \qquad \text{for } \alpha \in M(q,a).$$

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 $\log H \ll \log N$ and $\psi_0 = \int_0^1 \Psi(\alpha) d\alpha \ll Q^{-1}\Psi(0)$. Then

$$\int_M G(\alpha)\Psi(\alpha) d\alpha \ll N^{\varepsilon}\Psi(0). \tag{4.4}$$

First we replace f_2 by W_2. By Lemma 4.2 and (2.7),

$$\sup_{\alpha\in\mathfrak{M}} \Delta_k(\alpha) \ll n^{2/9+\varepsilon}, \tag{4.5}$$

hence by Lemmas 3.3, 3.6 and the Cauchy-Schwarz inequality,

$$\begin{aligned}\int_{\mathfrak{M}} |\Delta_2 F_1 F_2| d\alpha &\ll n^{2/9+\varepsilon}\left(\int_0^1 |F_1|^2\right)^{1/2}\left(\int_0^1 |F_2|^2\right)^{1/2}. \\ &\ll F(0) n^{-1-1/18-\nu/2+\varepsilon}\end{aligned} \tag{4.6}$$

To replace f_3 by W_3, we write $|W_2(\alpha)|^2 = G(\alpha)$. By Lemma 4.1, G satisfies the conditions of Lemma 4.3 with $N=n$. Let $F_3 = f_4 f_6 f_{12}$ and $K_3 = \{4,6,12\}$. Then clearly $|F_3(\alpha)|^2 = \sum_h \psi_h e(\alpha h)$, where ψ_h is the number of solutions of

$$h = \sum_{k\in K_3} (x_k^k - y_k^k)$$

with $P_k < x_k, y_k \leqslant 2P_k$. The proof of Lemma 3.3 implies that

$$\int_0^1 |F_3|^2 \ll F_3(0)^2 n^{-4/9+\varepsilon}, \tag{4.7}$$

hence it follows from Lemma 4.3 (with $Q = n^{4/9}$) that

$$\int_{\mathfrak{M}} |W_2 F_3|^2 d\alpha \ll n^{\varepsilon} F_3(0)^2. \tag{4.8}$$

From (4.5), (4.8), the Cauchy-Schwarz inequality, and Lemma 3.6 we have

$$\int_{\mathfrak{M}} |W_2 \Delta_3 F_3 F_2| \ll n^{2/9+\varepsilon}\left(\int_0^1 |F_2|^2\right)^{1/2} \left(n^{\varepsilon} F_3(0)^2\right)^{1/2} \ll F(0)\, n^{-1-\nu/2+\varepsilon}. \tag{4.9}$$

We now whittle down the major arcs before replacing the generating functions for the fourth and sixth powers. Write $\mathfrak{M} = \mathfrak{M}_1 \cup \mathfrak{M}_2$, where $\mathfrak{M}_1$ is the union of all $\mathfrak{M}(q,a)$ with $q \leqslant n^{1/3}$, and $\mathfrak{M}_2 = \mathfrak{M}\backslash\mathfrak{M}_1$. On $\mathfrak{M}_2$ we have, by Lemma 4.1,

$$|W_2(\alpha)W_3(\alpha)|^2 \ll \left(\frac{n}{q}\right)^{5/3}(1+n\beta)^{-4} \ll \frac{n^{13/9}}{q}(1+n\beta)^{-1}.$$

Applying (4.7), Lemmas 3.6 and 4.3, and the Cauchy-Schwarz inequality as before, we deduce that

$$\begin{aligned}\int_{\mathfrak{M}_2} |W_2W_3F_3F_2| &\ll \left(\int_{\mathfrak{M}_2} |W_2W_3F_3|^2\right)^{1/2}\left(\int_0^1 |F_2|^2\right)^{1/2} \\ &\ll n^{2/9}F_3(0)n^{-7/18-\nu/2}F_2(0) \\ &\ll n^{-1-\nu/2+\varepsilon}F(0).\end{aligned} \tag{4.10}$$

Now write $F_2 = F_4F_5F_6$, where $F_4 = f_{10}f_{13}f_{14}f_{15}f_{16}$, $F_5 = f_7f_9f_{11}$ and $F_6 = f_5f_8$. Applying Hölder's inequality and Lemma 3.5 yields

$$\int_0^1 |F_4|^2 \ll \prod_{k=10,13,14,15,16}\left(\int_0^1 |f_k|^{10}\right)^{1/5} \ll F_4^2(0)n^{-.3556}, \tag{4.11}$$

$$\int_0^1 |F_5|^2 \ll \prod_{k=7,9,11}\left(\int_0^1 |f_k|^{6}\right)^{1/3} \ll F_5^2(0)n^{-.3398}, \tag{4.12}$$

$$\int_0^1 |F_6|^6 \ll \left(\int_0^1 |f_5|^{12}\right)^{1/2}\left(\int_0^1 |f_8|^{12}\right)^{1/2} \ll F_6^6(0)n^{-.7511}. \tag{4.13}$$

By Lemmas 4.1 and 4.2 we have

$$|W_3|^3 \ll \frac{n}{q}(1+n\beta)^{-1}, \tag{4.14}$$

$$|W_4^6\Delta_6| \ll \frac{n^{3/2}}{q}(1+n^{8/9}\beta)^{-1} = \frac{n^{8/9}}{q}(1+n^{8/9}\beta)^{-1}n^{11/18}. \tag{4.15}$$

Combining (4.11) through (4.14) with Lemma 4.3 and Hölder's inequality gives

(4.16)

$$\begin{aligned}\int_{\mathfrak{M}_1} |W_2W_3\Delta_4f_6f_{12}F_2| \leqslant P_6P_{12}\sup_{\mathfrak{M}_1}|\Delta_4|F_5(0)^{1/3}\times \\ \left(\int_{\mathfrak{M}_1}|W_2F_4|^2\right)^{1/2}\left(\int_{\mathfrak{M}_1}|W_3^3F_5^2|\right)^{1/3}\left(\int_0^1|F_6|^6\right)^{1/6} \\ \ll F(0)n^{-\frac{1}{12}-\frac{1}{2}-\frac{1}{3}-\frac{1}{6}(.7511)} \ll F(0)n^{-1.04}.\end{aligned}$$

A very similar estimate, incorporating (4.15), produces

(4.17)

$$\begin{aligned}\int_{\mathfrak{M}_1} |W_2W_3W_4\Delta_6f_{12}F_2| \leqslant P_{12}\sup_{\mathfrak{M}_1}|\Delta_6|^{5/6}F_5(0)^{1/3}\times \\ \left(\int_{\mathfrak{M}_1}|W_2F_4|^2\right)^{1/2}\left(\int_{\mathfrak{M}_1}|W_3^3F_5^2|\right)^{1/3}\left(\int_{\mathfrak{M}_1}|W_4^6\Delta_6F_6^6|\right)^{1/6} \\ \ll F(0)n^{-\frac{1}{36}-\frac{1}{2}-\frac{1}{3}-(\frac{1}{4}-\frac{11}{108})+\varepsilon} \ll F(0)n^{-1-\frac{1}{108}+\varepsilon}\end{aligned}$$

Collecting together (4.6), (4.9), (4.10), (4.16) and (4.17) yields

$$\int_{\mathfrak{M}} F(\alpha)e(-n\alpha)d\alpha = \int_{\mathfrak{M}_1} W_2W_3W_4W_6f_{12}F_2e(-n\alpha)d\alpha + O(F(0)n^{-1-\nu/3}). \tag{4.18}$$

5. Completion of the analytic Argument

Writing $m = n - x_{12}^{12} - \sum_{k\in K_2} x_k^k$, the integral on the right hand side of (4.18) becomes

$$\sum_m \int_{\mathfrak{M}_1} W_2W_3W_4W_6e(-m\alpha)d\alpha,$$

where $\frac{15}{16}n \leqslant m \leqslant n$ by virtue of (2.1), (2.2) and (2.3).

First, we will replace each major arc with the entire unit interval, following Thanigasalam [11]. This requires several additional lemmas. The first is [15, Lemma 6.2] and the second is [17, Lemma 3] and follows from the proof of [15,Theorem 4.2].

Lemma 5.1. $w_k(\beta) \ll P_k(1 + P_k^k|\beta|)^{-1}$.

Lemma 5.2. *We have $q^{-1}S_k(q,a) \ll g_k(q)$, where $g_k(q)$ is the multiplicative function defined by*

$$g_k(p^{uk+v}) = \begin{cases} kp^{-u-1/2} & v = 1 \\ p^{-u-1} & 2 \leqslant v \leqslant k \end{cases}. \tag{5.2}$$

Furthermore, $q^{-1/2} \leqslant g_k(q) \ll q^{-1/k}$.

Lemma 5.3. *Let*

$$B(q) = \sum_{(a,q)=1} \frac{|S_2S_3S_4S_6(q,a)|}{q^4}.$$

Then

$$\sum_{q\leqslant Y} B(q) \ll (\log Y)^C$$

for some constant C.

Proof. Let $g(q) = qg_2(q)g_3(q)g_4(q)g_6(q)$. Since $g(q)$ is multiplicative and positive, we have by Lemma 5.2

$$\sum_{q\leqslant Y} B(q) \ll \sum_{q\leqslant Y} g(q) \leqslant \prod_{p\leqslant Y}\left(\sum_{h=0}^{\infty} g(p^h)\right).$$

By (5.2) it follows that $g(p) \ll p^{-1}$, $g(p^2) \ll p^{-2}$, $g(p^3) \ll p^{-3/2}$ and

$$\sum_{h=4}^{\infty} g(p^h) \ll \sum_{h=4}^{\infty} p^{-h/4} \ll p^{-1}.$$

Thus for some constants c_1 and c_2 we have

$$\sum B(q) \ll \prod (1 + c_1p^{-1}) \ll (\log Y)^{c_2}.$$

Since $q \leqslant n^{1/3}$, it follows that $\beta + a/q \notin \mathfrak{M}$ implies $|\beta| \geqslant q^{-1}n^{-5/9} \geqslant n^{-8/9}$. Consequently, by (2.7), (4.3), Lemmas 5.1 and 5.3,

(5.3)
$$\begin{aligned}\sum_{q\leqslant n^{1/3}} \sum_{(a,q)=1} \int_{[0,1]\setminus\mathfrak{M}(q,a)} & W_2W_3W_4W_6(\alpha,q,a)e(-m\alpha)d\alpha \\ &\ll \sum_{a,q} \frac{1}{q^4} P_2P_3P_4P_6|S_2S_3S_4S_6(q,a)| \int_{n^{4/9}/(qn)}^{\infty} \frac{d\beta}{\prod_k(1+P_k^k\beta)} \\ &\ll P_2P_3P_4P_6 \sum_{q,a} \frac{|S_2S_3S_4S_6(q,a)|}{q^4} \frac{(qn^{5/9})^3}{P_2^2P_3^3P_4^4P_6^6} \\ &\ll P_2P_3P_4P_6n^{-10/9} \sum_{q\leqslant n^{1/3}} \sum_{(a,q)=1} \frac{|S_2S_3S_4S_6(q,a)|}{q^4} \\ &\ll P_2P_3P_4P_6n^{-10/9+\varepsilon}.\end{aligned}$$

Finally, by (4.1), (4.2) and (4.3),

(5.4)
$$\sum_{q\leqslant n^{1/3}} \sum_{(a,q)=1} \int_0^1 W_2W_3W_4W_6(\alpha,q,a)e(-m\alpha)d\alpha = \mathfrak{S}(m,n^{1/3})I(m),$$

where

(5.5)
$$\mathfrak{S}(m,Y) = \sum_{q\leqslant Y} A(m,q),$$

(5.6)
$$A(m,q) = \sum_{(a,q)=1} \frac{1}{q^4} S_2S_3S_4S_6(q,a)e(-am/q),$$

(5.7)
$$I(m) = \frac{1}{144} \sum_{\substack{P_k^k<m_k\leqslant(2P_k)^k \\ m=m_2+m_3+m_4+m_6}} m_2^{-1/2}m_3^{-2/3}m_4^{-3/4}m_6^{-5/6}.$$

Combining (5.3) and (5.4) it follows that

(5.8)
$$\int_{\mathfrak{M}_1} W_2W_3W_4W_6(\alpha)e(-m\alpha)d\alpha = \mathfrak{S}(m,n^{1/3})I(m) + O(P_2P_3P_4P_6n^{-10/9+\varepsilon}).$$

6. The singular Series

Our treatment of the singular series is essentially that of Roth [7]. The fact that we replaced the fifth power by a sixth power has no effect on the applicability of the method.

Lemma 6.1. $I(m) \gg P_2P_3P_4P_6n^{-1}$.

Proof. From (5.7),

$$I(m) \gg P_2^{-1}P_3^{-2}P_4^{-3}P_6^{-5}I_1(m),$$

where $I_1(m)$ is the number of representations of m as $m = m_2+m_3+m_4+m_6$. If we take $\frac{1}{4}n \leqslant m_2 \leqslant \frac{1}{2}n$, $m-m_2-n^{8/9} \leqslant m_3 \leqslant m-m_2-\frac{1}{2}n^{8/9}$, $\frac{1}{8}n^{8/9} \leqslant m_4 \leqslant \frac{1}{4}n^{8/9}$, then there is a number $m_6 \in [P_6^6,(2P_6)^6]$ such that $m = m_2+m_3+m_4+m_6$. Therefore

$$I(m) \gg n^{-\frac{1}{2}-\frac{1}{3}-\frac{8}{9}(\frac{3}{4}+\frac{5}{6})+1+\frac{8}{9}+\frac{8}{9}} = n^{\frac{11}{54}} \gg P_2P_3P_4P_6n^{-1}.$$

Next define

$$\mathfrak{S}(m) = \sum_{q=1}^{\infty} A(m,q).$$

Lemma 6.2. $|\mathfrak{S}(m) - \mathfrak{S}(m,Y)| \ll Y^{-1/20}m^{\varepsilon}$.

Proof. By (5.6) we have $|A(m,q)| \leqslant B(q)$, and by the proof of Lemma 5.3 together with the estimates $g(p^4) \ll p^{-3/2}$ and $g(p^5) \ll p^{-2}$ it follows that

$$\sum_{h=2}^{\infty} |A(m,p^h)| \ll p^{-3/2}. \tag{6.1}$$

By the same argument given for [7,Lemma 20], we also have

$$|A(m,p)| \ll \begin{cases} p^{-1} & p|m \\ p^{-3/2} & p \nmid m \end{cases}. \tag{6.2}$$

The lemma now follows in the same manner as [7, Lemma 29].

Lemma 6.3. $\mathfrak{S}(m) \gg (\log\log m)^{-C}$, *where C is a positive constant.*

Proof. It follows from (6.1) and (6.2) that the sum in (5.5) is absolutely convergent, and since $A(m,q)$ is multiplicative in q, we have

$$\mathfrak{S}(m) = \prod_p \chi_p, \qquad \text{where } \chi_p = \sum_{h=0}^{\infty} A(m,p^h). \tag{6.3}$$

By (6.1) , (6.2) and the fact that $A(m,1)=1$ it follows that

$$|\chi_p - 1| \ll \begin{cases} p^{-1} & p|m \\ p^{-3/2} & p \nmid m \end{cases}. \tag{6.4}$$

This is a satisfactory estimate for large primes p. For small primes, we employ the following general result, which is an easy deduction from the results of [15, §2.6].

Lemma 6.4. *Let $k_1, k_2, \ldots, k_r$ be integers with $k_i \geqslant 2$ for each i. For each prime p and integer k, let $p^\tau || k$ and define*

$$\gamma(k,p) = \begin{cases} \tau + 2 & p = 2 \text{ and } \tau > 0 \\ \tau + 1 & \text{else} \end{cases}.$$

Fix a prime p and let $\gamma = \max(\gamma(k_1,p), \ldots, \gamma(k_r,p))$. If

$$\sum_{i=1}^{r} \frac{p-1}{p^{\gamma(k_i,p)}(k_i, p-1)} \geqslant 1, \tag{6.5}$$

and

$$\chi_p = \sum_{h=0}^{\infty} A(n, p^h), \qquad A(n,q) = q^{-r} \sum_{(a,q)=1} S_{k_1} \ldots S_{k_r}(q,a) e(-an/q),$$

then χ_p is real and $\chi_p \geqslant p^{-(r-1)\gamma}$.

To complete the proof of Lemma 6.3, take $k_1 = 2, k_2 = 3, k_3 = 4, k_4 = 6$. It is readily verified that (6.5) holds when $p = 2$ and $p = 3$. When $p \geqslant 5$, the left side of (6.5) is at least

$$\frac{p-1}{p}\left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{6}\right) \geqslant 1.$$

Since $\max_{p,i} \gamma(k_i,p) = 4$, we have $\chi(p) \geqslant p^{-12}$ for every prime p. By (6.4) there exist positive constants c_1, c_2, c_3 and c_4 such that

$$\begin{aligned} \mathfrak{S}(m) &= \prod_p \chi_p \\ &> \left(\prod_{p \leqslant c_3} p^{-12}\right)\left(\prod_{\substack{p > c_3 \\ p|m}} (1 - c_1 p^{-1})\right)\left(\prod_{p > c_3} (1 - c_2 p^{-3/2})\right) \\ &\gg (\log\log m)^{-c_4}. \end{aligned}$$

The theorem now follows from (3.3), (4.18), (5.1), (5.8) and Lemmas 6.1, 6.2 and 6.3.

7. Optimization of Hölder's Inequality

For every k, define $f_k = f_k(\alpha)$ as in (2.3) with $P_k = \frac{1}{2} n^{1/k}$. The works of Vaughan and Wooley [18, 19,20] yield mean value estimates of the form

$$\int_0^1 |f_k(\alpha)|^{2s} d\alpha \ll P_k^{\lambda(k,s)}, \tag{7.1}$$

for each natural number s. If h is an integer and $0 < \theta < 1$, we have by Hölder's inequality

$$\int_0^1 |f_k|^{2(h+\theta)} d\alpha \leqslant \left(\int_0^1 |f_k|^{2h} d\alpha\right)^{1-\theta} \left(\int_0^1 |f_k|^{2h+2} d\alpha\right)^{\theta}. \tag{7.2}$$

Thus, defining

$$\lambda(k, h+\theta) = (1-\theta)\lambda(k,h) + \theta\lambda(k,h+1) \tag{7.3}$$

extends (7.1) to all positive real s. If $x_1, \dots, x_r$ are positive numbers satisfying

$$x_1 + \cdots + x_r = 1 \tag{7.4}$$

then by Hölder's inequality we have

$$S = \int_0^1 |f_{k_1} \dots f_{k_r}|^2 \leqslant \prod_{i=1}^r \left(\int_0^1 |f_{k_i}|^{2/x_i}\right)^{x_i}. \tag{7.5}$$

It follows from (7.1), (7.2) and (7.5) that

$$S \ll n^{\phi}, \qquad \phi = \sum_{i=1}^r \frac{x_i \lambda(k_i, 1/x_i)}{k_i}. \tag{7.6}$$

For a given set of exponents $\{k_i\}$, our goal is to minimize ϕ subject to (7.4). To this end, define the functions

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

The Cauchy-Schwarz inequality gives

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

Combined with (7.3) this shows that each $\lambda(k_i, s)$ is convex as a function of s for $s > 0$. The following lemma shows that the convexity of $\lambda(k_i, s)$ implies the convexity of $g_i(x)$.

Lemma 7.1. *If $f(x)$ is convex for $x > 0$ then $xf(1/x)$ is convex for $x > 0$.*

Proof. Suppose $a > 0$, $y > 0$ and $0 \leqslant \lambda \leqslant 1$. It suffices to show

$$(a+\lambda y) f\left(\frac{1}{a+\lambda y}\right) \leqslant (1-\lambda) a f(1/a) + \lambda(a+y) f\left(\frac{1}{a+y}\right),$$

or

$$f\left(\frac{1}{a+\lambda y}\right) \leqslant \frac{a(1-\lambda)}{a+\lambda y} f\left(\frac{1}{a}\right) + \frac{\lambda(a+y)}{a+\lambda y} f\left(\frac{1}{a+y}\right). \tag{7.7}$$

But

$$\frac{1}{a+\lambda y} = \frac{a(1-\lambda)}{a+\lambda y}\frac{1}{a} + \frac{\lambda(a+y)}{a+\lambda y}\frac{1}{a+y},$$

and (7.7) follows from the convexity of f.

The next lemma shows how this convexity property may be exploited in finding the minimum of the sum in (7.6).

Lemma 7.2. *Suppose that $g_i(x)$ are continuous, convex, piecewise differentiable functions on $(0,\infty)$ for $1 \leqslant i \leqslant r$, and let D^+ and D^- denote the right and left differential operators, respectively. Subject to the constraint $\sum x_i = 1$, the sum $\sum g_i(x_i)$ is minimized whenever*

$$\min_i D^+ g_i(x_i) \geqslant \max_i D^- g_i(x_i), \tag{7.8}$$

Proof. Suppose $\sum_i y_i = 1$, and write $y_i = x_i + \delta_i$ for each i. Let $A = \{i : \delta_i \geqslant 0\}$ and $B = \{i : \delta_i < 0\}$. The hypotheses of the lemma imply that $D^+ g_i(x)$ and $D^- g_i(x)$ are monotone increasing. Thus by (7.8) and the fact that $\sum_i \delta_i = 0$ we have

$$\begin{aligned}
\sum_i g_i(y_i) &= \sum_i g_i(x_i) + \sum_{i \in A} \int_{x_i}^{y_i} D^+ g_i(x)dx - \sum_{i \in B} \int_{y_i}^{x_i} D^- g_i(x)dx \\
&\geqslant \sum_i g_i(x_i) + \sum_{i \in A} \delta_i D^+ g_i(x_i) + \sum_{i \in B} \delta_i D^- g_i(x_i) \\
&\geqslant \sum_i g_i(x_i) + \sum_i \delta_i \left(\min_i D^+ g_i(x_i)\right) \\
&= \sum_i g_i(x_i).
\end{aligned}$$

Designing an algorithm based on Lemma 7.2 is straightforward. Find i and j such that

$$D^+ g_i(x_i) < D^- g_j(x_j), \tag{7.9}$$

and set $x_i = x_i + \delta$ and $x_j = x_j - \delta$, where δ is the least positive number for which (7.9) fails to hold. In the actual algorithm, the i and j are chosen to maximize the difference between the left and right hand sides of (7.9). Also, because of the discontinuous nature of the derivatives of each $g_i(x)$, at most one of the numbers $1/x_i$ will not be integral, unless the numbers $\lambda(k,s)$ satisfy some unusual relations. As a final remark, when r is large and so are all of the k_i, choosing the numbers x_i so that $x_i k_i = x_j k_j$ for all i, j usually yields a value of ϕ in (7.6) which is very close to optimal. This choice of numbers x_i is used as a starting point in the algorithm.

8. Computations

In this section, we discuss some numerical calculations that were carried out with the purpose of determining the least s for which (1.1) is solvable for various ranges of n. Call this least value $s(n)$. Similar computations were carried out by Kløve [6] for $n \leqslant 250000$.

The first calculations were to determine $s(n)$ for each $n \leqslant 5 \cdot 10^7$. We have $s(n) \leqslant 5$ for all such n and $s(n) \leqslant 4$ for all but 22 values of n in this range, the largest being 26471. We conjecture that in fact $s(n) \leqslant 5$ for all n and $s(n) \leqslant 4$ for all $n > 26471$.

A general conjecture in Waring's problem (see the introduction to [15, Chapter 8]) implies that (1.1) holds for all large integers for $s = 3$. The next calculations were aimed at finding the probable value of the largest number which is not the sum of a square, cube and fourth power of nonnegative integers. Table 8.1 lists the number of integers E in the range $[N, N+10^6-1]$ which are not so representable.

N	E
10^6	168355
10^7	134642
10^8	100281
10^9	72854
10^{10}	49246
10^{11}	31307
10^{12}	18535
10^{13}	10609
10^{14}	5698
10^{15}	2629
10^{16}	1144
10^{17}	482
10^{18}	144

Table 8.1

From the table, then the density of numbers n near N with $s(n)>3$ appears to be decreasing faster than any power of N. It is also evident from the table that if there is a largest exceptional n, then that n is very large. If $E(x)$ denotes the number of exceptional n less than x, then Roth [8] showed that $E(x)=O(x(\log x)^{-1/20})$. The best result known is due to Brüdern [3], who has sharpened this to $E(x) \ll x^{1-1/14+\varepsilon}$.

9. Addendum

The author has recently improved the main theorem of this paper, namely showing that every sufficiently large number n has a representation

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

The significant new idea is an adaptation of the new iterative method [18,19,20] to obtain upper bounds for

$$\int_0^1 |f_h(\alpha) f_k^s(\alpha)|^2 d\alpha,$$

where f_h is an ordinary Weyl sum (as in (2.2)) and f_k is a "smooth" Weyl sum (as in (2.3)). Details will appear in a forthcoming paper.

References

1. J. Brüdern, *Sums of squares and higher powers, II*, J. London Math. Soc. **(2) 35** (1987), 244-250.

2. J. Brüdern, *A problem in additive number theory*, Math. Proc. Cambridge Philos. Soc. **103** (1988), no. 1, 27-33.
3. J. Brüdern, *Ternary problems of Waring's type*, Math.-Scand. **68** (1991), 27-45.
4. G. A. Freiman, *Solution to Waring's problem in a new form*, Uspekhi Mat. Nauk, no. 1, **4** (1949), no. 29, 193.
5. A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, Sem. Theor. Nombres Bordeaux (to appear).
6. Torleiv Kløve, *Representations of integers as sums of powers with increasing exponents*, Nordisk Tidskr. Informations behandling (BIT) **12** (1972), 342-346.
7. K. F. Roth, *A problem in additive number theory*, Proc. London Math. Soc. **(2) 53** (1951), 381-395.
8. K. F. Roth, *Proof that almost all positive integers are sums of a square, a positive cube and a fourth power*, J. London Math. Soc. **24** (1949), 4-13.
9. E. J. Scourfield, *A generalization of Waring's problem*, J. London Math. Soc. **35** (1960), 98-116.
10. K. Thanigasalam, *On additive number theory*, Acta Arith. **13** (1968), 237-258.
11. K. Thanigasalam, *On sums of powers and a related problem*, Acta Arith. **36** (1980), 125-141; *Addendum and corrigendum*, ibid **42** (1983), 425.
12. K. Thanigasalam, *On certain additive representations of integers*, Portugaliae Math. **42** (1983-1984), 447-465.
13. R. C. Vaughan, *On the representation of numbers as sums of powers of natural numbers*, Proc. London Math. Soc. (3) **21** (1970), 160-180.
14. R. C. Vaughan, *On sums of mixed powers*, J. London Math. Soc. (2) **3** (1971), 677-688.
15. R. C. Vaughan, *The Hardy-Littlewood method*, University Press, Cambridge, 1981.
16. R. C. Vaughan, *Some remarks on Weyl sums*, Topics in Classical Number Theory , Colloq. Math. Soc. Janos Bolyai, vol. 34, Elsevier North-Holland, 1984, pp. 1585-1602.
17. R. C. Vaughan, *On Waring's problem for smaller exponents*, Proc. London Math. Soc. **(3) 52** (1986), 445-463.
18. R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1-71.
19. R. C. Vaughan, *A new iterative method in Waring's problem, II*, J. London Math. Soc. **(2) 39** (1989), 219-230.
20. T. D. Wooley, *Large improvements in Waring's problem*, Annals of Math. **135** (1992), 131-164.

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

E-mail address: `ford@math.uiuc.edu`