ADDITIVE REPRESENTATION IN THIN SEQUENCES, VI: REPRESENTING PRIMES, AND RELATED PROBLEMS

J. BRÜDERN, K. KAWADA, AND T. D. WOOLEY

Abstract. We discuss the representation of primes, almost-primes, and related arithmetic sequences as sums of kth powers of natural numbers. In particular, we show that on GRH, there are infinitely many primes represented as the sum of $2\lceil 4k/3\rceil$ positive integral kth powers, and we prove unconditionally that infinitely many P_2-numbers are the sum of $2k+1$ positive integral kth powers. The sieve methods required to establish the latter conclusion demand that we investigate the distribution of sums of kth powers in arithmetic progressions, and our conclusions here may be of independent interest.

1. Introduction

In previous papers in this series, we investigated the number of elements in a given polynomial sequence that admit a representation in a certain prescribed form. The introductions of parts I and IV [2, 3] supply a discussion of the scope and the underlying ideas of our methods, so we content ourselves here with the remark that the aim of our analysis is to preserve the arithmetic structure of the sequence in which a representation property is tested. The purpose of the present note is to show that the technique of part IV in this series is applicable even when the sequences tested are rather denser than the thin polynomial sequences in our earlier communications. Rather than proceeding in undue generality, we make this notion more precise with a specific example. Let $P(k)$ denote the smallest integer s for which the set of numbers representable as the sum of s positive k-th powers contains infinitely many primes. Then since the work of Wooley [12] implies that all but $O\big(N(\log N)^{-2}\big)$ of the natural numbers not exceeding N are the sum of s positive k-th powers whenever $s\ge \frac{1}{2}k\log k+O(k\log\log k)$, it follows immediately that

$$P(k)\le \tfrac{1}{2}k\log k+O(k\log\log k).$$

This decidedly weak bound seems to be all that is currently known for larger k, although there is every expectation that $P(k)\le 3$ (indeed, Heath-Brown has recently announced a proof of the infinitude of primes of the form x^3+2y^3, and from this it follows that $P(3)=3$). If the Riemann Hypothesis is true for all Dirichlet L-functions (a conjecture hereafter referred to as GRH), we are able to reduce the order of magnitude of the aforementioned upper bound for $P(k)$.

Theorem 1. *On the assumption of GRH, one has $P(k)\le 2\lceil 4k/3\rceil$ for every natural number k. Moreover, if $\pi_{k,s}(N)$ denotes the number of primes not exceeding N that can be written as the sum of s k-th powers of natural numbers, then for $s\ge 2\lceil 4k/3\rceil$ one has $\pi_{k,s}(N)\gg N^{1-\theta(s)}$, where $\theta(s)=e^{1-2s/k}$.*

1991 *Mathematics Subject Classification.* 11N32, 11P05, 11P55 (11N36, 11P32).
The third author is a Packard Fellow and is supported in part by NSF grant DMS-9970440.

Unconditional results of the same strength remain a desideratum. When combined with the linear sieve, our methods yield a conclusion similar to Theorem 1, but with the primes replaced by the set of natural numbers with at most two prime factors.

Theorem 2. *Let $k \geq 3$ and $s \geq 2k+1$. Then there exist infinitely many natural numbers with at most two prime factors that are the sum of s k-th powers of natural numbers.*

Other sequences may be substituted for the primes without seriously affecting the argument. We mention in passing a result involving sums of two squares.

Theorem 3. *Let $k \geq 3$ and $s \geq 2\lceil (\frac{1}{4}+\frac{1}{2}\log 2)k\rceil + 2$. Then infinitely many of the numbers representable as the sum of two integral squares are the sum of s positive k-th powers. If $S_{k,s}(N)$ is the number of sums of two squares not exceeding N with this property, and $\theta(s)$ is defined as in Theorem 1, then $S_{k,s}(N) \gg N^{1-\theta(s)}$.*

Theorem 3 may be regarded as a special case of our final result, which concerns sums of l l-th powers in place of sums of two squares.

Theorem 4. *Let $l \geq 2$ and $k \geq 3$ be natural numbers. Then there exists a number $c(l)$ such that whenever $s \geq c(l)k$, there are infinitely many integers that are simultaneously the sum of l l-th powers and s k-th powers of natural numbers.*

It will be clear from the proofs below that our methods are capable of providing lower bounds, analogous to those stated for $\pi_{k,s}(N)$ and $S_{k,s}(N)$, in Theorems 2 and 4.

Throughout, ε will denote a sufficiently small positive number, and P will be a large real number. We use $\ll$ and $\gg$ to denote Vinogradov's notation, and write $e(z)$ for $e^{2\pi i z}$. Finally, we write $\lceil \alpha \rceil$ for the smallest integer exceeding α, and $\|\alpha\|$ for $\min_{y\in\mathbb{Z}} |\alpha - y|$.

2. The main argument

We proceed by describing the argument at the core of our method in abstracted form. Suppose that we are given a set $\mathcal{U} \subseteq \mathbb{N}$ and we wish to show that it contains many sums of s k-th powers. Our technique rests on two ingredients. First, we require a lower bound for the number of solutions of the equation

$$u = x_1^k + x_2^k + \ldots + x_s^k \tag{2.1}$$

with $x_i \in [1,P]\cap\mathbb{Z}$ (perhaps restricted to possess only small prime divisors) and $u \in \mathcal{U}$. This can be achieved by the circle method, and in favourable circumstances one can expect to obtain a lower bound of order UP^{s-k} for the number of solutions of (2.1), where $U = \text{card}\big(\mathcal{U} \cap [1,P^k]\big)$. If $\mathcal{U}$ has neat arithmetic properties, then it transpires that the latter lower bound may be obtained when $s > ck$, for a suitable constant c. In a second step, one aims to show that an individual element u has not too many representations in the form (2.1), and thereby one derives a lower bound for the number of elements in $\mathcal{U}$ that have a representation in the shape (2.1). This can be achieved classically via Cauchy's inequality, or by means of the tools introduced in part IV of this series.

We now set the scene for a result that covers the lower bound problem in sufficient generality for the proofs of Theorems 1, 3 and 4. Let

$$\mathcal{A}(P,R) = \{n \in [1,P]\cap\mathbb{Z} : p \text{ prime}, p|n \Rightarrow p \leq R\},$$

and define the exponential sums $g(\alpha)=g_k(\alpha;P,R)$ by

$$g(\alpha)=\sum_{x\in\mathcal{A}(P,R)} e(\alpha x^k).$$

For an integer $t\geq 1$, the real number Δ_t is called a permissible exponent if for each $\varepsilon>0$, there exists a positive number $\eta=\eta(\varepsilon)$ with the property that whenever $R\leq P^\eta$, one has

$$\int_0^1 \big|g(\alpha)\big|^{2t}d\alpha \ll P^{2t-k+\Delta_t+\varepsilon}. \tag{2.2}$$

Here we recall that the corollary to Theorem 2.1 of Wooley [13] shows that for each natural number t, there is a permissible exponent Δ_t satisfying the inequality

$$\Delta_t e^{\Delta_t/k}\leq ke^{1-2t/k}. \tag{2.3}$$

Next we require the singular series corresponding to a sum of s k-th powers, defined by

$$\mathfrak{S}_{k,s}(n)=\sum_{q=1}^\infty \sum_{\substack{a=1\\(a,q)=1}}^q q^{-s}S(q,a)^s e(-an/q),$$

where

$$S(q,a)=\sum_{r=1}^q e(ar^k/q).$$

For the moment, we are content to record the classical fact that $\mathfrak{S}_{k,s}(n)$ is absolutely convergent and non-negative for $s\geq 4$ (see Theorem 4.3 of Vaughan [9]).

In order to describe our next result, we must define a suitable Farey dissection of the unit interval. For $1\leq Q\leq \frac12 N^{1/2}$, let $\mathfrak{M}(Q)$ denote the union of the arcs $\mathfrak{M}(q,a;Q)=\{\alpha\in(0,1]:|q\alpha-a|\leq Q/N\}$ with $0\leq a\leq q\leq Q$ and $(a,q)=1$. We note that the arcs $\mathfrak{M}(q,a;Q)$ comprising the latter union are pairwise disjoint. Hence, we may define a function $\Upsilon:\mathbb{R}\to\mathbb{R}$ of period 1 which is given for $\alpha\in\mathfrak{M}(q,a;\frac12 N^{1/2})\subseteq \mathfrak{M}\big(\frac12 N^{1/2}\big)$ by

$$\Upsilon(\alpha)=\frac{q}{\varphi(q)}\big(q+N|q\alpha-a|\big)^{-1},$$

and is 0 on $(0,1]\backslash\mathfrak{M}(\frac12 N^{1/2})$.

On defining u_n to be the indicator function of the sequence $\mathcal{U}$, the scope of our methods is readily discerned from the following lemma.

Lemma 2.1. *Let u_n $(1\leq n\leq N)$ be non-negative real numbers, and let δ be a positive number. Suppose that the exponential sum*

$$U(\alpha)=\sum_{n\leq N} u_n e(\alpha n)$$

satisfies the inequality $U(\alpha)\ll U(0)\big(N^{-\delta}+\Upsilon(\alpha)\big)$ uniformly for $\alpha\in\mathbb{R}$. Suppose also that t is an integer exceeding $k/2$ for which there exist permissible exponents Δ_t and Δ_{t-1} satisfying $\Delta_t<\delta k$ and $\Delta_{t-1}\leq(1-\delta)k$. Then provided that $\eta>0$ is sufficiently small, one has

$$\int_0^1 g(\alpha)^{2t}U(-\alpha)d\alpha \gg \sum_{n\leq N} u_n\mathfrak{S}_{k,2t}(n)n^{2t/k-1}+O(U(0)N^{2t/k-1}(\log N)^{-1/(100k)}),$$

where we abbreviate $g_k(\alpha; N^{1/k}, N^\eta)$ *to* $g(\alpha)$.

We postpone the proof of this lemma to §4 and presently direct our attention to the second step in our argument. Let $r_{k,s}(u)$ denote the number of solutions of (2.1) with $x_j \in \mathcal{A}(N^{1/k}, N^\eta)$ $(1 \le j \le s)$, and define $w_{k,s}(u)$ to be 1 when $r_{k,s}(u) > 0$, and to be 0 otherwise. We put $s = 2t$, and recall the hypotheses and notation of the statement of Lemma 2.1. We suppose in addition that

$$\sum_{n \le N} u_n \mathfrak{S}_{k,2t}(n) n^{2t/k-1} \gg U(0) N^{2t/k-1}, \tag{2.4}$$

a lower bound that often holds in practise. The following two lemmata provide lower bounds of a type suitable for this second stage of our treatment.

Lemma 2.2. *In addition to the hypotheses of Lemma* 2.1*, suppose that* (2.4) *holds. Then*

$$\sum_{n \le N} u_n^2 w_{k,2t}(n) \gg N^{-1-\Delta_{2t}/k-\varepsilon} \Big(\sum_{n \le N} u_n \Big)^2.$$

Proof. On substituting (2.4) into the conclusion of Lemma 2.1, orthogonality yields

$$\sum_{n \le N} u_n r_{k,s}(n) = \int_0^1 g(\alpha)^s U(-\alpha)\, d\alpha \gg U(0) N^{s/k-1}. \tag{2.5}$$

By Cauchy's inequality, on the other hand, one has

$$\Big(\sum_{n \le N} u_n r_{k,s}(n) \Big)^2 \le \Big(\sum_{n \le N} u_n^2 w_{k,s}(n) \Big) \Big(\sum_{n \le N} r_{k,s}(n)^2 \Big). \tag{2.6}$$

But by orthogonality and (2.2), one finds that

$$\sum_{n \le N} r_{k,s}(n)^2 \le \int_0^1 |g(\alpha)|^{2s}\, d\alpha \ll N^{(2s-k+\Delta_s+\varepsilon)/k},$$

and so the desired conclusion follows immediately on substituting (2.5) into (2.6). □

Lemma 2.3. *In addition to the hypotheses of Lemma* 2.1*, suppose that* (2.4) *holds. Then*

$$\sum_{n \le N} u_n w_{k,2t}(n) \gg N^{-\Delta_t/k-\varepsilon} \sum_{n \le N} u_n.$$

Proof. By orthogonality and (2.2), we find that

$$\begin{aligned} \sum_{n \le N} u_n r_{k,2t}(n) &= \sum_{n \le N} u_n \int_0^1 g(\alpha)^{2t} e(-n\alpha)\, d\alpha \\ &\ll N^{(2t-k+\Delta_t+\varepsilon)/k} \sum_{n \le N} u_n w_{k,2t}(n), \end{aligned}$$

and so the desired conclusion is immediate from (2.5). □

3. Sieve-free conclusions

The rather abstract material of the previous section makes it easy to deduce those results mentioned in the introduction which do not depend on sieves. The proof of Theorem 1, which we now describe, can serve as a model.

Let $u_p = \log p$ when p is a prime, and otherwise let $u_n = 0$. Then, in the notation of Lemma 2.1, it readily follows from GRH that the approximation

$$U(\beta + a/q) = \frac{\mu(q)}{\varphi(q)} \sum_{n \le N} e(n\beta) + O\big((qN)^{1/2}(1 + N|\beta|)(\log N)^2\big) \tag{3.1}$$

holds for coprime $a \in \mathbb{Z}$ and $q \in \mathbb{N}$, and for all $\beta \in \mathbb{R}$ (see, for example, Lemma 2 of Brüdern and Perelli [4]). By Dirichlet's theorem, for any $\alpha \in \mathbb{R}$, there are coprime $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $1 \le q \le N^{2/3}$ and $|q\alpha - a| \le N^{-2/3}$. One now routinely deduces from (3.1) that $U(\alpha) \ll N\Upsilon(\alpha) + N^{5/6+\varepsilon}$, whereas the prime number theorem yields $U(0) \gg N$. Moreover, a simple calculation leads from (2.3) to the upper bounds $\Delta_t < k/6$ and $\Delta_{t-1} < 2k/5$ whenever $t \ge 4k/3$. It is now easily seen that the hypotheses of Lemma 2.1 are satisfied by choosing $\Delta_t/k < \delta < 1/6$. The condition (2.4) required in Lemma 2.2 is harmless. Indeed, the classical theory of the singular series shows that for $t \ge 4k/3$ one has $\mathfrak{S}_{k,2t}(n) \gg 1$ at least for $n \equiv 1 \pmod{4k}$, and hence (2.4) follows in the present situation from the prime number theorem for arithmetic progressions. From Lemma 2.2, we then infer the lower bound

$$\sum_{p \le N} (\log p)^2 w_{k,2t}(p) \gg N^{1-\Delta_{2t}/k-\varepsilon},$$

and Theorem 1 follows immediately for even values of $s = 2t \ge 2\lceil 4k/3 \rceil$, by invoking (2.3). We leave to the reader the routine modifications needed to cover odd values of s.

The proof of Theorem 3 follows the same pattern. Let $X = \left(\frac{1}{2}N\right)^{1/2}$ and write u_n for the number of solutions of $n = x_1^2 + x_2^2$ with $1 \le x_1, x_2 \le X$. Then $u_n = 0$ for $n > N$, and for $n \le \frac{1}{2}N$, one finds that u_n is equal to the number of solutions of the equation $n = x_1^2 + x_2^2$ with $x_1, x_2 \in \mathbb{N}$. In the notation of Lemma 2.1, one has

$$U(\alpha) = \Big(\sum_{1 \le x \le X} e(\alpha x^2) \Big)^2,$$

and on combining Theorem 4.1 and Lemma 4.6 of Vaughan [9], it is easily seen that $U(\alpha) \ll N\Upsilon(\alpha) + N^{1/2+\varepsilon}$. Moreover, one has $U(0) = [X]^2 \gg N$. Finally, we note that for $t > (\frac{1}{4} + \frac{1}{2}\log 2)k + 1$, it follows from (2.3) that $\Delta_t < k/2$ and $\Delta_{t-1} < k/2$, whence the hypotheses of Lemma 2.1 are satisfied by choosing δ in the range $\Delta_t/k < \delta < 1/2$. By summing over $n \equiv 1 \pmod{4k}$, we again see easily that (2.4) holds. Since $u_n = O(n^\varepsilon)$, we may now proceed as in the proof of Theorem 1 to complete the proof of Theorem 3.

For the proof of Theorem 4 we have recourse to an old-fashioned diminishing ranges trick essentially due to Hardy and Littlewood. For more details of the procedure described in this paragraph the uninitiated reader is referred to §5.4 of Vaughan [9]. In view of the conclusion of Theorem 3, we may suppose that l is a natural number with $l \ge 3$. Write

$$h(\alpha, X) = \sum_{X < x \le 2X} e(\alpha x^l).$$

Also, define

$$X_1 = \tfrac{1}{2}(N/l)^{1/l}, \quad X_j = \tfrac{1}{2}X_{j-1}^{1-1/l} \quad (2 \le j \le l),$$

and write

$$U(\alpha) = h(\alpha, X_1)h(\alpha, X_2)\dots h(\alpha, X_l).$$

We note that

$$U(0) \asymp X_1X_2\dots X_l \asymp N^{1-(1-1/l)^l},$$

where it is useful to observe that $1/4 \le (1-1/l)^l < e^{-1}$. In the notation of Lemma 2.1, we find that u_n is the number of solutions of the equation $n = x_1^l + \dots + x_l^l$ with $X_j < x_j \le 2X_j$ $(1 \le j \le l)$. Note that $u_n \le 1$ for $1 \le n \le N$, and that $u_n = 0$ for $n > N$.

When $a \in \mathbb{Z}$, $q \in \mathbb{N}$ and $\alpha \in \mathbb{R}$ satisfy $q \le X^{1/2}$, $(a, q) = 1$ and $|q\alpha - a| \le X^{1/2-l}$, one readily confirms from Theorem 4.1 and Lemma 6.3 of Vaughan [9] that

$$h(\alpha, X) \ll Xq^{-1/l}(1 + X^l|\alpha - a/q|)^{-1}.$$

Under the weaker hypotheses $q \le X$ and $|q\alpha - a| \le X^{1-l}$, the same argument nonetheless supplies the estimate $h(\alpha, X) \ll Xq^{-1/l}$. Applying the displayed inequality when $X = X_1$, and the weaker substitute when $X = X_j$ $(2 \le j \le l)$, one finds that $U(\alpha) \ll U(0)\Upsilon(\alpha)$ for $\alpha \in \mathfrak{M}(X_l)$. By Weyl's inequality, on the other hand (see Lemma 2.4 of [9]), there exists a positive number $\delta = \delta(l) < 1/2$ such that $U(\alpha) \ll U(0)N^{-\delta}$ for $\alpha \in [0,1) \setminus \mathfrak{M}(X_l)$. On combining the last two estimates, we find that the hypotheses of Lemma 2.1 concerning $U(\alpha)$ are satisfied. We observe next that, in view of (2.3), there is a positive number $c = c(l) > 2$ with the property that whenever $t > ck$, then one has $\Delta_t < \delta k$ and $\Delta_{t-1} < (1-\delta)k$. Also, by Theorem 4.6 of [9] one finds that $\mathfrak{S}_{k,2t}(n) \gg 1$ for $2t \ge 4k$, and hence the lower bound (2.4) also holds. We therefore conclude from Lemma 2.3 that

$$\sum_{n \le N} w_{l,l}(n)w_{k,2t}(n) \ge \sum_{n \le N} u_n w_{k,2t}(n) \gg U(0)N^{-\Delta_t/k-\varepsilon},$$

and this suffices to establish Theorem 4 when $s = 2t$ is even. Again, we leave to the reader the simple modifications required to accomodate odd values of s.

We conclude this section with a remark concerning norm forms $F \in \mathbb{Z}[X_1, \dots, X_d]$ associated with an algebraic number field of degree d. By modifying the argument applied in the proof of Theorem 3, one may establish that there exists a number $c(d)$ such that for all $d \ge 3$ and all $s \ge c(d)k$, there are infinitely many values in $F(\mathbb{Z}^d)$ that are the sum of s positive k-th powers. Unfortunately, a proper account of the complications accompanying such a generalisation of Theorem 3 demands more space than is available herein.

4. A pruning exercise

The sole purpose of this section is to prove Lemma 2.1. We recall the notation of the statement of that lemma, and write $P = N^{1/k}$ and $\mathfrak{N} = \mathfrak{M}\left((\log N)^{1/10}\right)$. Standard endgame technique from the Hardy-Littlewood method, involving Lemma 5.4 of Vaughan [8], shows that whenever $s > k$, there exists a positive number $c_s(\eta)$ such that

$$\int_{\mathfrak{N}} g(\alpha)^s e(-n\alpha)d\alpha = c_s(\eta)n^{s/k-1}\mathfrak{S}_{k,s}(n) + O\left(N^{s/k-1}(\log N)^{-1/(11k)}\right)$$

for each natural number n with $n \le N$. We take $s = 2t$, multiply by u_n and sum over integers n with $1 \le n \le N$. Thus it is evident that Lemma 2.1 follows provided we establish that the complementary integral over the set $\mathfrak{n} = [0,1) \setminus \mathfrak{N}$ is negligible. To this end, we engineer a succession of pruning processes that lead to the bound

$$\int_{\mathfrak{n}} |g(\alpha)^{2t} U(\alpha)| d\alpha \ll U(0) P^{2t-k} (\log N)^{-1/(100k)}. \tag{4.1}$$

Observe first that in view of the simple bound $\varphi(q) \gg q/\log\log q$, our hypothesis on $U(\alpha)$ implies that $U(\alpha) \ll U(0) N^{-\delta} \log N$ unless $\alpha \in \mathfrak{M}(N^\delta)$, in which case $U(\alpha) \ll U(0)\Upsilon(\alpha)$. It follows that

$$\int_{\mathfrak{n}} |g(\alpha)^{2t} U(\alpha)| d\alpha \ll U(0) N^{-\delta} \log N \int_0^1 |g(\alpha)|^{2t} d\alpha + U(0) \int_{\mathfrak{n} \cap \mathfrak{M}(N^\delta)} \Upsilon(\alpha) |g(\alpha)|^{2t} d\alpha.$$

Here, by (2.2) and the inequality $\Delta_t < \delta k$, the first term on the right hand side is majorised by the right hand side of (4.1). We now define the sets

$$\mathfrak{M}_1 = \mathfrak{M}(N^\delta), \quad \mathfrak{M}_2 = \mathfrak{M}(P^{1/10}), \quad \mathfrak{M}_3 = \mathfrak{M}((\log N)^A), \quad \mathfrak{M}_4 = \mathfrak{N},$$

where $A \geq 2k$ is a parameter to be chosen in due course, and we write also $\mathfrak{N}_j = \mathfrak{M}_j \setminus \mathfrak{M}_{j+1}$ $(1 \leq j \leq 3)$. In view of the definition of $\mathfrak{n}$, it is now apparent that the desired conclusion (4.1) follows immediately from the estimates

$$\int_{\mathfrak{N}_j} \Upsilon(\alpha) |g(\alpha)|^{2t} d\alpha \ll P^{2t-k} (\log N)^{-1/(100k)} \quad (1 \leq j \leq 3), \tag{4.2}$$

and these we presently seek to establish.

On recalling the above lower bound for $\varphi(q)$, we find from Lemma 2 of Brüdern [1] that whenever $Q \leq \frac{1}{2} N^{1/2}$, one has

$$\int_{\mathfrak{M}(Q)} \Upsilon(\alpha) |g(\alpha)|^{2t-2} d\alpha \ll N^{\varepsilon - 1} \Big(Q \int_0^1 |g(\alpha)|^{2t-2} d\alpha + P^{2t-2} \Big). \tag{4.3}$$

Also, as a consequence of Theorem 1.8 of Vaughan [8], there exists a positive number ξ with the property that $|g(\alpha)| \ll P^{1-\xi}$ whenever $\alpha \in \mathfrak{N}_1$. On taking $Q = N^\delta$, therefore, and noting the hypothesis $\Delta_{t-1} \leq (1-\delta)k$, we deduce from (4.3) that

$$\begin{aligned} \int_{\mathfrak{N}_1} \Upsilon(\alpha) |g(\alpha)|^{2t} d\alpha &\leq \Big(\sup_{\alpha \in \mathfrak{N}_1} |g(\alpha)| \Big)^2 \int_{\mathfrak{M}_1} \Upsilon(\alpha) |g(\alpha)|^{2t-2} d\alpha \\ &\ll (P^{1-\xi})^2 N^{\varepsilon-1} P^{2t-2+\varepsilon} \ll P^{2t-k-\xi}. \end{aligned}$$

This establishes the estimate (4.2) when $j = 1$.

Next we observe that a routine modification of the proof of Lemma 5.4 of Vaughan and Wooley [11], combined with the above lower bound for $\varphi(q)$, shows that whenever $t \geq [k/2] + 1$ and $2 \leq Q \leq P$, one has

$$\int_{\mathfrak{M}(Q)} \Upsilon(\alpha) |g(\alpha)|^{2t-2} d\alpha \ll (\log Q)^B P^{2t-2-k},$$

where B is a suitable positive number depending at most on k. Also, from Lemmata 7.2 and 8.5 of Vaughan and Wooley [10], one has $g(\alpha) \ll P(\log N)^{\varepsilon - A/k}$ for $\alpha \in \mathfrak{N}_2$. On taking $Q = P^{1/10}$ and A sufficiently large in terms of B, therefore, and recalling the hypothesis $t > k/2$, we find that (4.2) follows for $j = 2$ in much the same way as for the case $j = 1$.

The final pruning step requires the bound $|g(\alpha)|^k \ll \Upsilon(\alpha)P^k(\log P)^\varepsilon$, which is valid for $\alpha \in \mathfrak{N}_3$ by virtue of Lemma 8.5 of Vaughan and Wooley [10]. On recalling once more that $2t > k$, straightforward estimates that need no explanation here then show that

$$\int_{\mathfrak{N}_3} \Upsilon(\alpha)|g(\alpha)|^{2t} d\alpha \ll P^{2t}(\log P)^{2t\varepsilon} \int_{\mathfrak{N}_3} \Upsilon(\alpha)^{1+2t/k} d\alpha \\ \ll P^{2t}N^{-1}(\log N)^{-1/(11k)}.$$

We thus arrive at the estimate (4.2) for $j = 3$, so that in view of our earlier discussion, the proof of Lemma 2.1 is complete.

5. Sums of k-th powers in arithmetic progressions

Our proof of Theorem 2 invokes standard sieve theory, and this forces us to examine the level of distribution of sums of k-th powers. For this purpose we employ familiar diminishing ranges techniques. Although more modern technology involving smooth numbers would permit conclusions with fewer kth powers, the well-trodden path greatly simplifies our analysis. We make some comments concerning the use of smooth numbers, and a consequent refinement to Theorem 2, at the end of this section.

The Hardy-Littlewood diminishing ranges trick, already applied in the proof of Theorem 4, is now applied to k-th powers. Let k be a natural number with $k \geq 3$, and write

$$f(\alpha, X) = \sum_{X < x \leq 2X} e(\alpha x^k).$$

When t is a natural number, we write $s = 2t + 1$ and define

$$X_1 = \tfrac{1}{2}(N/s)^{1/k}, \quad X_j = \tfrac{1}{2}X_{j-1}^{1-1/k} \quad (j \geq 2).$$

Also, we put

$$F_t(\alpha) = f(\alpha, X_1)f(\alpha, X_2)\dots f(\alpha, X_t).$$

On considering the underlying diophantine equation, one discerns the bound

$$\int_0^1 \bigl|F_t(\alpha)\bigr|^2 d\alpha \ll \Xi, \tag{5.1}$$

where $\Xi = X_1X_2\dots X_t$ (see, for example, §5.4 of [9]). Next let $\varrho(n) = \varrho_{k,s}(n)$ denote the number of solutions of the equation

$$z^k + \sum_{j=1}^{t}(x_j^k + y_j^k) = n, \tag{5.2}$$

subject to

$$X_1 < z \leq 2X_1 \quad \text{and} \quad X_j < x_j, y_j \leq 2X_j \quad (1 \leq j \leq t). \tag{5.3}$$

In order to sieve the "weighted sequence" $\varrho(n)$, we seek an asymptotic formula for the sums $\sum_m \varrho(md)$, valid for as large a range for d as is feasible. In this context we note that $\varrho(n) = 0$ for $n > N$, so that the latter sum is finite. Observe also that this sum is equal to the number of solutions of the congruence

$$z^k + \sum_{j=1}^{t} (x_j^k + y_j^k) \equiv 0 \pmod{d}, \tag{5.4}$$

with z and x_j, y_j $(1 \le j \le t)$ satisfying (5.3). Write $v(d) = v_{k,s}(d)$ for the number of solutions of (5.4) with

$$1 \le z \le d \quad \text{and} \quad 1 \le x_j, y_j \le d \quad (1 \le j \le t).$$

Then on sorting x_j, y_j $(1 \le j \le t)$ and z into congruence classes modulo d, we find that whenever $1 \le d < X_t$, one has

$$\sum_{n \equiv 0 \pmod{d}} \varrho(n) = d^{-s} v(d) X_1 \Xi^2 + O\left(d^{1-s} v(d) X_1 X_t^{-1} \Xi^2\right). \tag{5.5}$$

By averaging over the modulus d, the range for d can be much extended.

Lemma 5.1. *Suppose that $0 < \delta < 1/2$, and that t is an integer with $t \ge k$ for which $(1 - 1/k)^t < \delta$. Then whenever $D \le N^{1-\delta}$, there exists a positive number η such that*

$$\sum_{d \le D} \Big| \sum_{n \equiv 0 \pmod{d}} \varrho(n) - d^{-s} v(d) X_1 \Xi^2 \Big| \ll X_1 \Xi^{2-\eta}.$$

Proof. Write

$$J(\alpha, X) = \sum_{1 \le n \le X} e(n\alpha) \quad \text{and} \quad J_d(\alpha) = J(d\alpha, N/d).$$

Then by orthogonality,

$$\sum_{n \equiv 0 \pmod{d}} \varrho(n) = \int_0^1 f(\alpha, X_1) F_t(\alpha)^2 J_d(-\alpha) d\alpha. \tag{5.6}$$

We observe that $J_d(\alpha) \ll \min\{N/d, \|d\alpha\|^{-1}\}$. Consequently, on combining Lemma 2.2 of [9] with a familiar argument (see, for example, Exercise 2 of Chapter 2 of [9]), we deduce that whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$ and $|\alpha - a/q| \le q^{-2}$, one has

$$\sum_{d \le D} |J_d(\alpha)| \ll \Big(\frac{N}{q + N|q\alpha - a|} + D + q + N|q\alpha - a| \Big) \log N.$$

By Dirichlet's approximation theorem, however, for each $\alpha \in [0, 1)$ there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$, $q \le \frac{1}{2} N^{1/2}$ and $|q\alpha - a| \le 2N^{-1/2}$, and thus the estimate

$$\sum_{d \le D} |J_d(\alpha)| \ll \left(N \Upsilon(\alpha) + D + N^{1/2}\right) \log N \tag{5.7}$$

holds uniformly for $\alpha \in [0, 1)$.

In order to make further progress, we define a Hardy-Littlewood dissection. Write $Z = \Xi^{2^{k+2}\eta}$, and suppose that η is so small that $Z \le X_t^{1/6}$. Define $\mathfrak{M} = \mathfrak{M}(Z)$ and $\mathfrak{m} = [0, 1) \setminus \mathfrak{M}$. We note that when $\alpha \notin \mathfrak{M}(N^\delta)$, one has $\Upsilon(\alpha) \ll N^{-\delta} \log N$. Also, in view of our hypotheses $D \le N^{1-\delta}$ and $(1 - 1/k)^t < \delta < 1/2$, it follows from the relation $\Xi \asymp N^{1-(1-1/k)^t}$ that

$$N^\varepsilon (D + N^{1/2}) \Xi \ll N^{1-\delta+\varepsilon} \Xi \ll \Xi^2 N^{(1-1/k)^t - \delta + \varepsilon} \ll \Xi^{2-\eta},$$

whenever $\eta > 0$ is sufficiently small. Consequently, on writing $\mathfrak{K} = \mathfrak{M}(N^\delta)\setminus\mathfrak{M}$ and noting the availability of a trivial estimate for $f(\alpha, X_1)$, one finds from (5.7) and (5.1) that

$$\begin{aligned}\sum_{d\le D}\Bigl|\int_{\mathfrak{m}} f(\alpha,X_1)F_t(\alpha)^2J_d(-\alpha)d\alpha\Bigr| \\ &\ll N^{1-\delta+\varepsilon}X_1\int_0^1|F_t(\alpha)|^2d\alpha + N^{1+\varepsilon}\int_{\mathfrak{K}}\Upsilon(\alpha)|f(\alpha,X_1)F_t(\alpha)^2|d\alpha \\ &\ll X_1\Xi^{2-\eta} + N^{1+\varepsilon}\int_{\mathfrak{K}}\Upsilon(\alpha)|f(\alpha,X_1)F_t(\alpha)^2|d\alpha. \qquad (5.8)\end{aligned}$$

We conclude our analysis of the minor arcs by noting first that Lemma 2 of Brüdern [1] in combination with (5.1) yields

$$\begin{aligned}\int_{\mathfrak{M}(N^\delta)}\Upsilon(\alpha)|F_t(\alpha)|^2d\alpha &\ll N^{\varepsilon-1}\Bigl(N^\delta\int_0^1|F_t(\alpha)|^2d\alpha + F_t(0)^2\Bigr)\\ &\ll N^{\varepsilon-1}\left(N^\delta\Xi + \Xi^2\right).\end{aligned}$$

But our hypotheses on t and δ ensure that

$$\delta - (1-(1-1/k)^t) = (1-1/k)^t - (1-\delta) < e^{-1} - 1/2 < 0,$$

whence the relation $\Xi \asymp N^{1-(1-1/k)^t}$ implies that $N^\delta \ll \Xi$. An application of Weyl's inequality therefore leads to the bound

$$\begin{aligned}\int_{\mathfrak{K}}\Upsilon(\alpha)|f(\alpha,X_1)F_t(\alpha)^2|d\alpha &\le \Bigl(\sup_{\alpha\in\mathfrak{m}}|f(\alpha,X_1)|\Bigr)\int_{\mathfrak{M}(N^\delta)}\Upsilon(\alpha)|F_t(\alpha)|^2d\alpha\\ &\ll (X_1\Xi^{-4\eta})N^{\varepsilon-1}\Xi^2 \ll N^{-1}X_1\Xi^{2-2\eta},\end{aligned}$$

provided only that η is a sufficiently small positive number. We may therefore conclude from (5.8) that

$$\sum_{d\le D}\Bigl|\int_{\mathfrak{m}} f(\alpha,X_1)F_t(\alpha)^2J_d(-\alpha)d\alpha\Bigr| \ll X_1\Xi^{2-\eta}. \qquad (5.9)$$

In the next step, we evaluate the contribution of the major arcs $\mathfrak{M}$ asymptotically. Let

$$w(\beta,X) = \int_X^{2X} e(\beta\gamma^k)d\gamma,$$

and put

$$W(\beta) = w(\beta,X_1)^3w(\beta,X_2)^2\dots w(\beta,X_t)^2.$$

We recall the definition of $S(q,a)$ from §2, and define the function $F^*(\alpha)$ for $\alpha\in\mathfrak{M}$ by taking

$$F^*(\alpha) = q^{-s}S(q,a)^sW(\alpha - a/q),$$

when $\alpha\in\mathfrak{M}(q,a;Z)\subseteq\mathfrak{M}$. It follows from Theorem 4.1 of [9] that whenever $\alpha\in\mathfrak{M}(q,a;Z)\subseteq\mathfrak{M}$, one has

$$f(\alpha,X_i) - q^{-1}S(q,a)w(\alpha-a/q;X_i) \ll Z^{1/2+\varepsilon}. \qquad (5.10)$$

Consequently, on making a trivial estimate for $J_d(\alpha)$, we find that

$$\sup_{\alpha\in\mathfrak{M}}\Bigl(|f(\alpha,X_1)F_t(\alpha)^2 - F^*(\alpha)|\sum_{d\le D}|J_d(-\alpha)|\Bigr) \ll N\Xi^2X_1X_t^{-1}Z^{1/2+\varepsilon}.$$

But the measure of $\mathfrak{M}$ is $O(Z^2N^{-1})$, and thus we deduce that

$$\sum_{d\le D}\Big|\int_{\mathfrak{M}}(f(\alpha,X_1)F_t(\alpha)^2-F^*(\alpha))J_d(-\alpha)d\alpha\Big|\ll \Xi^2X_1X_t^{-1/2}. \tag{5.11}$$

Let us now write

$$M(d)=\int_{\mathfrak{M}}F^*(\alpha)J_d(-\alpha)d\alpha.$$

Then it follows from (5.6), (5.9) and (5.11) that

$$\sum_{n\equiv 0\ (\mathrm{mod}\ d)}\varrho(n)=M(d)+R_1(d), \tag{5.12}$$

where, for $\eta>0$ sufficiently small,

$$\sum_{d\le D}|R_1(d)|\ll X_1\Xi^{2-\eta}. \tag{5.13}$$

Before evaluating the expression $M(d)$ more precisely, we rewrite it in the form

$$M(d)=\sum_{q\le Z}\sum_{\substack{a=1\\(a,q)=1}}^{q}q^{-s}S(q,a)^s\int_{-Z/(qN)}^{Z/(qN)}W(\beta)J_d(-\beta-a/q)d\beta, \tag{5.14}$$

and demonstrate that the terms in (5.14) with $q\nmid d$ make a negligible contribution. Observe first that when $(a,q)=1$ and $q\nmid d$, one has $\|da/q\|\ge q^{-1}\ge Z^{-1}$. Also, when $\eta>0$ is sufficiently small, we have $DZ^2\le N^{1-\eta}$, and thus for $\alpha\in\mathfrak{M}(q,a;Z)\subseteq\mathfrak{M}$ it follows that $|d(\alpha-a/q)|\le DZ/(qN)<(2Z)^{-1}$. We therefore arrive at the estimate

$$\sup_{|\beta|\le Z/(qN)}|J_d(\beta+a/q)|\ll\sup_{|\beta|\le Z/(qN)}\min\{N/d,\|d(\beta+a/q)\|^{-1}\}\ll Z, \tag{5.15}$$

valid for $0\le a\le q\le Z$, $(a,q)=1$ and $q\nmid d$. Next, on recalling the upper bound $w(\beta,X)\ll X(1+X^k|\beta|)^{-1}$ that is immediate on integrating by parts, we find that

$$\int_{-\infty}^{\infty}|W(\beta)|d\beta\ll X_1\Xi^2\int_{-\infty}^{\infty}(1+N|\beta|)^{-3}d\beta\ll X_1\Xi^2N^{-1}. \tag{5.16}$$

On combining (5.15) and (5.16) with Lemma 4.9 of [9], therefore, we deduce that for $t\ge k$, the contribution to the right hand side of (5.14) arising from those terms with $q\nmid d$ is majorised by

$$Z\int_{-\infty}^{\infty}|W(\beta)|d\beta\sum_{q\le Z}\sum_{\substack{a=1\\(a,q)=1}}^{q}|q^{-1}S(q,a)|^s\ll Z^{1+\varepsilon}X_1\Xi^2N^{-1}.$$

But when $q|d$, one has $J_d(\beta+a/q)=J_d(\beta)$, so that on writing

$$M^*(d)=\sum_{\substack{q\le Z\\q|d}}\sum_{\substack{a=1\\(a,q)=1}}^{q}q^{-s}S(q,a)^s\int_{-Z/(qN)}^{Z/(qN)}W(\beta)J_d(-\beta)d\beta, \tag{5.17}$$

we may conclude that

$$\sum_{d\le D}|M(d)-M^*(d)|\ll DZ^{1+\varepsilon}X_1\Xi^2N^{-1}\ll X_1\Xi^{2-\eta}. \tag{5.18}$$

Again applying only routine endgame technique in the Hardy-Littlewood method, we can complete the singular series and singular integral implicit in (5.17) to obtain

$$M^*(d) = \sum_{q|d} \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s} S(q,a)^s I(d) + O\left(\Xi^2 X_1 Z^{-1} d^{-1}\right),$$

where

$$I(d) = \int_{-\infty}^{\infty} W(\beta) J_d(-\beta) d\beta.$$

But Lemma 2.12 of [9] supplies the identity

$$\sum_{q|d} \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s} S(q,a)^s = d^{1-s} v_{k,s}(d), \tag{5.19}$$

and hence we infer that

$$\sum_{d\le D} \left| M^*(d) - d^{1-s} v_{k,s}(d) I(d) \right| \ll \Xi^2 Z^{-1} X_1 \log N \ll X_1 \Xi^{2-\eta}.$$

On combining this estimate with (5.12), (5.13) and (5.18), we may conclude thus far that

$$\sum_{d\le D} \Bigg| \sum_{n\equiv 0 \pmod d} \varrho(n) - d^{1-s} v_{k,s}(d) I(d) \Bigg| \ll X_1 \Xi^{2-\eta}. \tag{5.20}$$

It remains to evaluate $I(d)$. We note first that Euler's summation formula yields

$$\sum_{y\le Y} e(\alpha y) = \int_0^Y e(\alpha\gamma) d\gamma + O(1 + |\alpha| Y),$$

whence

$$J_d(\alpha) = d^{-1} J^*(\alpha) + O(1 + |\alpha| N),$$

where we write

$$J^*(\alpha) = \int_0^N e(\alpha\gamma) d\gamma.$$

But $W(\beta) \ll X_1 \Xi^2 (1 + N|\beta|)^{-3}$, so it follows that the estimate

$$I(d) = d^{-1} \int_{-\infty}^{\infty} W(\beta) J^*(-\beta) d\beta + O(\Xi^2 X_1 N^{-1})$$

holds uniformly in d. Next, on noting that $J^*(\beta) = (2\pi i\beta)^{-1}(e(\beta N) - 1)$, it follows by evaluating a simple contour integral that

$$\int_{-\infty}^{\infty} W(\beta) J^*(-\beta) d\beta = W(0) = X_1 \Xi^2,$$

and hence

$$I(d) = d^{-1} X_1 \Xi^2 + O(\Xi^2 X_1 N^{-1}).$$

Then since Theorem 4.2 of [9] coupled with (5.19) establishes that $d^{1-s} v_{k,s}(d) \ll 1$ whenever $t \ge k$, we conclude from (5.20) that

$$\sum_{d\le D} \Bigg| \sum_{n\equiv 0 \pmod d} \varrho(n) - d^{-s} v_{k,s}(d) X_1 \Xi^2 \Bigg| \ll \Xi^2 X_1 N^{-1} D + X_1 \Xi^{2-\eta}.$$

The conclusion of the lemma is now immediate from the estimate $DN^{-1} \ll \Xi^{-\eta}$. □

The argument required to establish Theorem 2 is now a simple exercise in sieve theory. It suffices to note that by (5.19) and Lemma 4.3 of [9], whenever π is prime and $t \geq k$, one has $v_{k,s}(\pi) = \pi^{s-1} + O(\pi^{s-2})$. We apply Lemma 5.1 with $D = N^{3/5}$, $\delta = 2/5$ and $t \geq k$. Then a standard version of the weighted linear sieve (see, for example, Greaves [7]) immediately implies that

$$\sum_{n \in \mathcal{N}} \varrho(n) \gg X_1 \Xi^{2-\varepsilon},$$

where $\mathcal{N}$ denotes the set of integers with at most two prime factors. But on recalling (5.1), we find that for each n in $\mathcal{N}$, one has

$$\varrho(n) = \int_0^1 f(\alpha, X_1) F_t(\alpha)^2 e(-n\alpha) d\alpha \leq \int_0^1 |f(\alpha, X_1) F_t(\alpha)^2| d\alpha \ll X_1 \Xi,$$

and hence

$$\sum_{\substack{n \in \mathcal{N} \\ \varrho(n) > 0}} 1 \gg \Xi^{1-\varepsilon}.$$

The conclusion of Theorem 2 is immediate.

We finish with some comments on the use of the set of smooth numbers $\mathcal{A}(P, P^\eta)$ and its brethren. The ranges for the variables $\mathbf{x}$ and $\mathbf{y}$ underlying the sums (5.2) are in diminishing ranges, and this permits strong control of the implicit exponential sums on major arcs of height up to a power of N via the relations (5.10). Unfortunately, if one were to place the variables $\mathbf{x}$ and $\mathbf{y}$ in the set $\mathcal{A}(N^{1/k}, N^\eta)$, surrogates for (5.10) are available for major arcs of height up to only a power of $\log N$, and this is inadequate for the purpose at hand. Such difficulties may be circumvented by use of the smooth sets $\mathcal{C}(N^{1/k}, N^\eta)$, where we define

$$\mathcal{C}(P, R) = \{lm \,:\, 1 \leq l \leq \sqrt{R},\, 1 \leq m \leq P/\sqrt{R},\, p|m \Rightarrow \sqrt{R} < p \leq R\}.$$

Here, as in the definition of $\mathcal{A}(P, R)$, the letter p denotes a prime number. We direct the reader to Brüdern and Wooley [5] (see especially §5) and [6] (see especially §8) for illustrative applications of such ideas. When $\mathbf{x}$ and $\mathbf{y}$ lie in such a set, an analogue of (5.10) holds for major arcs of height up to $N^{\eta/2}$ (see, for example, equations (8.11) and (8.12) of [6]). Moreover, available mean value estimates for the exponential sums

$$\tilde{g}(\alpha) = \sum_{x \in \mathcal{C}(P,R)} e(\alpha x^k)$$

retain the same strength as the estimates provided by (2.2) and (2.3). In this way, one may demonstrate that the conclusion of Theorem 2 remains valid with the condition $s \geq 2k+1$ replaced by

$$s \geq 2\lceil \tfrac{7}{11} k \rceil + 1$$

(here we make use of Greaves' weights [7] for detecting almost-primes).

References

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

[2] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, I: Waring's problem for cubes*, Ann. Sci. École Norm. Sup. (4) (to appear).

[3] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, IV: Lower bound methods*, (submitted).

[4] J. Brüdern and A. Perelli, *The addition of primes and powers*, Canad. J. Math. **48** (1996), 512–526.

[5] J. Brüdern and T. D. Wooley, *On Waring's problem for cubes and smooth Weyl sums*, Proc. London Math. Soc. (3) **82** (2001), 89–109.

[6] J. Brüdern and T. D. Wooley, *On Waring's problem: three cubes and a sixth power*, Nagoya Math. J. (to appear).

[7] G. Greaves, *The weighted linear sieve and Selberg's λ^2-method*, Acta Arith. **47** (1986), 71–96.

[8] R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.

[9] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd edn., Cambridge University Press, Cambridge 1997.

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

[11] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem, IV: Higher powers*, Acta Arith. **94** (2000), 203–285.

[12] T. D. Wooley, *Large improvements in Waring's problem*, Ann. of Math. (2) **135** (1992), 131–164.

[13] T. D. Wooley, *The application of a new mean value theorem to the fractional parts of polynomials*, Acta Arith. **65** (1993), 163–179.

JB: Mathematisches Institut A, Universität Stuttgart, D-70511 Stuttgart, Germany
E-mail address: bruedern@mathematik.uni-stuttgart.de

KK: Department of Mathematics, Faculty of Education, Iwate University, Morioka 020-8550, Japan
E-mail address: kawada@iwate-u.ac.jp

TDW: Department of Mathematics, University of Michigan, East Hall, 525 East University Avenue, Ann Arbor, MI 48109-1109, U.S.A.
E-mail address: wooley@math.lsa.umich.edu