ANALYTIC NUMBER THEORY: INTRODUCTION TO THE CIRCLE METHOD AND ITS APPLICATION, 2023

TREVOR D. WOOLEY

1. Introduction – motivation and history

The circle method was devised by Hardy and Ramanujan in 1918, with an important variant due to Hardy and Littlewood in 1920 known as *the Hardy-Littlewood method* (see [11] and [12], respectively). Hardy and Ramanujan were interested primarily in the partition function $p(n)$, which for each natural number n counts the number of ways of writing n in the shape

$$n = x_1 + \ldots + x_s,$$

with $s \in \mathbb{N}$ and $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_s$ all natural numbers[1]. Hardy and Littlewood were instead interested in Waring's problem.

Conjecture 1.1 (E. Waring, 1770)**.** *All natural numbers are the sum of at most* 4 *squares of natural numbers, or of at most* 9 *cubes of natural numbers, or of at most* 19 *fourth powers of natural numbers, and so on.*

In order to formulate a more precise statement, we introduce some notation.

Definition 1.2. Given $k \in \mathbb{N}$, we define $g(k)$ to be the least integer s having the property that all natural numbers are the sum of at most s positive integral k-th powers.

Thus, whenever $s \geqslant g(k)$, it follows that for each $n \in \mathbb{N}$ there exist non-negative integers $x_1, \ldots, x_s$ having the property that

$$n = x_1^k + \ldots + x_s^k.$$

A more precise version of Conjecture 1.1 may now be formulated.

Conjecture 1.3. *One has $g(2) = 4$, $g(3) = 9$, $g(4) = 19$, ..., and for each $k \in \mathbb{N}$ one has $g(k) < \infty$.*

Before further discussion of Hardy and Littlewood's ideas concerning the analysis of Waring's problem, we pause to consider what is known concerning this conjecture of Waring. The starting point is a theorem of Lagrange from 1770 showing that all positive integers are the sum of four squares of integers. Thus, for example, we have

$$2023 = 43^2 + 13^2 + 2^2 + 1^2.$$

Such representations need not be unique, even on ordering the summands, for we also have

$$2023 = 37^2 + 25^2 + 5^2 + 2^2.$$

[1] In this course the natural numbers $\mathbb{N}$ are defined to be the set of positive integers $\{1, 2, \ldots\}$.

Moreover, the integer 2023 is not the sum of 3 or fewer squares of natural numbers. The diligent reader might care to prove the latter assertion (hint: examine whether or not the congruence $x_1^2+x_2^2+x_3^2\equiv 2023 \pmod 8$ is soluble).

Investigations concerning larger exponents k culminated with the work of Hilbert [14] in 1909, who applied polynomial identities to show that for each $k\in\mathbb{N}$, one has $g(k)<\infty$. The method applied by Hilbert does not deliver satisfactorily explicit bounds on $g(k)$. It is worth noting that the representation of small numbers n leads to surprisingly large lower bounds on $g(k)$. Consider, for example, representations of the integer

$$n_0=2^k\left\lfloor (3/2)^k\right\rfloor -1,$$

in the shape $n_0=x_1^k+\ldots+x_s^k$, with $x_i\in\mathbb{N}$. Since $n_0<3^k$, one finds that $x_i\in\{1,2\}$ for each i, and hence the most efficient representation has the shape

$$n_0=2^k+\ldots+2^k+1^k+\ldots+1^k,$$

with $\lfloor (3/2)^k\rfloor -1$ copies of 2^k, and 2^k-1 copies of 1^k. Hence

$$g(k)\geqslant 2^k+\left\lfloor (3/2)^k\right\rfloor -2.$$

In fact, it is now known that $g(k)=2^k+\left\lfloor (3/2)^k\right\rfloor -2$, with the possible exception of a finite number of values of k (which almost surely do not occur – see Mahler [21]). It is known that if any exceptional exponent k occurs, then one necessarily has

$$2^k\left\{(3/2)^k\right\}+\left\lfloor (3/2)^k\right\rfloor >2^k, \tag{1.1}$$

and it was shown that this inequality fails for every exponent $k\leqslant 471,600,000$ (see Kubina and Wunderlich [16]). It is tempting to believe that progress on computational power in the last three decades should permit this inequality to be checked for very much larger values of k. Even in those exceptional circumstances where (1.1) holds, it is known that

$$g(k)=2^k+\left\lfloor (3/2)^k\right\rfloor+\left\lfloor (4/3)^k\right\rfloor -3$$

when $\left\lceil (4/3)^k\right\rceil\left\lceil (3/2)^k\right\rceil >2^k+1$, and that

$$g(k)=2^k+\left\lfloor (3/2)^k\right\rfloor+\left\lfloor (4/3)^k\right\rfloor -2$$

when $\left\lceil (4/3)^k\right\rceil\left\lceil (3/2)^k\right\rceil =2^k+1$.

Motivated by the difficulty of representing certain small integers, we may seek instead to understand the *typical* situation for large integers.

Definition 1.4. Given $k\in\mathbb{N}$, define $G(k)$ to be the least integer having the property that whenever $s\geqslant G(k)$, then all sufficiently large natural numbers are the sum of s positive integral k-th powers.

Thus, when $k\in\mathbb{N}$ and $s\geqslant G(k)$, there exists $N_0=N_0(s,k)$ such that, whenever $n\geqslant N_0$, then there exist $x_1,\ldots,x_s\in\mathbb{N}$ having the property that $n=x_1^k+\ldots+x_s^k$. A relatively easy exercise (see the first problem sheet) confirms that when $k\geqslant 2$, one has $G(k)\geqslant k+1$. In brief, the sharpest upper bounds currently available are as follows:

$G(2)=4$, a consequence of Lagrange's theorem from 1770;

$G(3)\leqslant 7$, due to Linnik, 1942;

$G(4)=16$, due to Davenport, 1939;

$G(5)\leqslant 17$, due to Vaughan and Wooley, 1995;

$G(6) \leqslant 24$, due to Vaughan and Wooley, 1994;

$G(7) \leqslant 31$, due to Wooley, 2016;

$G(8) \leqslant 39$, due to Wooley, 2016 (and it is known that $G(8) \geqslant 32$);

$G(9) \leqslant 47$, due to Wooley, 2016;

and so on (see [8, 19, 20, 24, 25, 31]). In general, for large values of k, it was shown 30 years ago that

$$G(k) \leqslant k(\log k + \log\log k + 2 + o(1)) \quad \text{(Wooley, 1992 and 1995)},$$

where $o(1) \to 0$ as $k \to \infty$ (see [27, 28]). Within the past year, this longstanding upper bound has been improved so that for all natural numbers k one has

$$G(k) \leqslant \lceil k(\log k + 4.20032) \rceil \quad \text{(Brüdern and Wooley [5])}.$$

Let us now return to Hardy and Littlewood in 1920, and indeed to Hardy and Ramanujan in 1918. Their idea was to write down a power series

$$g_k(z) = \sum_{m=1}^{\infty} z^{m^k}.$$

Note that this series is absolutely convergent for $|z| < 1$. If one now considers the expression $g_k(z)^s$, one sees that

$$\begin{aligned} g_k(z)^s &= \left(\sum_{m_1=1}^{\infty} z^{m_1^k}\right)\left(\sum_{m_2=1}^{\infty} z^{m_2^k}\right)\dots\left(\sum_{m_s=1}^{\infty} z^{m_s^k}\right) \\ &= \sum_{m_1=1}^{\infty}\dots\sum_{m_s=1}^{\infty} z^{m_1^k+\dots+m_s^k} \\ &= \sum_{n=1}^{\infty} R_{s,k}(n) z^n, \end{aligned}$$

where we write

$$R_{s,k}(n) = \text{card}\left\{m_1, \dots, m_s \in \mathbb{N} : m_1^k + \dots + m_s^k = n\right\}.$$

We can recover the coefficients $R_{s,k}(n)$ by employing Cauchy's integral formula to evaluate a suitable contour integral. Thus

$$R_{s,k}(n) = \frac{1}{2\pi i}\int_{\mathcal{C}} g_k(z)^s z^{-n-1}\,\mathrm{d}z,$$

in which $\mathcal{C}$ denotes a circular contour, centred at 0, and with radius r satisfying $0 < r < 1$. This was the age, after all, when methods from complex analysis seemed all-powerful, and any opportunity to wield such methods was difficult to resist!

When $k = 1$, the series in question is $g_1(z) = z/(1-z)$, and one is justified in being optimistic concerning the computation of $R_{s,k}(n)$. Hardy and Ramanujan were also able to make progress in the case $k = 2$. But how should one approach this computation when $k > 2$? This is where Hardy and Littlewood's innovations come into play.

In order to prepare the ground for an explanation of such innovations, it is expedient first to simplify the situation by making use of the observation made by I. M. Vinogradov in the 1930's to the effect that the infinite sums in question may be replaced by finite Fourier series without losing any information concerning $R_{s,k}(n)$. If one substitutes $z =$

$re(\theta)$, in which (as usual in analytic number theory) we write $e(\theta)$ for $e^{2\pi i\theta}$, then one sees initially that the series $g_k(z)$ can be replaced by an infinite Fourier series. Note that whenever $n = x_1^k + \ldots + x_s^k$, one must have $x_i \leqslant n^{1/k}$. With this observation in mind, we put $X = \lfloor n^{1/k} \rfloor$ and define

$$f_k(\theta) = \sum_{1 \leqslant x \leqslant X} e(\theta x^k).$$

We then have

$$\begin{aligned} f_k(\theta)^s &= \Bigg(\sum_{1 \leqslant x_1 \leqslant X} e(\theta x_1^k) \Bigg) \ldots \Bigg(\sum_{1 \leqslant x_s \leqslant X} e(\theta x_s^k) \Bigg) \\ &= \sum_{1 \leqslant x_1 \leqslant X} \ldots \sum_{1 \leqslant x_s \leqslant X} e(\theta(x_1^k + \ldots + x_s^k)) \\ &= \sum_{1 \leqslant m \leqslant sX^k} R^*_{s,k}(m) e(\theta m), \end{aligned}$$

where

$$R^*_{s,k}(m) = \mathrm{card}\left\{ m_1, \ldots, m_s \in \mathbb{N} \cap [1, X] : m_1^k + \ldots + m_s^k = m \right\}.$$

Thus, by applying the orthogonality relation

$$\int_0^1 e(\theta h) \, \mathrm{d}\theta = \begin{cases} 0, & \text{when } h \in \mathbb{Z} \setminus \{0\}, \\ 1, & \text{when } h = 0, \end{cases}$$

we deduce that

$$\int_0^1 f_k(\theta)^s e(-\theta n) \, \mathrm{d}\theta = \sum_{1 \leqslant m \leqslant sX^k} R^*_{s,k}(m) \int_0^1 e(\theta(m - n)) \, \mathrm{d}\theta = R_{s,k}(n).$$

In this guise, the key insight of Hardy and Littlewood may be explained as follows. When s is large enough in terms of k, there may be large contributions to the integral yielding $R_{s,k}(n)$ arising from those values of θ in small neighbourhoods of each rational number a/q in which $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a, q) = 1$ and q is small. Whilst this is certainly the case when $a = 0$, $q = 1$ and $0 \leqslant \theta \leqslant \frac{1}{8} X^{-k}$, in which case one has

$$|f_k(\theta)| \geqslant \mathrm{Re}\Bigg(\sum_{1 \leqslant x \leqslant X} e(1/8) \Bigg) \geqslant X \cos(\pi/4) = X/\sqrt{2},$$

similar observations hold in wider generality. Meanwhile, when θ is not closely approximated by a rational number with a small denominator, then the argument θx^k may be expected to be rather randomly distributed modulo 1, and hence the summands $e(\theta x^k)$ should exhibit plenty of cancellation when they are summed over x.

When $k = 1$, Hardy and Ramanujan were able to evaluate their generating functions asymptotically for all values of θ, and thereby obtain an asymptotic formula for $R_{s,k}(n)$. The method also applies even in the more delicate situation with $k = 2$. However, when $k \geqslant 3$, the generating function $f_k(\theta)$ could be evaluated only for θ lying on small neighbourhoods of rational numbers having small denominators. Hardy and Littlewood labelled such a set *the major arcs* in their application of the circle method. In order to motivate later discussion, consider a parameter δ with $0 < \delta < 1$ and the set

$$\mathfrak{M}_\delta = \bigcup_{\substack{0 \leqslant a \leqslant q \leqslant X^\delta \\ (a,q)=1}} \mathfrak{M}_\delta(q,a),$$

where

$$\mathfrak{M}_\delta(q,a) = \{\theta \in [0,1) : |\theta - a/q| \leqslant X^{\delta - k}\}.$$

Provided that $\delta < 1/3$ and X is sufficiently large, one sees that the *major arcs* $\mathfrak{M}_\delta(q,a)$ are distinct and non-overlapping for $0 \leqslant a \leqslant q \leqslant X^\delta$ and $(a,q) = 1$. Provided that δ is sufficiently small and $\theta \in \mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$, asymptotic formulae may be derived for the generating functions $f_k(\theta)$. Indeed, if we write

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q) \quad \text{and} \quad v(\beta) = \int_0^X e(\beta\gamma^k)\,\mathrm{d}\gamma,$$

then we shall see that

$$f_k(\theta) = q^{-1} S(q,a) v(\theta - a/q) + O(X^{2\delta}).$$

Thus, when s is large enough in terms of k, and again for small enough values of δ, one may obtain an asymptotic formula of the shape

$$\int_{\mathfrak{M}_\delta} f_k(\theta)^s e(-n\theta)\,\mathrm{d}\theta \sim \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n) n^{s/k-1}, \tag{1.2}$$

as $n \to \infty$. Here, we have written $\Gamma(\cdot)$ for the classical Gamma function defined for $\mathrm{Re}(z) > 0$ by putting

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t,$$

and $\mathfrak{S}_{s,k}(n)$ is a real number reflecting the local solubility behaviour of the equation $n = x_1^k + \ldots + x_s^k$ (that is, the solubility of congruences corresponding to this equation). We note in passing that in the contour integrals pursued by Hardy and Littlewood, the sets corresponding to the major arcs $\mathfrak{M}_\delta(q,a)$ were indeed arcs contained within the circular contour, and the moniker has stuck.

Since

$$R_{s,k}(n) = \int_0^1 f_k(\theta)^s e(-\theta n)\,\mathrm{d}\theta,$$

the asymptotic analysis of $R_{s,k}(n)$ will be completed by showing that the contribution arising from the set $\mathfrak{m}_\delta = [0,1) \setminus \mathfrak{M}_\delta$, the so-called *minor arcs*, is asymptotically smaller than the contribution (1.2) arising from the major arcs. In Hardy and Ramanujan's treatment of the cases $k = 1$ and 2, the minor arcs are absent from the analysis, but for $k \geqslant 3$ Hardy and Littlewood had somehow to estimate their contribution. Fortunately, they were able to make use of very recent work of Weyl concerning equidistribution of polynomials (see [26]). Thereby, they were able to show that when $\theta \in \mathfrak{m}_\delta$, then for each $\varepsilon > 0$ one has

$$f_k(\theta) = O(X^{1-\delta 2^{1-k}+\varepsilon}).$$

This permits one to show that for large enough values of s,

$$\int_{\mathfrak{m}_\delta} f_k(\theta)^s e(-\theta n)\,\mathrm{d}\theta = o(n^{s/k-1}),$$

whence

$$\begin{aligned} R_{s,k}(n) &= \int_{\mathfrak{M}_\delta} f_k(\theta)^s e(-\theta n)\,\mathrm{d}\theta + \int_{\mathfrak{m}_\delta} f_k(\theta)^s e(-\theta n)\,\mathrm{d}\theta \\ &\sim \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}\mathfrak{S}_{s,k}(n) n^{s/k-1}. \end{aligned}$$

Provided that $\mathfrak{S}_{s,k}(n)$ is bounded away from 0, one is able to infer that $R_{s,k}(n) \to \infty$ as $n \to \infty$, establishing that $G(k) \leqslant s$. Indeed, Hardy and Littlewood were initially able to establish such a conclusion whenever $s \geqslant (k-2)2^{k-1} + 5$.

In the next several sections, our task will be to justify this sketch analysis. Rather than following the initial analysis of Hardy and Littlewood too closely, we will instead pursue a simpler treatment originating with work of Hua from 1938 (see [15]). This argument permits the proof of an asymptotic formula for $R_{s,k}(n)$ whenever $s \geqslant 2^k + 1$, and in particular establishes the bound $G(k) \leqslant 2^k + 1$ $(k \geqslant 2)$.

Acknowledgements: This set of notes started life as a set of handwritten notes developed from courses given by the author at the University of Michigan and Harvard University. The present set of LaTexed notes were typed in real time, with no doubt many typos and oversights, at Purdue University in Spring 2023. I thank everyone who has notified me of typos and corrections in an earlier version, including (but not limited to) Yu-Ru Liu and Steve Fan.

2. A few notes on notation

We begin by recalling the *Bachmann-Landau* notation. When x is a real variable, the function $f(x)$ is complex valued, and $g(x)$ is a non-negative real valued function of x, we write $f(x) = O(g(x))$, as $x \to \infty$, when there exists a constant $C > 0$ having the property that, for all large enough values of x, one has

$$|f(x)| \leqslant Cg(x).$$

Equivalently, we may interpret $f(x) = O(g(x))$ as meaning that

$$\limsup_{x\to\infty} \frac{|f(x)|}{g(x)} < \infty.$$

We note that mention of the limiting environment $x \to \infty$ is often suppressed and taken as implicit. There are variants associated with limiting situations in which $x \to x_0$. In a similar manner, we write $f(x) = o(g(x))$ when

$$\lim_{x\to\infty} \frac{|f(x)|}{g(x)} = 0.$$

Vinogradov's notation seeks to summarise the Bachmann-Landau notation in a manner that is less awkward in multi-step arguments. In the scenario just introduced, we write $f(x) \ll g(x)$ to mean that $f(x) = O(g(x))$. Also, we write $f(x) \gg g(x)$ when both f and g are non-negative and $g(x) \ll f(x)$. Finally, we write $f(x) \asymp g(x)$ when one has both $f(x) \ll g(x)$ and $g(x) \ll f(x)$.

One should not confuse either notation with the statement that the functions $f(x)$ and $g(x)$ have the same order of growth, a notion employed in some other subdisciplines of mathematics. Thus, one has $3x+1 \ll x^2$ as $x \to \infty$, but not $x^2 \ll x$. Also, one has $x^2 \log x \gg x^2$ and $3x^2 + \log x \asymp x^2$. Finally, one has $3x^2+1 \asymp x^2 + \log x$ and $3x^2+1 = o(x^2 \log x)$ as $x \to \infty$.

In analytic number theory, it is nearly ubiquitous that a special convention is employed concerning the appearance of the letter ε. Thus, whenever a statement contains the letter ε, then we assert that the statement holds for all positive values of ε. The implicit quantifier renders the interpretation of ε as having a fixed value meaningless, and in this sense the value of ε may change from statement to statement. Thus, we may write a string of statements of the type: "we have $f(x) \ll x^\varepsilon \log x$, whence $f(x) \ll x^\varepsilon$" without further comment. If, on occasion, we wish to consider a fixed value of ε, then this is clearly stated in so many words so as to avoid ambiguity concerning this *ε-convention*.

Other standard notation from analytic number theory used in this course includes the following. When $\theta \in \mathbb{R}$, we write

$$\lfloor \theta \rfloor := \max_{\substack{n \in \mathbb{Z} \\ n \leqslant \theta}} n, \quad \lceil \theta \rceil := \min_{\substack{n \in \mathbb{Z} \\ n \geqslant \theta}} n, \quad \{\theta\} := \theta - \lfloor \theta \rfloor, \quad \|\theta\| := \min_{n \in \mathbb{Z}} |\theta - n|.$$

Also, when p and n are integers with $|p| > 1$, we write $p^h \| n$ when $p^h | n$ and $p^{h+1} \nmid n$ (that is, when p^h divides n and p^{h+1} does not divide n). Usually, we will employ this notation with p a (positive) prime number.

3. Weyl's inequality

We begin by considering pointwise bounds for the exponential sum

$$f(\alpha) = \sum_{1 \leqslant x \leqslant X} e(\alpha x^k)$$

with X large, seeking bounds of use on the minor arcs $\mathfrak{m}_\delta$. We begin with a consideration of the situation when $k = 1$.

Lemma 3.1. *Let X and Y be real numbers with $Y > 1$, and let $\alpha \in \mathbb{R}$. Then*

$$\sum_{X < x \leqslant X+Y} e(\alpha x) \ll \min\left\{Y, \|\alpha\|^{-1}\right\}. \tag{3.1}$$

We note that in the statement of this conclusion, we interpret $\min\{Y, \|0\|^{-1}\}$ as being equal to Y.

Proof. On utilising the trivial upper bound $|e(\alpha x)| \leqslant 1$ for each x, one sees that the left hand side of (3.1) is at most $Y+1$, and this bound already suffices to establish the desired conclusion when $\|\alpha\| = 0$. When $\alpha \neq 0$, meanwhile, by viewing the sum as a geometric progression, one finds that

$$\begin{aligned} \sum_{X < x \leqslant X+Y} e(\alpha x) &= \frac{e\left(\alpha \lfloor X+Y+1 \rfloor\right) - e\left(\alpha \lfloor X+1 \rfloor\right)}{e(\alpha) - 1} \\ &\ll |e(\alpha/2) - e(-\alpha/2)|^{-1} \\ &\ll |\sin(\pi\alpha)|^{-1}. \end{aligned}$$

The function $|\sin(\pi\alpha)|$ is periodic in α with period 1, and when $0 \leqslant \alpha \leqslant 1/2$, one may check that $2\alpha \leqslant \sin(\pi\alpha) \leqslant \pi\alpha$. We therefore infer that $|\sin(\pi\alpha)| \asymp \|\alpha\|$, whence

$$\sum_{X<x\leqslant X+Y} e(\alpha x) \ll \|\alpha\|^{-1}.$$

The conclusion of the lemma follows by combining this with the earlier trivial bound. $\square$

This lemma provides a means of estimating exponential sums of degree 1, but what about degrees k exceeding 1? An idea of Weyl (see [26]) permits one to reduce the degree inductively to obtain an estimate in terms of linear exponential sums. In order to outline this idea, consider a polynomial $\psi(x)$ with real coefficients and the exponential sum

$$T(X) = \sum_{1\leqslant x\leqslant X} e(\psi(x)).$$

Using the fact that the complex conjugate of $e(\psi(x))$ is equal to $e(-\psi(x))$, one finds that

$$\begin{aligned}|T(X)|^2 &= \sum_{1\leqslant y\leqslant X}\sum_{1\leqslant x\leqslant X} e\left(\psi(y)-\psi(x)\right)\\ &= \sum_{|h|<X}\sum_{\substack{1\leqslant x\leqslant X\\ 1\leqslant x+h\leqslant X}} e\left(\psi(x+h)-\psi(x)\right),\end{aligned}$$

in which we have made the change of variable $y = x+h$. But one has $\psi(x+h)-\psi(x) = h\psi_1(x;h)$, where $\psi_1(x;h)$ is a polynomial in x and h having degree $\deg(\psi)-1$ with respect to x. Thus, we have

$$|T(X)|^2 \leqslant \sum_{|h|<X}\left|\sum_{\substack{1\leqslant x\leqslant X\\ 1\leqslant x+h\leqslant X}} e\left(h\psi_1(x;h)\right)\right|,$$

with the inner exponential sum being one having a polynomial argument of degree one less than that of the original sum $T(X)$. This idea may now be used repeatedly to reduce the degree of the exponential sum under consideration until one is left in the linear situation amenable to Lemma 3.1.

A careful analysis of this idea is facilitated by the introduction of some standard notation. When $\psi(x)$ is a real-valued function of x, we denote by Δ_1 the forward difference operator defined via

$$\Delta_1(\psi(x);h) = \psi(x+h)-\psi(x).$$

We then define Δ_j for $j \geqslant 2$ recursively by means of the relation

$$\begin{aligned}\Delta_j(\psi(x);\mathbf{h}) &= \Delta_j(\psi(x);h_1,\dots,h_j)\\ &= \Delta_1\left(\Delta_{j-1}(\psi(x);h_1,\dots,h_{j-1});h_j\right).\end{aligned}$$

By convention, we take $\Delta_0(\psi(x);h) = \psi(x)$. One may verify that when $1 \leqslant j \leqslant k$, one has

$$\Delta_j(x^k;\mathbf{h}) = h_1\dots h_j p_j(x;h_1,\dots,h_j),$$

where p_j is a polynomial in x of degree $k-j$ with leading coefficient $k!/(k-j)!$. By the linearity of the operator Δ_j, one sees that

$$\Delta_j(a_kx^k+\ldots+a_1x;\mathbf{h}) = \sum_{i=1}^{k} a_i\Delta_j(x^i;\mathbf{h}),$$

and so one is able easily to infer the structure of expressions of the shape $\Delta_j(p(x);\mathbf{h})$, for polynomial arguments $p(x)$, by using what is known for the special case $p(x)=x^k$. We note, in particular, that when $j>\deg(p)$, then one has $\Delta_j(p(x);\mathbf{h})=0$.

Lemma 3.2 (Weyl differencing)**.** *Let $\psi(x)$ be a real-valued arithmetic function, and put*

$$F(\psi)=\sum_{1\leqslant x\leqslant X} e(\psi(x)).$$

Then for each natural number j, one has

$$|F(\psi)|^{2^j}\leqslant (2X)^{2^j-j-1}\sum_{|h_1|<X}\cdots\sum_{|h_j|<X}\sum_{x\in I_j(\mathbf{h})} e(\Delta_j(\psi(x);\mathbf{h})), \tag{3.2}$$

where $I_j(\mathbf{h})$ denotes the interval of integers defined by putting $I_0(h)=[1,X]$, and, when $j\geqslant 2$, by recursively setting

$$I_j(h_1,\dots,h_j)=I_{j-1}(h_1,\dots,h_{j-1})\cap\{x\in[1,X]:x+h_j\in I_{j-1}(h_1,\dots,h_{j-1})\}.$$

We note for future reference that the intervals occurring in the statement of this lemma satisfy the inclusions

$$I_j(h_1,\dots,h_j)\subseteq I_{j-1}(h_1,\dots,h_{j-1})\subseteq\ldots\subseteq I_1(h_1)\subseteq[1,X].$$

Proof. We proceed by induction. When $j=1$, one has

$$\begin{aligned}|F(\psi)|^2&=\sum_{1\leqslant x\leqslant X}\sum_{1\leqslant y\leqslant X} e(\psi(y)-\psi(x))\\&=\sum_{1\leqslant x\leqslant X}\sum_{1-x\leqslant h_1\leqslant X-x} e(\psi(x+h_1)-\psi(x))\\&=\sum_{|h_1|<X}\sum_{x\in I_1(h_1)} e(\Delta_1(\psi(x);h_1)),\end{aligned}$$

where $I_1(h_1)=[1,X]\cap[1-h_1,X-h_1]$. This confirms (3.2) when $j=1$.

Suppose now that (3.2) has been established for all integers j with $1\leqslant j<J$. Then, as a consequence of Cauchy's inequality, one finds that

$$|F(\psi)|^{2^J}=\left(|F(\psi)|^{2^{J-1}}\right)^2\leqslant\left((2X)^{2^{J-1}-J}\right)^2\left(\sum_{|h_1|<X}\cdots\sum_{|h_{J-1}|<X}1\right)\Xi(\psi),$$

where

$$\Xi(\psi)=\sum_{|h_1|<X}\cdots\sum_{|h_{J-1}|<X}\left|\sum_{x\in I_{J-1}(\mathbf{h})} e(\Delta_{J-1}(\psi(x);\mathbf{h}))\right|^2.$$

As in the case $j=1$, one discerns that

$$\left|\sum_{x\in I_{J-1}(\mathbf{h})} e\left(\Delta_{J-1}(\psi(x);\mathbf{h})\right)\right|^2=\sum_{|h_J|<X}\sum_{x\in I_J(\mathbf{h})} e\left(\Delta_1\left(\Delta_{J-1}(\psi(x);\mathbf{h});h_J\right)\right),$$

where

$$I_J(\mathbf{h},h_J)=I_{J-1}(\mathbf{h})\cap\{x\in[1,X]:x+h_J\in I_{J-1}(\mathbf{h})\}.$$

We therefore conclude that

$$|F(\psi)|^{2^J} \leqslant (2X)^{2^J-J-1} \sum_{|h_1|<X} \cdots \sum_{|h_J|<X} \sum_{x\in I_J(\mathbf{h})} e(\Delta_J(\psi(x);\mathbf{h})).$$

This confirms the inductive hypothesis (3.2) for $j=J$, and the conclusion of the lemma follows via induction. □

If we apply Weyl differencing $k-1$ times to the exponential sum

$$f(\alpha)=\sum_{1\leqslant x\leqslant X} e(\alpha x^k),$$

then we obtain a bound of the shape

$$|f(\alpha)|^{2^{k-1}} \ll X^{2^{k-1}-k} \sum_{\substack{h_1,\ldots,h_{k-1}\\ |h_i|<X}} \sum_{x\in I(\mathbf{h})} e\left(k!\alpha h_1\ldots h_{k-1}\left(x+\tfrac{1}{2}(h_1+\ldots+h_{k-1})\right)\right).$$

This gives us an exponential sum amenable to Lemma 3.1, but we will need to average over the products $h_1\ldots h_{k-1}$. This entails a discussion of the r-fold divisor function

$$\tau_r(n)=\sum_{\substack{d_i\in\mathbb{N}\\ d_1\ldots d_r=n}} 1$$

in the case $r=k-1$.

Lemma 3.3. *For each $r\in\mathbb{N}$ and $\varepsilon>0$, one has $\tau_r(n)\ll_{r,\varepsilon} n^\varepsilon$.*

Proof. We note that by considering the prime factorisations of n and d_i with $n=d_1\ldots d_r$, one obtains for each fixed $\varepsilon>0$ the relation

$$\frac{\tau_r(n)}{n^\varepsilon}=\prod_{\substack{p \text{ prime}\\ p^h\|n}} \frac{\tau_r(p^h)}{p^{\varepsilon h}} \leqslant \prod_{p^h\|n} \frac{(h+1)^{r-1}}{p^{\varepsilon h}}.$$

In the last inequality, we have made use of an upper bound for the number of ways of writing h in the shape $h=a_1+\ldots+a_r$ with $0\leqslant a_i\leqslant h$, this being connected to $\tau_r(p^h)$ via the implicit relation $p^{a_1}\ldots p^{a_r}=p^h$. Plainly, one can make use of a sharper bound here, but the crude bound that we have chosen to employ is sufficient for our purposes. For on noting the trivial inequality $h+1\leqslant 2^h$, one sees that

$$(h+1)^{r-1}p^{-\varepsilon h}\leqslant (2^r/p^\varepsilon)^h,$$

a quantity that is at most 1 for $p\geqslant 2^{r/\varepsilon}$. Meanwhile, the function $(h+1)^{r-1}/p^{\varepsilon h}$ is bounded above, uniformly in p, purely in terms of r and ε as h increases. Thus, one has

$$(h+1)^{r-1}p^{-\varepsilon h}\leqslant (h+1)^{r-1}2^{-\varepsilon h}<C(r,\varepsilon),$$

say. We consequently obtain the bound

$$\frac{\tau_r(n)}{n^\varepsilon}\leqslant \prod_{p<2^{r/\varepsilon}} C(r,\varepsilon)<C(r,\varepsilon)^{2^{r/\varepsilon}}\ll_{r,\varepsilon} 1.$$

Then $\tau_r(n)\ll_{r,\varepsilon} n^\varepsilon$, and the proof of the lemma is complete. □

Our final prerequisite from classical elementary number theory is Dirichlet's theorem on Diophantine approximation.

Lemma 3.4 (Dirichlet's approximation theorem)**.** *Let $\alpha \in \mathbb{R}$, and suppose that $X \geqslant 1$ is a real number. Then there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$ and $1 \leqslant q \leqslant X$ such that $|\alpha - a/q| \leqslant 1/(qX)$.*

Corollary 3.5. *When $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, there exist infinitely many rational numbers a/q with $a \in \mathbb{Z}$, $q \in \mathbb{N}$ and $(a, q) = 1$ such that $|\alpha - a/q| \leqslant q^{-2}$.*

Proof. Apply Lemma 3.4 to see that for each $X \geqslant 1$ one has $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$, $1 \leqslant q \leqslant X$ and

$$|\alpha - a/q| \leqslant 1/(qX) \leqslant 1/q^2.$$

If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then $|\alpha - a/q| > 0$, so we may fix a real number Y with $Y > |\alpha - a/q|^{-1}$. Applying Lemma 3.4 again, we see that there exist $b \in \mathbb{Z}$, $r \in \mathbb{N}$ with $(b, r) = 1$, $1 \leqslant r \leqslant Y$ and $|\alpha - b/r| \leqslant 1/r^2$. Moreover,

$$|\alpha - b/r| \leqslant 1/(rY) < |\alpha - a/q|,$$

so that $b/r \neq a/q$. Repeating this argument yields arbitrarily many distinct approximations to α of the desired type. □

The proof of Lemma 3.4. We apply the box principle. Let $N = \lfloor X \rfloor$, and consider the N real numbers $\alpha r - \lfloor \alpha r \rfloor$ for $1 \leqslant r \leqslant N$. All of these real numbers lie in the interval $[0, 1)$. If any of them lie in either $[0, 1/(N + 1))$ or $[N/(N + 1), 1)$, then (since $N + 1 > X$) we are done with $q = r/(r, b)$ and $a = b/(r, b)$, in which $b = \lfloor \alpha r \rfloor$ or $b = \lfloor \alpha r \rfloor + 1$. We may therefore suppose that all N of these real numbers lie in one of the $N - 1$ intervals

$$\left[\frac{j-1}{N+1}, \frac{j}{N+1}\right) \quad (2 \leqslant j \leqslant N).$$

It follows that two at least lie in some common such interval, say $\alpha u - \lfloor \alpha u \rfloor$ and $\alpha v - \lfloor \alpha v \rfloor$ with $u < v$. We then have

$$|\alpha(v - u) - (\lfloor \alpha v \rfloor - \lfloor \alpha u \rfloor)| < 1/(N + 1) < 1/X,$$

and the desired conclusion follows with $q = r/(r, b)$ and $a = b/(r, b)$, in which $r = v - u$ and $b = \lfloor \alpha v \rfloor - \lfloor \alpha u \rfloor$. □

We are now equipped to derive an averaged version of Lemma 3.1. Here, we proceed in slightly wider generality than is required for our purpose at hand.

Lemma 3.6. *Suppose that X and Y are real numbers with $X \geqslant 1$ and $Y \geqslant 1$. Suppose also that α and β are real numbers, and that there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$ and $|\alpha - a/q| \leqslant q^{-2}$. Then one has*

$$\sum_{1 \leqslant x \leqslant X} \min\{Y, \|\alpha x + \beta\|^{-1}\} \ll XY \left(q^{-1} + Y^{-1} + X^{-1} + q(XY)^{-1}\right) \log(2q).$$

Proof. We may suppose that $\alpha \in \mathbb{R}$, $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy the conditions $(a, q) = 1$ and $|\alpha - a/q| \leqslant 1/q^2$, conditions that may always be ensured by application of Dirichlet's theorem on Diophantine approximation. We write $\theta = \alpha - a/q$, so that $|\theta| \leqslant 1/q^2$. Since α is close to a/q, we would like to imagine that as x varies, the successive values of $\|\alpha x + \beta\|$ are of size roughly $1/q$, $2/q$, ..., in some order. However, it is possible that some values are duplicated, and then we must also take care of values of $\|\alpha x + \beta\|$ very close to 0.

Divide up the range of summation into intervals of integers of length $\lfloor q/2\rfloor + 1$, and consider a typical such interval, say

$$\mathcal{I} = \{Z, Z+1, \ldots, Z + \lfloor q/2\rfloor\}.$$

For any two distinct integers n_1 and n_2 with $n_2 < n_1$ lying in $\mathcal{I}$, we have

$$\|(\alpha n_1 + \beta) - (\alpha n_2 + \beta)\| = \|\alpha(n_1 - n_2)\| \geqslant \left\| \frac{a(n_1 - n_2)}{q} \right\| - |n_1 - n_2||\theta|.$$

Since $n_1 \ne n_2$ and $|n_1 - n_2| \leqslant q/2$, we have $q \nmid (n_1 - n_2)$. On recalling that $(a, q) = 1$, we therefore deduce that

$$\|(\alpha n_1 + \beta) - (\alpha n_2 + \beta)\| \geqslant \frac{1}{q} - \frac{q/2}{q^2} = \frac{1}{2q}.$$

We have shown that, modulo 1, the real numbers $\alpha n_1 + \beta$ and $\alpha n_2 + \beta$ are distinct, and spaced apart by a distance at least $(2q)^{-1}$. But then

$$\begin{aligned}\sum_{n\in\mathcal{I}} \min\{Y, \|\alpha n + \beta\|^{-1}\} &\leqslant Y + \sum_{0<|r|\leqslant q/2} \left|\frac{r}{2q}\right|^{-1} \\ &\leqslant Y + 4 \sum_{1\leqslant r\leqslant q/2} q/r \\ &\ll Y + q\log(2q).\end{aligned}$$

Observe next that there are at most $\lceil X/(q/2)\rceil$ intervals of the shape $\mathcal{I}$ required to cover all of the summands x with $1 \leqslant x \leqslant X$. It follows that

$$\sum_{1\leqslant x\leqslant X} \min\{Y, \|\alpha x + \beta\|^{-1}\} \ll (X/q + 1)(Y + q\log(2q)),$$

and the conclusion of the lemma follows. □

We may now execute the strategy that we have carefully prepared designed to estimate exponential sums with polynomial arguments.

Lemma 3.7 (Weyl's inequality). *Let $k \geqslant 2$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$ and $|\alpha_k - a/q| \leqslant q^{-2}$. Then*

$$\sum_{1\leqslant x\leqslant X} e(\alpha_1 x + \ldots + \alpha_k x^k) \ll X^{1+\varepsilon}\left(q^{-1} + X^{-1} + qX^{-k}\right)^{2^{1-k}}.$$

Proof. Write $\psi(x) = \alpha_1 x + \ldots + \alpha_k x^k$ and

$$F(\boldsymbol{\alpha}) = \sum_{1\leqslant x\leqslant X} e(\psi(x)).$$

Note first that when $q > X^k$, the desired estimate is trivial from the bound

$$|F(\boldsymbol{\alpha})| \leqslant \sum_{1\leqslant x\leqslant X} 1 \leqslant X.$$

Thus, we may suppose without loss that $q \leqslant X^k$.

Next, by applying the Weyl differencing lemma with $j = k-1$ (see Lemma 3.2), we obtain the bound

$$|F(\boldsymbol{\alpha})|^{2^{k-1}} \ll X^{2^{k-1}-k} \sum_{|h_1|<X} \cdots \sum_{|h_{k-1}|<X} \Upsilon(\mathbf{h}),$$

where we write

$$\Upsilon(\mathbf{h}) = \sum_{x \in I_{k-1}(\mathbf{h})} e(\Delta_{k-1}(\psi(x);\mathbf{h})),$$

and $I_{k-1}(\mathbf{h})$ is a suitable interval of integers contained in $[1,X]$. We note that

$$\Delta_{k-1}(\psi(x);\mathbf{h}) = k! h_1 \dots h_{k-1} x \alpha_k + \gamma,$$

where $\gamma = \gamma(\boldsymbol{\alpha};\mathbf{h})$ is independent of x. Thus, Lemma 3.1 delivers the bound

$$\Upsilon(\mathbf{h}) \ll \min\left\{X, \|k!h_1\dots h_{k-1}\alpha_k\|^{-1}\right\},$$

whence

$$|F(\boldsymbol{\alpha})|^{2^{k-1}} \ll X^{2^{k-1}-k} \sum_{|h_1|<X} \cdots \sum_{|h_{k-1}|<X} \min\left\{X, \|k!h_1\dots h_{k-1}\alpha_k\|^{-1}\right\}.$$

Accounting for the summands in which $h_1 \dots h_{k-1} = 0$, we are led from here to the bound

$$|F(\boldsymbol{\alpha})|^{2^{k-1}} \ll X^{2^{k-1}-k}\biggl(X^{k-1} + \sum_{1 \leqslant n \leqslant k!X^{k-1}} \tau_k(n) \min\left\{X, \|n\alpha_k\|^{-1}\right\}\biggr).$$

Invoking Lemma 3.6 and recalling our assumption that $q \leqslant X^k$, we therefore obtain the estimate

$$F(\boldsymbol{\alpha}) \ll X^{1-2^{1-k}} + X^{1+\varepsilon}\left(q^{-1} + X^{-1} + X^{1-k} + qX^{-k}\right)^{2^{1-k}},$$

and the conclusion of the lemma follows at once. □

We pause momentarily to reflect on the consequences of Weyl's inequality so far as the behaviour of the exponential sum

$$f(\alpha) = \sum_{1 \leqslant x \leqslant X} e(\alpha x^k) \tag{3.3}$$

is concerned for $\alpha \in \mathfrak{m}_\delta$. Given $\alpha \in [0,1)$, it is a consequence of Dirichlet's approximation theorem that there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $1 \leqslant q \leqslant X^{k-\delta}$, $(a,q) = 1$ and

$$|\alpha - a/q| \leqslant 1/(qX^{k-\delta}) \leqslant X^{\delta-k}.$$

If $q \leqslant X^\delta$, then we would have $\alpha \in \mathfrak{M}_\delta$. Thus, when $\alpha \in \mathfrak{m}_\delta$, we may suppose that $X^\delta < q \leqslant X^{k-\delta}$. We thus conclude from Weyl's inequality that, whenever $0 < \delta < 1$, one has

$$f(\alpha) \ll X^{1+\varepsilon}\left(X^{-\delta} + X^{-1} + X^{k-\delta}/X^k\right)^{2^{1-k}} \ll X^{1-\delta 2^{1-k}+\varepsilon}.$$

Provided that $s > (k/\delta)2^{k-1}$, we may conclude from this upper bound that

$$\begin{aligned}\left|\int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha\right| &\leqslant \Bigl(\sup_{\alpha \in \mathfrak{m}_\delta} |f(\alpha)|\Bigr)^s \int_{\mathfrak{m}_\delta} \mathrm{d}\alpha \\ &\ll \Bigl(X^{1-\delta 2^{1-k}+\varepsilon}\Bigr)^s = o(X^{s-k}).\end{aligned}$$

With such a value of s, it transpires that this minor arc estimate is small enough to establish an asymptotic formula for $R_{s,k}(n)$. It remains to handle the contribution of the major arcs, a project that requires much further analysis. However, a transference principle permits us to obtain useful bounds for $f(\alpha)$ on sets of major arcs with no additional effort.

Lemma 3.8. *Let θ, X, Y, Z be positive real numbers. Suppose that $\Psi : \mathbb{R} \to \mathbb{C}$ satisfies the property that whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q)=1$ and $|\alpha - a/q| \leqslant 1/q^2$, then*

$$\Psi(\alpha) \ll X(q^{-1}+Y^{-1}+qZ^{-1})^\theta. \tag{3.4}$$

Then, whenever $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ satisfy $(b,r)=1$, one has

$$\Psi(\alpha) \ll X(\lambda^{-1}+Y^{-1}+\lambda Z^{-1})^\theta, \tag{3.5}$$

where $\lambda = r + Z|r\alpha - b|$.

Proof. Suppose that $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ satisfy $(b,r)=1$. By Dirichlet's approximation theorem, there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $1 \leqslant q \leqslant 2r$, $(a,q)=1$ and $|q\alpha - a| \leqslant (2r)^{-1}$. Suppose in the first instance that $a/q \neq b/r$. Then

$$\frac{1}{qr} \leqslant \left| \frac{a}{q} - \frac{b}{r} \right| \leqslant \left| \alpha - \frac{a}{q} \right| + \left| \alpha - \frac{b}{r} \right| \leqslant \left| \alpha - \frac{b}{r} \right| + \frac{1}{2qr},$$

whence $q^{-1} \leqslant 2|r\alpha - b|$. It therefore follows from (3.4) that

$$\begin{aligned} \Psi(\alpha) &\ll X\left(|r\alpha - b| + Y^{-1} + rZ^{-1}\right)^\theta \\ &\ll X\left(Y^{-1} + \lambda Z^{-1}\right)^\theta, \end{aligned}$$

and (3.5) follows.

If, meanwhile, one has $a/q = b/r$, then since $(a,q)=(b,r)=1$, one has $q=r$ and $a=b$, whence $|r\alpha - b| \leqslant (2r)^{-1}$. Then if $\alpha = b/r$, we have $\lambda = r$, and (3.5) is again immediate from (3.4). When $\alpha \neq b/r$, on the other hand, we have the bounds $0 < |\alpha - b/r| \leqslant r^{-2}$. Applying Dirichlet's approximation theorem here, we obtain $a' \in \mathbb{Z}$ and $q' \in \mathbb{N}$ with $1 \leqslant q' \leqslant 2|r\alpha - b|^{-1}$, $(a',q')=1$ and $|q'\alpha - a'| \leqslant \frac{1}{2}|r\alpha - b|$. Now, if $a'/q' = b/r$, we again have $(a',q')=(b,r)=1$, and so $q'=r$ and $a'=b$. Thus

$$0 < |r\alpha - b| \leqslant \tfrac{1}{2}|r\alpha - b|,$$

yielding a contradiction. Then $a'/q' \neq b/r$, whence

$$\begin{aligned} \frac{1}{q'r} \leqslant \left| \frac{a'}{q'} - \frac{b}{r} \right| &\leqslant \left| \alpha - \frac{a'}{q'} \right| + \left| \alpha - \frac{b}{r} \right| \\ &\leqslant \left| \alpha - \frac{b}{r} \right| + (2q')^{-1}|r\alpha - b| \\ &\leqslant \left| \alpha - \frac{b}{r} \right| + \frac{1}{2q'r}. \end{aligned}$$

We therefore infer that $|r\alpha - b| \geqslant (2q')^{-1}$, and by applying (3.4) with a and q replaced by a' and q', we obtain the bound

$$\Psi(\alpha) \ll X\left(|r\alpha - b| + Y^{-1} + (Z|r\alpha - b|)^{-1}\right)^\theta.$$

As an alternative bound, since $|\alpha - b/r| \leqslant r^{-2}$, we may apply (3.4) to give the estimate

$$\Psi(\alpha) \ll X\left(r^{-1} + Y^{-1} + rZ^{-1}\right)^\theta.$$

Thus, in either case, one has

$$\Psi(\alpha) \ll X \left(\lambda^{-1} + Y^{-1} + \lambda Z^{-1}\right)^{\theta},$$

where $\lambda = r + Z|r\alpha - b|$. This completes the proof of the lemma. □

We recall the definition of the exponential sum $f(\alpha)$ from (3.3).

Corollary 3.9. *Suppose that $\alpha \in \mathbb{R}$. Then whenever $a \in \mathbb{Z}$ and $q \in \mathbb{Z}$ satisfy $(a, q) = 1$, one has*

$$f(\alpha) \ll X^{1+\varepsilon} \left(\frac{1}{q + X^k|q\alpha - a|} + X^{-1} + \frac{q + X^k|q\alpha - a|}{X^k} \right)^{2^{1-k}}.$$

Proof. Simply combine Lemmata 3.7 and 3.8. □

Estimates of the shape presented in Corollary 3.9 provide control of the exponential sums in question for α lying in parts of the unit interval that one might ordinarily construe as lying on sets of major arcs. In order to illustrate such notions, we offer a mean value estimate for an exponential sum over the entire unit interval.

Corollary 3.10. *Suppose that $k \geqslant 2$ and $s \geqslant k2^{k-1}$. Then one has*

$$\int_0^1 |f(\alpha)|^s \, \mathrm{d}\alpha \ll X^{s-k+\varepsilon}.$$

Proof. Suppose that $\alpha \in \mathbb{R}$. By Dirichlet's approximation theorem, there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $1 \leqslant q \leqslant X^{k-1}$, $(a, q) = 1$ and $|\alpha - a/q| \leqslant q^{-1}X^{1-k}$. By Corollary 3.9, one then has

$$\begin{aligned} f(\alpha) &\ll X^{1+\varepsilon} \left(\frac{1}{q + X^k|q\alpha - a|} + X^{-1} + \frac{X^{k-1} + X}{X^k} \right)^{2^{1-k}} \\ &\ll X^{1-2^{1-k}+\varepsilon} + \frac{X^{1+\varepsilon}}{(q + X^k|q\alpha - a|)^{2^{1-k}}}. \end{aligned}$$

Notice here that, when $q > X$, the second term on the right hand side is no larger than the first. Thus, on making use of the trivial inequality $|a + b|^s \ll |a|^s + |b|^s$, we deduce that

$$\int_0^1 |f(\alpha)|^s \, \mathrm{d}\alpha \ll (X^{1-2^{1-k}+\varepsilon})^s + I, \tag{3.6}$$

where

$$I = \sum_{1 \leqslant q \leqslant X} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{a/q - q^{-1}X^{1-k}}^{a/q + q^{-1}X^{1-k}} \left(\frac{X^{1+\varepsilon}}{(q + X^k|q\alpha - a|)^{2^{1-k}}} \right)^s \mathrm{d}\alpha.$$

We have

$$I \ll X^{s+\varepsilon} \sum_{1 \leqslant q \leqslant X} q^{1-k} \int_{-1/2}^{1/2} \frac{d\beta}{(1 + X^k|\beta|)^k} \ll X^{s-k+\varepsilon} \log X.$$

Thus, since $s(1 - 2^{1-k}) < s - k$, the conclusion of the lemma follows from (3.6). □

By combining the conclusion of Corollary 3.10 with Weyl's inequality, we may conclude thus far that when $s \geqslant k2^{k-1}+1$, one has

$$\begin{aligned}\left|\int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha\right| &\leqslant \Bigl(\sup_{\alpha\in\mathfrak{m}_\delta} |f(\alpha)|\Bigr)^{s-k2^{k-1}} \int_0^1 |f(\alpha)|^{k2^{k-1}}\,\mathrm{d}\alpha\\ &\ll \Bigl(X^{1-\delta 2^{1-k}+\varepsilon}\Bigr)^{s-k2^{k-1}} X^{k2^{k-1}-k+\varepsilon} = o(X^{s-k}).\end{aligned}$$

This provides a satisfactory treatment of the contribution of the minor arcs at the cost of $k2^{k-1}+1$ variables. In the next section we will exploit an idea of Hua so as to reduce the number of variables to 2^k+1.

It may be illuminating to explain that, in the context of exponential sums of the shape

$$f_k(\alpha) = \sum_{1\leqslant x\leqslant X} e(\alpha x^k),$$

one has a notion of *complexity* associated with each real argument α. To be precise, one can define the quantity $\mathfrak{c}(\alpha)=\mathfrak{c}_k(\alpha;X)$ by putting

$$\begin{aligned}\mathfrak{c}(\alpha) &= \inf\{q+X^k|q\alpha-a| : q\in\mathbb{N} \text{ and } a\in\mathbb{Z}\}\\ &= \inf_{q\in\mathbb{N}}\left(q+X^k\|q\alpha\|\right).\end{aligned}$$

Equipped with this notion, we see that the real numbers $\alpha\in[0,1)$ having small complexity $\mathfrak{c}(\alpha)$ define the major arcs, while $\alpha\in[0,1)$ of large complexity define the minor arcs. Notice that, as a consequence of Dirichlet's approximation theorem, given a real number α, there exists $q\in\mathbb{N}$ with $1\leqslant q\leqslant X^{k/2}$ for which $\|q\alpha\|\leqslant X^{-k/2}$. Hence all real numbers α satisfy the complexity constraint $\mathfrak{c}(\alpha)\leqslant 2X^{k/2}$. The enthusiastic reader may verify that there is a subset of $[0,1)$ of asymptotically full measure (as $X\to\infty$) for which $\mathfrak{c}(\alpha)\gg X^{k/2-\varepsilon}$. Classical technology in the circle method permits useful asymptotic analyses of exponential sums of degree k when the arguments α have complexity no larger than about X. More recent technology making use of Poisson summation permits this range to be extended almost as far as $\mathfrak{c}(\alpha)\leqslant X^2$. One may infer that direct approaches are available for linear, quadratic and cubic problems, while quartic and higher degree problems lie beyond reach of direct approaches. It is in the handling of points α having complexity in the range $X^2\leqslant\mathfrak{c}(\alpha)\leqslant 2X^{k/2}$ that the deepest problems in the circle method currently reside.

We finish this section by noting that, for small values of k, the sharpest available estimates of Weyl type take the shape

$$\sup_{\alpha\in\mathfrak{m}_1}|f_k(\alpha)| \ll X^{1-2^{1-k}}(\log X)^{A_k},$$

where A_k is a suitable positive number. This estimate, in which Weyl's factor of X^ε is replaced by $(\log X)^{A_k}$, is due to Vaughan [22, 23]. When $k\geqslant 6$, one has the superior estimate

$$\sup_{\alpha\in\mathfrak{m}_1}|f_k(\alpha)| \ll X^{1-1/(k(k-1))+\varepsilon},$$

which is a consequence of recent progress on Vinogradov's mean value theorem (see [3] and [32, 33]). The latter is a topic to which we shall return later in the course.

4. Hua's lemma

By combining the conclusion of Corollary 3.10 with Weyl's inequality, one finds that when $s \geqslant k2^{k-1}+1$, then for a suitable $\delta > 0$, one has

$$\int_{\mathfrak{m}_\delta} |f_k(\alpha)|^s \,\mathrm{d}\alpha \ll X^{s-k-\delta 2^{-k}}.$$

Hardy and Littlewood [12] obtained this estimate under the slightly less stringent hypothesis $s \geqslant (k-2)2^{k-1}+5$, and later Hua [15] reduced this bound on s to $s \geqslant 2^k+1$. This latter argument involves a Diophantine interpretation of the even moments of exponential sums. Throughout this section, we take X to be a large real number and write

$$f(\alpha) = \sum_{1 \leqslant x \leqslant X} e(\alpha x^k).$$

Lemma 4.1 (Hua's lemma)**.** *Suppose that $k \geqslant 2$ and $1 \leqslant j \leqslant k$. Then one has*

$$\int_0^1 |f(\alpha)|^{2^j} \,\mathrm{d}\alpha \ll X^{2^j-j+\varepsilon}.$$

Proof. We proceed by induction on j. When $j=1$, it follows via orthogonality that

$$\int_0^1 |f(\alpha)|^2 \,\mathrm{d}\alpha = \int_0^1 f(\alpha)f(-\alpha)\,\mathrm{d}\alpha = \mathrm{card}\{1 \leqslant x, y \leqslant X : x^k = y^k\} = \lfloor X \rfloor.$$

This confirms the claimed conclusion in the base case $j=1$.

Now suppose that the desired conclusion has been established already when $1 \leqslant j \leqslant J$ for some integer J satisfying $1 \leqslant J < k$. We apply the Weyl differencing lemma (see Lemma 3.2) to see that

$$|f(\alpha)|^{2^J} \leqslant (2X)^{2^J-J-1} \sum_{|h_1|<X} \cdots \sum_{|h_J|<X} \sum_{x \in I_J(\mathbf{h})} e\left(\alpha \Delta_J(x^k;\mathbf{h})\right),$$

where $I_J(\mathbf{h})$ is a suitable subinterval of $[1,X]$. It follows that

$$\begin{aligned}\int_0^1 |f(\alpha)|^{2^{J+1}} \,\mathrm{d}\alpha &= \int_0^1 f(\alpha)^{2^{J-1}} f(-\alpha)^{2^{J-1}} |f(\alpha)|^{2^J} \,\mathrm{d}\alpha \\ &\leqslant (2X)^{2^J-J-1} T, \end{aligned} \tag{4.1}$$

where

$$T = \sum_{|h_1|<X} \cdots \sum_{|h_J|<X} \sum_{x \in I_J(\mathbf{h})} \int_0^1 f(\alpha)^{2^{J-1}} f(-\alpha)^{2^{J-1}} e\left(\alpha \Delta_J(x^k;\mathbf{h})\right) \,\mathrm{d}\alpha.$$

By orthogonality, the expression T is bounded above by the number of integral solutions of the equation

$$\sum_{i=1}^{2^{J-1}} (u_i^k - v_i^k) = \Delta_J(x^k;\mathbf{h}),$$

with $1 \leqslant u_i, v_i \leqslant X$ $(1 \leqslant i \leqslant 2^{J-1})$, $1 \leqslant x \leqslant X$ and $|h_j| < X$ $(1 \leqslant j \leqslant J)$. Notice here that we have weakened the condition $x \in I_J(\mathbf{h})$ to $1 \leqslant x \leqslant X$, which simplifies the exposition but inflates the potential number of solutions to be counted in an inconsequential manner.

The solutions counted by T are of two types, and these we now address in turn. First, there are the solutions in which

$$\sum_{i=1}^{2^{J-1}}(u_i^k-v_i^k)=0.$$

In this situation, one has $\Delta_J(x^k;\mathbf{h})=0$. By orthogonality, the number of choices of $\mathbf{u}$ and $\mathbf{v}$ here is

$$\int_0^1 f(\alpha)^{2^{J-1}}f(-\alpha)^{2^{J-1}}\,\mathrm{d}\alpha=\int_0^1|f(\alpha)|^{2^J}\,\mathrm{d}\alpha\ll X^{2^J-J+\varepsilon},$$

by invoking the inductive hypothesis. Meanwhile, since $\Delta_J(x^k;\mathbf{h})=0$, we have

$$h_1\dots h_J\left(\frac{k!}{(k-J)!}x^{k-J}+\dots\right)=0.$$

So either $h_j=0$ for some $1\leqslant j\leqslant J$, or else x satisfies a polynomial equation determined by $h_1,\dots,h_J$. Then the total number of choices for x and $h_1,\dots,h_J$ is $O(X^J)$. The contribution of the solutions of this first type to T is consequently

$$\ll X^J\cdot X^{2^J-J+\varepsilon}\ll X^{2^J+\varepsilon}.$$

For the second class of solutions counted by T, one has

$$\sum_{i=1}^{2^{J-1}}(u_i^k-v_i^k)=N,$$

for some non-zero integer $N=N(\mathbf{u},\mathbf{v})$ with $|N|\ll X^k$. For each such choice of $\mathbf{u}$ and $\mathbf{v}$, we have

$$h_1\dots h_J\left(\frac{k!}{(k-J)!}x^{k-J}+\dots\right)=N,$$

and thus there are $O(\tau_{J+1}(N))\ll X^\varepsilon$ possible choices for $h_1,\dots,h_J$ and x. The contribution to T from this second class of solutions is therefore

$$\ll\sum_{\substack{1\leqslant u_i,v_i\leqslant X\\ 1\leqslant i\leqslant 2^{J-1}}}\tau_{J+1}(N(\mathbf{u},\mathbf{v}))\ll X^\varepsilon\sum_{\substack{1\leqslant u_i,v_i\leqslant X\\ 1\leqslant i\leqslant 2^{J-1}}}1\ll X^{2^J+\varepsilon}.$$

Combining these two estimates, we see that $T\ll X^{2^J+\varepsilon}$. Substituting this bound into (4.1), we conclude that

$$\int_0^1|f(\alpha)|^{2^{J+1}}\,\mathrm{d}\alpha\ll(2X)^{2^J-J-1}\cdot X^{2^J+\varepsilon}\ll X^{2^{J+1}-(J+1)+\varepsilon}.$$

This confirms the inductive hypothesis with $j=J+1$, and the conclusion of the lemma follows. □

We stress here the number theoretic nature of the proof of this lemma, for we are making use of prime factorisations via estimates for the divisor function.

Corollary 4.2. *When $s\geqslant 2^k+1$, one has*

$$\int_{\mathfrak{m}_\delta}f(\alpha)^se(-n\alpha)\,\mathrm{d}\alpha\ll X^{s-k-\delta 2^{-k}}.$$

Proof. By making use of Weyl's inequality in combination with Hua's lemma, one obtains

$$\begin{aligned}\left|\int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha\right| &\leqslant \left(\sup_{\alpha\in\mathfrak{m}_\delta} |f(\alpha)|\right)^{s-2^k} \int_0^1 |f(\alpha)|^{2^k}\,\mathrm{d}\alpha \\ &\ll (X^{1-\delta 2^{1-k}+\varepsilon})^{s-2^k} X^{2^k-k+\varepsilon} \\ &\ll X^{s-k-(s-2^k)\delta 2^{1-k}+s\varepsilon}.\end{aligned}$$

The conclusion of the corollary follows on recalling that $s \geqslant 2^k+1$. □

We finish this section by noting the sharpest estimates currently available for mean values of the exponential sum

$$f_k(\alpha; X) = \sum_{1\leqslant x\leqslant X} e(\alpha x^k).$$

Thus, one has

$$\int_0^1 |f_2(\alpha;X)|^4\,\mathrm{d}\alpha \asymp X^2 \log X,$$

and, when $k \geqslant 3$, Vaughan [22, 23] has established bounds of the shape

$$\int_{\mathfrak{m}_1} |f_k(\alpha;X)|^{2^k}\,\mathrm{d}\alpha \ll X^{2^k-k}(\log X)^{-A_k},$$

in which A_k is a suitable positive real number depending at most on k. These bounds have been superseded for all exponents k with $k \geqslant 4$ by developments in the orbit of recent work concerning Vinogradov's mean value theorem. Thus, the resolution of the main conjecture in Vinogradov's mean value theorem by Bourgain, Demeter and Guth [3] and Wooley [31, 33] may be applied in combination with earlier work of Wooley [30] to deliver the bound

$$\int_0^1 |f_k(\alpha)|^s\,\mathrm{d}\alpha \ll X^{s-k+\varepsilon},$$

whenever $s \geqslant s_0(k)$, where $s_0(4) = 14$, $s_0(5) < 23$,..., and in general

$$s_0(k) \leqslant k^2 - k + 2\lfloor\sqrt{2k+2}\rfloor - 1.$$

Details may be found in Wooley [33, Corollary 14.7] (a less precise result was stated by Bourgain [2] with a sketch of the ideas involved in its proof).

5. The analysis of the major arcs

We begin by recalling our goal and overall strategy, with the progress achieved to date. Our goal is to asymptotically evaluate the number $R_{s,k}(n)$ of representations of the large natural number n as the sum of s k-th powers of natural numbers, in the shape

$$x_1^k + \ldots + x_s^k = n.$$

This we achieve by applying Fourier analysis. By writing $X = n^{1/k}$ and

$$f(\alpha) = \sum_{1\leqslant x\leqslant X} e(\alpha x^k),$$

we find via orthogonality that

$$R_{s,k}(n) = \int_0^1 f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha.$$

We divide the interval of integration according to a Hardy-Littlewood dissection with *major arcs* $\mathfrak{M}_\delta$ equal to the union of the intervals

$$\mathfrak{M}_\delta(q,a) = \{\alpha \in [0,1) : |\alpha - a/q| \leqslant X^{\delta-k}\},$$

with $0 \leqslant a \leqslant q \leqslant X^\delta$ and $(a,q)=1$, and with *minor arcs* $\mathfrak{m}_\delta = [0,1] \setminus \mathfrak{M}_\delta$. We have not yet made a choice for the parameter δ initially introduced subject to the condition $0 < \delta < 1$. At this point, it makes sense to insist that the major arcs $\mathfrak{M}_\delta(q,a)$ are disjoint. Notice that, if some real number α lies in two distinct major arcs $\mathfrak{M}_\delta(q_1,a_1)$ and $\mathfrak{M}_\delta(q_2,a_2)$ lying in $\mathfrak{M}_\delta$, then by the triangle inequality, one has

$$\left|\frac{a_1}{q_1} - \frac{a_2}{q_2}\right| \leqslant \left|\alpha - \frac{a_1}{q_1}\right| + \left|\alpha - \frac{a_2}{q_2}\right| \leqslant 2X^{\delta-k}.$$

Thus, one finds that

$$\frac{1}{q_1q_2} \leqslant \left|\frac{a_1q_2 - a_2q_1}{q_1q_2}\right| \leqslant 2X^{\delta-k},$$

whence $1 \leqslant 2q_1q_2X^{\delta-k} \leqslant 2X^{3\delta-k}$. This is plainly impossible when $\delta < 1/3$ and X is large, as we henceforth assume. Subject to this latter condition the major arcs $\mathfrak{M}_\delta$ defined in this way are a disjoint union of the arcs $\mathfrak{M}_\delta(q,a)$.

We have shown thus far that when $s \geqslant 2^k+1$, one has

$$\int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha \ll X^{s-k-\delta 2^{-k}} = o(n^{s/k-1}).$$

Our goal is now to establish the lower bound

$$\int_{\mathfrak{M}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha \gg n^{s/k-1}.$$

By combining these major and minor arc contributions, we obtain the asymptotic relation

$$R_{s,k}(n) = \int_{\mathfrak{M}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha + \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha \gg n^{s/k-1} + o(n^{s/k-1}) \to \infty,$$

as $n \to \infty$. Thus, we will be equipped to infer that $G(k) \leqslant 2^k+1$.

Our first step towards this goal is to asymptotically evaluate $f(\alpha)$ in the situation that $\alpha \in \mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$. Here, we apply the mean value theorem. Write $\beta = \alpha - a/q$, so that $|\beta| \leqslant X^{\delta-k}$. By breaking the summand into arithmetic progressions modulo q, one discerns that

$$\begin{aligned}\sum_{1\leqslant x\leqslant X} e(\alpha x^k) &= \sum_{r=1}^{q} \sum_{(1-r)/q\leqslant y\leqslant (X-r)/q} e\left((\beta + a/q)(yq+r)^k\right)\\ &= \sum_{r=1}^{q} e(ar^k/q) \sum_{(1-r)/q\leqslant y\leqslant (X-r)/q} e\left(\beta(yq+r)^k\right). \qquad (5.1)\end{aligned}$$

Since β is small, we can hope to approximate the inner sum here by a smooth function with control of the accompanying error terms. By the mean value theorem, when $F(z)$ is a differentiable function on $[a,b]$ with $a<b$, one sees that

$$F(a)-F(b)=(a-b)F'(\xi),$$

for some $\xi\in(a,b)$. Also, trivially, one has

$$e(F(z))=\int_{-1/2}^{1/2}e(F(z))\,\mathrm{d}\eta.$$

Hence

$$\begin{aligned}\left|e(F(z))-\int_{-1/2}^{1/2}e(F(z+\eta))\,\mathrm{d}\eta\right|&\leqslant\sup_{|\eta|\leqslant 1/2}|e(F(z+\eta))-e(F(z))|\\&\ll\sup_{|\eta|\leqslant 1/2}|F'(z+\eta)|.\end{aligned}$$

Using this approximation, we may infer that

$$\begin{aligned}\sum_{(1-r)/q\leqslant y\leqslant(X-r)/q}e(\beta(yq+r)^k)-&\int_{-r/q}^{(X-r)/q}e(\beta(zq+r)^k)\,\mathrm{d}z\\&\ll 1+(X/q)\sup_{0\leqslant z\leqslant X/q}\left|k\beta q(qz+r)^{k-1}\right|\\&\ll 1+X^k|\beta|.\end{aligned}$$

Notice here that the error incorporated into the right hand side accounts for the initial and final half-intervals within the integral on the left hand side.

By substituting the last relation into (5.1), we deduce that

$$f(\alpha)=\sum_{r=1}^q e(ar^k/q)\left(\int_{-r/q}^{(X-r)/q}e(\beta(zq+r)^k)\,\mathrm{d}z+O(1+X^k|\beta|)\right),$$

so that

$$f(\alpha)-\sum_{r=1}^q e(ar^k/q)\int_{-r/q}^{(X-r)/q}e(\beta(zq+r)^k)\,\mathrm{d}z\ll q+X^k|q\beta|. \tag{5.2}$$

By the change of variable $\gamma=zq+r$, moreover, we have

$$\int_{-r/q}^{(X-r)/q}e(\beta(zq+r)^k)\,\mathrm{d}z=q^{-1}\int_0^X e(\beta\gamma^k)\,\mathrm{d}\gamma. \tag{5.3}$$

We summarise these deliberations in the form of a lemma, but first record some notation. When $a\in\mathbb{Z}$ and $q\in\mathbb{N}$, we write

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

and when $\beta\in\mathbb{R}$, we put

$$v(\beta)=\int_0^X e(\beta\gamma^k)\,\mathrm{d}\gamma.$$

Lemma 5.1. *Suppose that $\alpha\in\mathbb{R}$, $a\in\mathbb{Z}$ and $q\in\mathbb{N}$. Then one has*

$$f(\alpha)-q^{-1}S(q,a)v(\alpha-a/q)\ll q+X^k|q\alpha-a|.$$

Proof. The desired conclusion follows by substituting (5.3) into (5.2). □

Corollary 5.2. *When $\alpha \in \mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$, one has*

$$f(\alpha) - q^{-1}S(q,a)v(\alpha - a/q) \ll X^{2\delta}.$$

Proof. When $\alpha \in \mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$, one has

$$|q\alpha - a| = q|\alpha - a/q| \leqslant X^\delta \cdot X^{\delta-k},$$

whence $q + X^k|q\alpha - a| \ll X^{2\delta}$. The claimed bound now follows from Lemma 5.1. □

Notice that when $\delta < 1/2$, the estimate supplied by this corollary is already non-trivial. If, moreover, we take δ to be small, say $\delta = 1/100$, then this asymptotic relation for $f(\alpha)$ is extremely precise.

Let us now substitute the conclusion of Corollary 5.2 into the formula for the major arc contribution. We see that since

$$\mathfrak{M}_\delta = \bigcup_{\substack{0\leqslant a\leqslant q\leqslant X^\delta\\ (a,q)=1}} \mathfrak{M}_\delta(q,a),$$

then

$$\int_{\mathfrak{M}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha = \sum_{1\leqslant q\leqslant X^\delta} \sum_{\substack{a=1\\ (a,q)=1}}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} f(\beta + a/q)^s e(-n(\beta + a/q))\,\mathrm{d}\beta.$$

Focusing our attention on a real number α lying in $\mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$, we put

$$f^*(\alpha) = q^{-1}S(q,a)v(\alpha - a/q),$$

and write $E(\alpha) = f(\alpha) - f^*(\alpha)$. It follows from Corollary 5.2 that $E(\alpha) \ll X^{2\delta}$. Since

$$\begin{aligned} f(\alpha)^s - f^*(\alpha)^s &= (f(\alpha) - f^*(\alpha))(f(\alpha)^{s-1} + \ldots + f^*(\alpha)^{s-1}) \\ &\ll X^{s-1}|E(\alpha)| \ll X^{s-1+2\delta}, \end{aligned}$$

we obtain the asymptotic relation

$$\begin{aligned} \int_{\mathfrak{M}_\delta} f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha = &\sum_{1\leqslant q\leqslant X^\delta} \sum_{\substack{a=1\\ (a,q)=1}}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} \left(q^{-1}S(q,a)v(\beta)\right)^s e(-n(\beta + a/q))\,\mathrm{d}\beta \\ &+ \sum_{1\leqslant q\leqslant X^\delta} \sum_{\substack{a=1\\ (a,q)=1}}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} X^{s-1+2\delta}\,\mathrm{d}\alpha. \end{aligned} \tag{5.4}$$

The second term on the right hand side of (5.4) is

$$\ll X^{s-1+2\delta} \sum_{1\leqslant q\leqslant X^\delta} q \cdot X^{\delta-k} \ll X^{s-k-1+3\delta} \cdot X^{2\delta} \ll X^{s-k+(5\delta-1)}. \tag{5.5}$$

This is $o(X^{s-k})$ whenever $\delta < 1/5$. Turning to the first term on the right hand side of (5.4), we find that it factorises in the shape

$$\sum_{1\leqslant q\leqslant X^\delta} \sum_{\substack{a=1\\ (a,q)=1}}^{q} \left(q^{-1}S(q,a)\right)^s e(-na/q) \int_{-X^{\delta-k}}^{X^{\delta-k}} v(\beta)^s e(-\beta n)\,\mathrm{d}\beta. \tag{5.6}$$

We thus obtain an asymptotic formula which we presently record in the form of a lemma.

The introduction of additional notation eases our passage at this point. When Q is a positive real number, we define the *truncated singular series*

$$\mathfrak{S}_{s,k}(n;Q)=\sum_{1\leqslant q\leqslant Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\bigl(q^{-1}S(q,a)\bigr)^s e(-na/q)$$

and the *truncated singular integral*

$$J_{s,k}(n;Q)=\int_{-QX^{-k}}^{QX^{-k}}v(\beta)^s e(-\beta n)\,\mathrm{d}\beta.$$

Lemma 5.3. *When $0<\delta<1$, one has*

$$\int_{\mathfrak{M}_\delta}f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha=J_{s,k}(n;X^\delta)\mathfrak{S}_{s,k}(n;X^\delta)+O(X^{s-k+(5\delta-1)}).$$

Proof. The desired conclusion follows on substituting (5.5) and (5.6) into (5.4). □

Corollary 5.4. *When $s\geqslant 2^k+1$ and $0<\delta<1/5$, one has*

$$R_{s,k}(n)=J_{s,k}(n;X^\delta)\mathfrak{S}_{s,k}(n;X^\delta)+o(X^{s-k}),$$

in which $X=n^{1/k}$.

Proof. Since $[0,1)$ is the disjoint union of $\mathfrak{m}_\delta$ and $\mathfrak{M}_\delta$, one has

$$R_{s,k}(n)=\int_{\mathfrak{M}_\delta}f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha+\int_{\mathfrak{m}_\delta}f(\alpha)^s e(-n\alpha)\,\mathrm{d}\alpha,$$

and the conclusion follows from Corollary 4.2 and Lemma 5.3. □

Our objective is now to analyse the truncated singular series and singular integral, with the goal of showing that $\mathfrak{S}_{s,k}(n;X^\delta)\gg 1$ and $J_{s,k}(n;X^\delta)\gg X^{s-k}$, provided at least that s is large enough in terms of k.

6. The singular integral

We first consider the truncated singular integral $J_{s,k}(n;Q)$, our first step being to complete this integral to obtain the (complete) *singular integral*

$$J_{s,k}(n)=\int_{-\infty}^{\infty}v(\beta)^s e(-n\beta)\,\mathrm{d}\beta.$$

Here, we must consider the tails of the domain of integration.

Lemma 6.1. *Whenever $\beta\in\mathbb{R}$, one has*

$$v(\beta)\ll X(1+X^k|\beta|)^{-1/k}.$$

Proof. Recall that

$$v(\beta)=\int_0^X e(\beta\gamma^k)\,\mathrm{d}\gamma.$$

The estimate $|v(\beta)| \leqslant X$ is trivial. Also, since $|v(\beta)| = |v(-\beta)|$, we may assume henceforth that $\beta > X^{-k}$. Next, making the change of variable $u = \beta\gamma^k$, we find that when $\beta > 0$, one has

$$v(\beta) = k^{-1}\beta^{-1/k}\int_0^{\beta X^k} u^{-1+1/k}e(u)\,\mathrm{d}u,$$

whence

$$|v(\beta)| \leqslant k^{-1}\beta^{-1/k}\biggl|\int_0^{\beta X^k} u^{-1+1/k}e(u)\,\mathrm{d}u\biggr|. \tag{6.1}$$

Notice that $u^{-1+1/k}$ decreases monotonically to 0 as $u \to \infty$. Then it follows from Dirichlet's test for convergence of an infinite integral that the integral on the right hand side of (6.1) is uniformly bounded, and indeed

$$\biggl|\int_0^{\beta X^k} u^{-1+1/k}e(u)\,\mathrm{d}u\biggr| \leqslant \sup_{Y\geqslant 0}\biggl|\int_0^{Y} u^{-1+1/k}e(u)\,\mathrm{d}u\biggr| < \infty.$$

Note here that, when $0 < Y < 1$, we are also making use of the inequality

$$\biggl|\int_0^{Y} u^{-1+1/k}e(u)\,\mathrm{d}u\biggr| \leqslant \int_0^{Y} u^{-1+1/k}\,\mathrm{d}u \ll 1.$$

Hence we deduce that when $|\beta| > X^{-k}$, one has

$$v(\beta) \ll |\beta|^{-1/k} \ll X(1+X^k|\beta|)^{-1/k}.$$

The desired conclusion follows on combining this estimate with our earlier bound $|v(\beta)| \leqslant X$, applied in circumstances wherein $|\beta| \leqslant X^{-k}$. □

Corollary 6.2. *Suppose that $s \geqslant k+1$. Then the singular series $J_{s,k}(n)$ converges absolutely, and moreover,*

$$J_{s,k}(n;Q) - J_{s,k}(n) \ll X^{s-k}Q^{-1/k}.$$

Proof. By applying Lemma 6.1, one sees that

$$J_{s,k}(n) \ll \int_{-\infty}^{\infty} \frac{X^s}{(1+X^k|\beta|)^{s/k}}\,\mathrm{d}\beta \ll X^s\int_0^{\infty}\frac{\mathrm{d}\beta}{(1+X^k\beta)^{1+1/k}} \ll X^{s-k}.$$

Thus, the integral defining $J_{s,k}(n)$ is indeed absolutely convergent, and the singular integral exists. Moreover, and similarly,

$$J_{s,k}(n;Q) - J_{s,k}(n) \ll \int_{QX^{-k}}^{\infty}\frac{X^s}{(1+X^k\beta)^{1+1/k}} \ll X^{s-k}Q^{-1/k}.$$

This completes the proof of the lemma. □

The evaluation of the singular integral $J_{s,k}(n)$ is an exercise in classical Fourier analysis.

Lemma 6.3. *When $s \geqslant k+1$, one has*

$$J_{s,k}(n) = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1},$$

in which $\Gamma(z)$ denotes the familiar Γ-function defined by

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\,\mathrm{d}t \quad (\mathrm{Re}(z) > 0).$$

Proof. We begin by observing that

$$\begin{aligned} J_{s,k}(n) &= \lim_{B\to\infty} \int_{-B}^{B} v(\beta)^s e(-\beta n)\,\mathrm{d}\beta \\ &= \lim_{B\to\infty} \int_{-B}^{B} \int_{[0,X]^s} e(\beta(\gamma_1^k + \ldots + \gamma_s^k - n))\,\mathrm{d}\boldsymbol{\gamma}\,\mathrm{d}\beta \\ &= \lim_{B\to\infty} \int_{[0,X]^s} \int_{-B}^{B} e(\beta(\gamma_1^k + \ldots + \gamma_s^k - n))\,\mathrm{d}\beta\,\mathrm{d}\boldsymbol{\gamma}. \end{aligned}$$

We make use of the observation that when $\phi \neq 0$, one has

$$\int_{-B}^{B} e(\beta\phi)\,\mathrm{d}\beta = \frac{\sin(2\pi B\phi)}{\pi\phi}.$$

Adopting the convention that when $\phi = 0$, we are to interpret the right hand side of this formula to be $2B$, we obtain the relation

$$J_{s,k}(n) = \lim_{B\to\infty} \int_{[0,X]^s} \frac{\sin(2\pi B(\gamma_1^k + \ldots + \gamma_s^k - n))}{\pi(\gamma_1^k + \ldots + \gamma_s^k - n)}\,\mathrm{d}\boldsymbol{\gamma}.$$

We aim to transform this into a linear problem. To this end, we substitute $u_i = \gamma_i^k$ $(1 \leqslant i \leqslant s)$, and recall that $n = X^k$. Thus, we obtain

$$J_{s,k}(n) = k^{-s} \lim_{B\to\infty} I(B),$$

where we write

$$I(B) = \int_{[0,n]^s} \frac{\sin(2\pi B(u_1 + \ldots + u_s - n))}{\pi(u_1 + \ldots + u_s - n)} (u_1 \ldots u_s)^{-1+1/k}\,\mathrm{d}\mathbf{u}.$$

A further substitution reduces our task to one of evaluating an integral in just one variable. We put $v = u_1 + \ldots + u_s$ and make the change of variable $(u_1, \ldots, u_s) \mapsto (u_1, \ldots, u_{s-1}, v)$, obtaining the relation

$$I(B) = \int_0^{sn} \Psi(v) \frac{\sin(2\pi B(v - n))}{\pi(v - n)}\,\mathrm{d}v,$$

in which

$$\Psi(v) = \int_{\mathfrak{B}(v)} (u_1 \ldots u_{s-1})^{-1+1/k} (v - u_1 - \ldots - u_{s-1})^{-1+1/k}\,\mathrm{d}u_1 \ldots \mathrm{d}u_{s-1},$$

and

$$\mathfrak{B}(v) = \left\{(u_1, \ldots, u_{s-1}) \in [0,n]^{s-1} : 0 \leqslant v - u_1 - \ldots - u_{s-1} \leqslant n\right\}.$$

Notice that the condition on $u_1, \ldots, u_{s-1}$ in the definition of $\mathfrak{B}(v)$ may be rephrased as $v - n \leqslant u_1 + \ldots + u_{s-1} \leqslant v$. Since $\Psi(v)$ is a function of bounded variation, it follows from Fourier's Integral Theorem that since $n \in (0, sn)$, one has

$$\lim_{B\to\infty} I(B) = \Psi(n) = \int_{\mathfrak{B}(n)} (u_1 \ldots u_{s-1})^{-1+1/k} (n - u_1 - \ldots - u_{s-1})^{-1+1/k}\,\mathrm{d}\mathbf{u}.$$

Note that

$$\mathfrak{B}(n) = \{(u_1, \ldots, u_{s-1}) \in [0,n]^{s-1} : 0 \leqslant u_1 + \ldots + u_{s-1} \leqslant n\}.$$

Thus

$$J_{s,k}(n)=k^{-s}\Psi(n)=k^{-s}\underset{u_1+\ldots+u_{s-1}\leqslant n}{\int_0^n\ldots\int_0^n}(u_1\ldots u_{s-1})^{-1+1/k}(n-u_1-\ldots-u_{s-1})^{-1+1/k}\,\mathrm{d}\mathbf{u}.$$

We now apply induction to show that

$$J_{s,k}(n)=\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1}.$$

First, when $s=2$, we have

$$\begin{aligned}J_{2,k}(n)&=k^{-2}\int_0^n u_1^{-1+1/k}(n-u_1)^{-1+1/k}\,\mathrm{d}u_1\\&=k^{-2}n^{-1+2/k}\int_0^1 v^{-1+1/k}(1-v)^{-1+1/k}\,\mathrm{d}v.\end{aligned}$$

Thus, on recalling the classical Beta function, we obtain the formula

$$J_{2,k}(n)=k^{-2}n^{-1+2/k}\mathrm{B}(1/k,1/k)=k^{-2}n^{-1+2/k}\frac{\Gamma(1/k)^2}{\Gamma(2/k)}=\frac{\Gamma(1+1/k)^2}{\Gamma(2/k)}n^{-1+2/k}.$$

Thus, the inductive hypothesis holds for $s=2$.

Suppose now that the inductive hypothesis holds for $s=t$. Then we have

$$\begin{aligned}J_{t+1,k}(n)&=k^{-1}\int_0^n u_t^{-1+1/k}J_{t,k}(n-u_t)\,\mathrm{d}u_t\\&=k^{-1}\frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}\int_0^n u_t^{-1+1/k}(n-u_t)^{-1+t/k}\,\mathrm{d}u_t.\end{aligned}$$

Recalling once again the classical Beta function, we see that

$$\begin{aligned}J_{t+1,k}(n)&=k^{-1}\frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}n^{-1+(t+1)/k}\mathrm{B}(1/k,t/k)\\&=k^{-1}\frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}n^{-1+(t+1)/k}\frac{\Gamma(1/k)\Gamma(t/k)}{\Gamma((t+1)/k)}\\&=\frac{\Gamma(1+1/k)^{t+1}}{\Gamma((t+1)/k)}n^{-1+(t+1)/k}.\end{aligned}$$

This confirms the inductive hypothesis with t replaced by $t+1$. We have therefore shown that whenever $s\geqslant k+1$, one has

$$J_{s,k}(n)=\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1}.$$

□

Corollary 6.4. *Suppose that $s\geqslant k+1$. Then one has*

$$J_{s,k}(n;Q)=\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1}+O(n^{s/k-1}Q^{-1/k}),$$

as $Q\to\infty$.

Proof. Substitute the conclusion of Lemma 6.3 into Corollary 6.2, and recall that $X=n^{1/k}$. □

7. The singular series, I: convergence of sums and products

We next consider the truncated singular series $\mathfrak{S}_{s,k}(n;Q)$. Our first step is to complete this series to obtain the (complete) *singular series*

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

Again, we must consider the tail of the infinite sum.

Lemma 7.1. *Whenever $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ satisfy $(a,q)=1$, one has $S(q,a)\ll q^{1-2^{1-k}+\varepsilon}$.*

Proof. We apply Weyl's inequality (Lemma 3.7) with $\alpha_k=a/q$ and $X=q$ to obtain

$$\sum_{r=1}^{q}e(ar^k/q)\ll q^{1+\varepsilon}\left(q^{-1}+q^{-1}+q^{1-k}\right)^{2^{1-k}}.$$

□

This was a cheap estimate. Later on we shall obtain the estimate $S(q,a)\ll q^{1-1/k}$, and indeed even sharper estimates are available if one exploits the prime factorisation of q more carefully.

By applying Lemma 7.1 to estimate the tail of the truncated singular series, we see that

$$\sum_{q>Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|(q^{-1}S(q,a))^s e(-na/q)\right|\ll\sum_{q>Q}\phi(q)\left(q^{\varepsilon-2^{1-k}}\right)^s.$$

Thus, when $s\geqslant 2^k+1$, we deduce that

$$\sum_{q>Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|(q^{-1}S(q,a))^s e(-na/q)\right|\ll\sum_{q>Q}q^{\varepsilon-1-2^{1-k}}\ll Q^{-2^{-k}}.$$

It follows that the infinite series $\mathfrak{S}_{s,k}(n)$ converges absolutely under these conditions, and moreover that

$$\mathfrak{S}_{s,k}(n)-\mathfrak{S}_{s,k}(n;Q)\ll Q^{-2^{-k}}.$$

Notice that this estimate is uniform in n. We summarise these conclusions in the form of a lemma.

Lemma 7.2. *Suppose that $s\geqslant 2^k+1$. Then $\mathfrak{S}_{s,k}(n)$ converges absolutely, and*

$$\mathfrak{S}_{s,k}(n)-\mathfrak{S}_{s,k}(n;Q)\ll Q^{-2^{-k}},$$

uniformly in n.

We shall see shortly that there is a close connection between the singular series $\mathfrak{S}_{s,k}(n)$ and the number of solutions of the congruence

$$x_1^k+\ldots+x_s^k\equiv n\ (\mathrm{mod}\ q),$$

as q varies. This suggests a multiplicative theme.

Lemma 7.3. *Suppose that $(a,q)=(b,r)=(q,r)=1$. Then one has the quasi-multiplicative relation*

$$S(qr,ar+bq)=S(q,a)S(r,b).$$

Proof. Each residue m modulo qr with $1\leqslant m\leqslant qr$ is in bijective correspondence with a pair (t,u) with $1\leqslant t\leqslant q$ and $1\leqslant u\leqslant r$, with $m\equiv tr+uq\pmod{qr}$. Indeed, if we write $\overline{q}$ for any integer congruent to the multiplicative inverse of q modulo r, and $\overline{r}$ for any integer congruent to the multiplicative inverse of r modulo q, then we have $m\equiv(m\overline{r})r+(m\overline{q})q\pmod{qr}$, and this claimed bijection becomes transparent. This is, of course, an application of the Chinese Remainder Theorem. Thus, we see that

$$\begin{aligned}S(qr,ar+bq)&=\sum_{m=1}^{qr}e\left(\frac{ar+bq}{qr}m^k\right)\\&=\sum_{t=1}^{q}\sum_{u=1}^{r}e\left(\frac{(ar+bq)(tr+uq)^k}{qr}\right)\\&=\sum_{t=1}^{q}\sum_{u=1}^{r}e\left(\frac{a}{q}(tr)^k+\frac{b}{r}(uq)^k\right).\end{aligned}$$

By the change of variable $tr\mapsto t'\pmod{q}$ and $uq\mapsto u'\pmod{r}$, bijective owing to the coprimality of q and r, we obtain the relation

$$S(qr,ar+bq)=\left(\sum_{v=1}^{q}e(av^k/q)\right)\left(\sum_{w=1}^{r}e(bw^k/r)\right)=S(q,a)S(r,b).$$

This completes the proof of the lemma. □

Now define the quantity

$$A(q,n)=\sum_{\substack{a=1\\(a,q)=1}}^{q}\left(q^{-1}S(q,a)\right)^s e(-na/q).$$

Lemma 7.4. *The quantity $A(q,n)$ is a multiplicative function of q.*

Proof. Suppose that $(q,r)=1$. Then by the Chinese Remainder Theorem, there is a bijection between the residue classes a modulo qr with $(a,qr)=1$, and the ordered pairs (b,c) with b modulo q and c modulo r satisfying $(b,q)=(c,r)=1$, via the relation $a\equiv br+cq\pmod{qr}$. Thus, we obtain

$$\begin{aligned}A(qr,n)&=\sum_{\substack{a=1\\(a,qr)=1}}^{qr}\left((qr)^{-1}S(qr,a)\right)^s e(-na/qr)\\&=\sum_{\substack{b=1\\(b,q)=1}}^{q}\sum_{\substack{c=1\\(c,r)=1}}^{r}\left((qr)^{-1}S(qr,br+cq)\right)^s e\left(-\frac{br+cq}{qr}n\right).\end{aligned}$$

By applying Lemma 7.3, we infer that

$$\begin{aligned} A(qr,n) &= \sum_{\substack{b=1\\(b,q)=1}}^{q} \sum_{\substack{c=1\\(c,r)=1}}^{r} \left(q^{-1}S(q,b)\right)^s \left(r^{-1}S(r,c)\right)^s e(-bn/q)e(-cn/r) \\ &= A(q,n)A(r,n). \end{aligned}$$

Since $A(1,n)=1$, this confirms the multiplicative property for $A(q,n)$ and completes the proof of the lemma. □

Observe that

$$\mathfrak{S}_{s,k}(n) = \sum_{q=1}^{\infty} A(q,n).$$

The multiplicativity of $A(q,n)$ therefore suggests that $\mathfrak{S}_{s,k}(n)$ should factor as a product over prime numbers p of the p-adic densities

$$\sigma(p) = \sum_{h=0}^{\infty} A(p^h,n).$$

Theorem 7.5. *Suppose that $s \geqslant 2^k+1$. Then the following hold:*

(i) *the series $\sigma(p)$ converges absolutely, and one has*

$$|\sigma(p)-1| \ll p^{-1-2^{-k}};$$

(ii) *the infinite product*

$$\prod_{p\ \mathit{prime}} \sigma(p)$$

converges absolutely;

(iii) *one has $\mathfrak{S}_{s,k}(n) = \prod_p \sigma(p)$;*

(iv) *there exists a natural number $C=C(k)$ with the property that*

$$1/2 < \prod_{p \geqslant C(k)} \sigma(p) < 3/2.$$

Proof. We begin by establishing (i). We recall from Lemma 7.1 that whenever $(a,p)=1$, one has

$$S(p^h,a) \ll (p^h)^{1-2^{1-k}+\varepsilon}.$$

Then, whenever $s \geqslant 2^k+1$, one finds that

$$A(p^h,n) = \sum_{\substack{a=1\\(a,p)=1}}^{p^h} \left(p^{-h}S(p^h,a)\right)^s e(-na/p^h) \ll p^{h(1-s2^{1-k})+\varepsilon} \ll p^{-h(1+2^{-k})}.$$

Hence

$$\sigma(p)-1 = \sum_{h=1}^{\infty} A(p^h,n) \ll \sum_{h=1}^{\infty} p^{-h(1+2^{-k})} \ll p^{-1-2^{-k}}.$$

Thus $\sigma(p)$ converges absolutely, and one has $|\sigma(p)-1| \ll p^{-1-2^{-k}}$.

We next turn to the proof of (ii). We begin by noting that from the conclusion of part (i), there is a positive number $B = B(k)$ with the property that $|\sigma(p) - 1| \leqslant Bp^{-1-2^{-k}}$. Hence, whenever p is sufficiently large in terms of p, one sees that

$$\log(1 + |\sigma(p) - 1|) \leqslant \log(1 + Bp^{-1-2^{-k}}) \leqslant Bp^{-1-2^{-k-1}},$$

whence

$$\sum_{p \text{ prime}} \log(1 + |\sigma(p) - 1|) \ll B \sum_p p^{-1-2^{-k-1}} \ll 1.$$

Thus we deduce that the infinite product $\prod_p \sigma(p)$ converges absolutely.

The proof of (iii) employs the multiplicative property of $A(q, n)$ established in Lemma 7.4. One finds that

$$\mathfrak{S}_{s,k}(n) = \sum_{q=1}^{\infty} A(q, n) = \sum_{q=1}^{\infty} \prod_{p^h \| q} A(p^h, n).$$

Then since $\prod_p \sigma(p)$ converges absolutely as a product, and $\sum_{q=1}^{\infty} A(q, n)$ converges absolutely as a sum, we may rearrange summands to deduce that

$$\mathfrak{S}_{s,k}(n) = \prod_p \sum_{h=0}^{\infty} A(p^h, n) = \prod_p \sigma(p).$$

Finally, we establish (iv). We begin by observing that from part (i), it follows that whenever p is sufficiently large in terms of k, one has

$$1 - p^{-1-2^{-k}} \leqslant \sigma(p) \leqslant 1 + p^{-1-2^{-k}}.$$

Hence, provided that $C = C(k)$ is sufficiently large, one finds that

$$\left| \prod_{p \geqslant C(k)} \sigma(p) - 1 \right| \leqslant \sum_{n \geqslant C(k)} n^{-1-2^{-k}} \ll C(k)^{-2^{-k}}.$$

Then, again if $C(k)$ is chosen sufficiently large in terms of k, we may infer that

$$\left| \prod_{p \geqslant C(k)} \sigma(p) - 1 \right| < 1/2,$$

and we conclude that

$$1/2 < \prod_{p \geqslant C(k)} \sigma(p) < 3/2.$$

The final conclusion of the theorem therefore follows, and the proof of the theorem is complete. □

8. Local solubility and p-adic densities

We would like to show that $\mathfrak{S}_{s,k}(n) \gg 1$ (uniformly in n). At this stage, we can at least show that the latter is the case provided that $\sigma(p) > 0$ for $p \leqslant C(k)$ with sufficient uniformity in n. We establish this conclusion by relating $\sigma(p)$ to the density of solutions of a congruence associated to the original problem of Waring-type.

When $q \in \mathbb{N}$, we put

$$M_n(q) = \text{card}\{\mathbf{m} \in (\mathbb{Z}/q\mathbb{Z})^s : m_1^k + \ldots + m_s^k = n\}.$$

Lemma 8.1. *For each natural number q, one has*

$$\sum_{d|q} A(d,n) = q^{1-s} M_n(q).$$

Proof. We make use of the orthogonality relation

$$q^{-1}\sum_{r=1}^{q} e(hr/q) = \begin{cases} 1, & \text{when } q|h, \\ 0, & \text{when } q \nmid h. \end{cases}$$

Then

$$M_n(q) = q^{-1}\sum_{r=1}^{q}\left(\sum_{m_1=1}^{q}\cdots\sum_{m_s=1}^{q} e\left(r(m_1^k+\ldots+m_s^k-n)/q\right)\right).$$

Classifying the values of r according to their common factors q/d with q, we obtain the relation

$$\begin{aligned} M_n(q) &= q^{-1}\sum_{d|q}\sum_{\substack{a=1\\(a,d)=1}}^{d}(q/d)^s\sum_{m_1=1}^{d}\cdots\sum_{m_s=1}^{d} e\left(a(m_1^k+\ldots+m_s^k-n)/d\right)\\ &= q^{-1}\sum_{d|q} q^s \sum_{\substack{a=1\\(a,d)=1}}^{d}\left(d^{-1}S(d,a)\right)^s e(-na/d)\\ &= q^{s-1}\sum_{d|q} A(d,n). \end{aligned}$$

Hence

$$\sum_{d|q} A(d,n) = q^{1-s} M_n(q),$$

and the proof of the lemma is complete. □

Corollary 8.2. *For each prime number p, one has*

$$\sigma(p) = \lim_{h\to\infty} p^{h(1-s)} M_n(p^h).$$

Proof. Take $q = p^h$ in Lemma 8.1 to obtain the relation

$$\sum_{l=0}^{h} A(p^l,n) = (p^h)^{1-s} M_n(p^h).$$

On taking the limit as $h \to \infty$, the left hand side converges (absolutely) to $\sigma(p)$. □

We now seek to show that for the small primes p with $p < C(k)$, and for all large enough values of h, one has $M_n(p^h) \gg p^{h(s-1)}$. From this we will deduce that $\sigma(p) > 0$, and the desired conclusion $\mathfrak{S}_{s,k}(n) \gg 1$ follows from Theorem 7.5(iv).

We begin by recalling some elementary number theory. We define $\tau = \tau(p,k)$ via the relation $p^\tau \| k$, and then define $\gamma = \gamma(p,k)$ by means of the relation

$$\gamma = \begin{cases} \tau+1, & \text{when } p > 2, \text{ or when } p = 2 \text{ and } \tau = 0, \\ \tau+2, & \text{when } p = 2 \text{ and } \tau > 0. \end{cases}$$

The following lemma describes the structure of k-th powers in $\mathbb{Z}/p^h\mathbb{Z}$ for each prime number p. In a sense, this result is a surrogate for Hensel's lemma.

Lemma 8.3. *Suppose that p is a prime number, and that $(a,p)=1$. Suppose also that the congruence $x^k \equiv a \pmod{p^\gamma}$ is soluble. Then whenever $h \geqslant \gamma$, the congruence $x^k \equiv a \pmod{p^h}$ also possesses a solution.*

Proof. Suppose that $x^k \equiv a \pmod{p^\gamma}$ is soluble and $(a,p)=1$. We suppose first that p is an odd prime. Then there is a primitive root g modulo p^2. It follows that g is a primitive root modulo p^r for all natural numbers r. Let h be an integer with $h \geqslant \gamma$, and let u be the integer with $1 \leqslant u \leqslant \phi(p^h)$ for which $g^u \equiv a \pmod{p^h}$. Then $g^u \equiv a \pmod{p^\gamma}$, and so we deduce from the relation $x^k \equiv a \pmod{p^\gamma}$ that

$$(g^u)^{\phi(p^\gamma)/(k,\phi(p^\gamma))} \equiv a^{\phi(p^\gamma)/(k,\phi(p^\gamma))} \equiv (x^{\phi(p^\gamma)})^{k/(k,\phi(p^\gamma))} \equiv 1 \pmod{p^\gamma}.$$

Hence, since g has order $\phi(p^\gamma)$ modulo p^γ, it follows that $(k,\phi(p^\gamma))$ divides u. Notice that the definition of τ and γ ensures that

$$(k,\phi(p^\gamma)) = (k, p^\tau(p-1)) = (k,\phi(p^h)).$$

Thus $(k,\phi(p^h))$ divides u. We put $l = k/(k,\phi(p^h))$. Then $(l,\phi(p^h))=1$, so there exists an integer m with $lm \equiv 1 \pmod{\phi(p^h)}$. Putting $r = u/(k,\phi(p^h))$, we deduce that

$$a \equiv g^u = (g^r)^{(k,\phi(p^h))} \equiv (g^{rm})^{l(k,\phi(p^h))} \equiv (g^{rm})^k \pmod{p^h}.$$

We therefore conclude that the congruence $y^k \equiv a \pmod{p^h}$ has the solution $y = g^{rm}$.

When $p=2$ and $\tau = 0$, one has $(k,\phi(2^h)) = (k,2^{h-1}) = 1$ for every exponent h, and in such circumstances the solubility of the congruence $x^k \equiv a \pmod{2^h}$ is equivalent to the solubility of a linear congruence. Indeed, there exists an integer r with $rk \equiv 1 \pmod{\phi(2^h)}$, whence

$$x \equiv (x^k)^r \equiv a^r \pmod{2^h}.$$

Finally, suppose that $p=2$ and $\tau > 0$. In this situation, there are integers u and v with $v \in \{0,1\}$ and $1 \leqslant u \leqslant 2^{h-2}$ for which

$$a \equiv (-1)^v 5^u \pmod{2^h}.$$

We note that 5 has order 2^{r-2} modulo 2^r for $r \geqslant 2$. Then we may proceed as in the case that p is odd. We suppose that $a \equiv x^k \pmod{2^\gamma}$. Hence $(-1)^v 5^u \equiv x^k \pmod{2^\gamma}$. We may find integers w and z with $x \equiv (-1)^w 5^z \pmod{2^\gamma}$. Since $2^\tau | k$, it follows that $x^k \equiv (5^z)^k \equiv 1 \pmod{2^\gamma}$, and thus $(-1)^v 5^u \equiv 1 \pmod{2^\gamma}$. We deduce that $v=0$ and $2^{\gamma-2}|u$, say $u = 2^\tau r$ for some integer r. Writing $k = l2^\tau$, we have $2 \nmid l$, and hence there exists an integer m with $lm \equiv 1 \pmod{\phi(2^h)}$. We may thus conclude that

$$a \equiv (-1)^v 5^u = (5^r)^{2^\tau} \equiv (5^{rm})^{l2^\tau} = (5^{rm})^k \pmod{2^h}.$$

Thus, the congruence $y^k \equiv a \pmod{2^h}$ has the solution $y = 5^{rm}$. This completes the proof of the lemma in the final case. □

Now denote by $M_n^*(q)$ the number of solutions of the congruence

$$x_1^k + \ldots + x_s^k \equiv n \pmod{q},$$

with $(x_1,q)=1$ and $1 \leqslant x_i \leqslant q$ $(1 \leqslant i \leqslant s)$.

Lemma 8.4. *Suppose that $M_n^*(p^\gamma) \geqslant 1$. Then whenever $h \geqslant \gamma$, one has*

$$M_n(p^h) \geqslant (p^{h-\gamma})^{s-1}.$$

Proof. Suppose, as we may, that $y_1, \ldots, y_s$ are integers with $1 \leqslant y_i \leqslant p^\gamma$ $(1 \leqslant i \leqslant s)$ and $(y_1, p) = 1$ such that

$$y_1^k + \ldots + y_s^k \equiv n \pmod{p^\gamma}.$$

Let x_i be any integer with $1 \leqslant x_i \leqslant p^h$ and $x_i \equiv y_i \pmod{p^\gamma}$, for $2 \leqslant i \leqslant s$. There are $(p^{h-\gamma})^{s-1}$ possible such choices for $x_2, \ldots, x_s$. We have

$$n - x_2^k - \ldots - x_s^k \equiv n - y_2^k - \ldots - y_s^k \equiv y_1^k \pmod{p^\gamma}.$$

The left hand side of this congruence is a k-th power modulo p^γ, and hence also a k-th power modulo p^h (as a consequence of Lemma 8.3). Thus, there is an integer x_1 with $1 \leqslant x_1 \leqslant p^h$ such that

$$n - x_2^k - \ldots - x_s^k \equiv x_1^k \pmod{p^h}.$$

In particular, we may conclude that $M_n(p^h) \geqslant (p^{h-\gamma})^{s-1}$. □

This lemma reduces the problem of showing that $\mathfrak{S}_{s,k}(n) \gg 1$ to a problem concerning the solubility of congruences modulo p^γ, for finitely many prime numbers p. In order to address this task, we begin with an auxiliary lemma of combinatorial flavor. This is a lemma that is periodically "rediscovered", and is often attributed to Cauchy (1813) and Davenport (1935) (see [6] and [7]). In this context, when $\mathcal{A}$ and $\mathcal{B}$ are sets of integers, we write

$$\mathcal{A} + \mathcal{B} = \{a + b \pmod{q} : a \in \mathcal{A},\, b \in \mathcal{B}\}.$$

Lemma 8.5. *Suppose that $\mathcal{A}$ is a set of r residue classes modulo q, and that $\mathcal{B}$ is a set of s such classes. Suppose further that $0 \in \mathcal{B}$, and that whenever $b \in \mathcal{B}$ and $q \nmid b$, one has $(b, q) = 1$. Then*

$$\operatorname{card}(\mathcal{A} + \mathcal{B}) \geqslant \min\{q, r + s - 1\}.$$

Proof. When $r + s - 1 > q$, then we may delete $s - (q - r + 1)$ elements from $\mathcal{B} \setminus \{0\}$ to obtain a new set $\mathcal{B}_1^*$. The set $\mathcal{B}_1 = \mathcal{B}_1^* \cup \{0\}$ is a subset of $\mathcal{B}$ having $s_1 = q - r + 1$ elements, but now with $r + s_1 - 1 \leqslant q$. So the general case will follow from that subject to the hypothesis $r + s - 1 \leqslant q$. We may assume, moreover, that $r < q$, for if $r = q$ we have $\mathcal{A} + \mathcal{B} = \{0, 1, \ldots, q - 1\}$, and we are done.

We proceed by induction on s, starting with the trivial case $s = 1$ in which $\mathcal{A} + \mathcal{B} = \mathcal{A} + \{0\} = \mathcal{A}$. Since in this case we have $\operatorname{card}(\mathcal{A} + \mathcal{B}) = r = r + s - 1$, the inductive hypothesis holds when $s = 1$. We may therefore suppose that $s > 1$, and that the desired conclusion holds whenever $\operatorname{card}(\mathcal{B}) < s$. Suppose in the first instance that for every $a \in \mathcal{A}$ and $b \in \mathcal{B}$, one has $a + b \in \mathcal{A}$. Then for all $b \in \mathcal{B}$ one has

$$\sum_{a \in \mathcal{A}} (a + b) \equiv \sum_{a \in \mathcal{A}} a \pmod{q},$$

whence $rb \equiv 0 \pmod{q}$. But $r < q$, and so $(b, q) > 1$ for all $b \in \mathcal{B}$. This contradicts the hypotheses of the statement of the lemma, and thus we conclude that there exists $c \in \mathcal{A}$ and $b \in \mathcal{B}$ for which $c + b \notin \mathcal{A}$.

We define

$$\mathcal{C} = \{b \in \mathcal{B} \,:\, c + b \notin \mathcal{A}\}, \quad \mathcal{A}_1 = \mathcal{A} \cup \{\{c\} + \mathcal{C}\} \quad \text{and} \quad \mathcal{B}_1 = \mathcal{B} \setminus \mathcal{C}.$$

Notice here that $c+0\in\mathcal{A}$, and so $\mathcal{C}\subsetneq\mathcal{B}$. Then $1\leqslant \mathrm{card}(\mathcal{B}_1)<s$, and

$$\mathrm{card}(\mathcal{A}_1)+\mathrm{card}(\mathcal{B}_1)=(\mathrm{card}(\mathcal{A})+\mathrm{card}(\mathcal{C}))+(\mathrm{card}(\mathcal{B})-\mathrm{card}(\mathcal{C}))=r+s.$$

By the inductive hypothesis, one has

$$\mathrm{card}(\mathcal{A}_1+\mathcal{B}_1)\geqslant \min\{q,r+s-1\}.$$

Also,

$$\begin{aligned}\mathcal{A}_1+\mathcal{B}_1&=(\mathcal{A}+\mathcal{B}_1)\cup((\{c\}+\mathcal{C})+\mathcal{B}_1)\\&=(\mathcal{A}+\mathcal{B}_1)\cup((\{c\}+\mathcal{B}_1)+\mathcal{C}).\end{aligned}$$

But $\{c\}+\mathcal{B}_1\subseteq\mathcal{A}$ and $\mathcal{C}\subset\mathcal{B}$, and thus $\mathcal{A}_1+\mathcal{B}_1\subseteq\mathcal{A}+\mathcal{B}$. Therefore, we have

$$\mathrm{card}(\mathcal{A}+\mathcal{B})\geqslant\mathrm{card}(\mathcal{A}_1+\mathcal{B}_1)\geqslant\min\{q,r+s-1\}.$$

This establishes the inductive hypothesis when $\mathrm{card}(\mathcal{B})=s$, and this completes the proof of the lemma. □

This result may be applied to sets of k-th powers modulo q, when q is a prime power.

Lemma 8.6. *Suppose that $s\geqslant s_0(k)$, where*

$$s_0(k)=\begin{cases}2^{\tau+2}, & \text{when } \gamma=\tau+2 \text{ and } k>2,\\ 5, & \text{when } p=k=2,\\ \dfrac{p}{p-1}(k,p^\tau(p-1)), & \text{when } \gamma=\tau+1.\end{cases}$$

Then $M_n^*(p^\gamma)\geqslant 1$.

Proof. Suppose first that $\gamma=\tau+1$. We apply Lemma 8.5 with

$$\mathcal{A}=\{x^k \ (\mathrm{mod}\ p^\gamma): 1\leqslant x\leqslant p^\gamma \text{ and } (x,p)=1\}\quad\text{and}\quad \mathcal{B}=\mathcal{A}\cup\{0\}.$$

Thus

$$\mathrm{card}(\mathcal{A})=\frac{\phi(p^\gamma)}{(k,\phi(p^\gamma))}\quad\text{and}\quad \mathrm{card}(\mathcal{B})=1+\frac{\phi(p^\gamma)}{(k,\phi(p^\gamma))}.$$

Write

$$t=\left\lceil\frac{p}{p-1}\,(k,p^\tau(p-1))\right\rceil.$$

Then we find by induction on t that

$$\mathrm{card}\left(\mathcal{A}+(t-1)\mathcal{B}\right)\geqslant\min\left\{p^\gamma,\frac{t\phi(p^\gamma)}{(k,\phi(p^\gamma))}\right\}=p^\gamma.$$

Thus $\mathcal{A}+(t-1)\mathcal{B}$ contains all residues modulo p^γ, and we are forced to conclude that the congruence

$$x_1^k+\cdots+x_t^k\equiv n \ (\mathrm{mod}\ p^\gamma)$$

possesses a solution with $(x_1,p)=1$, as desired.

Suppose next that $p=2$ and $\gamma=\tau+2$. Then we put $t=2^{\tau+2}=2^\gamma$, and solve the congruence

$$x_1^k+\ldots+x_t^k\equiv n \ (\mathrm{mod}\ 2^\gamma)$$

by putting $x_i=1$ for $1\leqslant i\leqslant r$ and $x_i=0$ for $r<i\leqslant t$, where r is the integer with $1\leqslant r\leqslant 2^\gamma$ for which one has $n\equiv r$ (mod 2^γ).

Finally, when $p = 2$ and $k = 2$, one can solve the congruence

$$x_1^2 + \cdots + x_5^2 \equiv n \pmod 8$$

with $x_1 = 1$ and $x_2, \ldots, x_5 \in \{0, 1, 2\}$, as one can verify on the back of an envelope.

In all of these cases, we see that $M_n^*(p^\gamma) \geqslant 1$, and thus the proof of the lemma is complete. □

We are now equipped to wrap up the discussion of the singular series. Lemma 8.6 shows that $M_n^*(p^\gamma) \geqslant 1$ provided only that when $k \neq 2$ one has $s \geqslant 4k$, and that when $k = 2$ one has $s \geqslant 5$. Indeed, when k is not a power of 2, then the lower bound $s \geqslant \frac{3}{2}k$ suffices for this conclusion. We therefore deduce from Lemma 8.4 that when $s \geqslant 2^k + 1$, one has

$$M_n(p^h) \geqslant p^{(h-\gamma)(s-1)}.$$

Corollary 8.2 thence confirms that

$$\sigma(p) = \lim_{h \to \infty} p^{h(1-s)} M_n(p^h) \geqslant p^{-\gamma(s-1)} > 0,$$

and thereby we conclude via Theorem 7.5 that for a suitable positive number $C = C(k)$,

$$\mathfrak{S}_{s,k}(n) = \Big(\prod_{p<C} \sigma(p) \Big) \Big(\prod_{p \geqslant C} \sigma(p) \Big) > \tfrac{1}{2} \prod_{p<C} \sigma(p) > 0.$$

We summarise this discussion in the form of a corollary.

Corollary 8.7. *When $s \geqslant 2^k + 1$, one has $1 \ll_{s,k} \mathfrak{S}_{s,k}(n) \ll_{s,k} 1$.*

Proof. One has only to combine the conclusion of the above discussion with that of Lemma 7.2. □

9. The asymptotic formula in Waring's problem

We have shown in Corollary 5.4 that when $s \geqslant 2^k + 1$ and $0 < \delta < 1/5$, one has

$$R_{s,k}(n) = J_{s,k}(n; X^\delta) \mathfrak{S}_{s,k}(n; X^\delta) + o(X^{s-k}).$$

Also, Corollary 6.4 shows that

$$J_{s,k}(n; X^\delta) = \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} n^{s/k-1} + O(n^{s/k-1-\delta/k^2})$$

and Lemma 7.2 and Theorem 7.5 demonstrate that

$$\mathfrak{S}_{s,k}(n; X^\delta) = \mathfrak{S}_{s,k}(n) + O(n^{-\delta 2^{-k}/k}),$$

where $1 \ll \mathfrak{S}_{s,k}(n) \ll 1$. Thus we conclude that when $s \geqslant 2^k + 1$, one has

$$R_{s,k}(n) = \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} n^{s/k-1} \mathfrak{S}_{s,k}(n) + o(n^{s/k-1}) \gg n^{s/k-1}.$$

Then $R_{s,k}(n) \to \infty$ as $n \to \infty$, whence $G(k) \leqslant 2^k + 1$.

One may speculate concerning the extent to which this conclusion might be improved in terms of the number of variables required to ensure its validity. This prompts the following conjecture.

Conjecture 9.1. *When $k \geqslant 3$, one has*

$$\int_0^1 |f(\alpha)|^{2s}\,\mathrm{d}\alpha \ll X^s + X^{2s-k}.$$

The validity of this conjecture would imply the anticipated asymptotic formula for $s \geqslant 2k+1$, as a consequence of which one would have

$$G(k) = 4k \quad \text{for} \quad k = 2^r,$$

and

$$G(k) \leqslant 2k+1 \quad \text{when} \quad k \neq 2^r.$$

It is expected that one should have $G(k) = k+1$ unless there are congruence obstructions to the solubility of the equation $x_1^k + \ldots + x_s^k = n$.

It is useful to introduce some additional notation at this point. When $k \in \mathbb{N}$, we denote by $\widetilde{G}(k)$ the least integer t with the property that whenever $s \geqslant t$, then one has the anticipated asymptotic formula

$$R_{s,k}(n) = \mathfrak{S}_{s,k}(n)\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1} + o(n^{s/k-1}).$$

We have shown that $\widetilde{G}(k) \leqslant 2^k+1$. In the current state of knowledge, one has the bound

$$\widetilde{G}(k) \leqslant 2^k \quad (k \geqslant 3),$$

as a consequence of work of Vaughan from 1986 (see [22, 23]). For larger values of k, this work is inferior to earlier bounds originating with the work of Vinogradov in the mid-1930s. This work shows that $\widetilde{G}(k) \leqslant (C+o(1))k^2\log k$ for an appropriate value of C, with permissible choices for C having been progressively reduced over the years. Most recently, with the proof of the main conjecture in Vinogradov's mean value theorem (see [3, 32, 33], one improves on this earlier work of Vaughan for all $k \geqslant 4$. Thus, via [30], one obtains

$$\widetilde{G}(4) \leqslant 15, \quad \widetilde{G}(5) \leqslant 23, \ldots, \quad \widetilde{G}(k) \leqslant k^2 - k + 2\lfloor\sqrt{2k+2}\rfloor - 1.$$

See [2] for a sketch of how to obtain a slightly weaker bound than that provided in which the reader must avoid a misleading hint and solve a non-trivial problem, and [33, section 14] for full details.

If one is not concerned with the asymptotic formula for $R_{s,k}(n)$, but merely with the existence statement $R_{s,k}(n) \geqslant 1$ for large n, then sharper results are available. Indeed, we have reported on these bounds for $G(k)$ in §1.

10. The Hasse Principle for diagonal forms, I: preliminaries

When $k \in \mathbb{N}$, $s \geqslant 2^k+1$ and $c_1, \ldots, c_s$ are fixed non-zero integers, we consider the diagonal form

$$\phi(\mathbf{x}) = c_1x_1^k + \ldots + c_sx_s^k,$$

and the solubility of the equation $\phi(\mathbf{x}) = 0$. Here, since all such equations possess the trivial solution $\mathbf{x} = \mathbf{0}$, we shall be interested in solutions with $\mathbf{x} \in \mathbb{Z}^s \setminus \{\mathbf{0}\}$. Our goal, for the purposes of this class, is to show that the methods already developed apply to resolve questions of this type concerning this Diophantine equation $\phi(\mathbf{x}) = 0$.

We will suppose that the equation $\phi(\mathbf{x}) = 0$ has a projective solution everywhere locally. Thus:

(i) there exists $\mathbf{z}^{(\infty)} \in \mathbb{R}^s \setminus \{\mathbf{0}\}$ with the property that $\phi(\mathbf{z}^{(\infty)}) = 0$;
(ii) for each prime p, there exists $\mathbf{z}^{(p)} \in \mathbb{Q}_p^s \setminus \{\mathbf{0}\}$ with the property that $\phi(\mathbf{z}^{(p)}) = 0$.

We note that condition (ii) here is equivalent to the statement that, for each prime number p, and all natural numbers h, there exists some $\mathbf{y} \in \mathbb{Z}^s$ with

$$\phi(\mathbf{y}) \equiv 0 \pmod{p^h},$$

subject to the additional condition that $(y_j, p) = 1$ for some index j with $1 \leqslant j \leqslant s$.

Definition 10.1. We say that *the Hasse Principle* holds for the equation $\phi(\mathbf{x}) = 0$ if, whenever $\phi(\mathbf{x}) = 0$ has a projective solution everywhere locally, then the equation $\phi(\mathbf{x}) = 0$ has a solution $\mathbf{x} \in \mathbb{Z}^s \setminus \{\mathbf{0}\}$.

Notice that solubility over $\mathbb{R}$ here is easy to characterise. This is assured when $s \geqslant 2$ provided either that k is odd, or else k is even and not all of the coefficients c_i have the same sign. Solubility over $\mathbb{Q}_p$ is more subtle. We will make use of a theorem of Davenport and Lewis from 1963 showing that solubility over $\mathbb{Q}_p$ is assured for every prime p provided that $s \geqslant k^2 + 1$ (see [9]). As an alternative to this result, one could apply the Cauchy-Davenport Lemma to establish a similar conclusion, though subject to the condition $s \geqslant 4k^2$.

There is continuing and active interest in more general results. Consider the homogeneous polynomial

$$F(x_1, \ldots, x_s) = \sum_{\substack{i_1,\ldots,i_s \geqslant 0 \\ i_1+\ldots+i_s=k}} c_{i_1,\ldots,i_s} x_1^{i_1} \ldots x_s^{i_s},$$

with $c_{\mathbf{i}} \in \mathbb{Z}$ fixed, and not all 0. Then it is known that when $k \geqslant 2$, the equation $F(\mathbf{x}) = 0$ has a non-trivial solution over $\mathbb{Q}_p$ whenever $s \geqslant k^{2^k}$ (see [29]). For $k = 2, 3$ one has sharper conclusions. Thus a non-trivial solution exists when $k = 2$ for $s \geqslant 5$ (a classical result of Hasse [13]), and when $k = 3$ such holds for $s \geqslant 10$ (see Lewis [17] and Demyanov [10]). There exist infinitely many exponents k and forms F in s variables failing to possess non-trivial solutions, meanwhile, with s as large as

$$\exp\left(\frac{k}{(\log k)(\log\log k)\ldots(\log_r k)(\log_{r+1} k)^{1+\varepsilon}}\right),$$

in which ε is any positive number, and r is any fixed natural number. A conclusion of this type was obtained more or less simultaneously by Arkhipov and Karatsuba [1], Lewis and Montgomery [18] and Brownawell [4].

Let us return now to the topic of diagonal forms. Before embarking on our application of the Hardy-Littlewood (circle) method, we make some preliminary simplifications. Fix $s \geqslant 2^k + 1$ and $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$. Let $\boldsymbol{\zeta} \in \mathbb{R}^s \setminus \{\mathbf{0}\}$ satisfy the equation $\phi(\boldsymbol{\zeta}) = 0$. By the homogeneity of ϕ, we may rescale so as to assume without loss of generality that $|\zeta_i| < 1/2$ for $1 \leqslant i \leqslant s$. We suppose also that for each prime number p and each $h \in \mathbb{N}$, there is a solution $\mathbf{z} = \mathbf{z}^{(p,h)} \in \mathbb{Z}^s$ of the congruence

$$c_1 z_1^k + \ldots + c_s z_s^k \equiv 0 \pmod{p^h},$$

in which $p \nmid z_i$ for some $1 \leqslant i \leqslant s$.

Our goal is to evaluate the number $N_\phi(X)$ of integral solutions $\mathbf{x} \in [-X,X]^s \cap \mathbb{Z}^s$ of the equation $\phi(\mathbf{x}) = 0$, in which $\phi(\mathbf{x}) = c_1x_1^k + \ldots + c_sx_s^k$. Write

$$f(\alpha) = \sum_{|x| \leqslant X} e(\alpha x^k).$$

Then by orthogonality one has

$$N_\phi(X) = \int_0^1 f(c_1\alpha)\ldots f(c_s\alpha)\,\mathrm{d}\alpha.$$

We will apply the Hardy-Littlewood method to show that $N_\phi(X) \to \infty$ as $X \to \infty$, whence there exists $\mathbf{x} \in \mathbb{Z}^s \setminus \{\mathbf{0}\}$ with $\phi(\mathbf{x}) = 0$. The Hasse Principle follows for the equation $\phi(\mathbf{x}) = 0$.

We require a Hardy-Littlewood dissection, and this entails a slight modification of that which we have previously employed. We define

$$\mathfrak{M}_\delta(q,a) = \{\alpha \in [0,1) : |\alpha - a/q| \leqslant X^{\delta-k}\},$$

and take $\mathfrak{M}_\delta$ to be the union of the arcs $\mathfrak{M}_\delta(q,a)$ with $0 \leqslant a \leqslant q \leqslant X^\delta$ and $(a,q) = 1$. We will take δ to be a real number with $0 < \delta < 1/5$. Finally, we put $\mathfrak{m}_\delta = [0,1) \setminus \mathfrak{M}_\delta$ for the corresponding set of minor arcs.

11. The Hasse Principle for diagonal forms, II: the minor arcs

The minor arc contribution is rapidly reduced to an estimate of the type previously encountered. By applying Hölder's inequality, we obtain

$$\left| \int_{\mathfrak{m}_\delta} f(c_1\alpha)\ldots f(c_s\alpha)\,\mathrm{d}\alpha \right| \leqslant \prod_{i=1}^s \biggl(\int_{\mathfrak{m}_\delta} |f(c_i\alpha)|^s\,\mathrm{d}\alpha \biggr)^{1/s}. \tag{11.1}$$

Moreover, for each index i one finds that

$$\int_{\mathfrak{m}_\delta} |f(c_i\alpha)|^s\,\mathrm{d}\alpha \leqslant \Bigl(\sup_{\alpha\in\mathfrak{m}_\delta} |f(c_i\alpha)| \Bigr)^{s-2^k} \int_0^1 |f(c_i\alpha)|^{2^k}\,\mathrm{d}\alpha. \tag{11.2}$$

By a change of variable, invoking periodicity modulo 1 of $f(\beta)$ and Hua's Lemma, we discern that

$$\int_0^1 |f(c_i\alpha)|^{2^k}\,\mathrm{d}\alpha = |c_i|^{-1} \int_0^{|c_i|} |f(\beta)|^{2^k}\,\mathrm{d}\beta = \int_0^1 |f(\beta)|^{2^k}\,\mathrm{d}\beta \ll X^{2^k-k+\varepsilon}. \tag{11.3}$$

The estimation of $\sup\limits_{\alpha\in\mathfrak{m}_\delta} |f(c_i\alpha)|$, meanwhile, is achieved through Weyl's inequality. It is convenient at this point to write

$$H = \max_{1\leqslant i\leqslant s} |c_i|.$$

Suppose that $\alpha \in \mathfrak{m}_\delta$. By Dirichlet's approximation theorem, there exists $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $1 \leqslant q \leqslant X^{k-\delta}$, $(a,q) = 1$ and $|c_i\alpha - a/q| \leqslant X^{\delta-k}$. Put

$$b = \frac{a}{(c_i,a)} \cdot \frac{|c_i|}{c_i} \quad \text{and} \quad r = \frac{|c_i|q}{(c_i,a)},$$

so that $|\alpha - b/r| \leqslant X^{\delta-k}$ with $1 \leqslant r \leqslant Hq \leqslant HX^{k-\delta}$ and $(b,r) = 1$. If one were to have $1 \leqslant r \leqslant X^\delta$, then α would lie on $\mathfrak{M}_\delta$. Thus we may assume that $r > X^\delta$, whence $H^{-1}X^\delta < q \leqslant X^{k-\delta}$. We therefore deduce from Weyl's inequality that

$$\begin{aligned} f(c_i\alpha) &\ll X^{1+\varepsilon}\left(r^{-1} + X^{-1} + rX^{-k}\right)^{2^{1-k}} \\ &\ll X^{1+\varepsilon}\left(HX^{-\delta} + X^{-1} + X^{-\delta}\right)^{2^{1-k}}, \end{aligned}$$

whence

$$\sup_{\alpha\in\mathfrak{m}_\delta} |f(c_i\alpha)| \ll_{\mathbf{c}} X^{1-\delta 2^{1-k}+\varepsilon}. \tag{11.4}$$

By substituting (11.3) and (11.4) into (11.2), we deduce that when $s \geqslant 2^k + 1$,

$$\int_{\mathfrak{m}_\delta} |f(c_i\alpha)|^s \,\mathrm{d}\alpha \ll (X^{1-\delta 2^{1-k}+\varepsilon})^{s-2^k} X^{2^k-k+\varepsilon} \ll X^{s-k-\delta 2^{-k}}.$$

Hence, on recalling (11.1), we obtain the bound

$$\int_{\mathfrak{m}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha \ll X^{s-k-\delta 2^{-k}}. \tag{11.5}$$

12. The Hasse Principle for diagonal forms, III: the major arcs

We first obtain an asymptotic formula for each exponential sum $f(c_i\alpha)$. We have

$$f(\alpha) = \sum_{|x|\leqslant X} e(\alpha x^k) = 1 + \sum_{1\leqslant x\leqslant X} e(\alpha x^k) + \sum_{1\leqslant x\leqslant X} e(\alpha(-x)^k).$$

Thus, as a consequence of Lemma 5.1, we see that whenever $\alpha \in \mathbb{R}$, $a \in \mathbb{Z}$ and $q \in \mathbb{N}$, one has

$$f(\alpha) - q^{-1}S(q,a)v(\alpha - a/q) \ll q + X^k|q\alpha - a|,$$

where now we write

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q) \quad \text{and} \quad v(\beta) = \int_{-X}^{X} e(\beta\gamma^k)\,\mathrm{d}\gamma.$$

Thus

$$f(c_i\alpha) - q^{-1}S(q,c_ia)v(c_i(\alpha - a/q)) \ll q + |c_i|X^k|q\alpha - a|.$$

Define the function $\widetilde{f}_i(\alpha)$ to be $q^{-1}S(q,c_ia)v(c_i(\alpha - a/q))$ when $\alpha \in \mathfrak{M}_\delta(q,a) \subseteq \mathfrak{M}_\delta$, and otherwise take $\widetilde{f}_i(\alpha) = 0$. Then it follows that

$$\sup_{\alpha\in\mathfrak{M}_\delta} |f(c_i\alpha) - \widetilde{f}_i(\alpha)| \ll_{\mathbf{c}} X^{2\delta},$$

and hence

$$\sup_{\alpha\in\mathfrak{M}_\delta} |f(c_1\alpha)\dots f(c_s\alpha) - \widetilde{f}_1(\alpha)\dots\widetilde{f}_s(\alpha)| \ll X^{s-1+2\delta}.$$

We may therefore infer that

$$\begin{aligned} \int_{\mathfrak{M}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha - \int_{\mathfrak{M}_\delta} \widetilde{f}_1(\alpha)\dots\widetilde{f}_s(\alpha)\,\mathrm{d}\alpha &\ll X^{s-1+2\delta}\mathrm{mes}(\mathfrak{M}_\delta) \\ &\ll X^{s-k+(5\delta-1)} = o(X^{s-k}). \end{aligned}$$

On noting that

$$\int_{\mathfrak{M}_\delta} \widetilde{f}_1(\alpha)\dots\widetilde{f}_s(\alpha)\,\mathrm{d}\alpha = J(X^\delta)\mathfrak{S}(X^\delta),$$

where

$$J(Q) = \int_{-QX^{-k}}^{QX^k} v(c_1\beta)\dots v(c_s\beta)\,\mathrm{d}\beta$$

and

$$\mathfrak{S}(Q) = \sum_{q=1}^{Q} \sum_{\substack{a=1\\(a,q)=1}}^{q} \prod_{i=1}^{s} \left(q^{-1}S(q,c_ia)\right),$$

we may conclude as follows.

Lemma 12.1. *One has*

$$\int_{\mathfrak{M}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha = J(X^\delta)\mathfrak{S}(X^\delta) + o(X^{s-k}).$$

It remains now only to complete the truncated singular integral and singular series, and to analyse the latter quantities. One can apply Fourier's Integral Theorem to asymptotically evaluate the completed singular integral

$$J = \int_{-\infty}^{\infty} v(c_1\beta)\dots v(c_s\beta)\,\mathrm{d}\beta,$$

though we will take the opportunity momentarily to explore a route that offers an intuitively more concrete formulation. For now, we concentrate on the singular series. One has

$$S(q,c_ia) = \sum_{r=1}^{q} e(c_iar^k/q) = (q,c_i)S\left(\frac{q}{(q,c_i)}, \frac{c_ia}{(q,c_i)}\right),$$

in which the last two arguments are pairwise coprime. We thus deduce from Lemma 7.1 that

$$S(q,c_ia) \ll (q,c_i)\left(\frac{q}{(q,c_i)}\right)^{1-2^{1-k}+\varepsilon} \ll_{\mathbf{c}} q^{1-2^{1-k}+\varepsilon}.$$

We write

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{a=1\\(a,q)=1}}^{q} \prod_{i=1}^{s} \left(q^{-1}S(q,c_i)\right).$$

Then it follows that whenever $s \geqslant 2^k+1$, then

$$\mathfrak{S} - \mathfrak{S}(Q) \ll \sum_{q>Q}\sum_{a=1}^{q} \left(q^{\varepsilon-2^{1-k}}\right)^s \ll Q^{-2^{-k}},$$

so that $\mathfrak{S}$ is absolutely convergent and $\mathfrak{S} - \mathfrak{S}(X^\delta) \ll X^{-\delta 2^{-k}}$.

Next we observe that the argument of the proof of Lemma 7.3 shows that for each index i, whenever $(a,q)=(b,r)=(q,r)=1$, one has

$$S(qr, c_i(ar+bq)) = S(q,c_ia)S(r,c_ib).$$

Write

$$A(q)=\sum_{\substack{a=1\\(a,q)=1}}^{q}\prod_{i=1}^{s}\left(q^{-1}S(q,c_ia)\right).$$

Then, just as in the proof of Lemma 7.4, one finds that $A(q)$ is a multiplicative function of q. Next we put

$$\sigma(p)=\sum_{h=0}^{\infty}A(p^h),$$

Then we deduce, as in the proof of Theorem 7.5, that when $s\geqslant 2^k+1$ the infinite sum $\sigma(p)$ converges absolutely, with

$$|\sigma(p)-1|\ll_{\mathbf{c}} p^{-1-2^{-k}},$$

that $\prod_p \sigma(p)$ converges absolutely with $\prod_p \sigma(p)=\mathfrak{S}$, and that there is a natural number $C=C(s,k,\mathbf{c})$ having the property that

$$1/2<\prod_{p\geqslant C}\sigma(p)<3/2.$$

Our next step is to analyse the p-adic densities $\sigma(p)$ when p is small. Put

$$M(q)=\text{card}\{\mathbf{m}\in(\mathbb{Z}/q\mathbb{Z})^s : c_1m_1^k+\ldots+c_sm_s^k=0\}.$$

Then, just as in the proof of Lemma 8.1, we find that

$$\sum_{d|q}A(d)=q^{1-s}M(q),$$

whence

$$\sigma(p)=\lim_{h\to\infty}p^{h(1-s)}M(p^h).$$

In order to obtain a lower bound for $\sigma(p)$, we employ our hypothesis concerning the local solubility of the equation $\phi(\mathbf{x})=0$. Define κ by putting

$$\kappa=\max_{1\leqslant i\leqslant s}\{m\in\mathbb{Z} : p^m|c_i\},$$

and put $\nu=\kappa+\gamma$, where $\gamma=\gamma(p,k)$ is defined as in the preamble to Lemma 8.3. We may suppose that there is a solution of the congruence

$$c_1y_1^k+\ldots+c_sy_s^k\equiv 0 \pmod{p^\nu},$$

in which $p\nmid y_i$ for some index i with $1\leqslant i\leqslant s$. By relabelling variables, if necessary, we may suppose that this index is $i=1$.

Consider an integer h with $h\geqslant\nu$, and suppose that $p^\mu\|c_1$, noting that $\mu\leqslant\kappa$. We fix a choice of $x_2,\ldots,x_s$ with $1\leqslant x_j\leqslant p^h$ and $x_j\equiv y_j \pmod{p^\nu}$ $(1\leqslant i\leqslant s)$. There are $(p^{h-\nu})^{s-1}$ such choices. We then have

$$-c_1^{-1}(c_2x_2^k+\ldots+c_sx_s^k)\equiv -c_1^{-1}(c_2y_2^k+\ldots+c_sy_s^k)\equiv y_1^k \pmod{p^{\nu-\mu}},$$

and hence

$$-c_1^{-1}(c_2x_2^k+\ldots+c_sx_s^k)\equiv y_1^k \pmod{p^\gamma}.$$

Thus, since the left hand side of this last congruence is a k-th power modulo p^γ, it follows from Lemma 8.3 that the left hand side is also a k-th power modulo p^h. Hence the congruence

$$-(c_2x_2^k+\ldots+c_sx_s^k)\equiv c_1x_1^k \pmod{p^h}$$

is soluble, and we may conclude that $M(p^h)\geqslant (p^{h-\nu})^{s-1}$. In particular, we have

$$\sigma(p)=\lim_{h\to\infty}p^{h(1-s)}M(p^h)\geqslant p^{-\nu(s-1)},$$

whence $\sigma(p)>0$ for all prime numbers p. From here we discern that

$$\mathfrak{S}=\Bigl(\prod_{p<C}\sigma(p)\Bigr)\Bigl(\prod_{p\geqslant C}\sigma(p)\Bigr)>\frac{1}{2}\prod_{p<C}\sigma(p)\gg 1,$$

so that in fact, subject to the local solubility hypothesis and the assumption $s\geqslant 2^k+1$, one has $1\ll\mathfrak{S}\ll 1$.

It remains to consider the singular integral. With this in mind, we observe initially that for each index i,

$$\begin{aligned}v(c_i\beta)&=\int_0^X e(c_i\beta\gamma^k)\,\mathrm{d}\gamma+\int_0^X e(c_i\beta(-\gamma)^k)\,\mathrm{d}\gamma\\&\ll\frac{X}{(1+X^k|c_i\beta|)^{1/k}}\ll_{\mathbf{c}}\frac{X}{(1+X^k|\beta|)^{1/k}}.\end{aligned}$$

Thus, defining the (complete) singular integral by means of the relation

$$J=\int_{-\infty}^\infty v(c_1\beta)\ldots v(c_s\beta)\,\mathrm{d}\beta,$$

we find that whenever $s\geqslant k+1$, one has

$$J-J(Q)\ll\int_{QX^{-k}}^\infty\frac{X^s}{(1+X^k\beta)^{s/k}}\,\mathrm{d}\beta\ll X^{s-k}Q^{-1/k}.$$

In particular, the singular integral J is absolutely convergent.

A slightly more robust treatment than in our discussion of Waring's problem is more illuminating in general than our application there of Fourier's Integral Theorem. We begin by observing that, by a change of variable,

$$J=\int_{-\infty}^\infty v(c_1\beta)\ldots v(c_s\beta)\,\mathrm{d}\beta=X^{s-k}I,$$

where

$$I=\int_{-\infty}^\infty v_1(c_1\beta)\ldots v_1(c_s\beta)\,\mathrm{d}\beta,$$

in which

$$v_1(\beta)=\int_{-1}^1 e(\beta\gamma^k)\,\mathrm{d}\gamma.$$

When $0<\eta\leqslant 1$, we define the auxiliary function

$$w_\eta(\beta)=\eta\Bigl(\frac{\sin(\pi\eta\beta)}{\pi\eta\beta}\Bigr)^2.$$

This function has Fourier transform

$$\widehat{w}_\eta(\gamma) = \int_{-\infty}^{\infty} w_\eta(\beta) e(-\beta\gamma)\,\mathrm{d}\beta = \max\{0, 1 - |\gamma|/\eta\}. \tag{12.1}$$

Notice here that this integral converges absolutely. One can apply the formula (12.1) to construct a continuous approximation to the indicator function of an interval. When $0 < \delta < \eta$, we define

$$W_{\eta,\delta}(\gamma) = \begin{cases} 1, & \text{when } |\gamma| \leqslant \eta, \\ 1 - \dfrac{|\gamma| - \eta}{\delta}, & \text{when } \eta < |\gamma| < \eta + \delta, \\ 0, & \text{when } |\gamma| \geqslant \eta + \delta. \end{cases}$$

We put

$$W_\eta^+(\gamma) = W_{\eta,\eta^2}(\gamma) \quad \text{and} \quad W_\eta^-(\gamma) = W_{\eta-\eta^2,\eta^2}(\gamma).$$

Then $W_\eta^+(\gamma)$ and $W_\eta^-(\gamma)$ supply upper and lower bounds for the function

$$\begin{cases} 1, & \text{when } |\gamma| \leqslant \eta, \\ 0, & \text{when } |\gamma| > \eta. \end{cases}$$

Our formula (12.1) shows that

$$W_{\eta,\delta}(\gamma) = (1 + \eta/\delta)\widehat{w}_{\eta+\delta}(\gamma) - (\eta/\delta)\widehat{w}_\eta(\gamma),$$

whence

$$W_\eta^+(\gamma) = (1 + \eta^{-1})\widehat{w}_{\eta+\eta^2}(\gamma) - \eta^{-1}\widehat{w}_\eta(\gamma)$$

and

$$W_\eta^-(\gamma) = \eta^{-1}\widehat{w}_\eta(\gamma) + (1 - \eta^{-1})\widehat{w}_{\eta-\eta^2}(\gamma).$$

Write $M_\infty(\eta)$ for the volume of the subset of $[-1,1]^s$ defined by the inequality $|\phi(\boldsymbol{\xi})| < \eta$. Then our discussion shows that

$$\int_{[-1,1]^s} W_\eta^-(\phi(\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi} \leqslant M_\infty(\eta) \leqslant \int_{[-1,1]^s} W_\eta^+(\phi(\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi}.$$

We now turn our attention to the Fourier side. Define

$$U(\eta) = \int_{[-1,1]^s} \eta^{-1}\widehat{w}_\eta(\phi(\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi}.$$

Lemma 12.2. *Let η be a real number with $0 < \eta < 1$. Suppose that η_1 is a real number with $|\eta_1 - \eta| \leqslant \eta^2$. Then*

$$U(\eta_1) = I + O(\eta^{1/(2k)}).$$

Proof. Put

$$K(\beta) = \eta_1^{-1} w_{\eta_1}(\beta).$$

Then by interchanging orders of integration, it follows that

$$U(\eta_1) = \int_{-\infty}^{\infty} v_1(c_1\beta)\dots v_1(c_s\beta) K(\beta)\,\mathrm{d}\beta,$$

whence

$$U(\eta_1) - I = \int_{-\infty}^{\infty} v_1(c_1\beta)\dots v_1(c_s\beta)\,(K(\beta) - 1)\,\mathrm{d}\beta. \tag{12.2}$$

We want to show that the right hand side in (12.2) is small. Put $\mathfrak{D} = [-\eta^{-1/2}, \eta^{-1/2}]$ and $\mathfrak{E} = \mathbb{R} \setminus \mathfrak{D}$. Then from the power series expansion of $w_\eta(\beta)$, we have

$$0 \leqslant 1 - K(\beta) \ll \min\{1, \eta^2\beta^2\}.$$

The absolute convergence of the integral I ensures that the contribution arising from integrating over $\mathfrak{D}$ on the right hand side of (12.2) is at most

$$\sup_{\beta\in\mathfrak{D}} |1 - K(\beta)| \int_{-\infty}^{\infty} v_1(c_1\beta)\dots v_1(c_s\beta)\,\mathrm{d}\beta \ll \eta.$$

The contribution arising from $\mathfrak{E}$, meanwhile, is bounded above by

$$\int_{\mathfrak{E}} (1+|\beta|)^{-s/k}\,\mathrm{d}\beta \ll (\eta^{-1/2})^{-1/k} = \eta^{1/(2k)}.$$

Thus we infer that

$$U(\eta_1) - I \ll \eta^{1/(2k)}.$$

This completes the proof of the lemma. □

By using the definitions of $W_\eta^{\pm}(\gamma)$, we obtain

$$\begin{aligned}\int_{[-1,1]^s} W_\eta^+(\phi(\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi} &= \left((\eta+\eta^2)(1+1/\eta) - \eta(1/\eta)\right)\left(I + O(\eta^{1/(2k)})\right)\\ &= \left(2\eta + O(\eta^2)\right)\left(I + O(\eta^{1/(2k)})\right)\end{aligned}$$

and

$$\begin{aligned}\int_{[-1,1]^s} W_\eta^-(\phi(\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi} &= \left(\eta(1/\eta) + (\eta-\eta^2)(1-1/\eta)\right)\left(I + O(\eta^{1/(2k)})\right)\\ &= \left(2\eta + O(\eta^2)\right)\left(I + O(\eta^{1/(2k)})\right).\end{aligned}$$

We therefore conclude from that

$$(2\eta)^{-1}M_\infty(\eta) = I + O(\eta^{1/(2k)}),$$

whence

$$\lim_{\eta\to 0+} (2\eta)^{-1}M_\infty(\eta) = I = \int_{-\infty}^{\infty} v_1(c_1\beta)\dots v_1(c_s)\beta)\,\mathrm{d}\beta.$$

In particular, one has $J = \sigma_\infty X^{s-k}$, where σ_∞ is the Siegel volume

$$\sigma_\infty = \lim_{\eta\to 0+} (2\eta)^{-1}M_\infty(\eta).$$

Notice that $M_\infty(\eta)$ is the volume of an η-neighborhood of the hypersurface $\phi(\boldsymbol{\xi}) = 0$. By hypothesis, the equation $\phi(\mathbf{z}) = 0$ has a real point $\mathbf{z} = \mathbf{z}^{(\infty)} \in [-\frac{1}{2}, \frac{1}{2}]^s$. Hence, an application of the Implicit Function Theorem shows that there is a positive number ω with the property that in a ω-neighborhood D of $\mathbf{z}^{(\infty)}$, this hypersurface is well-approximated by a hyperplane, and thus an η-neighborhood of D has volume $\gg_\omega \eta$. In particular, one has $\sigma_\infty > 0$, whence $J \gg X^{s-k}$. We therefore conclude that when $s \geqslant 2^k + 1$, one has

$$\begin{aligned}\int_{\mathfrak{M}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha &= \mathfrak{S}(X^\delta)J(X^\delta) + o(X^{s-k})\\ &= \mathfrak{S}J + o(X^{s-k})\\ &= \sigma_\infty\Big(\prod_p \sigma_p\Big)X^{s-k} + o(X^{s-k}).\end{aligned}$$

Finally,

$$\begin{aligned}N_\phi(X) &= \int_{\mathfrak{M}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha + \int_{\mathfrak{m}_\delta} f(c_1\alpha)\dots f(c_s\alpha)\,\mathrm{d}\alpha\\ &= \sigma_\infty\Big(\prod_p \sigma_p\Big)X^{s-k} + o(X^{s-k}),\end{aligned}$$

whence $N_\phi(X) \to \infty$ as $X \to \infty$. Then the equation $c_1x_1^k + \ldots + c_sx_s^k = 0$ has a non-trivial integral solution.

13. Analogues of Hua's lemma: diminishing ranges

We now turn to the task of reducing the number of variables required to successfully apply the circle method. By incorporating the refined analysis of the major arcs available from the fourth problem set, it is apparent that only $4k$ variables are required to asymptotically estimate the major arc contribution. Our goal then is to refine the minor arc analysis in a comparable manner. With this end in mind, we restrict the variables in an attempt to make the minor arc contribution easier to estimate, in the hope that the major arc analysis remains under our control.

Let n be a large natural number, write $X = n^{1/k}$, and consider real numbers X_i $(0 \leqslant i \leqslant s)$ defined by putting

$$X_0 = \tfrac{1}{20}X \quad \text{and} \quad X_{i+1} = \tfrac{1}{2}X_i^{1-1/k} \quad (i \geqslant 0).$$

We consider the number $\rho(n)$ of representations of the integer n in the shape

$$n = \sum_{i=1}^{t} x_i^k + \sum_{j=0}^{s-1}(y_j^k + z_j^k),$$

with $1 \leqslant x_i \leqslant X$ $(1 \leqslant i \leqslant t)$ and $X_j < y_j, z_j \leqslant 2X_j$ $(j \geqslant 0)$. If we can show that $\rho(n) \to \infty$ as $n \to \infty$, then it follows that $G(k) \leqslant 2s + t$. We may then seek to optimise the choices for s and t so as to minimise $G(k)$.

Write

$$\begin{aligned}f(\alpha) &= \sum_{1\leqslant x\leqslant X} e(\alpha x^k),\\ g_j(\alpha) &= \sum_{X_j<y\leqslant 2X_j} e(\alpha y^k) \quad (0 \leqslant j \leqslant s-1),\\ G(\alpha) &= \prod_{j=0}^{s-1} g_j(\alpha).\end{aligned}$$

Then it follows via orthogonality that

$$\rho(n)=\int_0^1 f(\alpha)^t G(\alpha)^2 e(-n\alpha)\,\mathrm{d}\alpha.$$

It is useful for future reference to observe that a trivial estimate yields

$$|G(\alpha)|\leqslant G(0)=X_0X_1\dots X_{s-1}\asymp\prod_{j=0}^{s-1}X^{(1-1/k)^j}=X^{k(1-(1-1/k)^s)}.$$

Lemma 13.1 (Diminishing ranges estimate)**.** *One has*

$$\int_0^1|G(\alpha)|^2\,\mathrm{d}\alpha\ll X^{k-k(1-1/k)^s}.$$

In order to have reasonable prospects for successfully applying the circle method, we should seek to save a factor X^k over the trivial count for the number of choices for the underlying variables, which suggests that we should aim for a bound $G(0)^2X^{-k}$ in the above lemma. The argument of the proof of the lemma in fact shows that

$$\int_0^1|G(\alpha)|^2\,\mathrm{d}\alpha\asymp G(0)^2X^{-k+k(1-1/k)^s}.$$

Since $k(1-1/k)^s\leqslant ke^{-s/k}$, this estimate gets exponentially close (in the exponent) to the required bound as s grows.

Proof of Lemma 13.1. By orthogonality, one sees that the integral in question is equal to the number of solutions of the equation

$$y_0^k-z_0^k=\sum_{j=1}^{s-1}(y_j^k-z_j^k),\tag{13.1}$$

with $X_j<y_j,z_j\leqslant 2X_j$ $(0\leqslant j\leqslant s-1)$. However, we have

$$\biggl|\sum_{j=1}^{s-1}(y_j^k-z_j^k)\biggr|<(2X_1)^k+o(X_1^k)=(1+o(1))X_0^{k-1},$$

whereas, when $y_0\neq z_0$, one has

$$|y_0^k-z_0^k|=|y_0-z_0|\cdot|y_0^{k-1}+y_0^{k-2}z_0+\ldots+z_0^{k-1}|\geqslant kX_0^{k-1}.$$

Thus, when $y_0\neq z_0$, we have

$$|y_0^k-z_0^k|>\biggl|\sum_{j=1}^{s-1}(y_j^k-z_j^k)\biggr|,$$

contradicting the validity of the equation (13.1). We are therefore forced to conclude that $y_0=z_0$, whence

$$\int_0^1|G(\alpha)|^2\,\mathrm{d}\alpha\leqslant X_0\int_0^1|g_1(\alpha)\dots g_{s-1}(\alpha)|^2\,\mathrm{d}\alpha.$$

The mean value on the right hand side here has the same shape as that on the left hand side, though with $s-1$ pairs of generating functions in place of s. Thus, an inductive argument confirms that

$$\int_0^1 |G(\alpha)|^2\,\mathrm{d}\alpha \leqslant X_0X_1\dots X_{s-1} \asymp X^{k-k(1-1/k)^s}.$$

This completes the proof of the lemma. □

This lemma serves as a substitute for Hua's lemma in our analysis of $\rho(n)$, though it is far more efficient. In order to make further progress, we require a Hardy-Littlewood dissection of the unit interval. Here it is expedient to make use of our earlier work. Take δ to be a positive number with $\delta<1/5$, and define $\mathfrak{M}_\delta$ to be the union of the intervals

$$\mathfrak{M}_\delta(q,a)=\{\alpha\in[0,1):|\alpha-a/q|\leqslant X^{\delta-k}\}$$

with $0\leqslant a\leqslant q\leqslant X^\delta$ and $(a,q)=1$. We then put $\mathfrak{m}_\delta=[0,1)\setminus\mathfrak{M}_\delta$.

Our analysis of the minor arcs makes use of Weyl's inequality just as in Lemma 3.7 and its sequel. Thus, we have

$$\sup_{\alpha\in\mathfrak{m}_\delta}|f(\alpha)|\ll X^{1-\delta 2^{1-k}+\varepsilon},$$

whence

$$\begin{aligned}\int_{\mathfrak{m}_\delta} f(\alpha)^tG(\alpha)^2e(-n\alpha)\,\mathrm{d}\alpha &\leqslant \Big(\sup_{\alpha\in\mathfrak{m}_\delta}|f(\alpha)|\Big)^t\int_0^1|G(\alpha)|^2\,\mathrm{d}\alpha\\ &\ll \Big(X^{1-\delta 2^{1-k}+\varepsilon}\Big)^tX^{k-k(1-1/k)^s}.\end{aligned}$$

It follows that, for some positive number τ, one has

$$\int_{\mathfrak{m}_\delta} f(\alpha)^tG(\alpha)^2e(-n\alpha)\,\mathrm{d}\alpha\ll G(0)^2X^{t-k-k\tau}\asymp G(0)^2n^{t/k-1-\tau},$$

provided only that

$$t\delta 2^{1-k}>k(1-1/k)^s. \tag{13.2}$$

We shall presently establish the major arc estimate

$$\int_{\mathfrak{M}_\delta} f(\alpha)^tG(\alpha)^2e(-n\alpha)\,\mathrm{d}\alpha\gg G(0)^2n^{t/k-1},$$

subject to the condition that $t\geqslant 4k$. Since this quantity tends to ∞ as $n\to\infty$, it follows that

$$G(k)\leqslant\inf_{s\geqslant 0}\left(2s+\max\left\{4k,\left\lceil\frac{k(1-1/k)^s}{\delta 2^{1-k}}\right\rceil\right\}\right).$$

By varying the parameter s, one may minimise the quantity central to the right hand side here, though this is a little messy. Observe that $(1-1/k)^s\leqslant e^{-s/k}$. Thus, taking $\delta=1/8$ in order to simplify the associated computations, we see that

$$2s+\left\lceil\frac{k(1-1/k)^s}{\delta 2^{1-k}}\right\rceil<2s+\lceil 8k2^{k-1}e^{-s/k}\rceil.$$

If we differentiate the smooth function underlying the right hand quantity here, then the optimal choice for s is seen to be approximated by the solution of the equation

$2=2^{k+2}e^{-s/k}$, so that s is approximately $k\log(2^{k+1})=k(k+1)\log 2$. Motivated by this observation, we put $s=\lceil k^2\log 2\rceil$ and set

$$t=\left\lceil \frac{k(1-1/k)^s}{\delta 2^{1-k}}\right\rceil \leqslant \lceil k2^{k+2}\cdot 2^{-k}\rceil = 4k.$$

Thus we see that

$$G(k)\leqslant 2\lceil k^2\log 2\rceil + 4k < (2\log 2)k^2+4k+2.$$

Let us return to the analysis of the major arc contribution. We have

$$\int_{\mathfrak{M}_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha)\,\mathrm{d}\alpha = \sum_{\mathbf{y},\mathbf{z}} \int_{\mathfrak{M}_\delta} f(\alpha)^t e(-N\alpha)\,\mathrm{d}\alpha,$$

where

$$N=n-\sum_{j=0}^{s-1}(y_j^k+z_j^k),$$

and the summation over $\mathbf{y},\mathbf{z}$ is over $X_j<y_j,z_j\leqslant 2X_j$ $(0\leqslant j\leqslant s-1)$. We note that

$$\sum_{j=0}^{s-1}(y_j^k+z_j^k)\leqslant 2(2X_0)^k+O(X_1^k)<10^{1-k}n,$$

whence

$$N=n-\sum_{j=0}^{s-1}(y_j^k+z_j^k)>n/2.$$

Also, since $\delta<1/5$ and $t\geqslant 4k$, our major arc analysis (using the estimate $S(q,a)\ll q^{1-1/k+\varepsilon}$ for $(q,a)=1$ from the third problem set) shows that

$$\int_{\mathfrak{M}_\delta} f(\alpha)^t e(-N\alpha)\,\mathrm{d}\alpha = \frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}\mathfrak{S}_{t,k}(N)N^{t/k-1}+o(N^{t/k-1})\gg n^{t/k-1}.$$

Hence

$$\int_{\mathfrak{M}_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha)\,\mathrm{d}\alpha \gg n^{t/k-1}\sum_{\mathbf{y},\mathbf{z}}1 = n^{t/k-1}G(0)^2.$$

This justifies our earlier claim.

We may summarise these deliberations as follows.

Theorem 13.2. *When $k\geqslant 2$, one has*

$$G(k)<(2\log 2)k^2+4k+2.$$

Although this result can be refined in various ways, that already presented represents a significant improvement on Hua's bound $G(k)\leqslant 2^k+1$. One may check on the back of an envelope that it improves on the latter bound for $k\geqslant 7$.

References

[1] G. I. Arkhipov and A. A. Karatsuba, *Local representation of zero by a form*, Izv. Akad. Nauk SSSR Ser. Mat. **45** (1981), no. 5, 948–961.

[2] J. Bourgain, *On the Vinogradov integral*, Proc. Steklov Inst. Math. **296** (2017), no. 1, 30–40.

[3] J. Bourgain, C. Demeter and L. Guth, *Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three*, Ann. of Math. (2) **184** (2016), no. 2, 633–682.

[4] W. D. Brownawell, *On p-adic zeros of forms*, J. Number Theory **18** (1984), no. 3, 342–349.

[5] J. Brüdern and T. D. Wooley, *On Waring's problem for larger powers*, submitted, 28pp; arxiv:2211.10380.

[6] A. L. Cauchy, *Recherches sur les nombres*, J. École Polytech. **9** (1813), 99–116.

[7] H. Davenport, *On the addition of residue classes*, J. London Math. Soc. **10** (1935), 30–32.

[8] H. Davenport, *On Waring's problem for fourth powers*, Ann. of Math. **40** (1939), 731–747.

[9] H. Davenport and D. J. Lewis, *Homogeneous additive equations*, Proc. Royal Soc. London, Series A **274** (1963), 443–460.

[10] V. B. Demyanov, *On cubic forms in discretely normed fields*, Doklady Akad. Nauk SSSR (N.S.) **74** (1950), 889–891.

[11] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. (2) **17** (1918), 75–115.

[12] G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum": I. A new solution of Waring's problem*, Göttingen Nachrichten (1920), 33-54.

[13] H. Hasse, *Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen*, J. Reine Angew. Math. **152** (1923), 129–148.

[14] D. Hilbert, *Beweis für Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl nter Potenzen (Waringsche Problem)*, Math. Ann. **67** (1909), 281–300.

[15] L.-K. Hua, *On Waring's problem*, Quart. J. Math. **9** (1938), 199–202.

[16] J. M. Kubina and M. C. Wunderlich, *Extending Waring's conjecture to* $471,600,000$, Math. Comp. **55** (1990), no. 192, 815–820.

[17] D. J. Lewis, *Cubic homogeneous polynomials over p-adic number fields*, Ann. of Math. (2) **56** (1952), 473–478.

[18] D. J. Lewis and H. L. Montgomery, *On zeros of p-adic forms*, Michigan Math. J. **30** (1983), no. 1, 83–87.

[19] Yu. V. Linnik, *On the representation of large numbers as sums of seven cubes*, Dokl. Akad. Nauk SSSR **35** (1942), 162.

[20] Yu. V. Linnik, *On the representation of large numbers as sums of seven cubes*, Mat. Sbornik **12** (1943), 218–224.

[21] K. Mahler, *On the fractional parts of the powers of a rational number II*, Mathematika **4** (1957), 122–124.

[22] R. C. Vaughan, *On Waring's problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.

[23] R. C. Vaughan, *On Waring's problem for smaller exponents*, Proc. London Math. Soc. (3) **52** (1986), 445–463.

[24] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem, II: sixth powers*, Duke Math. J. **76** (1994), 683–710.

[25] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), no. 2, 147–240.

[26] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. **77** (1916), 313–352.

[27] T. D. Wooley, *Large improvements in Waring's problem*, Ann. of Math. (2) **135** (1992), no. 1, 131–164.

[28] T. D. Wooley, *New estimates for smooth Weyl sums*, J. London Math. Soc. **51** (1995), 1–13.

[29] T. D. Wooley, *On the local solubility of Diophantine systems*, Compositio Math. **111** (1998), 149–165.

[30] T. D. Wooley, *The asymptotic formula in Waring's problem*, Int. Math. Res. Not. IMRN **2012** (2012), no. 7, 1485–1504.

[31] T. D. Wooley, *On Waring's problem for intermediate powers*, Acta Arith. **176** (2016), no. 3, 241–247.

[32] T. D. Wooley, *The cubic case of the main conjecture in Vinogradov's mean value theorem*, Adv. Math. **294** (2016), 532–561.

[33] T. D. Wooley, *Nested efficient congruencing and relatives of Vinogradov's mean value theorem*, Proc. London Math. Soc. (3) **118** (2019), no. 4, 942–1016.

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

E-mail address: twooley@purdue.edu