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 \( x_1 \geq x_2 \geq \ldots \geq x_s \) 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 \( s \geq g(k) \), 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 \( g(2) = 4 \), \( g(3) = 9 \), \( g(4) = 19 \), \( \ldots \), 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

\[
2020 = 44^2 + 8^2 + 4^2 + 2^2.
\]

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

\[
2020 = 42^2 + 16^2 + 0^2 + 0^2,
\]

so that 2020 is the sum of just two squares of natural numbers. Likewise, one has

\[
2023 = 37^2 + 25^2 + 5^2 + 2^2.
\]

\(^1\)In this course the natural numbers \( \mathbb{N} \) are defined to be the set of positive integers \( \{1, 2, \ldots\} \).
and yet 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 $n_0 = 2^k \cdot \left\lfloor \frac{3}{2}k \right\rfloor - 1$, 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 $\left\lfloor \frac{3}{2}k \right\rfloor - 1$ copies of $2^k$, and $2^k - 1$ copies of $1^k$. Hence

$$g(k) \geq 2^k + \left\lfloor \frac{3}{2}k \right\rfloor - 2.$$

In fact, it is now known that $g(k) = 2^k + \left\lfloor \frac{3}{2}k \right\rfloor - 2$, 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

$$2^k \left\{ \left\lfloor \frac{3}{2}k \right\rfloor + \left\lfloor \frac{4}{3}k \right\rfloor > 2^k, \quad \text{(1.1)} \right.$$  

and it was shown that this inequality fails for every exponent $k \leq 471,600,000$ (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

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

when $\left\lfloor \frac{4}{3}k \right\rfloor \left\lfloor \frac{3}{2}k \right\rfloor > 2^k + 1$, and that

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

when $\left\lfloor \frac{4}{3}k \right\rfloor \left\lfloor \frac{3}{2}k \right\rfloor = 2^k + 1$.

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 $s$ having the property that whenever $s \geq G(k)$, then all sufficiently large natural numbers are the sum of $s$ positive integral $k$-th powers.

Thus, when $k \in \mathbb{N}$ and $s \geq G(k)$, there exists $N_0 = N_0(s, k)$ such that, whenever $n \geq N_0$, 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 $k \geq 2$, one has $G(k) \geq k + 1$. In brief, the sharpest upper bounds currently available are as follows:

- $G(2) = 4$, a consequence of Lagrange’s theorem from 1770;
- $G(3) \leq 7$, due to Linnik, 1942;
- $G(4) = 16$, due to Davenport, 1939;
- $G(5) \leq 17$, due to Vaughan and Wooley, 1995;
ARITHMETIC HARMONIC ANALYSIS

\[ G(6) \leq 24, \text{ due to Vaughan and Wooley, 1994;} \]
\[ G(7) \leq 31, \text{ due to Wooley, 2016;} \]
\[ G(8) \leq 39, \text{ due to Wooley, 2016 (and it is known that } G(8) \geq 32) \]
\[ G(9) \leq 47, \text{ due to Wooley, 2016;} \]
and so on (see [7, 19, 20, 25, 26, 34]). In general, for large values of \( k \), it has been shown that
\[ G(k) \leq k(\log k + \log \log k + 2 + o(1)) \text{ (Wooley, 1992 and 1995),} \]
where \( o(1) \to 0 \) as \( k \to \infty \) (see [30, 31]).

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
\[ g_k(z)^s = \left( \sum_{m_1=1}^{\infty} z^{m_1^k} \right) \left( \sum_{m_2=1}^{\infty} z^{m_2^k} \right) \cdots \left( \sum_{m_s=1}^{\infty} z^{m_s^k} \right) \]
\[ = \sum_{m_1=1}^{\infty} \ldots \sum_{m_s=1}^{\infty} z^{m_1^k + \ldots + m_s^k} \]
\[ = \sum_{n=1}^{\infty} R_{s,k}(n) z^n, \]
where we write
\[ R_{s,k}(n) = \text{card } \{ m_1, \ldots, m_s \in \mathbb{N} : m_1^k + \ldots + m_s^k = n \}. \]
We can recover the coefficients \( R_{s,k}(n) \) by employing Cauchy’s integral formula to evaluate a suitable contour integral. Thus
\[ R_{s,k}(n) = \frac{1}{2\pi i} \int_C g_k(z)^s z^{-n-1} \, dz, \]
in which \( C \) 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 \( z = re^{i\theta}, \) 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 \( x_i \leq n^{1/k} \). With this observation in mind, we put \( X = \lfloor n^{1/k} \rfloor \) and define
\[
f_k(\theta) = \sum_{1 \leq x \leq X} e(\theta x^k).
\]
We then have
\[
f_k(\theta)^* = \left( \sum_{1 \leq x_1 \leq X} e(\theta x_1^k) \right) \ldots \left( \sum_{1 \leq x_s \leq X} e(\theta x_s^k) \right)
= \sum_{1 \leq x_1 \leq X} \ldots \sum_{1 \leq x_s \leq X} e(\theta(x_1^k + \ldots + x_s^k))
= \sum_{1 \leq m \leq X^k} R_{s,k}^*(m) e(\theta m),
\]
where
\[
R_{s,k}^*(m) = \text{card}\{m_1, \ldots, m_s \in \mathbb{N} \cap [1, X] : m_1^k + \ldots + m_s^k = m\}.
\]
Thus, by applying the orthogonality relation
\[
\int_0^1 e(\theta h) \, d\theta = \begin{cases} 0, & \text{when } h \in \mathbb{Z} \setminus \{0\}, \\ 1, & \text{when } h = 0, \end{cases}
\]
we deduce that
\[
\int_0^1 f_k(\theta)^* e(-\theta n) \, d\theta = \sum_{1 \leq m \leq X^k} R_{s,k}^*(m) \int_0^1 e(\theta(m - n)) \, d\theta = R_{s,k}(n).
\]
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 \( \theta \) 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 \( 0 \leq \theta \leq 1/8X^{-k} \), in which case one has
\[
|f_k(\theta)| \geq \text{Re} \left( \sum_{1 \leq x \leq X} e(1/8) \right) \geq X \cos(\pi/4) = X/\sqrt{2},
\]
similar observations hold in wider generality. Meanwhile, when \( \theta \) is not closely approximated by a rational number with a small denominator, then the argument \( \theta x^k \) 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 \( \theta \), 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 \( k \geq 3 \), the generating function \( f_k(\theta) \) could be evaluated only for \( \theta \) lying on small neighbourhoods of rational numbers having small denominators. Hardy and Littlewood labelled such a set \( \text{the major arcs} \) in their application of the circle method. In order to motivate later discussion, consider a parameter \( \delta \) with \( 0 < \delta < 1 \) and the set
\[
\mathcal{M}_\delta = \bigcup_{0 \leq a < q \leq X^\delta \atop (a, q) = 1} \mathcal{M}_\delta(q, a),
\]
where

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

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

\[ S(q,a) = \sum_{r=1}^{q} e(ar^k/q) \quad \text{and} \quad v(\beta) = \int_{0}^{X} e(\beta \gamma^k) \, d\gamma, \]

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 \( \delta \), one may obtain an asymptotic formula of the shape

\[ \int_{\mathfrak{m}_\delta} f_k(\theta)^s e(-n\theta) \, d\theta \sim \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n)n^{s/k-1}, \quad (1.2) \]

as \( n \to \infty \). Here, we have written \( \Gamma(\cdot) \) for the classical Gamma function, 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 \( \mathcal{M}_\delta(q,a) \) were indeed arcs contained within the circular contour, and the moniker has stuck.

Since

\[ R_{s,k}(n) = \int_{0}^{1} f_k(\theta)^s e(-\theta n) \, d\theta, \]

the asymptotic analysis of \( R_{s,k}(n) \) will be completed by showing that the contribution arising from the set \( \mathfrak{m}_\delta = [0,1) \setminus \mathcal{M}_\delta \), 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 \( k \geq 3 \) 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 [29]). 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 \( s \),

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

whence

\[ R_{s,k}(n) = \int_{\mathfrak{m}_\delta} f_k(\theta)^s e(-\theta n) \, d\theta + \int_{\mathfrak{m}_\delta} f_k(\theta)^s e(-\theta n) \, d\theta \sim \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n)n^{s/k-1}. \]
Provided that $\mathcal{G}_{s,k}(n)$ is bounded away from 0, one is able to infer that $R_{s,k}(n) \to \infty$ as $n \to \infty$, establishing that $G(k) \leq s$. Indeed, Hardy and Littlewood were initially able to establish such a conclusion whenever $s \geq (k - 2)2^{k-1} + 5$.

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 $s \geq 2^k + 1$, and in particular establishes the bound $G(k) \leq 2^k + 1$ ($k \geq 2$).

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 2020. I thank everyone who has notified me of typos and corrections, 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

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

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

$$\lim_{x \to x_0} \frac{|f(x)|}{g(x)} = 0.$$

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 $\varepsilon$. Thus, whenever a statement contains the letter $\varepsilon$, then we assert that the statement holds for all positive values of $\varepsilon$. The implicit quantifier renders the interpretation of $\varepsilon$ as having a fixed value meaningless, and in this sense the value of $\varepsilon$ 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 $\varepsilon$, then this is clearly stated in so many words so as to avoid ambiguity concerning this $\varepsilon$-convention.

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

$$\lfloor \theta \rfloor := \max_{n \in \mathbb{Z}, n \leq \theta} n,$$

$$\lceil \theta \rceil := \min_{n \in \mathbb{Z}, n \geq \theta} n,$$

$$\{ \theta \} := \theta - \lfloor \theta \rfloor,$$

$$\| \theta \| := \min_{n \in \mathbb{Z}} |\theta - n|.$$

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

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

with $X$ large, seeking bounds of use on the minor arcs $m_\delta$. 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

$$\sum_{X < x \leq X + Y} e(\alpha x) \ll \min \{ Y, \| \alpha \|^{-1} \}. \quad (3.1)$$

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 $|e(\alpha x)| \leq 1$ for each $x$, one sees that the left hand side of (3.1) is at most $Y$, 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

$$\sum_{X < x \leq X + Y} e(\alpha x) = \frac{e(\alpha [X + Y + 1]) - e(\alpha [X + 1])}{e(\alpha) - 1} \ll |e(\alpha/2) - e(-\alpha/2)|^{-1} \ll |\sin(\pi \alpha)|^{-1}.$$

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

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

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

This lemma provides a means of estimating exponential sums of degree 1, but what about degrees $k$ exceeding 1? An idea of Weyl (see [29]) 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

\[
T(X) = \sum_{1 \leq x \leq X} e(\psi(x)).
\]

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

\[
|T(X)|^2 = \sum_{1 \leq y \leq X} \sum_{1 \leq x \leq X} e(\psi(y) - \psi(x))
= \sum_{|h|<X} \sum_{1 \leq x \leq X} e(\psi(x+h) - \psi(x)),
\]

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

\[
|T(X)|^2 \leq \sum_{|h|<X} \left| \sum_{1 \leq x \leq X} e(h\psi_1(x; h)) \right|,
\]

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 \( \Delta_1 \) the forward difference operator defined via

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

We then define \( \Delta_j \) for \( j \geq 2 \) recursively by means of the relation

\[
\Delta_j(\psi(x); h) = \Delta_j(\psi(x); h_1, \ldots, h_j)
= \Delta_1(\Delta_{j-1}(\psi(x); h_1, \ldots, h_{j-1}); h_j).
\]

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

\[
\Delta_j(x^k; h) = h_1 \ldots h_j p_j(h; h_1, \ldots, h_j),
\]

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

\[
\Delta_j(a_k x^k + \ldots + a_1 x; h) = \sum_{i=1}^{k} a_i \Delta_j(x^i; h),
\]

and so one is able easily to infer the structure of expressions of the shape \( \Delta_j(p(x); h) \), 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 \( \Delta_j(p(x); h) = 0 \).

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

\[
F(\psi) = \sum_{1 \leq x \leq X} e(\psi(x)).
\]
Then for each natural number \( j \), one has
\[
|F(\psi)|^{2j} \leq (2X)^{2j-1} \sum_{|h_1|<X} \cdots \sum_{|h_j|<X} \sum_{x \in I_j(h)} e(\Delta_j(\psi(x); h)),
\]
where \( I_j(h) \) denotes the interval of integers defined by putting \( I_0(h) = [1, X] \), and, when \( j \geq 2 \), by recursively setting
\[
I_j(h_1, \ldots, h_j) = I_{j-1}(h_1, \ldots, h_{j-1}) \cap \{ x \in [1, X] : x + h_j \in I_{j-1}(h_1, \ldots, h_{j-1}) \}.
\]

We note for future reference that the intervals occurring in the statement of this lemma satisfy the inclusions
\[
I_j(h_1, \ldots, h_j) \subseteq I_{j-1}(h_1, \ldots, h_{j-1}) \subseteq \ldots \subseteq I_1(h_1) \subseteq [1, X].
\]

Proof. We proceed by induction. When \( j = 1 \), one has
\[
|F(\psi)|^2 = \sum_{1 \leq x \leq X} \sum_{1 \leq y \leq X} e(\psi(y) - \psi(x))
\]
\[
\quad = \sum_{1 \leq x \leq X} \sum_{1-x \leq h \leq X-x} e(\psi(x + h) - \psi(x))
\]
\[
\quad = \sum_{|h_1|<X} \sum_{x \in I_1(h_1)} e(\Delta_1(\psi(x); h_1)),
\]
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 \( 1 \leq j < J \). Then, as a consequence of Cauchy’s inequality, one finds that
\[
|F(\psi)|^{2j} = (|F(\psi)|^{2j-1})^2 \leq (2X)^{2j-1-1} \left( \sum_{|h_1|<X} \cdots \sum_{|h_{j-1}|<X} 1 \right) \Xi(\psi),
\]
where
\[
\Xi(\psi) = \sum_{|h_1|<X} \cdots \sum_{|h_{j-1}|<X} \left| \sum_{x \in I_{j-1}(h)} e(\Delta_{j-1}(\psi(x); h)) \right|^2.
\]
As in the case \( j = 1 \), one discerns that
\[
\left| \sum_{x \in I_{j-1}(h)} e(\Delta_{j-1}(\psi(x); h)) \right|^2 = \sum_{|h_j|<X} \sum_{x \in I_{j}(h)} e(\Delta_1(\Delta_{j-1}(\psi(x); h); h_j)),
\]
where
\[
I_{j}(h, h_J) = I_{j-1}(h) \cap \{ x \in [1, X] : x + h_J \in I_{j-1}(h) \}.
\]
We therefore conclude that
\[
|F(\psi)|^{2j} \leq (2X)^{2j-1} \sum_{|h_1|<X} \cdots \sum_{|h_j|<X} \sum_{x \in I_{j}(h)} e(\Delta_{j}(\psi(x); h)).
\]
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
\[ f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k), \]
then we obtain a bound of the shape
\[ |f(\alpha)|^{2k-1} \ll X^{2k-1-k} \sum_{h_1, \ldots, h_{k-1} \in I(h)} \sum_{|h_i| < X} e\left( k! \alpha h_1 \ldots h_{k-1} \left( x + \frac{1}{2} (h_1 + \ldots + h_{k-1}) \right) \right). \]
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
\[ \tau_r(n) = \sum_{d_i \in \mathbb{N}} 1 \]
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
\[ \frac{\tau_r(n)}{n^\varepsilon} = \prod_{p \text{ prime}} \frac{\tau_r(p^h)}{p^{h\varepsilon}} \leq \prod_{p \| n, p^h} \frac{(h + 1)^{r-1}}{p^{h\varepsilon}}. \]
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 $0 \leq a_i \leq h$, 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 $h + 1 \leq 2^h$, one sees that
\[ (h + 1)^{r-1} p^{-h} \leq \left( 2^{r/\varepsilon} \right)^h, \]
a quantity that is at most 1 for $p \geq 2^{r/\varepsilon}$. Meanwhile, the function $(h + 1)^{r-1}/p^{h\varepsilon}$ is bounded above, uniformly in $p$, purely in terms of $r$ and $\varepsilon$ as $h$ increases. Thus, one has
\[ (h + 1)^{r-1} p^{-h} \leq (h + 1)^{r-1} 2^{-h} < C(r, \varepsilon), \]
say. We consequently obtain the bound
\[ \frac{\tau_r(n)}{n^\varepsilon} \leq \prod_{p < 2^{r/\varepsilon}} C(r, \varepsilon) < C(r, \varepsilon)^{2^{r/\varepsilon}} \ll_{r, \varepsilon} 1. \]
Then $\tau_r(n) \ll_{r, \varepsilon} n^{\varepsilon}$, and the proof of the lemma is complete. \hfill \qed

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 $X \geq 1$ is a real number. Then there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a,q) = 1$ and $1 \leq q \leq X$ such that $|\alpha - a/q| \leq 1/(qX)$.*

**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 $|\alpha - a/q| \leq q^{-2}$.
Applying Lemma 3.4 again, we see that there exist \( b \) so that

\[
q = \frac{b}{r, b} \quad (2 \leq j \leq N).
\]

It follows that two at least lie in some common such interval, say \( \alpha u - [\alpha u] \) and \( \alpha v - [\alpha v] \) with \( u < v \). We then have

\[
|\alpha(v - u) - ([\alpha v] - [\alpha u])| < 1/(N + 1) < 1/X,
\]

and the desired conclusion follows with \( q = r/(r, b) \) and \( a = b/(r, b) \), in which \( r = v - u \) and \( b = [\alpha v] - [\alpha u] \).

The proof of Lemma 3.4. We apply the box principle. Let \( N = |X| \), and consider the \( N \) real numbers \( \alpha r - \lfloor\alpha r\rfloor \) for \( 1 \leq r \leq N \). 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 \( b = [\alpha r] \) or \( b = [\alpha r] + 1 \). We may therefore suppose that all \( N \) of these real numbers lie in one of the \( N - 1 \) intervals

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

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 \( X \geq 1 \) and \( Y \geq 1 \). Suppose also that \( \alpha \) and \( \beta \) are real numbers, and that there exist \( a \in \mathbb{Z} \) and \( q \in \mathbb{N} \) with \( (a, q) = 1 \) and \( |\alpha - a/q| \leq q^{-2} \). Then one has

\[
\sum_{1 \leq x \leq X} \min\{Y, \|\alpha x + \beta\|^{-1}\} \ll XY (q^{-1} + Y^{-1} + X^{-1} + q(XY)^{-1}) \log(2q).
\]

**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 \( |\alpha - a/q| \leq 1/q^2 \), conditions that may always be ensured by application of Dirichlet’s theorem on Diophantine approximation. We write \( \theta = \alpha - a/q \), so that \( |\theta| \leq 1/q^2 \). Since \( \alpha \) 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 \( I \), we have
\[
\| (\alpha n_1 + \beta) - (\alpha n_2 + \beta) \| = \| \alpha (n_1 - n_2) \| \geq \frac{|a(n_1 - n_2)|}{q} - |n_1 - n_2| |\theta|.
\]
Since \( n_1 \neq n_2 \) and \( |n_1 - n_2| \leq q/2 \), we have \( q \nmid (n_1 - n_2) \). On recalling that \((a, q) = 1\), we therefore deduce that
\[
\| (\alpha n_1 + \beta) - (\alpha n_2 + \beta) \| \geq \frac{1}{q} - \frac{q/2}{q^2} = \frac{1}{2q}.
\]
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
\[
\sum_{n \in I} \min \{ Y, \| \alpha n + \beta \|^{-1} \} \ll Y + \sum_{0 < |r| \leq q/2} \left| \frac{r}{2q} \right|^{-1} \ll Y + 4 \sum_{1 \leq r \leq q/2} q/r \ll Y + q \log(2q).
\]
Observe next that there are at most \( \lceil X/(q/2) \rceil \) intervals of the shape \( I \) required to cover all of the summands \( x \) with \( 1 \leq x \leq X \). It follows that
\[
\sum_{1 \leq x \leq X} \min \{ Y, \| \alpha x + \beta \|^{-1} \} \ll (X/q + 1)(Y + q \log(2q)),
\]
and the conclusion of the lemma follows. \( \square \)

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 \( k \geq 2 \) 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 \( |\alpha_k - a/q| \leq q^{-2} \). Then
\[
\sum_{1 \leq x \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k) \ll X^{1+\varepsilon} (q^{-1} + X^{-1} + qX^{-k})^{2^{1-k}}.
\]

**Proof.** Write \( \psi(x) = \alpha_1 x + \ldots + \alpha_k x^k \) and
\[
F(\alpha) = \sum_{1 \leq x \leq X} e(\psi(x)).
\]
Note first that when \( q > X^k \), the desired estimate is trivial from the bound
\[
|F(\alpha)| \leq \sum_{1 \leq x \leq X} 1 \leq X.
\]
Thus, we may suppose without loss that \( q \leq X^k \).

Next, by applying the Weyl differencing lemma with \( j = k - 1 \) (see Lemma 3.2), we obtain the bound
\[
|F(\alpha)|^{2^{k-1}} \ll X^{2^{k-1}-k} \sum_{|h_1| < X} \ldots \sum_{|h_{k-1}| < X} \delta(h),
\]
where we write
\[ \Upsilon(h) = \sum_{x \in I_{k-1}(h)} e(\Delta_{k-1}(\psi(x); h)), \]
and \( I_{k-1}(h) \) is a suitable interval of integers contained in \([1, X]\). We note that
\[ \Delta_{k-1}(\psi(x); h) = k!h_1 \ldots h_{k-1} x \alpha_k + \gamma, \]
where \( \gamma = \gamma(\alpha; h) \) is independent of \( x \). Thus, Lemma 3.1 delivers the bound
\[ \Upsilon(h) \ll \min \{X, \|k!h_1 \ldots h_{k-1} \alpha_k\|^{-1}\}, \]
whence
\[ |F(\alpha)|^{2k-1} \ll X^{2k-1-k} \sum_{|h_1| < X} \cdots \sum_{|h_{k-1}| < X} \min \{X, \|k!h_1 \ldots h_{k-1} \alpha_k\|^{-1}\}. \]
Accounting for the summands in which \( h_1 \ldots h_{k-1} = 0 \), we are led from here to the bound
\[ |F(\alpha)|^{2k-1} \ll X^{2k-1-k} \left(X^{k-1} + \sum_{1 \leq n \leq k!X^{k-1}} \tau_k(n) \min \{X, \|n \alpha_k\|^{-1}\}\right). \]
Invoking Lemma 3.6 and recalling our assumption that \( q \ll X^k \), we therefore obtain the estimate
\[ F(\alpha) \ll X^{1-2^{1-k}} + X^{1+\varepsilon} \left(q^{1-1} + X^{-1} + X^{1-k} + qX^{-k}\right)^{2^{1-k}}, \]
and the conclusion of the lemma follows at once. \( \square \)

We pause momentarily to reflect on the consequences of Weyl’s inequality so far as the behaviour of the exponential sum
\[ f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k) \tag{3.3} \]
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 \( 1 \leq q \leq X^{k-\delta} \), \((a, q) = 1\) and
\[ |\alpha - a/q| \ll 1/(q X^{k-\delta}) \ll X^{\delta-k}. \]
If \( q \ll X^\delta \), then we would have \( \alpha \in \mathfrak{m}_\delta \). Thus, when \( \alpha \in \mathfrak{m}_\delta \), we may suppose that \( X^{\delta} < q \ll X^{k-\delta} \). 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
\[ \left| \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \right| \ll \left( \sup_{\alpha \in \mathfrak{m}_\delta} |f(\alpha)| \right)^s \int_{\mathfrak{m}_\delta} \, d\alpha \ll \left(X^{1-\delta 2^{1-k}+\varepsilon}\right)^s = o(X^{s-k}). \]
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 $\theta, X, Y, Z$ 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 $|\alpha - a/q| \leq 1/q^2$, then
\[ \Psi(\alpha) \ll X(q^{-1} + Y^{-1} + qZ^{-1})^\theta. \] (3.4)
Then, whenever $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ satisfy $(b, r) = 1$, one has
\[ \Psi(\alpha) \ll X(\lambda^{-1} + Y^{-1} + \lambda Z^{-1})^\theta, \] (3.5)
where $\lambda = r + Z|\alpha - b|$.

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 $1 \leq q \leq 2r$, $(a, q) = 1$ and $|qa - a| \leq (2r)^{-1}$. Suppose in the first instance that $a/q \neq b/r$. Then
\[ \frac{1}{qr} \leq \left| \frac{a}{q} - \frac{b}{r} \right| \leq \left| \frac{a}{q} - \frac{b}{r} \right| + \left| \frac{a}{q} - \frac{b}{r} \right| \leq \left| \frac{a}{q} - \frac{b}{r} \right| + \frac{1}{2qr}, \]
whence $q^{-1} \leq 2|\alpha - b|$. It therefore follows from (3.4) that
\[ \Psi(\alpha) \ll X \left( |\alpha - b| + Y^{-1} + rZ^{-1} \right)^\theta \ll X \left( Y^{-1} + \lambda Z^{-1} \right)^\theta, \]
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 $|\alpha - b| \leq (2r)^{-1}$. 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 $0 < |\alpha - b/r| \leq r^{-2}$. Applying Dirichlet’s approximation theorem here, we obtain $a' \in \mathbb{Z}$ and $q' \in \mathbb{N}$ with $1 \leq q' \leq 2|\alpha - b|^{-1}$, $(a', q') = 1$ and $|q'a - a'| \leq \frac{1}{2}|\alpha - b|$. Now, if $a'/q' = b/r$, we again have $(a', q') = (b, r) = 1$, and so $q' = r$ and $a' = b$. Thus
\[ 0 < |\alpha - b| \leq \frac{1}{2}|\alpha - b|, \]
yielding a contradiction. Then $a'/q' \neq b/r$, whence
\[
\frac{1}{q'r} \leq \left| \frac{a'}{q'} - \frac{b}{r} \right| \leq \left| \alpha - \frac{a'}{q'} \right| + \left| \alpha - \frac{b}{r} \right| \\
\leq \left| \alpha - \frac{b}{r} \right| + (2q')^{-1}|\alpha - b| \\
\leq \left| \alpha - \frac{b}{r} \right| + \frac{1}{2q'r}.
\]
We therefore infer that $|\alpha - b| \geq (2q')^{-1}$, and by applying (3.4) with $a$ and $q$ replaced by $a'$ and $q'$, we obtain the bound
\[ \Psi(\alpha) \ll X \left( |\alpha - b| + Y^{-1} + (Z|\alpha - b|)^{-1} \right)^\theta. \]
As an alternative bound, since $|\alpha - b/r| \leq r^{-2}$, 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 $\lambda = r + Z|\alpha - b|$. This completes the proof of the lemma. \qed
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. \hfill \Box

Estimates of the shape presented in Corollary 3.9 provide control of the exponential sums in question for $\alpha$ 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 $k \geq 2$ and $s \geq k2^{k-1}$. Then one has

$$\int_{0}^{1} |f(\alpha)|^s \, d\alpha \ll X^{s-k+\varepsilon}.$$  

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

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

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

$$\int_{0}^{1} |f(\alpha)|^s \, d\alpha \ll (X^{1-2^{1-k}+\varepsilon})^s + I,$$  

where

$$I = \sum_{1 \leq q \leq X} \sum_{(a, q) = 1} \frac{1}{q} \sum_{a = 1}^{q} \int_{a/q - q^{-1}X^{1-k}}^{a/q + q^{-1}X^{1-k}} \frac{X^{1+\varepsilon}}{(q + X^k|q\alpha - a|)^{2^{1-k}}} \, d\alpha.$$  

We have

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

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

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

$$\left| \int_{m_{\delta}} f(\alpha)^s e(-n\alpha) \, d\alpha \right| \leq \left( \sup_{\alpha \in m_{\delta}} |f(\alpha)| \right)^{s-k2^{k-1}} \int_{0}^{1} |f(\alpha)|^{k2^{k-1}} \, d\alpha \ll \left( X^{1-\delta2^{1-k}+\varepsilon} \right)^{s-k2^{k-1}} X^{k2^{k-1} - k+\varepsilon} = o(X^{s-k}).$$
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

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

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

$$c(\alpha) = \inf \{ q + X^k|q\alpha - a| : q \in \mathbb{N} \text{ and } a \in \mathbb{Z} \} = \inf_{q \in \mathbb{N}} (q + X^k\|q\alpha\|).$$

Equipped with this notion, we see that the real numbers $\alpha \in [0, 1)$ having small complexity $c(\alpha)$ 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 $\alpha$, there exists $q \in \mathbb{N}$ with $1 \leq q \leq X^{k/2}$ for which $\|q\alpha\| \leq X^{-k/2}$. Hence all real numbers $\alpha$ satisfy the complexity constraint $c(\alpha) \leq 2X^{k/2}$. The enthusiastic reader may verify that there is a subset of $[0, 1)$ of asymptotically full measure (as $X \to \infty$) for which $c(\alpha) \gg X^{k/2-\varepsilon}$. Classical technology in the circle method permits useful asymptotic analyses of exponential sums of degree $k$ when the arguments $\alpha$ 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 $c(\alpha) \leq X^2$. 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 $\alpha$ having complexity in the range $X^2 \leq c(\alpha) \leq 2X^{k/2}$ 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

$$\sup_{\alpha \in m_1} |f_k(\alpha)| \ll X^{1-2^{-k}} (\log X)^{A_k},$$

where $A_k$ is a suitable positive number. This estimate, in which Weyl’s factor of $X^\varepsilon$ is replaced by $(\log X)^{A_k}$, is due to Vaughan [23, 24]. When $k \geq 6$, one has the superior estimate

$$\sup_{\alpha \in m_1} |f_k(\alpha)| \ll X^{1-1/(k(k-1)) + \varepsilon},$$

which is a consequence of recent progress on Vinogradov’s mean value theorem (see [3] and [36]). The latter is a topic to which we shall return later in the course.

4. HUA’s LEwMA

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

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

Hardy and Littlewood [12] obtained this estimate under the slightly less stringent hypothesis $s \geq (k-2)2^{k-1} + 5$, and later Hua [15] reduced this bound on $s$ to $s \geq 2^k + 1$. 
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

\[
f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k).
\]

**Lemma 4.1** (Hua’s lemma). Suppose that \(k \geq 2\) and \(1 \leq j \leq k\). Then one has

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

**Proof.** We proceed by induction on \(j\). When \(j = 1\), it follows via orthogonality that

\[
\int_0^1 |f(\alpha)|^2 \, d\alpha = \int_0^1 f(\alpha) f(-\alpha) \, d\alpha = \text{card}\{1 \leq x, y \leq X : x^k = y^k\} = [X].
\]

This confirms the claimed conclusion in the base case \(j = 1\).

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

\[
|f(\alpha)|^{2j} \leq (2X)^{2j-J-1} \sum_{|h_1| < X} \cdots \sum_{|h_J| < X} \sum_{x \in I_J(h)} e(\alpha \Delta_J(x^k; h)),
\]

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

\[
\int_0^1 |f(\alpha)|^{2j+1} \, d\alpha = \int_0^1 f(\alpha)^{2j-1} f(-\alpha)^{2j-1} |f(\alpha)|^{2j} \, d\alpha \\
\leq (2X)^{2j-J-1} T,
\]

where

\[
T = \sum_{|h_1| < X} \cdots \sum_{|h_J| < X} \sum_{x \in I_J(h)} \int_0^1 f(\alpha)^{2j-1} f(-\alpha)^{2j-1} e(\alpha \Delta_J(x^k; h)) \, d\alpha.
\]

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

\[
\sum_{i=1}^{2j-1} (u_i^k - v_i^k) = \Delta_J(x^k; h),
\]

with \(1 \leq u_i, v_i \leq X\) (\(1 \leq i \leq 2^{J-1}\)), \(1 \leq x \leq X\) and \(|h_j| < X\) (\(1 \leq j \leq J\)). Notice here that we have weakened the condition \(x \in I_J(h)\) to \(1 \leq x \leq X\), 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

\[
\sum_{i=1}^{2j-1} (u_i^k - v_i^k) = 0.
\]
In this situation, one has $\Delta_j(x^k; h) = 0$. By orthogonality, the number of choices of $u$ and $v$ here is

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

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

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

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

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

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 $N = N(u, v)$ with $|N| \ll X^k$. For each such choice of $u$ and $v$, we have

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

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

$$\ll \sum_{1 \leq u_i, v_i \leq X} \tau_{J+1}(N(u, v)) \ll X^\epsilon \sum_{1 \leq u_i, v_i \leq X} 1 \ll X^{2^J+\epsilon}.$$

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

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

This confirms the inductive hypothesis with $j = J + 1$, and the conclusion of the lemma follows. \(\square\)

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 $s \geq 2^k + 1$, one has

$$\int_{m_{\delta}} f(\alpha)^s e(-n\alpha) \, d\alpha \ll X^{s-k-\delta 2^{-k}}.$$
Proof. By making use of Weyl’s inequality in combination with Hua’s lemma, one obtains
\[
\left| \int_{m_{s}} f(\alpha)^s e(-n\alpha) \, d\alpha \right| \leq \left( \sup_{\alpha \in m_{s}} |f(\alpha)| \right)^{s-2k} \int_{0}^{1} |f(\alpha)|^{2k} \, d\alpha \\
\ll (X^{1-\delta_{2}^{1-k}+\varepsilon})^{s-2k} X^{2k-k+\varepsilon} \\
\ll X^{s-k-(s-2k)\delta_{2}^{1-k}+s\varepsilon}.
\]
The conclusion of the corollary follows on recalling that \( s \geq 2k + 1 \). \qed

We finish this section by noting the sharpest estimates currently available for mean values of the exponential sum
\[
f_k(\alpha; X) = \sum_{1 \leq x \leq X} e(\alpha x^k).
\]
Thus, one has
\[
\int_{0}^{1} |f_2(\alpha; X)|^4 \, d\alpha \asymp X^2 \log X,
\]
and, when \( k \geq 3 \), Vaughan [23, 24] has established bounds of the shape
\[
\int_{m_{1}} |f_k(\alpha; X)|^{2k} \, d\alpha \ll X^{2k-k}(\log X)^{-A_k},
\]
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 \( k \geq 4 \) 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 [34, 36] may be applied in combination with earlier work of Wooley [33] to deliver the bound
\[
\int_{0}^{1} |f_k(\alpha)|^s \, d\alpha \ll X^{s-k+\varepsilon},
\]
whenever \( s \geq s_0(k) \), where \( s_0(4) = 14 \), \( s_0(5) < 23 \),..., and in general
\[
s_0(k) \leq k^2 - k + 2\lfloor \sqrt{2k+2} \rfloor - 1.
\]
Details may be found in Wooley [36, 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
\[
f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k),
\]
we find via orthogonality that

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

We divide the interval of integration according to a Hardy-Littlewood dissection with major arcs \( M_\delta \) equal to the union of the intervals

\[ M_\delta(q,a) = \{ \alpha \in [0,1) : |\alpha - a/q| \leq X^{\delta - k} \}, \]

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

\[ \frac{|a_1 - a_2|}{q_1 - q_2} \leq |\alpha - a_1/q_1| + |\alpha - a_2/q_2| \leq 2X^{\delta - k}. \]

Thus, one finds that

\[ \frac{1}{q_1 q_2} \leq \left| \frac{a_1 q_2 - a_2 q_1}{q_1 q_2} \right| \leq 2X^{\delta - k}, \]

whence \( 1 \leq 2q_1 q_2 X^{\delta - k} \leq 2X^{3\delta - k} \). This is plainly impossible when \( \delta < 1/3 \) and \( X \) is large, as we henceforth assume. Subject to this latter condition the major arcs \( M_\delta \) defined in this way are a disjoint union of the arcs \( M_\delta(q,a) \).

We have shown thus far that when \( s \geq 2k + 1 \), one has

\[ \int_{m_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \ll X^{s - k - 2 - k} = o(n^{s/k-1}). \]

Our goal is now to establish the lower bound

\[ \int_{M_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \gg n^{s/k-1}. \]

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

\[ R_{s,k}(n) = \int_{M_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha + \int_{m_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \gg n^{s/k-1} + o(n^{s/k-1}) \to \infty, \]

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

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

\[ \sum_{1 \leq x \leq X} e(\alpha x^k) = \sum_{r=1}^q \sum_{(1-r)/q \leq y \leq (X-r)/q} e \left( (\beta + a/q)(yq + r)^k \right) \]

\[ = \sum_{r=1}^q e(ar^k/q) \sum_{(1-r)/q \leq y \leq (X-r)/q} e \left( \beta yq + r^k \right). \] *(5.1)*
Since \( \beta \) 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
\[
e(F(z)) = \int_{-1/2}^{1/2} e(F(z)) \, d\eta.
\]
Hence
\[
\left| e(F(z)) - \int_{-1/2}^{1/2} e(F(z + \eta)) \, d\eta \right| \leq \sup_{|\eta| \leq 1/2} |e(F(z + \eta)) - e(F(z))| \ll \sup_{|\eta| \leq 1/2} |F'(z + \eta)|.
\]
Using this approximation, we may infer that
\[
\sum_{r=1}^{q} e(\beta(yq + r)^k) - \int_{-r/q}^{(X-r)/q} e(\beta(zq + r)^k) \, dz
\ll 1 + (X/q) \sup_{0 \leq z \leq X/q} |k\beta q(zq + r)^{k-1}|
\ll 1 + X^k|\beta|.
\]
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
\[
f(\alpha) = \sum_{r=1}^{q} e(ar^k/q) \left( \int_{-r/q}^{(X-r)/q} e(\beta(zq + r)^k) \, dz + O(1 + X^k|\beta|) \right),
\]
so that
\[
f(\alpha) - \sum_{r=1}^{q} e(ar^k/q) \int_{-r/q}^{(X-r)/q} e(\beta(zq + r)^k) \, dz \ll q + X^k|q\beta|.
\] (5.2)
By the change of variable \( \gamma = zq + r \), moreover, we have
\[
\int_{-r/q}^{(X-r)/q} e(\beta(zq + r)^k) \, dz = q^{-1} \int_{0}^{X} e(\beta\gamma^k) \, d\gamma.
\] (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
\[
v(\beta) = \int_{0}^{X} e(\beta\gamma^k) \, d\gamma.
\]

**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 \( \alpha \in \mathcal{M}_\delta(q,a) \subseteq \mathcal{M}_\delta \), one has

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

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

\[
 |q\alpha - a| = q|\alpha - a/q| \leq X^\delta \cdot X^{\delta-k},
\]

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 \( \delta \) 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

\[
 \mathcal{M}_\delta = \bigcup_{0 \leq a, q \leq X^\delta} \mathcal{M}_\delta(q,a),
\]

then

\[
 \int_{\mathcal{M}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha = \sum_{1 \leq q \leq X^\delta} \sum_{a=1}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} f(\beta + a/q)^s e(-n(\beta + a/q)) \, d\beta.
\]

Focusing our attention on a real number \( \alpha \) lying in \( \mathcal{M}_\delta(q,a) \subseteq \mathcal{M}_\delta \), 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

\[
 f(\alpha)^s - f^*(\alpha)^s = (f(\alpha) - f^*(\alpha))(f(\alpha)^{s-1} + \ldots + f^*(\alpha)^{s-1}) \ll X^{s-1}|E(\alpha)| \ll X^{s-1+2\delta},
\]

we obtain the asymptotic relation

\[
 \int_{\mathcal{M}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha = \sum_{1 \leq q \leq X^\delta} \sum_{a=1}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} \left( q^{-1}S(q,a)v(\beta) \right)^s e(-n(\beta + a/q)) \, d\beta \\
+ \sum_{1 \leq q \leq X^\delta} \sum_{a=1}^{q} \int_{-X^{\delta-k}}^{X^{\delta-k}} X^{s-1+2\delta} \, d\alpha. \tag{5.4}
\]

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

\[
 \ll X^{s-1+2\delta} \sum_{1 \leq q \leq X^\delta} q \cdot X^{\delta-k} \ll X^{s-k-1+3\delta} \cdot X^{2\delta} \ll X^{s-k+(5\delta-1)} \tag{5.5}
\]

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

\[
 \sum_{1 \leq q \leq X^\delta} \sum_{a=1}^{q} (q^{-1}S(q,a))^s e(-na/q) \int_{-X^{\delta-k}}^{X^{\delta-k}} v(\beta)^s e(-\beta n) \, d\beta. \tag{5.6}
\]
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

\[
\mathfrak{S}_{s,k}(n; Q) = \sum_{1 \leq q \leq Q} \sum_{a=1 \atop (a,q)=1}^q (q^{-1} S(q,a))^s e(-na/q)
\]

and the truncated singular integral

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

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

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

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

**Corollary 5.4.** When \( s \geq 2k + 1 \) 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

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

and the conclusion follows from Corollary 4.2 and Lemma 5.3. \( \square \)

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

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

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

\[
v(\beta) = \int_{0}^{X} e(\beta \gamma^k) \, d\gamma.
\]
The estimate $|v(\beta)| \leq X$ 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

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

whence

$$|v(\beta)| \leq k^{-1} \beta^{-1/k} \left| \int_0^{\beta X^k} u^{-1+1/k} e(u) \, du \right|.$$  \hfill (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

$$\left| \int_0^{\beta X^k} u^{-1+1/k} e(u) \, du \right| \leq \sup_{Y \geq 0} \left| \int_0^Y u^{-1+1/k} e(u) \, du \right| < \infty.$$

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

$$\left| \int_0^Y u^{-1+1/k} e(u) \, du \right| \leq \int_0^Y u^{-1+1/k} \, du \ll 1.$$

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 $|v(\beta)| \leq X$, applied in circumstances wherein $|\beta| \leq X^{-k}$.

\begin{corollary}
Suppose that $s \geq k + 1$. Then the singular series $J_{s,k}(n)$ converges absolutely, and moreover,

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

\end{corollary}

\begin{proof}  
By applying Lemma 6.1, one sees that

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

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

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

This completes the proof of the lemma.  \hfill \Box

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

\begin{lemma}
When $s \geq k + 1$, 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 $\Gamma$-function defined by

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

\end{lemma}
Proof. We begin by observing that
\[ J_{s,k}(n) = \lim_{B \to \infty} \int_{-B}^{B} v(\beta)^{s} e(-\beta n) d\beta \]
\[ = \lim_{B \to \infty} \int_{-B}^{B} \int_{[0,X]^s} e(\beta(\gamma_1^k + \ldots + \gamma_s^k - n)) d\gamma d\beta \]
\[ = \lim_{B \to \infty} \int_{[0,X]^s} \int_{-B}^{B} e(\beta(\gamma_1^k + \ldots + \gamma_s^k - n)) d\beta d\gamma. \]

We make use of the observation that when \( \phi \neq 0 \), one has
\[ \int_{-B}^{B} e(\beta \phi) d\beta = \frac{\sin(2\pi B\phi)}{\pi \phi}. \]

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

We aim to transform this into a linear problem. To this end, we substitute \( u_i = \gamma_i^k \) \((1 \leq i \leq s)\), and recall that \( n = X^k \). Thus, we obtain
\[ J_{s,k}(n) = k^{-s} \lim_{B \to \infty} I(B), \]

where we write
\[ I(B) = \int_{[0,n]^s} \frac{\sin(2\pi B(u_1 + \ldots + u_s - n))}{\pi(u_1 + \ldots + u_s - n)} (u_1 \ldots u_s)^{-1+1/k} du. \]

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
\[ I(B) = \int_{0}^{sn} \Psi(v) \frac{\sin(2\pi B(v - n))}{\pi(v - n)} dv, \]

in which
\[ \Psi(v) = \int_{B(v)} (u_1 \ldots u_{s-1})^{-1+1/k} (v - u_1 - \ldots - u_{s-1})^{-1+1/k} du_1 \ldots du_{s-1}, \]

and
\[ B(v) = \{(u_1, \ldots, u_{s-1}) \in [0,n]^{s-1} : 0 \leq v - u_1 - \ldots - u_{s-1} \leq n\}. \]

Notice that the condition on \( u_1, \ldots, u_{s-1} \) in the definition of \( B(v) \) may be rephrased as \( v - n \leq u_1 + \ldots + u_{s-1} \leq v \). Since \( \Psi(v) \) is a function of bounded variation, it follows from Fourier’s Integral Theorem that since \( n \in (0, sn) \), one has
\[ \lim_{B \to \infty} I(B) = \Psi(n) = \int_{B(n)} (u_1 \ldots u_{s-1})^{-1+1/k} (n - u_1 - \ldots - u_{s-1})^{-1+1/k} du. \]

Note that
\[ B(n) = \{(u_1, \ldots, u_{s-1}) \in [0,n]^{s-1} : 0 \leq u_1 + \ldots + u_{s-1} \leq n\}. \]
Thus
\[ J_{s,k}(n) = k^{-s} \Psi(n) = k^{-s} \int_0^n \ldots \int_0^n (u_1 \ldots u_{s-1})^{-1+1/k}(n-u_1-\ldots-u_{s-1})^{-1+1/k} \, du. \]

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
\[
J_{2,k}(n) = k^{-2} \int_0^n u_1^{-1+1/k}(n-u_1)^{-1+1/k} \, du_1 \\
= k^{-2} n^{-1+2/k} \int_0^1 v^{-1+1/k}(1-v)^{-1+1/k} \, dv.
\]

Thus, on recalling the classical Beta function, we obtain the formula
\[
J_{2,k}(n) = \frac{k^{-2} n^{-1+2/k} B(1/k, 1/k)}{k^{-2} n^{-1+2/k} B(1/k, 1/k)} = \frac{\Gamma(1 + 1/k)^2}{\Gamma(2/k)} n^{-1+2/k}.
\]

Thus, the inductive hypothesis holds for \( s = 2 \).

Suppose now that the inductive hypothesis holds for \( s = t \). Then we have
\[
J_{t+1,k}(n) = k^{-1} \int_0^n u_t^{-1+1/k} J_{t,k}(n-u_t) \, du_t \\
= k^{-1} \frac{\Gamma(1 + 1/k)^t}{\Gamma(t/k)} \int_0^n u_t^{-1+1/k}(n-u_t)^{-1+t/k} \, du_t.
\]

Recalling once again the classical Beta function, we see that
\[
J_{t+1,k}(n) = k^{-1} \frac{\Gamma(1 + 1/k)^t}{\Gamma(t/k)} n^{-1+(t+1)/k} B(t/k, t/k) \\
= k^{-1} \frac{\Gamma(1 + 1/k)^t}{\Gamma(t/k)} n^{-1+(t+1)/k} \frac{\Gamma(1/k) \Gamma(t/k)}{\Gamma((t+1)/k)} \\
= \frac{\Gamma(1 + 1/k)^t}{\Gamma((t+1)/k)} n^{-1+(t+1)/k}.
\]

This confirms the inductive hypothesis with \( t \) replaced by \( t+1 \). We have therefore shown that whenever \( s \geq k+1 \), one has
\[ J_{s,k}(n) = \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} n^{s/k-1}. \]

\( \square \)

**Corollary 6.4.** Suppose that \( s \geq k + 1 \). 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 \( X = n^{1/k} \). \( \square \)
7. The singular series, I: convergence of sums and products

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

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

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

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

□

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

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

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

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

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

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

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

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

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

uniformly in $n$.

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

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

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. Making use of the bijection between the residue classes \(m\) modulo \(qr\), and the ordered pairs of residue classes \((t, u)\) with \(t\) modulo \(q\) and \(u\) modulo \(r\), suggested by the Chinese Remainder Theorem via the relation \(m \equiv (ar + bq)(tr + uq)^k \mod qr\), we see that

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

By the change of variable \(tr \mapsto t' \mod q\) and \(uq \mapsto u' \mod r\), bijective owing to the coprimality of \(q\) and \(r\), we obtain the relation

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

This completes the proof of the lemma. \(\square\)

Now define the quantity

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

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 \(a \equiv br + cq \mod qr\). Thus, we obtain

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

By applying Lemma 7.3, we infer that

\[ A(qr, n) = \sum_{\substack{b=1 \atop (b, q)=1}}^{q} \sum_{\substack{c=1 \atop (c, r)=1}}^{r} \left(q^{-1}S(q, b)\right)^s \left(r^{-1}S(r, c)\right)^s e(-bn/q)e(-cn/r) = A(q, n)A(r, n). \]
Since $A(1, n) = 1$, this confirms the multiplicative property for $A(q, n)$ and completes the proof of the lemma.

Observe that

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

The multiplicativity of $A(q, n)$ therefore suggests that $\mathcal{G}_{s,k}(n)$ 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 $s \geq 2^k + 1$. 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

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

converges absolutely;

(iii) one has $\mathcal{G}_{s,k}(n) = \prod_p \sigma(p)$;

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

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

**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 $s \geq 2^k + 1$, one finds that

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

Hence

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

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 $|\sigma(p) - 1| \ll Bp^{-1 - 2^{-k}}$. Hence, whenever $p$ is sufficiently large in terms of $p$, one sees that

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

whence

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

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
\[ S_{s, k}(n) = \sum_{q=1}^{\infty} A(q, n) = \sum_{q=1}^{\infty} \prod_{p \parallel q} A(p^h, n). \]
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
\[ S_{s, k}(n) = \prod_p \sum_{h=0}^{\infty} A(p^h, n) = \prod_p \sigma(p). \]
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
\[ 1 - p^{-1 - 2^{-k}} \leq \sigma(p) \leq 1 + p^{-1 - 2^{-k}}. \]
Hence, provided that $C = C(k)$ is sufficiently large, one finds that
\[ \left| \prod_{p \geq C(k)} \sigma(p) - 1 \right| \leq \sum_{n \geq C(k)} n^{-1 - 2^{-k}} \ll C(k)^{-2^{-k}}. \]
Then, again if $C(k)$ is chosen sufficiently large in terms of $k$, we may infer that
\[ \left| \prod_{p \geq C(k)} \sigma(p) - 1 \right| < 1/2, \]
and we conclude that
\[ 1/2 < \prod_{p \geq C(k)} \sigma(p) < 3/2. \]
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 $S_{s, k}(n) \gg 1$ (uniformly in $n$). At this stage, we can at least show that the latter is the case provided that $\sigma(p) > 0$ for $p \leq C(k)$ 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
\[ M_n(q) = \text{card}\{ m \in (\mathbb{Z}/q\mathbb{Z})^s : m_1^k + \ldots + m_s^k = n \}. \]

Lemma 8.1. For each natural number $q$, one has
\[ \sum_{d \mid q} A(d, n) = q^{1-s} M_n(q). \]

Proof. We make use of the orthogonality relation
\[ q^{-1} \sum_{r=1}^{q} e(hr/q) = \begin{cases} 1, & \text{when } q \mid h, \\ 0, & \text{when } q \nmid h. \end{cases} \]
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
\[ M_n(q) = q^{-1} \sum_{d|q} \left( \sum_{a=1}^{q/d} (q/d)^s \sum_{m_1=1}^{d} \cdots \sum_{m_s=1}^{d} e \left( a(m_1^k + \ldots + m_s^k - n)/d \right) \right) \]
\[ = q^{-1} \sum_{d|q} q^s \sum_{a=1}^{d} (d^{-1}S(d,a))^s e(-na/d) \]
\[ = q^{s-1} \sum_{d|q} A(d, n). \]

Hence
\[ \sum_{d|q} A(d, n) = q^{1-s}M_n(q), \]
and the proof of the lemma is complete. \( \square \)

**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 \( \sigma(p). \) \( \square \)

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 \( S_{s,k}(n) \gg 1 \) follows from Theorem 7.5(iv).

We begin by recalling some elementary number theory. We define \( \tau = \tau(p, k) \) via the relation \( p^r \parallel k \), 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 \( x^k \equiv a \pmod{p^\tau} \) is soluble. Then whenever \( h \geq \gamma \), 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 \( h \geq \gamma \), and let \( u \) be
the integer with \( 1 \leq u \leq \phi(p^h) \) 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 \( p^\gamma \), it follows that \((k,\phi(p^\gamma)) \) divides \( u \). Notice that the definition of \( \tau \) and \( \gamma \) ensures that
\[
(k, \phi(p^\gamma)) = (k, p^\gamma(p - 1)) = (k, \phi(p^h)).
\]
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
\[
a \equiv g^u = (g^r)^{(k,\phi(p^h))} \equiv (g^{rm})^{(k,\phi(p^h))} \equiv (g^{rm})^k \pmod{p^h}.
\]
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 \( 1 \leq u \leq 2^{h-2} \) for which
\[
a \equiv (-1)^v 5^u \pmod{2^h}.
\]
We note that \( 5 \) has order \( 2^{\tau - 2} \) modulo \( 2^r \) for \( r \geq 2 \). Then we may proceed as in the case that \( p \) is odd. We suppose that \( a \equiv x^k \pmod{2^\tau} \). Hence \((-1)^v 5^u \equiv x^k \pmod{2^\tau} \). We may find integers \( w \) and \( z \) with \( x \equiv (-1)^w 5^z \pmod{2^\tau} \). Since \( 2^\tau \mid k \), it follows that \( x^k \equiv (5^z)^k \equiv 1 \pmod{2^\tau} \), and thus \((-1)^v 5^u \equiv 1 \pmod{2^\tau} \). We deduce that \( v = 0 \) and \( 2^{\tau - 2} \mid u \), say \( u = 2^r u \) for some integer \( r \). Writing \( k = l 2^\tau \), 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
\[
a \equiv (-1)^v 5^u = (5^r)^{2^r} \equiv (5^{rm})^{2^r} = (5^{rm})^k \pmod{2^h}.
\]
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 \( 1 \leq x_i \leq q \) \((1 \leq i \leq s)\).

**Lemma 8.4.** Suppose that \( M_n^*(p^\gamma) \geq 1 \). Then whenever \( h \geq \gamma \), one has
\[
M_n(p^h) \geq (p^{h-\gamma})^{s-1}.
\]

**Proof.** Suppose, as we may, that \( y_1, \ldots, y_s \) are integers with \( 1 \leq y_i \leq p^\gamma \) \((1 \leq i \leq s)\) 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 $1 \leq x_i \leq p^h$ and $x_i \equiv y_i \pmod{p^\gamma}$, for $2 \leq i \leq s$. 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 $p^\gamma$, 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 $1 \leq x_1 \leq p^h$ such that

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

In particular, we may conclude that $M_n(p^h) \geq (p^{h+\gamma})^{s-1}$. \hfill $\square$

This lemma reduces the problem of showing that $\mathfrak{G}_{s,k}(n) \gg 1$ to a problem concerning the solubility of congruences modulo $p^\gamma$, 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 [5] and [6]). In this context, when $A$ and $B$ are sets of integers, we write

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

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

$$\text{card}(A + B) \geq \min\{q, r + s - 1\}.$$ 

**Proof.** When $r + s - 1 > q$, then we may delete $s - (q - r + 1)$ elements from $B \setminus \{0\}$ to obtain a new set $B_1^*$. The set $B_1 = B_1^* \cup \{0\}$ is a subset of $B$ having $s_1 = q - r + 1$ elements, but now with $r + s_1 - 1 \leq q$. So the general case will follow from that subject to the hypothesis $r + s - 1 \leq q$. We may assume, moreover, that $r < q$, for if $r = q$ we have $A + B = \{0, 1, \ldots, q - 1\}$, and we are done.

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

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

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

We define

$$C = \{b \in B : c + b \notin A\}, \quad A_1 = A \cup \{c\} + C \quad \text{and} \quad B_1 = B \setminus C.$$ 

Notice here that $c + 0 \in A$, and so $C \subseteq B$. Then $1 \leq \text{card}(B_1) < s$, and

$$\text{card}(A_1) + \text{card}(B_1) = (\text{card}(A) + \text{card}(C)) + (\text{card}(B) - \text{card}(C)) = r + s.$$ 

By the inductive hypothesis, one has

$$\text{card}(A_1 + B_1) \geq \min\{q, r + s - 1\}.$$
Also,
\[ \mathcal{A}_1 + \mathcal{B}_1 = (\mathcal{A} + \mathcal{B}_1) \cup ((\{c\} + \mathcal{C}) + \mathcal{B}_1) = (\mathcal{A} + \mathcal{B}_1) \cup ((\{c\} + \mathcal{B}_1) + \mathcal{C}). \]

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
\[ \text{card}(\mathcal{A} + \mathcal{B}) \geq \text{card}(\mathcal{A}_1 + \mathcal{B}_1) \geq \min\{q, r + s - 1\}. \]

This establishes the inductive hypothesis when \( \text{card}(\mathcal{B}) = s \), 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 \( s \geq s_0(k) \), where
\[
\begin{align*}
    s_0(k) &= \begin{cases} 
        2^{\tau+2}, & \text{when } \gamma = \tau + 2 \text{ and } k > 2, \\
        5, & \text{when } p = k = 2, \\
        \frac{p}{p-1} (k, p^\gamma(p-1)), & \text{when } \gamma = \tau + 1.
    \end{cases}
\end{align*}
\]

Then \( M_n^*(p\gamma) \geq 1 \).

**Proof.** Suppose first that \( \gamma = \tau + 1 \). We apply Lemma 8.5 with \( \mathcal{A} = \{x^k \mod p^\gamma : 1 \leq x \leq p^\gamma \text{ and } (x, p) = 1\} \) and \( \mathcal{B} = \mathcal{A} \cup \{0\} \).

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

Write
\[ t = \left\lceil \frac{p}{p-1} (k, p^\gamma(p-1)) \right\rceil. \]

Then we find by induction on \( t \) that
\[ \text{card} (\mathcal{A} + (t - 1)\mathcal{B}) \geq \min \left\{ p^\gamma, \frac{t\phi(p^\gamma)}{(k, \phi(p^\gamma))} \right\} = p^\gamma. \]

Thus \( \mathcal{A} + (t - 1)\mathcal{B} \) contains all residues modulo \( p^\gamma \), and we are forced to conclude that the congruence
\[ x_1^k + \cdots + x_t^k \equiv n \mod p^\gamma \]
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
\[ x_1^k + \cdots + x_t^k \equiv n \mod 2^\gamma \]
by putting \( x_i = 1 \) for \( 1 \leq i \leq r \) and \( x_i = 0 \) for \( r < i \leq t \), where \( r \) is the integer with \( 1 \leq r \leq 2^\gamma \) for which one has \( n \equiv r \mod 2^\gamma \).

Finally, when \( p = 2 \) and \( k = 2 \), one can solve the congruence
\[ x_1^2 + \cdots + x_5^2 \equiv n \mod 8 \]
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 \( M_n^*(p^\gamma) \geq 1 \), and thus the proof of the lemma is complete. \[ \square \]
We are now equipped to wrap up the discussion of the singular series. Lemma 8.6 shows that $M_n^*(p^h) \geq 1$ provided only that when $k \neq 2$ one has $s \geq 4k$, and that when $k = 2$ one has $s \geq 5$. Indeed, when $k$ is not a power of 2, then the lower bound $s \geq 3^2k$ suffices for this conclusion. We therefore deduce from Lemma 8.4 that when $s \geq 2^k + 1$, one has

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

Corollary 8.2 thence confirms that

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

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

$$\mathcal{G}_{s,k}(n) = \left(\prod_{p < C} \sigma(p)\right)\left(\prod_{p \geq C} \sigma(p)\right) > \frac{1}{2} \prod_{p < C} \sigma(p) > 0.$$

We summarise this discussion in the form of a corollary.

**Corollary 8.7.** When $s \geq 2^k + 1$, one has $1 \ll_{s,k} \mathcal{G}_{s,k}(n) \ll_{s,k} 1$.

*Proof.* One has only to combine the conclusion of the above discussion with that of Lemma 7.2. $\square$

## 9. The asymptotic formula in Waring’s problem

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

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

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

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

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

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

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

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 $k \geq 3$, one has

$$\int_0^1 |f(\alpha)|^{2s} d\alpha \ll X^s + X^{2s-k}.$$
The validity of this conjecture would imply the anticipated asymptotic formula for \( s \geq 2k + 1 \), as a consequence of which one would have

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

and

\[
G(k) \leq 2k + 1 \quad \text{when} \quad k \neq 2^r.
\]

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 \( \tilde{G}(k) \) the least integer \( t \) with the property that whenever \( s \geq t \), 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 \( \tilde{G}(k) \leq 2^k + 1 \). In the current state of knowledge, one has the bound

\[
\tilde{G}(k) \leq 2^k \quad (k \geq 3),
\]

as a consequence of work of Vaughan from 1986 (see [23, 24]). 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 \( \tilde{G}(k) \leq (C + o(1))k^2 \log k \) 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, 35, 36]), one improves on this earlier work of Vaughan for all \( k \geq 4 \). Thus, via [33], one obtains

\[
\tilde{G}(4) \leq 15, \quad \tilde{G}(5) \leq 23, \ldots, \quad \tilde{G}(k) \leq k^2 - k + 2[\sqrt{2k+2}] - 1.
\]

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 [36, 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 \( R_{s,k}(n) \geq 1 \) 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 \( k \in \mathbb{N} \), \( s \geq 2^k + 1 \) and \( c_1, \ldots, c_s \) are fixed non-zero integers, we consider the diagonal form

\[
\phi(x) = c_1 x_1^k + \ldots + c_s x_s^k,
\]

and the solubility of the equation \( \phi(x) = 0 \). Here, since all such equations possess the trivial solution \( x = 0 \), we shall be interested in solutions with \( x \in \mathbb{Z}^s \setminus \{0\} \). 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 \( \phi(x) = 0 \).

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

(i) there exists \( z^{(\infty)} \in \mathbb{R}^s \setminus \{0\} \) with the property that \( \phi(z^{(\infty)}) = 0 \);
(ii) for each prime \( p \), there exists \( z^{(p)} \in \mathbb{Q}_p^s \setminus \{0\} \) with the property that \( \phi(z^{(p)}) = 0 \).
subject to the additional condition that \((y_j, p) = 1\) for some index \(j\) with \(1 \leq j \leq s\).

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

Notice that solubility over \(\mathbb{R}\) here is easy to characterise. This is assured when \(s \geq 2\) 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 \(s \geq k^2 + 1\) (see [8]). As an alternative to this result, one could apply the Cauchy-Davenport Lemma to establish a similar conclusion, though subject to the condition \(s \geq 4k^2\).

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

\[
F(x_1, \ldots, x_s) = \sum_{i_1, \ldots, i_s \geq 0 \atop i_1 + \ldots + i_s = k} c_{i_1, \ldots, i_s} x_1^{i_1} \cdots x_s^{i_s},
\]

with \(c_1 \in \mathbb{Z}\) fixed, and not all 0. Then it is known that when \(k \geq 2\), the equation \(F(x) = 0\) has a non-trivial solution over \(\mathbb{Q}_p\) whenever \(s \geq k^2\) (see [32]). For \(k = 2, 3\) one has sharper conclusions. Thus a non-trivial solution exists when \(k = 2\) for \(s \geq 5\) (a classical result of Hasse [13]), and when \(k = 3\) such holds for \(s \geq 10\) (see Lewis [17] and Demyanov [9]). There exist infinitely many exponents \(k\) and forms \(F\) in \(s\) variables failing to possess non-trivial solutions, meanwhile, with \(s\) as large as

\[
\exp \left( \frac{k}{(\log k)(\log \log k) \cdots (\log_r k)(\log_{r+1} k)^{1+\varepsilon}} \right),
\]

in which \(\varepsilon\) 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 \(s \geq 2k + 1\) and \(c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}\). Let \(\zeta \in \mathbb{R}^s \setminus \{0\}\) satisfy the equation \(\phi(\zeta) = 0\). By the homogeneity of \(\phi\), we may rescale so as to assume without loss of generality that \(|\zeta_i| < 1/2\) for \(1 \leq i \leq s\). We suppose also that for each prime number \(p\) and each \(h \in \mathbb{N}\), there is a solution \(z = z^{(p,h)} \in \mathbb{Z}^s\) 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 \(1 \leq i \leq s\).

Our goal is to evaluate the number \(N_{\phi}(X)\) of integral solutions \(x \in [-X, X]^s \cap \mathbb{Z}^s\) of the equation \(\phi(x) = 0\), in which \(\phi(x) = c_1 x_1^k + \ldots + c_s x_s^k\). Write

\[
f(\alpha) = \sum_{|z| \leq X} e(\alpha x^k).
\]
Then by orthogonality one has
\[ N_\phi(X) = \int_0^1 f(c_1 \alpha) \ldots f(c_s \alpha) \, d\alpha. \]

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

We require a Hardy-Littlewood dissection, and this entails a slight modification of that which we have previously employed. We define
\[ \mathcal{M}_\delta(q, a) = \{ \alpha \in [0, 1) : |\alpha - a/q| \leq X^{\delta-k} \}, \]
and take \( \mathcal{M}_\delta \) to be the union of the arcs \( \mathcal{M}_\delta(q, a) \) with \( 0 \leq a \leq q \leq X^\delta \) and \((a, q) = 1\). We will take \( \delta \) to be a real number with \( 0 < \delta < 1/5 \). Finally, we put \( m_\delta = [0, 1) \setminus \mathcal{M}_\delta \) 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
\[ \left| \int_{m_\delta} f(c_1 \alpha) \ldots f(c_s \alpha) \, d\alpha \right| \leq \prod_{i=1}^s \left( \int_{m_\delta} |f(c_i \alpha)|^s \, d\alpha \right)^{1/s}. \quad (11.1) \]
Moreover, for each index \( i \) one finds that
\[ \int_{m_\delta} |f(c_i \alpha)|^s \, d\alpha \leq \left( \sup_{\alpha \in m_\delta} |f(c_i \alpha)| \right)^{s-2k} \int_0^1 |f(c_i \alpha)|^{2k} \, d\alpha. \quad (11.2) \]

By a change of variable, invoking periodicity modulo 1 of \( f(\beta) \) and Hua’s Lemma, we discern that
\[ \int_0^1 |f(c_i \alpha)|^{2k} \, d\alpha = |c_i|^{-1} \int_0^{|c_i|} |f(\beta)|^{2k} \, d\beta = \int_0^1 |f(\beta)|^{2k} \, d\beta \ll X^{2k-k+\varepsilon}. \quad (11.3) \]

The estimation of \( \sup_{\alpha \in m_\delta} |f(c_i \alpha)| \), meanwhile, is achieved through Weyl’s inequality. It is convenient at this point to write
\[ H = \max_{1 \leq i \leq s} |c_i|. \]
Suppose that \( \alpha \in m_\delta \). By Dirichlet’s approximation theorem, there exists \( a \in \mathbb{Z} \) and \( q \in \mathbb{N} \) with \( 1 \leq q \leq X^{k-\delta} \), \((a, q) = 1\) and \( |c_i \alpha - a/q| \leq X^{\delta-k} \). Put
\[ b = \frac{a}{(c_i, a)} , \quad \frac{|c_i|}{c_i} \quad \text{and} \quad r = \frac{|c_i| q}{(c_i, a)} , \]
so that \( |\alpha - b/r| \leq X^{\delta-k} \) with \( 1 \leq r \leq Hq \leq HX^{k-\delta} \) and \((b, r) = 1\). If one were to have \( 1 \leq r \leq X^\delta \), then \( \alpha \) would lie on \( \mathcal{M}_\delta \). Thus we may assume that \( r > X^\delta \), whence \( H^{-1}X^\delta < q \leq X^{k-\delta} \). We therefore deduce from Weyl’s inequality that
\[ f(c_i \alpha) \ll X^{1+\varepsilon} \left( r^{-1} + X^{-1} + rX^{-k} \right)^{2^{1-k}} \ll X^{1+\varepsilon} \left( HX^{-\delta} + X^{-1} + X^{-\delta} \right)^{2^{1-k}} , \]
whence
\[ \sup_{\alpha \in \mathcal{M}_3} |f(c_1 \alpha)| \ll \epsilon X^{1-\delta 2^{-k}+\varepsilon}. \tag{11.4} \]

By substituting (11.3) and (11.4) into (11.2), we deduce that when \( s \geq 2^k + 1, \)
\[ \int_{\mathcal{M}_3} |f(c_1 \alpha)|^s \, d\alpha \ll (X^{1-\delta 2^{-k}+\varepsilon})^{s-2^k} X^{2^k-k+\varepsilon} \ll X^{s-k-\delta 2^{-k}}. \]

Hence, on recalling (11.1), we obtain the bound
\[ \int_{\mathcal{M}_3} f(c_1 \alpha) \ldots f(c_s \alpha) \, d\alpha \ll X^{s-k-\delta 2^{-k}}. \tag{11.5} \]

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
\[ f(\alpha) = \sum_{|x| \leq X} e(\alpha x^k) = 1 + \sum_{1 \leq x \leq X} e(\alpha x^k) + \sum_{1 \leq x \leq X} e(\alpha(-x)^k). \]

Thus, as a consequence of Lemma 5.1, we see that whenever \( \alpha \in \mathbb{R}, a \in \mathbb{Z} \) and \( q \in \mathbb{N}, \)
\[ f(\alpha) - q^{-1} S(q, a) v(\alpha - a/q) \ll q + X^k |q\alpha - a|, \]
where now we write
\[ S(q, a) = \sum_{r=1}^{q} e(ar^k/q) \quad \text{and} \quad v(\beta) = \int_{-X}^{X} e(\beta \gamma^k) \, d\gamma. \]

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

Define the function \( \tilde{f}_i(\alpha) \) to be \( q^{-1} S(q, c_i a) v(c_i (\alpha - a/q)) \) when \( \alpha \in \mathcal{M}_3(q, a) \subseteq \mathcal{M}_3, \)
and otherwise take \( \tilde{f}_i(\alpha) = 0 \). Then it follows that
\[ \sup_{\alpha \in \mathcal{M}_3} |f(c_i \alpha) - \tilde{f}_i(\alpha)| \ll \epsilon X^{2\delta}, \]
and hence
\[ \sup_{\alpha \in \mathcal{M}_3} |f(c_1 \alpha) \ldots f(c_s \alpha) - \tilde{f}_1(\alpha) \ldots \tilde{f}_s(\alpha)| \ll X^{s-1+2\delta}. \]

We may therefore infer that
\[ \int_{\mathcal{M}_3} f(c_1 \alpha) \ldots f(c_s \alpha) \, d\alpha - \int_{\mathcal{M}_3} \tilde{f}_1(\alpha) \ldots \tilde{f}_s(\alpha) \, d\alpha \ll X^{s-1+2\delta} \operatorname{mes}(\mathcal{M}_3) \]
\[ \ll X^{s-k+(5\delta-1)} = o(X^{s-k}). \]

On noting that
\[ \int_{\mathcal{M}_3} \tilde{f}_1(\alpha) \ldots \tilde{f}_s(\alpha) \, d\alpha = J(X^\delta) \mathcal{G}(X^\delta), \]
where
\[ J(Q) = \int_{-Q X^{-k}}^{Q X^k} v(c_1 \beta) \ldots v(c_s \beta) \, d\beta \]
and
\[ \mathcal{G}(Q) = \sum_{q=1}^{Q} \sum_{a=1}^{q} \prod_{i=1}^{s} (q^{-1} S(q, c_i a)) , \]
we may conclude as follows.

**Lemma 12.1.** One has
\[ \int_{\mathbb{R}^s} f(c_1 \alpha) \ldots f(c_s \alpha) \, d\alpha = J(X^\delta) \mathcal{G}(X^\delta) + o(X^{-k}) . \]

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
\[ J = \int_{-\infty}^{\infty} v(c_1 \beta) \ldots v(c_s \beta) \, d\beta , \]
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
\[ S(q, c_i a) = \sum_{r=1}^{q} e(c_i a r^k / q) = (q, c_i) S \left( \frac{q}{(q, c_i)}, \frac{c_i a}{(q, c_i)} \right) , \]
in which the last two arguments are pairwise coprime. We thus deduce from Lemma 7.1 that
\[ S(q, c_i a) \ll (q, c_i) \left( \frac{q}{(q, c_i)} \right)^{1-2^{-k}+\varepsilon} \ll q^{1-2^{-k}+\varepsilon} . \]

We write
\[ \mathcal{G} = \sum_{q=1}^{\infty} \sum_{a=1}^{q} \prod_{i=1}^{s} (q^{-1} S(q, c_i)) . \]

Then it follows that whenever \( s \geq 2^k + 1 \), then
\[ \mathcal{G} - \mathcal{G}(Q) \ll \sum_{q>Q} \sum_{a=1}^{q} \left( q^{-2^{-k}} \right)^s \ll Q^{-2^{-k}} , \]
so that \( \mathcal{G} \) is absolutely convergent and \( \mathcal{G} - \mathcal{G}(X^\delta) \ll X^{-\delta 2^{-k}} \).

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
\[ S(q, c_i (a r + b q)) = S(q, c_i a) S(r, c_i b) . \]

Write
\[ A(q) = \prod_{\substack{a=1 \\
(a,q)=1}}^{q} \prod_{i=1}^{s} (q^{-1} S(q, c_i a)) . \]

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 \( s \geq 2^k + 1 \) the infinite sum \( \sigma(p) \) converges absolutely, with
\[
|\sigma(p) - 1| \ll p^{-1-2^{-k}},
\]
that \( \prod_p \sigma(p) \) converges absolutely with \( \prod_p \sigma(p) = \mathcal{G} \), and that there is a natural number \( C = C(s, k, c) \) having the property that
\[
1/2 < \prod_{p \geq C} \sigma(p) < 3/2.
\]

Our next step is to analyse the \( p \)-adic densities \( \sigma(p) \) when \( p \) is small. Put
\[
M(q) = \text{card}\{m \in (\mathbb{Z}/q\mathbb{Z})^* : c_1 m_1^k + \ldots + c_s m_s^k = 0\}.
\]

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 \( \phi(x) = 0 \). Define \( \kappa \) by putting
\[
\kappa = \max_{1 \leq i \leq s} \{m \in \mathbb{Z} : p^m | c_i\},
\]
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
\[
c_1 y_1^k + \ldots + c_s y_s^k \equiv 0 \pmod{p^\nu},
\]
in which \( p \nmid y_i \) for some index \( i \) with \( 1 \leq i \leq s \). By relabelling variables, if necessary, we may suppose that this index is \( i = 1 \).

Consider an integer \( h \) with \( h \geq \nu \), and suppose that \( p^\mu | c_1 \), noting that \( \mu \leq \kappa \). We fix a choice of \( x_2, \ldots, x_s \) with \( 1 \leq x_j \leq p^h \) and \( x_j \equiv y_j \pmod{p^\nu} \) (\( 1 \leq i \leq s \)). There are \( (p^{h-\nu})^{s-1} \) such choices. We then have
\[
-c_1^{-1}(c_2 x_2^k + \ldots + c_s x_s^k) \equiv -c_1^{-1}(c_2 y_2^k + \ldots + c_s y_s^k) \equiv y_1^k \pmod{p^\nu},
\]
and hence
\[
-c_1^{-1}(c_2 x_2^k + \ldots + c_s x_s^k) \equiv y_1^k \pmod{p^\gamma}.
\]
Thus, since the left hand side of this last congruence is a \( k \)-th power modulo \( p^\gamma \), it follows from Lemma 8.3 that the left hand side is also a \( k \)-th power modulo \( p^h \). Hence the congruence
\[
-(c_2 x_2^k + \ldots + c_s x_s^k) \equiv c_1 x_1^k \pmod{p^h}
\]
is soluble, and we may conclude that \( M(p^h) \geq (p^{h-\nu})^{s-1} \). In particular, we have
\[
\sigma(p) = \lim_{h \to \infty} p^{h(1-s)} M(p^h) \geq p^{-\nu(s-1)},
\]
whence \( \sigma(p) > 0 \) for all prime numbers \( p \). From here we discern that
\[
\mathcal{G} = \left( \prod_{p < C} \sigma(p) \right) \left( \prod_{p \geq C} \sigma(p) \right) > \frac{1}{2} \prod_{p < C} \sigma(p) \gg 1,
\]
so that in fact, subject to the local solubility hypothesis and the assumption \( s \geq 2^k + 1 \),
one has \( 1 \ll \mathcal{G} \ll 1 \).

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

\[
v(c_i \beta) = \int_0^X e(c_i \beta \gamma^k) \, d\gamma + \int_0^X e(c_i \beta (-\gamma)^k) \, d\gamma \ll \frac{X}{(1 + X^k |c_i \beta|)^{1/k}} \ll \frac{X}{(1 + X^k |\beta|)^{1/k}}.
\]

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

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

we find that whenever \( s \geq k + 1 \), one has

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

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,

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

where

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

in which

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

When \( 0 < \eta \leq 1 \), we define the auxiliary function

\[
w_\eta(\beta) = \eta \left( \frac{\sin(\pi \eta \beta)}{\pi \eta \beta} \right)^2.
\]

This function has Fourier transform

\[
\hat{w}_\eta(\gamma) = \int_{-\infty}^{\infty} w_\eta(\beta) e(-\beta \gamma) \, d\beta = \max\{0, 1 - |\gamma|/\eta\}.
\]

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

\[
W_{\eta,\delta}(\gamma) = \begin{cases} 1, & \text{when } |\gamma| \leq \eta, \\ 1 - \frac{|\gamma| - \eta}{\delta}, & \text{when } \eta < |\gamma| < \eta + \delta, \\ 0, & \text{when } |\gamma| \geq \eta + \delta.
\end{cases}
\]

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
\[
\begin{cases}
1, & \text{when } |\gamma| \leq \eta, \\
0, & \text{when } |\gamma| > \eta.
\end{cases}
\]

Our formula (12.1) shows that
\[
W_{\eta,\delta}(\gamma) = (1 + \eta/\delta)\hat{w}_{\eta+\delta}(\gamma) - (\eta/\delta)\hat{w}_{\eta}(\gamma)
\]
whence
\[
W^+_\eta(\gamma) = (1 + \eta^{-1})\hat{w}_{\eta+\eta^2}(\gamma) - \eta^{-1}\hat{w}_{\eta}(\gamma)
\]
and
\[
W^-_\eta(\gamma) = \eta^{-1}\hat{w}_{\eta}(\gamma) + (1 - \eta^{-1})\hat{w}_{\eta-\eta^2}(\gamma).
\]

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

We now turn our attention to the Fourier side. Define
\[
U(\eta) = \int_{[-1,1]} \eta^{-1}\hat{w}_{\eta}(\phi(\xi)) \, d\xi.
\]

**Lemma 12.2.** Let $\eta$ be a real number with $0 < \eta < 1$. Suppose that $\eta_1$ is a real number with $|\eta_1 - \eta| \leq \eta^2$. 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
\[
U(\eta_1) = \int_{-\infty}^{\infty} v_1(c_1\beta) \ldots v_1(c_s\beta) K(\beta) \, d\beta,
\]
whence
\[
U(\eta_1) - I = \int_{-\infty}^{\infty} v_1(c_1\beta) \ldots v_1(c_s\beta) (K(\beta) - 1) \, d\beta. \tag{12.2}
\]

We want to show that the right hand side in (12.2) is small. Put $\mathcal{D} = [-\eta^{-1/2}, \eta^{-1/2}]$ and $\mathcal{E} = \mathbb{R} \setminus \mathcal{D}$. Then from the power series expansion of $w_{\eta}(\beta)$, we have
\[
0 \leq 1 - K(\beta) \ll \min\{1, \eta^2\beta^2\}.
\]
The absolute convergence of the integral $I$ ensures that the contribution arising from integrating over $\mathcal{D}$ on the right hand side of (12.2) is at most
\[
\sup_{\beta \in \mathcal{D}} |1 - K(\beta)| \int_{-\infty}^{\infty} v_1(c_1\beta) \ldots v_1(c_s\beta) \, d\beta \ll \eta.
\]
The contribution arising from $\mathcal{E}$, meanwhile, is bounded above by
\[
\int_{\mathcal{E}} (1 + |\beta|)^{-s/k} \, d\beta \ll (\eta^{-1/2})^{-1/k} = \eta^{1/(2k)}.
\]
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

$$
\int_{[-1,1]^s} W_\eta^+(\phi(\xi)) \, d\xi = ((\eta + \eta^2)(1 + 1/\eta) - \eta(1/\eta)) \, (I + O(\eta^{1/(2k)}))
= (2\eta + O(\eta^2)) \, (I + O(\eta^{1/(2k)}))
$$

and

$$
\int_{[-1,1]^s} W_\eta^-(\phi(\xi)) \, d\xi = (\eta(1/\eta) + (\eta - \eta^2)(1 - 1/\eta)) \, (I + O(\eta^{1/(2k)}))
= (2\eta + O(\eta^2)) \, (I + O(\eta^{1/(2k)})).
$$

We therefore conclude from that

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

whence

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

In particular, one has $J = \sigma_\infty X^{s-k}$, where $\sigma_\infty$ 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 $\eta$-neighborhood of the hypersurface $\phi(\xi) = 0$. By hypothesis, the equation $\phi(z) = 0$ has a real point $z = z(\infty) \in [-1/2, 1/2]^s$. Hence, an application of the Implicit Function Theorem shows that there is a positive number $\omega$ with the property that in a $\omega$-neighborhood $D$ of $z(\infty)$, this hypersurface is well-approximated by a hyperplane, and thus an $\eta$-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 $s \geq 2^k + 1$, one has

$$
\int_{\mathbb{R}^s} f(c_1\alpha) \ldots f(c_s\alpha) \, d\alpha = \mathcal{G}(X^{\delta})J(X^{\delta}) + o(X^{s-k})
= \mathcal{G}J + o(X^{s-k})
= \sigma_\infty \left( \prod_p \sigma_p \right) X^{s-k} + o(X^{s-k}).
$$

Finally,

$$
N_\phi(X) = \int_{\mathbb{R}^s} f(c_1\alpha) \ldots f(c_s\alpha) \, d\alpha + \int_{\mathbb{R}^s} f(c_1\alpha) \ldots f(c_s\alpha) \, d\alpha
= \sigma_\infty \left( \prod_p \sigma_p \right) X^{s-k} + o(X^{s-k}),
$$

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 $4^k$ 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 $X_i$ $(0 \leq i \leq s)$ defined by putting

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

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

Write

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

$$g_j(\alpha) = \sum_{X_j < y \leq 2X_j} e(\alpha y^k) \quad (0 \leq j \leq s-1),$$

$$G(\alpha) = \prod_{j=0}^{s-1} g_j(\alpha).$$

Then it follows via orthogonality that

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

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

$$|G(\alpha)| \leq G(0) = X_0 X_1 \ldots X_{s-1} \sim \prod_{j=0}^{s-1} X^{(1-1/k)^j} = X^{k(1-(1-1/k)^s)}.$$

Lemma 13.1 (Diminishing ranges estimate). One has

$$\int_0^1 |G(\alpha)|^2 d\alpha \ll X^{2-k(1-1/k)^s}.$$
above lemma. The argument of the proof of the lemma in fact shows that
\[ \int_0^1 |G(\alpha)|^2 \, d\alpha \asymp G(0)^2 X^{-k+k(1-1/k)^s}. \]

Since \( k(1-1/k)^s \leq ke^{-s/k} \), 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
\[ y^k_0 - z^k_0 = \sum_{j=1}^{s-1} (y^k_j - z^k_j), \] (13.1)
with \( X_j < y_j, z_j \leq 2X_j \) (\( 0 \leq j \leq s-1 \)). However, we have
\[ \left| \sum_{j=1}^{s-1} (y^k_j - z^k_j) \right| < (2X_1)^k + o(X_1^k) = (1 + o(1))X_0^k, \]
whereas, when \( y_0 \neq z_0 \), one has
\[ |y^k_0 - z^k_0| = |y_0 - z_0| \cdot |y^{k-1}_0 + y^{k-2}_0 z_0 + \ldots + z^{k-1}_0| \geq kX_0^{k-1}. \]
Thus, when \( y_0 \neq z_0 \), we have
\[ |y^k_0 - z^k_0| > \left| \sum_{j=1}^{s-1} (y^k_j - z^k_j) \right|, \]
contradicting the validity of the equation (13.1). We are therefore forced to conclude that \( y_0 = z_0 \), whence
\[ \int_0^1 |G(\alpha)|^2 \, d\alpha \leq X_0 \int_0^1 |g_1(\alpha) \ldots g_{s-1}(\alpha)|^2 \, d\alpha. \]
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
\[ \int_0^1 |G(\alpha)|^2 \, d\alpha \leq X_0 X_1 \ldots X_{s-1} \asymp X^{k-k(1-1/k)^s}. \]
This completes the proof of the lemma. \( \square \)

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 \( \delta \) to be a positive number with \( \delta < 1/5 \), and define \( M_\delta \) to be the union of the intervals
\[ M_\delta(q,a) = \{ \alpha \in [0,1) : |\alpha - a/q| \leq X^{\delta-k} \} \]
with \( 0 \leq a \leq q \leq X^\delta \) and \( (a,q) = 1 \). We then put \( m_\delta = [0,1] \setminus M_\delta \).

Our analysis of the minor arcs makes use of Weyl’s inequality just as in Lemma 3.7 and its sequel. Thus, we have
\[ \sup_{\alpha \in m_\delta} |f(\alpha)| \ll X^{1-\delta^2-k+\varepsilon}, \]
whence
\[
\int_{m_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha) \, d\alpha \leq \left( \sup_{\alpha \in m_\delta} |f(\alpha)| \right)^t \int_0^1 |G(\alpha)|^2 \, d\alpha \\
\ll \left( X^{1-\delta 2^{1-k}+\varepsilon} \right)^t X^{k-k(1-1/k)^s}.
\]
It follows that, for some positive number $\tau$, one has
\[
\int_{m_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha) \, d\alpha \ll G(0)^2 X^t - k - k\tau \ll G(0)^2 n^t/k - 1 - \tau,
\]
provided only that
\[
t\delta 2^{1-k} > k(1 - 1/k)^s. \quad (13.2)
\]
We shall presently establish the major arc estimate
\[
\int_{m_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha) \, d\alpha \gg G(0)^2 n^t/k - 1,
\]
subject to the condition that $t \geq 4k$. Since this quantity tends to $\infty$ as $n \to \infty$, it follows that
\[
G(k) \leq \inf_{s \geq 0} \left( 2s + \max \left\{ 4k, \left[ \frac{k(1 - 1/k)^s}{\delta 2^{1-k}} \right] \right\} \right).
\]
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 $(1 - 1/k)^s \leq e^{-s/k}$. Thus, taking $\delta = 1/8$ in order to simplify the associated computations, we see that
\[
2s + \left[ \frac{k(1 - 1/k)^s}{\delta 2^{1-k}} \right] < 2s + \left[ 8k 2^{k-1} e^{-s/k} \right].
\]
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 $s = [k^2 \log 2]$ and set
\[
t = \left[ \frac{k(1 - 1/k)^s}{\delta 2^{1-k}} \right] \leq [k 2^{k+2} 2^{-k}] = 4k.
\]
Thus we see that
\[
G(k) \leq 2[k^2 \log 2] + 4k < (2 \log 2) k^2 + 4k + 2.
\]
Let us return to the analysis of the major arc contribution. We have
\[
\int_{m_\delta} f(\alpha)^t G(\alpha)^2 e(-n\alpha) \, d\alpha = \sum_{y,z} \int_{m_\delta} f(\alpha)^t e(-N\alpha) \, d\alpha,
\]
where
\[
N = n - \sum_{j=0}^{s-1} (y_j^k + z_j^k),
\]
and the summation over $y, z$ is over $X_j < y_j, z_j \leq 2X_j$ ($0 \leq j \leq s - 1$). We note that
\[
\sum_{j=0}^{s-1} (y_j^k + z_j^k) \leq 2(2X_0)^k + O(X_1^k) < 10^{1-k} n,
\]
whence
\[ N = n - \sum_{j=0}^{s-1} (y_j^k + z_j^k) > n/2. \]

Also, since \( \delta < 1/5 \) and \( t \geq 4k \), our major arc analysis (using the estimate \( S(q,a) \ll q^{1-1/k+\varepsilon} \) for \((q,a) = 1\) from the fourth problem set) shows that
\[
\int_{\mathfrak{M}_s} f(\alpha)^t e(-N\alpha) \, d\alpha = \frac{\Gamma(1+1/k)^t}{\Gamma(t/k)} \mathfrak{S}_{t,k}(N) N^{t/k-1} + o(N^{t/k-1}) \gg n^{t/k-1}.
\]
Hence
\[
\int_{\mathfrak{M}_s} f(\alpha)^t G(\alpha)^2 e(-n\alpha) \, d\alpha \gg n^{t/k-1} \sum_{y,z} 1 = n^{t/k-1} G(0)^2.
\]
This justifies our earlier claim.

We may summarise these deliberations as follows.

**Theorem 13.2.** When \( k \geq 2 \), 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 \( G(k) \leq 2^k + 1 \). One may check on the back of an envelope that it improves on the latter bound for \( k \geq 7 \).

14. **Vinogradov’s mean value theorem**

We now discuss a different strategy for estimating the exponential sum over \( k \)-th powers that provides analogues of Hua’s lemma and of Weyl’s inequality. When \( k \in \mathbb{N} \) and \( \alpha \in \mathbb{R}^k \), we define the exponential sum \( f(\alpha) = f_k(\alpha; X) \) by putting
\[ f_k(\alpha; X) = \sum_{1 \leq i \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k). \]

We then define the mean value
\[ J_{s,k}(X) = \int_{[0,1]^k} |f_k(\alpha; X)|^{2s} \, d\alpha. \]

By orthogonality, this mean value counts the number of integral solutions of the system of equations
\[ \sum_{i=1}^{s} (x_i^j - y_i^j) = 0 \quad (1 \leq j \leq k), \] (14.1)
with \( 1 \leq x_i, y_i \leq X \) (\( 1 \leq i \leq s \)). Estimates for \( J_{s,k}(X) \) have, for the past eight decades, gone by the name of “Vinogradov’s mean value theorem”.

Before proceeding further, it seems appropriate to ask ourselves what we should expect concerning the size of \( J_{s,k}(X) \). There are two immediate lower bounds to note. First, let \( T_s(X) \) denote the number of integers \( x, y \) in which \( 1 \leq x_i, y_i \leq X \) (\( 1 \leq i \leq s \)) and \( \{x_1, \ldots, x_s\} = \{y_1, \ldots, y_s\} \). Then on considering the solutions \( x, y \) of (14.1) counted by \( J_{s,k}(X) \) in which \( \{x_1, \ldots, x_s\} = \{y_1, \ldots, y_s\} \), we see that
\[ J_{s,k}(X) \geq T_s(X) \sim s!X^s. \]
When $s$ is large in terms of $k$, there may be many more than these diagonal solutions. Let $c = (100k)^{-1}$ and denote by $\mathcal{B}$ the set $[0, cX^{-1}) \times [0, cX^{-2}) \times \ldots \times [0, cX^{-k})$.

Then whenever $\alpha \in \mathcal{B}$, one sees that 
\[
|f_k(\alpha; X)| = \left| \sum_{1 \leq x \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k) \right| \geq \sum_{1 \leq x \leq X} \cos \left( \frac{2\pi}{100} \right).
\]

Thus
\[
J_{s,k}(X) \geq \int_{\mathcal{B}} |f_k(\alpha; X)|^{2s} \, d\alpha \gg X^{2s} \text{mes}(\mathcal{B}) \gg X^{2s-k(k+1)/2}.
\]

By combining these estimates we obtain the lower bound
\[
J_{s,k}(X) \gg X^s + X^{2s-k(k+1)/2}.
\]

This prompts the formulation of a conjecture concerning the size of $J_{s,k}(X)$.

**Conjecture.** (The Main Conjecture on Vinogradov’s mean value theorem) For each $\varepsilon > 0$, one has
\[
J_{s,k}(X) \ll X^{s+\varepsilon} + X^{2s-k(k+1)/2}.
\]

This conjecture has very recently become a theorem. In the cases $k = 2$ (and $k = 1$ of course), it is classical. When $k = 3$ it was proved by Wooley, and when $k > 3$ it was proved by Bourgain, Demeter and Guth, and via a different method, by Wooley (see [3, 35, 36]). In this course we will aim for a weaker classical conclusion that nonetheless makes the principal consequences accessible.

**Theorem 14.1** (Vinogradov’s mean value theorem). Suppose that $k \geq 3$, and that $s = rk$ for some $r \in \mathbb{N}$. Then, for all large values of $X$ one has
\[
J_{s,k}(X) \ll_{s,k} X^{k(2s-k(k+1)+\Delta_{s,k}}, \tag{14.2}
\]
where $\Delta_{s,k} = \frac{1}{2}k^2(1 - 1/k)^r$.

Note that when $s$ is a multiple of $k$, one has $\Delta_{s,k} \leq \frac{1}{2}k^2 e^{-r/k} = \frac{1}{2}k^2 e^{-s/k^2}$. Thus, when $s \geq 10k^2 \log k$, say, one has $\Delta_{s,k} \leq 1/k^8$, and the theorem provides a good approximation to the Main Conjecture in VMVT.

Before establishing this version of VMVT, we make some observations concerning $J_{s,k}(X)$ for small values of $k$. Note that the polynomials $t_j(x) = x_1^j + \ldots + x_k^j \ (1 \leq j \leq k)$ are symmetric polynomials in $x_1, \ldots, x_k$, and hence can be expressed as linear combinations of products of the elementary symmetric polynomials
\[
\sigma_1(x) = x_1 + x_2 + \ldots + x_k,
\]
\[
\sigma_2(x) = x_1 x_2 + \ldots + x_{s-1} x_s,
\]
and so on, up to
\[
\sigma_k(x) = x_1 x_2 \ldots x_k.
\]
Likewise, the elementary symmetric polynomials can be expressed as linear combinations of products of the polynomials $t_j(x)$ for $1 \leq j \leq k$. Thus, by applying the Newton-Girard formulae concerning the roots of polynomials, we see that whenever

$$\sum_{i=1}^{k} x_i^j = \sum_{i=1}^{k} y_i^j \quad (1 \leq j \leq k),$$

then necessarily one has the polynomial identity

$$\prod_{i=1}^{k} (t - x_i) = \prod_{j=1}^{k} (t - y_j).$$

By putting $t = x_u$ for any given index $u$ with $1 \leq u \leq k$, we see that

$$\prod_{j=1}^{k} (x_u - y_j) = 0,$$

and hence $y_v = x_u$ for some index $v$ with $1 \leq v \leq k$. Proceeding inductively, we discern that

$$\{x_1, \ldots, x_k\} = \{y_1, \ldots, y_k\},$$

and hence

$$J_{k,k}(X) = T_k(X) \sim k!X^k.$$

It is conjectured that $J_{s,k}(X) = T_s(X)$ even for values of $s$ slightly larger than $k$ when $k$ is sufficiently large, but little is known in this direction beyond the situation with $s = k + 1$. Thus, one has the asymptotic formula

$$J_{3,2}(X) = \frac{18}{\pi^2}X^3\log X + O(X^3)$$

(see Rogovskaya [22]), and conclusions of Vaughan and Wooley for larger values of $k$ showing that

$$X^2\log X \ll J_{4,3}(X) - T_4(X) \ll X^{10/3} (\log X)^{35}$$

and

$$J_{k+1,k}(X) - T_{k+1}(X) \ll X^{\sqrt{4k+5}}$$

(see [27, 28]).

The proof of Theorem 14.1. We prove the theorem inductively. In order to establish the base case of the induction, we observe that by orthogonality one has

$$J_{k,k}(X) = \sum_{x,y} 1,$$

where $1 \leq x, y \leq X$ and

$$\sum_{i=1}^{k} x_i^j = \sum_{i=1}^{k} y_i^j \quad (1 \leq j \leq k).$$

We have seen that this system of equations implies that

$$\{x_1, \ldots, x_k\} = \{y_1, \ldots, y_k\},$$

whence

$$J_{k,k}(X) \ll X^k = X^{k - \frac{1}{2}k(k+1) + \Delta_{k,k}},$$
with \( \Delta_{k,k} = \frac{1}{2}k(k + 1) - k = \frac{1}{2}k^2(1 - 1/k) \). Thus the inductive hypothesis (14.2) holds with \( s = rk \) when \( r = 1 \).

Suppose now that the inductive hypothesis holds when \( s = r'k \) whenever \( r' < r \) for a fixed \( r \geq 2 \), and put \( s = rk \). By orthogonality we again see that \( J_{s,k}(X) \) counts the number of integral solutions of the system of equations

\[
\sum_{i=1}^{s} x_i^j = \sum_{i=1}^{s} y_i^j \quad (1 \leq j \leq k),
\]

with \( 1 \leq x_i, y_i \leq X \). We classify the solutions of (14.3) counted by \( J_{s,k}(X) \) into two types.

(i) Let \( I_1 \) denote the contribution to \( J_{s,k}(X) \) arising from those solutions with \( x_i = x_h \) for some \( 1 \leq i < h \leq k \). In this case we may relabel variables so that \( i = 1 \) and \( h = 2 \) at the cost of inflating our estimates by a permutation factor. Thus

\[
I_1 \ll s,k \sum_{x,y} 1,
\]

where the summation over \( x_1, x_3, \ldots, x_s \) and \( y_1, \ldots, y_s \) is subject to the conditions \( 1 \leq x, y \leq X \) and

\[
2x_1^j + \sum_{i=3}^{s} x_i^j = \sum_{i=1}^{s} y_i^j \quad (1 \leq j \leq k).
\]

By orthogonality in combination with an application of Hölder’s inequality, therefore, one has

\[
I_1 \ll \int_{[0,1]^k} |f(2\alpha)f(\alpha)^{s-2}f(-\alpha)^s| \, d\alpha
\]

\[
\ll \left( \int_{[0,1]^k} |f(2\alpha)|^s \, d\alpha \right)^{1/s} \left( \int_{[0,1]^k} |f(\alpha)|^{2s} \, d\alpha \right)^{1-1/s}
\]

\[
\leq (X^s)^{1/s} (J_{s,k}(X))^{1-1/s}.
\]

(ii) Let \( I_2 \) denote the contribution to \( J_{s,k}(X) \) arising from those solutions in which \( x_i = x_h \) for no \( i \) and \( h \) with \( 1 \leq i < h \leq k \). Notice that, should one have the relation \( I_2 < I_1 \), then

\[
J_{s,k}(X) = I_1 + I_2 < 2I_1 \ll X J_{s,k}(X)^{1-1/s},
\]

whence \( J_{s,k}(X) \ll X^s \). Since this estimate is more than sufficient to establish the inductive hypothesis when \( r \geq 2 \), we may therefore assume henceforth that \( I_2 \geq I_1 \), and hence that \( J_{s,k}(X) \leq 2I_2 \).

Consider a typical solution \( x, y \) of (14.3) counted by \( I_2 \), and write

\[
\Xi(x) = \prod_{1 \leq i < h \leq k} (x_i - x_h).
\]

Then whenever \( 1 \leq x_i \leq X \), one has \( 1 \leq |\Xi(x)| \leq X^{k(k-1)} \). The number of prime numbers \( p \) with \( X^{1/k} < p \leq 2X^{1/k} \) dividing \( \Xi(x) \) is therefore at most

\[
\frac{\log |\Xi(x)|}{\log(X^{1/k})} < \frac{k(k-1) \log X}{(\log X) / k} < k^3.
\]
Taking $\mathcal{P}_0$ to be the set of prime numbers $p$ with $X^{1/k} < p \leq 2X^{1/k}$, it follows from a weak version of the prime number theorem that when $X$ is large, one has
\[ \text{card}(\mathcal{P}_0) \gg X^{1/k}/(\log X) > k^3. \]

Now write $\mathcal{P}$ for the set of the smallest $k^3$ prime numbers lying in $\mathcal{P}_0$. Then we see that, given any solution $x, y$ of \((14.3)\) with $1 \leq x, y \leq X$ counted by $I_2$, there must exist a prime number $p \in \mathcal{P}$ for which $p \nmid \Xi(x)$, and hence $x_i \equiv x_h \pmod{p}$ for no $i$ and $h$ with $1 \leq i < h \leq k$. Write $J_{s,k}(X; p)$ for the contribution to $I_2$ arising from the solutions $x, y$ of \((14.3)\) with $1 \leq x, y \leq X$, and with $x_i \not\equiv x_h \pmod{p}$ $(1 \leq i < h \leq k)$. Then it follows from this discussion that
\[ I_2 \leq \sum_{p \in \mathcal{P}} J_{s,k}(X; p) \ll_k \max_{p \in \mathcal{P}} J_{s,k}(X; p). \]

Let $p$ be the prime number corresponding to the maximum in the last relation. We now divide the variables implicit in the exponential sums underlying the mean value $J_{s,k}(X; p)$ into arithmetic progressions modulo $p$. Write
\[ g(\alpha; \xi) = \sum_{1 \leq x \leq X, x \equiv \xi \pmod{p}} e(\alpha_1 x + \ldots + \alpha_k x^k). \]

Also, let $\mathcal{A}$ denote the set of $k$-tuples $\xi = (\xi_1, \ldots, \xi_k)$, with $0 < \xi_r \leq p$, and with $\xi_i = \xi_h$ for no $i$ and $h$ with $1 \leq i < h \leq k$. Then by orthogonality and the triangle inequality, we have
\[ J_{s,k}(X; p) \leq \int_{[0, 1)^k} \left| \sum_{\xi \in \mathcal{A}} g(\alpha; \xi_1) \ldots g(\alpha; \xi_k) \right| |f(\alpha)|^{2s-k} \, d\alpha. \]

Next, by invoking Schwarz’s inequality, we find that $I_2 \ll I_3^{1/2} I_4^{1/2}$, where
\[ I_3 = \int_{[0, 1)^k} \left| \sum_{\xi \in \mathcal{A}} g(\alpha; \xi_1) \ldots g(\alpha; \xi_k) \right|^2 |f(\alpha)|^{2s-2k} \, d\alpha \tag{14.4} \]
and
\[ I_4 = \int_{[0, 1)^k} |f(\alpha)|^{2s} \, d\alpha = J_{s,k}(X). \]

Consequently, in the scenario currently under consideration, we have
\[ J_{s,k}(X) \ll I_3^{1/2} J_{s,k}(X)^{1/2}, \]
whence $J_{s,k}(X) \ll I_3$. An application of Hölder’s inequality reveals that
\[ |f(\alpha)|^{2s-2k} = \left| \sum_{0 \leq \eta < p} g(\alpha; \eta) \right|^{2s-2k} \leq p^{2s-2k-1} \sum_{0 \leq \eta < p} |g(\alpha; \eta)|^{2s-2k}. \]

Then it follows from \((14.4)\) that there is a unique integer $\eta$ with $0 \leq \eta < p$ for which
\[ J_{s,k}(X) \ll p^{2s-2k} \int_{[0, 1)^k} \left| \sum_{\xi \in \mathcal{A}} g(\alpha; \xi_1) \ldots g(\alpha; \xi_k) \right|^2 |g(\alpha; \eta)|^{2s-2k} \, d\alpha. \]
By orthogonality, the integral on the right hand side of the last relation counts the
number of integral solutions of the system
\[ \sum_{i=1}^{k} (x_i^j - y_i^j) = \sum_{h=1}^{s-k} \left( (pu_h + \eta)^j - (pv_h + \eta)^j \right) \quad (1 \leq j \leq k), \tag{14.5} \]
with
\[ 1 \leq x_i, y_i, pu_h + \eta, pv_h + \eta \leq X, \]
and with \(x_1, \ldots, x_k\) distinct modulo \(p\), and likewise \(y_1, \ldots, y_k\) distinct modulo \(p\). By
applying the binomial theorem within (14.5) and taking appropriate linear combinations
of the equations, we see that (14.5) is equivalent to the system of equations
\[ \sum_{i=1}^{k} \left( (x_i - \eta)^j - (y_i - \eta)^j \right) = p^j \sum_{h=1}^{s-k} (u_h^j - v_h^j) \quad (1 \leq j \leq k). \]
It follows, in particular, that
\[ \sum_{i=1}^{k} (x_i - \eta)^j \equiv \sum_{i=1}^{k} (y_i - \eta)^j \pmod{p^j} \quad (1 \leq j \leq k), \]
so that there are integers \(h_j\) with \(1 \leq h_j \leq p^{k-j} \quad (1 \leq j \leq k)\) having the property that
\[ \sum_{i=1}^{k} (x_i - \eta)^j \equiv h_j p^j + \sum_{i=1}^{k} (y_i - \eta)^j \pmod{p^k} \quad (1 \leq j \leq k). \]
Consider any fixed choice for \(h\) and \(y\). There are at most \((1 \cdot p \cdot p^2 \cdot \ldots \cdot p^{k-1}) \cdot X^k\) such
choices. Then, for some integers \(l_1, \ldots, l_k\), depending at most on \(h\) and \(y\), we have
\[ \sum_{i=1}^{k} (x_i - \eta)^j \equiv l_j \pmod{p^j} \quad (1 \leq j \leq k). \tag{14.6} \]
If \(x\) and \(x'\) are two solutions of this system of congruences, then by the Newton-Girard
formulae we have
\[ \prod_{i=1}^{k} (t - (x_i - \eta)) \equiv \prod_{h=1}^{k} (t - (x'_h - \eta)) \pmod{p^k}, \]
considered as polynomials in \(t\). Thus, on putting \(t = x'_j - \eta\) for a fixed index \(j\) with
\(1 \leq j \leq k\), we see that
\[ \prod_{i=1}^{k} (x'_j - x_i) \equiv 0 \pmod{p^k}. \]
Recall that \(x_1, \ldots, x_k\) are distinct modulo \(p\). Then we may conclude that \(x_i \equiv x'_j \pmod{p^k}\)
for some index \(i\) with \(1 \leq i \leq k\). Proceeding inductively, we discern that
\[ \{x_1, \ldots, x_k\} \equiv \{x'_1, \ldots, x'_k\} \pmod{p^k}, \]
so that there are at most \(k!\) solutions of the congruence (14.6) modulo \(p^k\). But \(p^k > X\),
and so the integers \(x_1, \ldots, x_k\) are determined uniquely by these congruences. We thus
conclude that the number of choices for \(x\) and \(y\) is at most \(k!p^{k(k-1)/2}X^k\).
Fix any one such choice of $x$ and $y$, and write
\[ n_j = p^{-j} \sum_{i=1}^{k} ((x_i - \eta)^j - (y_i - \eta)^j) \quad (1 \leq j \leq k). \]

Then $u$ and $v$ are determined by the equations
\[ \sum_{h=1}^{s-k} (u_j^h - v_j^h) = n_j \quad (1 \leq j \leq k). \]

By orthogonality, the number of such solutions is
\[
\int_{[0,1)^k} |1 + f_k(\alpha; (X - \eta)/p)|^{2s-2k} e(-\alpha_1 n_1 - \ldots - \alpha_k n_k) \, d\alpha \\
\leq \int_{[0,1)^k} |1 + f_k(\alpha; (X - \eta)/p)|^{2s-2k} \, d\alpha \\
\ll 1 + \int_{[0,1)^k} |f_k(\alpha; (X - \eta)/p)|^{2s-2k} \, d\alpha \\
\ll 1 + J_{s-k,k}(X/p).
\]

At this point, we have shown in this case that
\[ J_{s,k}(X) \ll p^{2s-2k} \cdot p^{k(k-1)/2} X^k J_{s-k,k}(X^{1-1/k}). \]

Making use of our inductive hypothesis, we therefore deduce that
\[ J_{s,k}(X) \ll X^k (X^{1/k})^{2s-2k+k(k-1)/2} (X^{1-1/k})^{2s-2k-k(k+1)/2+\Delta_{s-k,k}} \\
\ll X^{2s-k(k+1)/2+\Delta_{s,k}}, \]

where $\Delta_{s,k} = \Delta_{s-k,k}(1 - 1/k)$. This confirms the inductive hypothesis with $s$ in place of $s - k$, and the conclusion of the theorem follows. \qed

15. An analogue of Hua's lemma

It is straightforward to estimate the mean value
\[ I_{s,k}(X) = \int_0^1 \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{2s} \, d\alpha \]
in terms of the auxiliary mean value
\[ J_{s,k}(X) = \int_{[0,1)^k} \left| \sum_{1 \leq x \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k) \right|^{2s} \, d\alpha, \]
as we now demonstrate. We observe that, by orthogonality, the mean value $I_{s,k}(X)$ counts the number of integral solutions of the equation
\[ x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k, \]
with \(1 \leq x, y \leq X\). This is equal to the number of solutions of the system

\[
\sum_{i=1}^{s} (x_i^k - y_i^k) = 0
\]

\[
\sum_{i=1}^{s} (x_i^j - y_i^j) = h_j \quad (1 \leq j \leq k - 1),
\]

(15.1)

with \(1 \leq x, y \leq X\) and \(|h_j| \leq sX^j\) \((1 \leq j \leq k - 1)\). In order to confirm this observation, one has merely to observe that when \(1 \leq x, y \leq X\), the left hand side polynomials satisfy

\[
\left| \sum_{i=1}^{s} (x_i^j - y_i^j) \right| \leq sX^j,
\]

so that the equations (15.1) in fact impose no condition whatsoever.

We now apply orthogonality again to deduce that

\[
I_{s,k}(X) \leq \sum_{|h_1| \leq sX} \ldots \sum_{|h_{k-1}| \leq sX^{k-1}} \int_{[0,1)^k} |f_{\alpha}(\alpha; X)|^{2s} e(-h_1\alpha_1 - \ldots - h_{k-1}\alpha_{k-1}) \, d\alpha
\]

\[
\ll X^{k(k-1)/2} \int_{[0,1)^k} |f_{\alpha}(\alpha; X)|^{2s} \, d\alpha
\]

\[
\asymp X^{k(k-1)/2} J_{s,k}(X).
\]

We summarise these deliberations in the form of a theorem.

**Theorem 15.1.** One has

\[
\int_{0}^{1} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{2s} \, d\alpha \ll X^{k(k-1)/2} J_{s,k}(X).
\]

Thus, when \(s = rk\) with \(r \in \mathbb{N}\), one has

\[
\int_{0}^{1} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{2s} \, d\alpha \ll X^{2s-k+\frac{1}{2}k^2(1-1/k)^r}.
\]

**Proof.** The second conclusion is a consequence of Vinogradov’s mean value theorem in the shape of Theorem 14.1.

We will make use of this estimate later. For now, we remark that the proof of the Main Conjecture in Vinogradov’s mean value theorem shows that

\[
J_{s,k}(X) \ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}
\]

for \(s \geq k(k+1)/2\), whence we deduce from Theorem 15.1 that

\[
\int_{0}^{1} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{2s} \, d\alpha \ll X^{2s-k+\varepsilon}
\]

for \(s \geq k(k+1)/2\). More is known. When \(1 \leq r \leq k\), one can show that

\[
\int_{0}^{1} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{r(r+1)} \, d\alpha \ll X^{r^2+\varepsilon}
\]
(see Bourgain [2] for a sketch with details that might charitably be described as “obscure”, and [36] for a complete proof). Also, in the notation of which we have previously made use, one has

$$\int \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right|^{2s} \, d\alpha \ll X^{2s-k-1+\varepsilon}$$

for $s \geq k(k+1)/2$, as a consequence of [33] in combination with the now proved Main Conjecture in Vinogradov’s mean value theorem.

16. The large sieve inequality

We begin with an analytic version of the large sieve inequality, with a proof based on the Gallagher-Sobolev inequality.

**Lemma 16.1 (Gallagher-Sobolev inequality).** Suppose that $g : [0, 1) \to \mathbb{R}$ has continuous first derivatives on $[0, 1)$. Then whenever $\alpha \in [0, 1)$ and $\delta > 0$ satisfies the condition that $\alpha \pm \delta/2 \in [0, 1)$, one has

$$|g(\alpha)| \leq \int_{\alpha-\delta/2}^{\alpha+\delta/2} \left( \frac{1}{\delta} |g(u)| + |g'(u)| \right) \, du.$$

**Proof.** We begin with the proof of a simplified version of the lemma that we later rescale in order to obtain the desired conclusion in full. Suppose that $f : [0, 1] \to \mathbb{R}$ has continuous first derivative on $[0, 1)$. Observe that when $x \in [0, 1)$, it follows via integration by parts that one has

$$\int_{0}^{x} uf'(u) \, du = [uf(u)]_{0}^{x} - \int_{0}^{x} f(u) \, du = xf(x) - \int_{0}^{x} f(u) \, du$$

and

$$\int_{x}^{1} (u-1)f'(u) \, du = [(u-1)f(u)]_{x}^{1} - \int_{x}^{1} f(u) \, du = -(x-1)f(x) - \int_{x}^{1} f(u) \, du.$$

By combining these two identities through addition, one therefore discerns that

$$\int_{0}^{x} uf'(u) \, du + \int_{x}^{1} (u-1)f'(u) \, du + \int_{0}^{1} f(u) \, du = f(x),$$

whence

$$|f(x)| \leq \int_{0}^{1} |f(u)| \, du + \int_{0}^{x} |f'(u)| \, du + \int_{x}^{1} |f(u)| \, du$$

$$= \int_{0}^{1} |f(u)| \, du + \int_{0}^{1} |f'(u)| \, du.$$

By putting $f(u) = g(\alpha - \frac{1}{2}\delta u + \delta u)$ and then substituting $v = \alpha - \frac{1}{2}\delta + \delta v$, we see that

$$|g(\alpha)| \leq \int_{0}^{1} |g(\alpha - \frac{1}{2}\delta + \delta u)| \, du + \int_{0}^{1} \delta |g'(\alpha - \frac{1}{2}\delta + \delta u)| \, du$$

$$= \int_{\alpha-\delta/2}^{\alpha+\delta/2} \frac{1}{\delta} |g(v)| \, dv + \int_{\alpha-\delta/2}^{\alpha+\delta/2} |g'(v)| \, dv,$$

and the conclusion of the lemma follows. \qed
Lemma 16.2 (Large sieve inequality). Suppose that $\delta > 0$ and that $\Gamma = \{\gamma_1, \ldots, \gamma_R\}$ is a non-empty set of real numbers with the property that the real numbers $\gamma_i$ are spaced apart by at least $\delta$ modulo 1. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of complex numbers, and write

$$S(\beta) = \sum_{1 \leq n \leq N} a_n e(n\beta).$$

Then one has

$$\sum_{\gamma \in \Gamma} |S(\gamma)|^2 \ll (N + \delta^{-1}) \sum_{1 \leq n \leq N} |a_n|^2.$$

Here, for the purposes of clarity, we note that the spacing condition on the $\gamma_i$ ensures that the sets

$$\Omega(\gamma_i) = \{\beta \in [0, 1) : \|\beta - \gamma_i\| < \delta\}$$

are pairwise disjoint for $1 \leq i \leq R$. We remark also that the sharpest versions of the large sieve inequality (due to Selberg, and Montgomery and Vaughan) take the shape

$$\sum_{\gamma \in \Gamma} |S(\gamma)|^2 \leq (N + \delta^{-1} - 1) \sum_{1 \leq n \leq N} |a_n|^2.$$

The proof of Lemma 16.2. We apply the Gallagher-Sobolev inequality with $g(\alpha) = S(\alpha)^2$. Thus we obtain the bound

$$\sum_{\gamma \in \Gamma} |S(\gamma)|^2 \leq \sum_{\gamma \in \Gamma} \int_{\gamma - \delta/2}^{\gamma + \delta/2} \frac{1}{\delta} |S(\alpha)|^2 + 2|S(\alpha)S'(\alpha)| \, d\alpha.$$

On making use of the spacing condition on the real numbers $\gamma \in \Gamma$, we thus deduce that

$$\sum_{\gamma \in \Gamma} |S(\gamma)|^2 \leq \frac{1}{\delta} \int_0^1 |S(\alpha)|^2 \, d\alpha + 2 \int_0^1 |S(\alpha)S'(\alpha)| \, d\alpha.$$

By Parseval’s identity, one has

$$\int_0^1 |S(\alpha)|^2 \, d\alpha = \sum_{1 \leq n \leq N} |a_n|^2.$$

Meanwhile, by applying Schwarz’s inequality in combination with Parseval’s identity, we obtain the estimate

$$\int_0^1 |S(\alpha)S'(\alpha)| \, d\alpha \left( \int_0^1 |S(\alpha)|^2 \, d\alpha \right)^{1/2} \left( \int_0^1 |S'(\alpha)|^2 \, d\alpha \right)^{1/2}
\leq \left( \sum_{1 \leq n \leq N} |a_n|^2 \right)^{1/2} \left( \sum_{1 \leq n \leq N} 2\pi n|a_n|^2 \right)^{1/2}
\leq 2\pi N \sum_{1 \leq n \leq N} |a_n|^2.$$

Thus we conclude that

$$\sum_{\gamma \in \Gamma} |S(\gamma)|^2 \leq \left( \frac{1}{\delta} + 4\pi N \right) \sum_{1 \leq n \leq N} |a_n|^2.$$

This complete the proof of the lemma. $\square$
17. Vinogradov’s analogue of Weyl’s inequality

Our idea is to bound
\[ f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k) \]
in terms of an expression of the shape
\[ \sum_{1 \leq y \leq Y} \sum_{p \in P} b_y e(\alpha p^k y), \]
for suitable weights \( b_y \), and then study the spacing of the points \( \alpha p^k \) modulo 1 in preparation for an application of the large sieve inequality. We begin with an auxiliary lemma, the discussion of which is facilitated by the introduction of some additional notation. Thus, we write
\[ \psi(x; \alpha) = \alpha_1 x + \ldots + \alpha_k x^k, \]
and note that, equipped with this notation, one has
\[ f_k(\alpha; X) = \sum_{1 \leq x \leq X} e(\psi(x; \alpha)). \]

**Lemma 17.1.** Suppose that \( \mathcal{A} \subseteq [1, X] \cap \mathbb{Z} \) satisfies \( \text{card}(\mathcal{A}) = Y \). Then one has
\[ f_k(\alpha; X) \ll Y^{-1} \log(2X) \sum_{1 \leq x \leq 2X} \sup_{\beta \in [0,1]} |g(\alpha, \beta; x)|, \]
where we write
\[ g(\alpha, \beta; x) = \sum_{y \in \mathcal{A}} e(\psi(x - y; \alpha) + \beta y). \]

**Proof.** For each \( y \in \mathcal{A} \), we have
\[ f_k(\alpha; X) = \sum_{1 \leq y \leq x \leq X + y} e(\psi(x - y; \alpha)), \]
whence
\[ Yf_k(\alpha; X) = \sum_{y \in \mathcal{A}} \sum_{1 \leq y \leq x \leq X + y} e(\psi(x - y; \alpha)) \]
\[ = \sum_{1 \leq x \leq 2X} \sum_{y \in \mathcal{A}} e(\psi(x - y; \alpha)). \]
But
\[ \sum_{y \in \mathcal{A}} e(\psi(x - y; \alpha)) = \int_0^1 \sum_{y \in \mathcal{A}} e(\psi(x - y; \alpha) + \beta y) \sum_{x - X \leq u \leq x - 1} e(-\beta u) \, d\beta \]
\[ \ll \int_0^1 |g(\alpha, \beta; x)| \min\{X, \|\beta\|^{-1}\} \, d\beta \]
\[ \ll \sup_{\beta \in [0,1]} |g(\alpha, \beta; x)| \int_0^1 \min\{X, \|\beta\|^{-1}\} \, d\beta \]
\[ \ll \sup_{\beta \in [0,1]} |g(\alpha, \beta; x)| \log(2X). \]
Thus, we conclude that
\[ f_k(\alpha; X) \ll Y^{-1} \log(2X) \sup_{1 \leq x \leq 2X} |g(\alpha, \beta; x)|. \]

\[ \square \]

We have now bounded \(|f_k(\alpha; X)|\) in terms of an average over a potentially more amenable sum \(|g(\alpha, \beta; x)|\) in which the summands are as well-behaved as we choose. Before embarking on the next lemma, we record some additional notation with which conveniently to describe bounds for \(J_{s,k}(X)\). We say that the exponent \(\Delta_{s,k}\) is permissible when, for all large real numbers \(Z\), one has
\[ J_{s,k}(Z) \ll Z^{2s-\frac{1}{2}(k+1)+\Delta_{s,k}}. \]

Thus, we have shown that the exponent \(\Delta_{s,k} = \frac{1}{2} k^2 (1 - 1/k)^{s/k}\) is permissible whenever \(s = rk\) for some \(r \in \mathbb{N}\).

**Lemma 17.2.** Let \(M\) be a real number satisfying \(1 \leq M \leq X\). Also, let \(\alpha_k \in \mathbb{R}\), and suppose that \(a \in \mathbb{Z}\) and \(q \in \mathbb{N}\) satisfy \((a,q) = 1\), \(|q \alpha_k - a| \leq M^{-k}\), and either \(|q \alpha_k - a| > MX^{-k}\) or \(q > M\). Then, for each \(s \in \mathbb{N}\) and any permissible exponent \(\Delta_{s,k}\), one has
\[ f_k(\alpha; X) \ll q^s X^{1+\varepsilon} \left(M^{-1}(X/M)^{\Delta_{s,k}} (1 + q(X/M)^{-k})\right)^{1/(2s)}. \]

**Proof.** Let \(\mathcal{P}\) denote the set of prime numbers \(p\) with
\[ 2^{-1/(k-1)} M < p < 2^{-1/k} M. \]

We apply Lemma 17.1 with
\[ \mathcal{A} = \{n \in \mathbb{Z} \cap [1, X]: n = py, p \in \mathcal{P} \text{ and } 1 \leq y \leq X/M\}. \]

Note that \(\text{card}(\mathcal{A}) \gg (X/M)(M/\log M) \gg X^{1-\varepsilon}\). Thus we deduce that
\[ f_k(\alpha; X) \ll X^{\varepsilon^{-1}} \sup_{1 \leq x \leq 2X} \left| \sum_{\beta \in [0,1]} \sum_{a \in \mathcal{A}} e(\psi(x - w; \alpha) + \beta w) \right|. \]

If we now expand \(\psi(x - w; \alpha)\) as a polynomial in \(w\), we deduce that there exist real numbers \(\gamma_1, \ldots, \gamma_{k-1}\) depending on some fixed value of \(x\) with \(1 \leq x \leq 2X\), as well as a fixed number \(\beta \in [0,1)\), such that
\[ f_k(\alpha; X) \ll X^\varepsilon \left| \sum_{w \in \mathcal{A}} e(\alpha_k w^k + \gamma_{k-1} w^{k-1} + \ldots + \gamma_1 w) \right|. \]

An application of Hölder’s inequality yields the bound
\[ |f_k(\alpha; X)|^{2s} \ll X^{\varepsilon} \left| \sum_{p \in \mathcal{P}} \sum_{1 \leq y \leq X/M} e(\alpha_k (py)^k + \gamma_{k-1} (py)^{k-1} + \ldots + \gamma_1 (py)) \right|^{2s} \]
\[ \ll X^{\varepsilon} M^{2s-1} \left| \sum_{p \in \mathcal{P}} \sum_{1 \leq y \leq X/M} e(\alpha_k (py)^k + \gamma_{k-1} (py)^{k-1} + \ldots + \gamma_1 (py)) \right|^{2s}. \]

\[ (17.1) \]

We next classify the primes \(p \in \mathcal{P}\) in order to facilitate the exploitation of a spacing condition. When \((h,q) = 1\), the number of solutions of the congruence \(x^k \equiv h \pmod q\)
with \(1 \leq x \leq q\) is \(O(q^\epsilon)\). In order to confirm this claim, note that for each prime power \(p^j\), when \((h, p) = 1\), the number of solutions of the congruence \(x^k \equiv h \pmod{p^j}\) with \(1 \leq x \leq p^j\) is at most \(2k\). Thus, it follows from the Chinese Remainder Theorem that there are at most

\[
O((2k)^{\omega(q)}) \ll q^\epsilon
\]

possible solutions of the congruence \(x^k \equiv h \pmod{q}\) with \(1 \leq x \leq q\) subject to \((h, q) = 1\).

We deduce that there is an integer \(L\) with \(1 \leq L \ll q^\epsilon\) having the property that the primes \(p \in \mathcal{P}\) can be partitioned into \(L\) subsets \(\mathcal{P}_1, \ldots, \mathcal{P}_L\) satisfying the property that whenever \(1 \leq j \leq L\) and \(p_1, p_2 \in \mathcal{P}_j\), then \(p_1^k \equiv p_2^k \pmod{q}\) if and only if \(p_1 \equiv p_2 \pmod{q}\). As a consequence of (17.1), we now perceive that there is an index \(j\) with \(1 \leq j \leq L\) for which

\[
|f_k(\alpha; X)|^{2s} \ll (qX)^{\epsilon} M^{2s-1} \sum_{p \in \mathcal{P}_j} \left| \sum_{1 \leq y \leq X/M} e(\alpha_k(py)^k + \gamma_{k-1}(py)^{k-1} + \ldots + \gamma_1(py)) \right|^{2s}.
\]

Next we observe that

\[
\left( \sum_{1 \leq y \leq X/M} e(\alpha_k(py)^k + \gamma_{k-1}(py)^{k-1} + \ldots + \gamma_1(py)) \right)^s
= \sum_{1 \leq y_1, \ldots, y_s \leq X/M} e(\alpha_k p^k(y_1^k + \ldots + y_s^k) + \gamma_{k-1} p^{k-1}(y_1^{k-1} + \ldots + y_s^{k-1}) + \ldots + \gamma_1 p(y_1 + \ldots + y_s))
= \sum_{1 \leq h_1 \leq s(X/M)} \ldots \sum_{1 \leq h_k \leq s(X/M)^k} b(h)e(\alpha_k p^k h_k + \gamma_{k-1} p^{k-1} h_{k-1} + \ldots + \gamma_1 ph_1),
\]

where \(b(h)\) denotes the number of integral solutions of the system of equations

\[
\sum_{i=1}^s y_i^j = h_j \quad (1 \leq j \leq k),
\]

with \(1 \leq y_1, \ldots, y_s \leq X/M\). Thus, if we write \(H_j = s(X/M)^j\) \((1 \leq j \leq k)\) and

\[
\phi(\gamma; h; p) = \gamma_{k-1} h_{k-1} p^{k-1} + \ldots + \gamma_1 h_1 p,
\]

then we deduce that

\[
|f_k(\alpha; X)|^{2s} \ll (qX)^{\epsilon} M^{2s-1} \left( \sum_{h_1=1}^{H_1} \ldots \sum_{h_{k-1}=1}^{H_{k-1}} \sum_{h_k=1}^{H_k} b(h)e(\alpha_k p^k h_k + \phi(\gamma; h; p)) \right)^2.
\]

By applying Cauchy’s inequality, therefore, we infer that

\[
|f_k(\alpha; X)|^{2s} \ll (qX)^{\epsilon} M^{2s-1} \left( \sum_{h_1=1}^{H_1} \ldots \sum_{h_{k-1}=1}^{H_{k-1}} \sum_{h_k=1}^{H_k} b(h)e(\alpha_k p^k h_k) \right)^2
\ll (qX)^{\epsilon} M^{2s-1}(X/M)^{k(k-1)/2} \sum_{h_1=1}^{H_1} \ldots \sum_{h_{k-1}=1}^{H_{k-1}} S,
\]

where

\[
S = \sum_{p \in \mathcal{P}_j} \left| \sum_{h_k=1}^{H_k} b(h)e(\alpha_k p^k h_k) \right|^2.
\]
We now seek to apply the large sieve inequality to estimate $S$. When $p_1, p_2 \in \mathcal{P}_j$ and $p_1 \not\equiv p_2 \pmod{q}$, we have $p_1^k \not\equiv p_2^k \pmod{q}$. Then since $|q\alpha_k - a| \leq M^{-k}$, we see that
\[
\|\alpha_k(p_1^k - p_2^k)\| \geq \left\| \frac{a}{q}(p_1^k - p_2^k) \right\| - \frac{1}{q}M^k \geq \frac{1}{q} - \frac{1}{2q} = \frac{1}{2q}.
\]
(17.2)

We divide into cases:

(i) Suppose first that $q > M$. Then the elements of $\mathcal{P}_j$ are distinct modulo $q$, and it follows from (17.2) that for $p \in \mathcal{P}_j$, the points $\alpha_k p^k$ are spaced apart modulo 1 by at least $\frac{1}{2}q^{-1}$.

(ii) Suppose that $q \leq M$. Then our hypotheses allow us to assume that $|q\alpha_k - a| > M X^{-k}$. Then, given any two distinct elements $p_1, p_2 \in \mathcal{P}_j$ with $p_1 \not\equiv p_2 \pmod{q}$, we may conclude as in case (i) that $\alpha_k p_1^k$ and $\alpha_k p_2^k$ are spaced apart modulo 1 by at least $\frac{1}{2}q^{-1}$. Otherwise, when $p_1 \equiv p_2 \pmod{q}$ but $p_1 \not\equiv p_2$, we have $q|(p_1 - p_2)$ and
\[
\|\alpha_k(p_1^k - p_2^k)\| = \|\alpha_k(a/q)(p_1^k - p_2^k)\| = |\alpha_k - a/q| \cdot |p_1^k - p_2^k| \geq |\alpha_k - a/q|(2^{-1/(k-1)}M)^{k-1}q > \frac{1}{2}(X/M)^{-k}.
\]

Thus, for $p \in \mathcal{P}_j$, the points $\alpha_k p^k$ are spaced apart modulo 1 by at least
\[
\frac{1}{2} \min\{q^{-1}, (X/M)^{-k}\}.
\]

Applying now the large sieve inequality, we therefore discern that
\[
\sum_p \left| \sum_{h_k=1}^{H_k} b(h)e(\alpha_k p^k h_k) \right|^2 \ll (q + (X/M)^k) \sum_{h_k=1}^{H_k} |b(h)|^2,
\]
whence
\[
|f_k(\alpha; X)|^{2s} \ll (qX)^{\varepsilon} M^{2s-1}(X/M)^{k(k-1)/2} \left( \sum_{h_1=1}^{H_1} \ldots \sum_{h_k=1}^{H_k} |b(h)|^2 \right) (q + (X/M)^k)
\]
\[
\ll (qX)^{\varepsilon} M^{2s-1}(X/M)^{k(k-1)/2} J_{s,k}(X/M)(q + (X/M)^k)
\]
\[
\ll (qX)^{\varepsilon} M^{2s-1}(X/M)^{k(k-1)/2}(X/M)^{2s-\frac{1}{2}k(k+1)+\Delta_{s,k}}(q + (X/M)^k).
\]

Thus
\[
|f_k(\alpha; X)|^{2s} \ll (qX)^{\varepsilon} \left( M^{2s-1}(X/M)^{2s+\Delta_{s,k}}(1 + q(X/M)^{-k}) \right),
\]
which leads us to the conclusion
\[
|f_k(\alpha; X)| \ll (qX)^{\varepsilon} X \left( M^{-1}(X/M)^{\Delta_{s,k}}(1 + q(X/M)^{-k}) \right)^{1/(2s)}.
\]

This completes the proof of the lemma. \qed

Before discussing the consequences of this new bound of Weyl-type, we pause to recall our earlier notation concerning Hardy-Littlewood dissections. Let $\delta$ be a real number with $0 < \delta < 1$, and take $\mathcal{M}_\delta$ to be the union of the intervals
\[
\mathcal{M}_\delta(q, a) = \{ \alpha \in [0, 1) : |\alpha - a/q| \leq X^{\delta-k} \},
\]
with $0 \leq a < q \leq X^\delta$ and $(a, q) = 1$. Also, write $\mathcal{M}_\delta = [0, 1) \setminus \mathcal{M}_\delta$.
Corollary 17.3. When $0 < \delta \leq 1/2$ and the exponent $\Delta_{s,k}$ is permissible, one has
\[
\sup_{\alpha \in \mathfrak{m}_\delta} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right| \ll X^{1-\sigma+\varepsilon},
\]
where
\[
\sigma = \frac{\delta - (1 - \delta) \Delta_{s,k}}{2s}.
\]

Proof. We apply Lemma 17.2 with $M = X^\delta$. By Dirichlet’s approximation theorem, given $\alpha \in \mathfrak{m}_\delta$, there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $0 \leq a \leq q \leq X^{(1-\delta)}$, $(a, q) = 1$ and $|qa - a| \leq X^{k(\delta-1)}$. If one were to have both $q \leq X^\delta$ and $|q\alpha - a| \leq X^{\delta - k}$, then necessarily $\alpha \in \mathfrak{m}_\delta$. Thus $q > M$ or $|q\alpha - a| > MX^{-k}$, and so it follows from Lemma 17.2 that
\[
\left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right| \ll q^{\varepsilon} X^{1+\varepsilon} \left( X^{-\delta} (X^{1-\delta})^{\Delta_{s,k}} (1 + qX^{k(\delta-1)}) \right)^{1/(2s)}
\]
\[
\ll X^{1-\sigma+2\varepsilon},
\]
where we have written
\[
\sigma = \frac{\delta - \Delta_{s,k}(1 - \delta)}{2s}.
\]
This completes the proof of the corollary. \hfill \square

Observe that when $\delta = 1/2$, this conclusion shows that
\[
\sup_{\alpha \in \mathfrak{m}_{1/2}} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right| \ll X^{1-\sigma(k)+\varepsilon},
\]
where
\[
\sigma(k) = \frac{1 - \Delta_{s,k}}{4s} \geqslant \sup_{r \in \mathbb{N}} \frac{1 - \frac{1}{2} k^2 (1 - 1/k)^r}{4rk}.
\]
In fact, with additional work, one can obtain a bound of the above shape with the exponent $\sigma(k)$ replaced by
\[
\sigma(k) = \sup_{s \geq 1} \frac{1 - \Delta_{s,k-1}}{2s}.
\]
By applying the (now proven) Main Conjecture in Vinogradov’s Mean Value Theorem, this gives the permissible exponent
\[
\sigma(k) = \frac{1}{k(k-1)},
\]
which is superior to Weyl’s inequality for $k \geq 6$. We’ll opt for a weaker bound within the scope of what has been established within the course.

Corollary 17.4. For large values of $k$, one has
\[
\sup_{\alpha \in \mathfrak{m}_{1/8}} \left| \sum_{1 \leq x \leq X} e(\alpha x^k) \right| \ll X^{1-\tau+\varepsilon},
\]
where
\[
\tau = \frac{1}{32k^2 (\log k + O(\log \log k))}.
\]
Proof. We apply Corollary 17.3 with $\delta = 1/8$ and take $s = k\lceil 2k(\log k + \log \log k) \rceil$. Thus we obtain the permissible exponent

$$\tau = \frac{1}{8} - \frac{7}{16} k^2 (1 - 1/k) 2^{k(\log k + \log \log k)}$$

$$\geq \frac{1}{8} - \frac{7}{16} k^2 e^{-2 \log k - 2 \log \log k}$$

$$\geq \frac{1}{4k^2(\log k + O(\log \log k))}.$$ To obtain the permissible exponent $\tau = 1/8 - 7/16 k^2 (1 - 1/k) 2^{k(\log k + \log \log k)}$, we take $s = k\lceil k(1 - 1/k) 2^{k(\log k + \log \log k)} \rceil$. Thus we obtain the permissible exponent

$$\tau = \frac{1}{8} - \frac{7}{16} k^2 e^{-2 \log k - 2 \log \log k}$$

$$\geq \frac{1}{4k^2(\log k + O(\log \log k))}.$$ We now obtain

$$G(k) \leq \inf_{s \geq 0} \left( 2s + \max \left\{ 4k, \left[ \frac{k(1 - 1/k)^s}{\delta 2^{1-k}} \right] \right\} \right).$$

Taking $s = \lceil 2k(\log k + \log \log k) \rceil$, we see that

$$1 + 32k^3(\log k + O(\log \log k)) e^{-s/k} \leq 1 + \frac{32k^3(\log k + O(\log \log k))}{k^2(\log k)^2}$$

$$= 32k/ \log k + O(k \log \log k/(\log k)^2),$$

and hence

$$G(k) \leq 2\lceil 2k(\log k + \log \log k) \rceil + 4k$$

$$= 4k(\log k + O(\log \log k)).$$

This new bound of Weyl-type has immediate applications to Waring’s problem, upon exploiting our previous work.

**Theorem 17.5.** When $k$ is large, one has

$$G(k) \leq 4k(\log k + O(\log \log k)).$$

**Proof.** We apply the argument underlying the proof of Theorem 13.2 above, though now with the exponent

$$\tau = \frac{1}{32k^2(\log k + O(\log \log k))}$$

in place of Weyl’s exponent $\delta 2^{1-k}$. In Theorem 13.2 we showed that

$$G(k) \leq \inf_{s \geq 0} \left( 2s + \max \left\{ 4k, \left[ \frac{k(1 - 1/k)^s}{\delta 2^{1-k}} \right] \right\} \right).$$

We now obtain

$$G(k) \leq \inf_{s \geq 0} \left( 2s + \max \left\{ 4k, 1 + 32k^3(\log k + O(\log \log k)) e^{-s/k} \right\} \right).$$

Taking $s = \lceil 2k(\log k + \log \log k) \rceil$, we see that

$$1 + 32k^3(\log k + O(\log \log k)) e^{-s/k} \leq 1 + \frac{32k^3(\log k + O(\log \log k))}{k^2(\log k)^2}$$

$$= 32k/ \log k + O(k \log \log k/(\log k)^2),$$

and hence

$$G(k) \leq 2\lceil 2k(\log k + \log \log k) \rceil + 4k$$

$$= 4k(\log k + O(\log \log k)).$$

Similarly, we can improve our bound for $\tilde{G}(k)$, the least number $s$ such that the anticipated asymptotic formula in Waring’s problem holds, namely

$$R_{s,k}(n) = \text{card} \{ x \in \mathbb{N}^s : x_1^k + \ldots + x_s^k = n \}$$

$$= \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \left( \mathcal{S}_{s,k}(n) + o(1) \right) n^{s/k-1}.$$
Proof. As usual, we write
\[ f(\alpha) = \sum_{1 \leq x, \alpha \in X} e^{(\alpha x^k)}. \]
Then whenever \( r, t \in \mathbb{N} \) and \( s = rk \), we find by taking \( \delta = 1/8 \) that
\[
\int_{m} |f(\alpha)|^{2s+t} \, d\alpha \ll \left( \sup_{\alpha \in m} |f(\alpha)| \right)^t \int_{0}^{1} |f(\alpha)|^{2s} \, d\alpha \ll \left( X^{1-1/(32k^2(\log k + O(\log \log k)))} \right)^t X^{2s-k + \frac{1}{2}k^2(1-1/k)^r} = o(X^{2s+t-k}),
\]
provided only that
\[
\frac{t}{32k^2(\log k + O(\log \log k))} > \frac{1}{2}k^2 e^{-r/k}.
\]
This is equivalent to the condition that
\[
t > 16k^4(\log k + O(\log \log k))e^{-r/k}.
\]
Thus, on taking \( r = \lceil 2k(\log k + \log \log k) \rceil \), we find that our condition is satisfied provided that
\[
t > 16k^4(\log k + O(\log \log k))e^{-2(\log k + \log \log k)} = 16k^2 / \log k + O(k^2(\log k)/(\log k)^2).
\]
With these choices of \( r \) and \( t \), we have
\[
2s + t = 2rk + t \leq 4k^2(\log k + \log \log k) + O(k^2/\log k)
\]
\[
= 4k^2(\log k + O(\log \log k)).
\]
Applying the Hardy-Littlewood method with \( X = n^{1/k} \), we find that
\[
R_{2s+t,k}(n) = \int_{m} f(\alpha)^{2s+t} e(-n\alpha) \, d\alpha + \int_{m} f(\alpha)^{2s+t} e(-n\alpha) \, d\alpha = o(X^{2s+t-k}) + \frac{\Gamma(1+1/k)^{2s+t}}{\Gamma((2s+t)/k)} S_{2s+t,k}(n)n^{(2s+t)/k-1} + o(n^{(2s+t)/k-1})
\]
\[
= \frac{\Gamma(1+1/k)^{2s+t}}{\Gamma((2s+t)/k)} (S_{2s+t,k}(n) + o(1)) n^{(2s+t)/k-1}.
\]
Thus
\[
\tilde{G}(k) \leq 2s + t = 4k^2(\log k + O(\log \log k)).
\]
\[ \square \]
As we have previously pointed out, so far as the sharpest known bounds for \( G(k) \) and \( \tilde{G}(k) \) for large values of \( k \) are concerned, one has
\[
G(k) \leq k(\log k + \log \log k + 2 + o(1))
\]
and
\[
\tilde{G}(k) \leq k^2 - k + 2\lfloor \sqrt{2k+2} \rfloor - 1.
\]
We now consider the solubility of Diophantine inequalities (inequalities to be solved over the integers), focusing on the example

\[ |\lambda_1 x_1^k + \ldots + \lambda_s x_s^k| < \varepsilon, \]

with \( \lambda_i \in \mathbb{R} \setminus \{0\}, \varepsilon > 0 \) fixed, and \( x_i \in \mathbb{Z} \). Notice that when the \( \lambda_i \) are all in rational ratio, then we may multiply through by an appropriate real number \( \rho > 0 \) to obtain an inequality of the shape

\[ |c_1 x_1^k + \ldots + c_s x_s^k| < \rho \varepsilon, \]

and this is (morally speaking, at least when \( \varepsilon \) is sufficiently small) equivalent to the equation \( c_1 x_1^k + \ldots + c_s x_s^k = 0 \).

**Theorem 18.1** (Davenport and Heilbronn, 1946). Suppose that \( s \geq 2^k + 1 \), and that \( \lambda_1, \ldots, \lambda_s \in \mathbb{R} \setminus \{0\} \) are coefficients not all in rational ratio, and not all of the same sign in case \( k \) is even. Then for every positive number \( \varepsilon \), there exist infinitely many integral \( s \)-tuples \( x \) with \( |\lambda_1 x_1^k + \ldots + \lambda_s x_s^k| < \varepsilon \).

Before advancing to consider the proof of this theorem, we make some simplifying observations:

(i) If the theorem holds with \( \varepsilon = 1 \), then it holds for arbitrary \( \varepsilon \), for we can replace \( \lambda_i \) by \( \lambda_i/\varepsilon \) for each \( i \), and then consider

\[ \left| \frac{\lambda_1}{\varepsilon} x_1^k + \ldots + \frac{\lambda_s}{\varepsilon} x_s^k \right| < 1. \]

(ii) If \( k \) is odd, we can replace \( x_i^k \) by \((-x_i)^k\), if necessary, and thus assume that not all of the \( \lambda_j \) have the same sign.

(iii) By relabelling variables, if necessary, we may suppose that \( \lambda_1/\lambda_2 \not\in \mathbb{Q} \).

(iv) If \( \lambda_1/\lambda_2 > 0 \), then we may find \( \lambda_j \) with \( \lambda_1/\lambda_j < 0 \). If \( \lambda_1/\lambda_j \in \mathbb{Q} \) and \( \lambda_2/\lambda_j \in \mathbb{Q} \), then we have \( \lambda_1/\lambda_2 \in \mathbb{Q} \), contradicting the observation (iii). Thus \( \lambda_2/\lambda_j \not\in \mathbb{Q} \) in this situation, and

\[ \frac{\lambda_2}{\lambda_j} = \frac{\lambda_2}{\lambda_1} \cdot \frac{\lambda_1}{\lambda_j} < 0. \]

Then in any case we may suppose that \( \lambda_1/\lambda_2 \not\in \mathbb{Q} \) and \( \lambda_1/\lambda_2 < 0 \).

We now prepare an analogue of the orthogonality relation for inequalities. This runs along similar lines to our treatment of the singular series for diagonal equations.

Observe that when \( \omega \neq 0 \), the Fourier transform of the function

\[ G(t) = \begin{cases} 1 - |t|, & \text{when } |t| < 1, \\ 0, & \text{when } |t| \geq 1, \end{cases} \]
is given by
\[
\tilde{G}(\omega) = \int_0^1 (1 - t)e(t\omega) \, dt + \int_{-1}^0 (1 + t)e(t\omega) \, dt
\]
\[
= \left[ \frac{(1 - t)e(t\omega)}{2\pi i\omega} \right]_0^1 + \int_0^1 \frac{e(t\omega)}{2\pi i\omega} \, dt + \left[ \frac{(1 + t)e(t\omega)}{2\pi i\omega} \right]_{-1}^0 - \int_{-1}^0 \frac{e(t\omega)}{2\pi i\omega} \, dt
\]
\[
= -\frac{1}{2\pi i\omega} + \frac{e(\omega) - 1}{(2\pi i\omega)^2} + \frac{1}{2\pi i\omega} - \frac{1 - e(-\omega)}{(2\pi i\omega)^2},
\]
whence
\[
\tilde{G}(\omega) = \frac{2 - 2\cos(2\pi\omega)}{4\pi^2\omega^2} = \left( \frac{\sin(\pi\omega)}{\pi\omega} \right)^2.
\]
When \( \omega = 0 \) we obtain \( \tilde{G}(\omega) = 1 \).

Write
\[
K(\alpha) = \left( \frac{\sin(\pi\alpha)}{\pi\alpha} \right)^2.
\]
Then it follows that
\[
\int_{-\infty}^\infty e(\alpha\beta)K(\alpha) \, d\alpha = \max\{1 - |\beta|, 0\}.
\]
This is not quite an orthogonality-based detector, but it does give a lower bound for the characteristic function on \([-1, 1]\), and it has superior convergence properties to the kernel
\[
K_0(\alpha) = \frac{\sin(2\pi\alpha)}{\pi\alpha}.
\]

We may now say more concerning our strategy for establishing Theorem 18.1. As usual, we put
\[
f(\alpha) = \sum_{1 \leq x \leq X} e(\alpha x^k),
\]
and then write
\[
f_j(\alpha) = f(\lambda_j\alpha) \quad (1 \leq j \leq s).
\]
We seek a lower bound for the integral
\[
R_s(X) = \int_{-\infty}^\infty f_1(\alpha) \cdots f_s(\alpha) K(\alpha) \, d\alpha,
\]
valid for all large \( X \). Here it is worth remarking that Davenport and Heilbronn proved a lower bound valid only for an infinite sequence of values of \( X \), say \( X = X_k \ (k \in \mathbb{N}) \), with \( X_k \to \infty \) as \( k \to \infty \). Since
\[
R_s(X) = \sum_{1 \leq x_1 \leq X} \cdots \sum_{1 \leq x_s \leq X} \max\{0, 1 - |\lambda_1 x_1^k + \cdots + \lambda_s x_s^k|\},
\]
then whenever \( R_s(X) > 0 \), one necessarily has an integral solution of the inequality \( |\lambda_1 x_1^k + \cdots + \lambda_s x_s^k| < 1 \). Indeed, one has
\[
R_s^*(X) = \sum_{1 \leq x_1 \leq X} \cdots \sum_{1 \leq x_s \leq X} 1 \geq R_s(X).
\]
We will prove the lower bound on $R_s(X)$ contained in the following theorem due to Freeman [10].

**Theorem 18.2.** Suppose that $s > 2^k$, and that $\lambda_1, \ldots, \lambda_s$ satisfy the hypotheses of the statement of Theorem 18.1. Then for all large $X$ one has $R_s(X) \gg_{\lambda,s,k} X^{s-k}$.

19. The Hardy-Littlewood dissection

The precise definition of our dissection of the real line will be elucidated in due course. For the moment it may be useful to think temporarily of a division into three subsets:

(i) (major arc) the set of $\alpha \in \mathbb{R}$ with $\alpha$ very close to 0, say $|\alpha| \leq X^{\delta-k}$. It is to be hoped that the dominant contribution arises from this major arc.

(ii) (trivial arcs) the set of $\alpha \in \mathbb{R}$ with $|\alpha|$ very large, say $|\alpha| \geq X^\delta$. These arcs give a small contribution, owing to the estimate $K(\alpha) \ll |\alpha|^{-2}$.

(iii) (minor arcs) the set of $\alpha \in \mathbb{R}$ of moderate size, say $X^{\delta-k} < |\alpha| < X^\delta$. Here, we must make use of the irrationality of $\lambda_1/\lambda_2$ in order to show that $|f(\lambda_1\alpha)f(\lambda_2\alpha)|$ is small.

In order to put this dissection and the associated strategy into action, we must make some subtle adjustments, and it is here that the innovations of Freeman (inspired by work of Bent’kus and Götze) play a role.

**Lemma 19.1.** Suppose that $X$ is sufficiently large, and that $|f(\alpha)| \geq \gamma X$, where

$$X^{-2^{-k-1}} \leq \gamma \leq 1.$$  

Then there exist $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ with $(a, q) = 1$, and absolute constants $c_0$ and $c_1$, with

$$q \leq c_0 \gamma^{-2k} \quad \text{and} \quad |q\alpha - a| \leq c_1 \gamma^{-2k} X^{-k}.$$

**Proof.** Suppose first that there exist $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ with $(b, r) = 1$, $1 \leq r \leq X^{1/2}$ and $|r\alpha - b| \leq X^{1/2-k}$. Then our discussion of the major arc analysis in the circle method shows that

$$f(\alpha) - r^{-1}S(r, b)v(\alpha - b/r) \ll r + X^k|r\alpha - b| \ll X^{1/2},$$

whence

$$f(\alpha) \ll \epsilon X(r + X^k|r\alpha - b|)^{-1/k} + X^{1/2}.$$

If we take $\epsilon = 1/(2k)$ here, then the constant becomes absolute and we see that the hypothesis $|f(\alpha)| \geq \gamma X$ implies that for a suitable positive number $A_1$, one has

$$\gamma X \leq |f(\alpha)| \leq A_1 (X(r + X^k|r\alpha - b|)^{-1/(2k)} + X^{1/2}) \leq 2A_1 X(r + X^k|r\alpha - b|)^{-1/(2k)},$$

whence

$$r + X^k|r\alpha - b| \leq (2A_1/\gamma)^{2k}.$$

Consequently, one has

$$r \leq (2A_1)^{2k}\gamma^{-2k} \quad \text{and} \quad |r\alpha - b| \leq (2A_1)^{2k}\gamma^{-2k} X^{-k}. \quad (19.1)$$

When no such integers $b$ and $r$ exist, meanwhile, it follows from Dirichlet’s approximation theorem nonetheless that one can find $r \in \mathbb{N}$ with $1 \leq r \leq X^{k-1/2}$, $(a, r) = 1$ and $|r\alpha - b| \leq X^{1/2-k}$. One must then have $r > X^{1/2}$, for otherwise we would be in the situation just examined. In the present circumstances, it follows from Weyl’s inequality that

$$f(\alpha) \ll X^{1+\epsilon}(r^{-1} + X^{-1} + rX^{-k})^{21-k} \ll X^{1-2^{-k+\epsilon}} \ll X^{1-(2/3)2^{-k}}.$$
Thus, when $X$ is large enough, one has
\[ |f(\alpha)| < \frac{1}{2} X^{1-2^{-k}} \leq 1/2 \gamma X. \]
This contradicts the hypotheses of the statement of the lemma, and hence we are forced to conclude that (19.1) holds.

This lemma shows that large values of Weyl sums imply good rational approximations to the associated coefficients. We can exploit this observation to reveal the consequences of the ratio $\lambda_1/\lambda_2$ being irrational.

**Lemma 19.2.** Suppose that $S$ and $T$ are fixed real numbers with $0 < S \leq 1 \leq T$. Then one has
\[
\lim_{X \to \infty} \sup_{S \leq |\alpha| \leq T} \frac{f_1(\alpha; X) f_2(\alpha; X)}{X^2} = 0.
\]

**Proof.** If the desired conclusion fails, then we can find $\varepsilon > 0$, and a sequence of positive numbers $(X_n)_{n=1}^{\infty}$, tending monotonically to $+\infty$, with a corresponding sequence of real numbers $(\alpha_n)_{n=1}^{\infty}$, satisfying
\[ S \leq |\alpha_n| \leq T \ (n \geq 1), \]
such that
\[ |f_1(\alpha_n; X_n) f_2(\alpha_n; X_n)| \geq \varepsilon X_n^2 \]
for all $n \in \mathbb{N}$. But since $|f_i(\alpha; X)| \leq X$, it follows that $|f_i(\alpha_n; X_n)| \geq \varepsilon X_n$ for $i = 1$ and 2. Whenever $n$ is large enough that $X_n \geq \varepsilon^{-2^{k+1}}$, we may apply Lemma 19.1 with $\gamma = \varepsilon$.

Thus we obtain integers $q_{in} \in \mathbb{N}$ and $a_{in} \in \mathbb{Z}$ with $(q_{in}, a_{in}) = 1,$
\[ 1 \leq q_{in} \leq c_0 \varepsilon^{-2k} \quad \text{and} \quad |\lambda_i \alpha_n q_{in} - a_{in}| \leq c_1 \varepsilon^{-2k} X_n^{-k}. \]
Consequently, when $n$ is sufficiently large, one has
\[ |a_{in}| \leq c_0 \varepsilon^{-2k} |\lambda_i| T + c_1 \varepsilon^{-2k} X_n^{-k} \ll \varepsilon 1 \]
and
\[ |q_{in}| \leq c_0 \varepsilon^{-2k} \ll \varepsilon 1. \]
There are therefore only finitely many possible 4-tuples $(a_{1n}, q_{1n}, a_{2n}, q_{2n})$. Hence there must be one of these 4-tuples, say $(a_1, q_1, a_2, q_2)$, that occurs for infinitely many values of $n$, say for the sequence $(n_i)_{i=1}^{\infty}$, and these indices are associated with the subsequence $(\alpha_{n_i})_{i=1}^{\infty}$. The compactness of $[S, T]$ then ensures that this subsequence $(\alpha_{n_i})_{i=1}^{\infty}$ contains a convergent sub-subsequence $(\alpha_{n_{i_t}})_{t=1}^{\infty}$ with some non-zero limit $\alpha_0$. Furthermore, for every value of $t$, one has
\[ |\lambda_i \alpha_{n_{i_t}} q_i - a_i| \leq c_1 \varepsilon^{-2k} X_{n_{i_t}}^{-k} \quad (i = 1, 2). \]
Thus, on taking the limit as $t \to \infty$, we deduce that
\[ \lambda_i \alpha_0 q_i = a_i \quad (i = 1, 2), \]
whence
\[ \frac{\lambda_1}{\lambda_2} = \frac{a_1 q_2}{a_2 q_1} \in \mathbb{Q}. \]
This contradicts our assumption that \( \lambda_1/\lambda_2 \notin \mathbb{Q} \), and hence we deduce that for each \( \alpha \in \mathbb{R} \) with \( S \leq |\alpha| \leq T \), one has
\[
\lim_{X \to +\infty} \frac{|f_1(\alpha; X)f_2(\alpha; X)|}{X^2} = 0.
\]

The conclusion of the lemma is now immediate. \( \square \)

We now reach the point at which we are able to define the functions used to define the minor arcs.

**Lemma 19.3.** There exist positive functions \( S(X) \) and \( T(X) \), depending only on \( \lambda_1 \) and \( \lambda_2 \), such that
\[
\lim_{X \to +\infty} S(X) = 0 \quad \text{and} \quad \lim_{X \to +\infty} T(X) = +\infty,
\]
and such that
\[
\lim_{X \to +\infty} \sup_{S(X) \leq |\alpha| \leq T(X)} \frac{|f_1(\alpha; X)f_2(\alpha; X)|}{X^2} = 0.
\]

**Proof.** For every positive integer \( m \), it follows from Lemma 19.2 that there is a real number \( X_m > 0 \) such that
\[
\frac{|f_1(\alpha; X)f_2(\alpha; X)|}{X^2} \leq \frac{1}{m}
\]
whenever \( X \geq X_m \) and \( 1/m \leq |\alpha| \leq m \). We may plainly assume that the sequence \( (X_m) \) is non-decreasing. We define the functions \( S(X) \) and \( T(X) \) for \( X_m \leq X < X_{m+1} \) by putting
\[
S(X) = 1/m \quad \text{and} \quad T(X) = m.
\]
Then one sees that whenever \( X \geq X_m \) and \( S(X) \leq |\alpha| \leq T(X) \), one has
\[
\frac{|f_1(\alpha; X)f_2(\alpha; X)|}{X^2} \leq \frac{1}{m}.
\]

The conclusion of the lemma follows immediately. \( \square \)

We are now equipped, at last, to define our dissection of the real line. We define the *trivial arcs*
\[
t = \{ \alpha \in \mathbb{R} : |\alpha| > T(X) \},
\]
the corresponding *major arc*
\[
\mathcal{M} = \{ \alpha \in \mathbb{R} : |\alpha| \leq \lambda^{-1}X^{1/8-k} \},
\]
in which we write \( \lambda = \max_{1 \leq i \leq s} |\lambda_i| \), and the associated *minor arcs*
\[
\mathcal{m} = \{ \alpha \in \mathbb{R} : \lambda^{-1}X^{1/8-k} < |\alpha| \leq T(X) \}.
\]
Thus the real line \( \mathbb{R} \) is the disjoint union \( \mathbb{R} = t \cup \mathcal{M} \cup \mathcal{m} \).
20. The trivial arcs

We begin with an auxiliary lemma.

**Lemma 20.1.** Suppose that $\delta > 0$ and $s \geq 2^k + \delta$. Then one has

\[
\int_0^1 |f(\alpha)|^s \, d\alpha \ll X^{s-k}.
\]

**Proof.** We have

\[
\int_{m_1/s} |f(\alpha)|^s \, d\alpha \ll \left( \sup_{\alpha \in m_1/s} |f(\alpha)| \right)^{s-2^k} \int_0^1 |f(\alpha)|^{2^k} \, d\alpha \\
\ll (X^{1-2^{-k-2+\varepsilon}})^{s-2^k} X^{2^k - k + \varepsilon} = o(X^{s-k}).
\]

Also, since when $\alpha \in M_{1/8}(q, a) \subseteq M_1$, we have

\[
f(\alpha) = q^{-1} S(q, a) v(\alpha - a/q) + O(X^{1/4}),
\]

we see (making use of the homework exercises) that

\[
f(\alpha) \ll \frac{X}{(q + X^k |q\alpha - a|)^{s-1/k}},
\]

whence

\[
\int_{\mathcal{M}_{1/8}} |f(\alpha)|^s \, d\alpha \ll X^s \sum_{0 \leq a \leq q \leq X^{1/8}} q^{1-s/k+\varepsilon} \int_{-\infty}^{\infty} \frac{d\beta}{(1 + |\beta| X^k)^{s/k-\varepsilon}} \\
\ll X^{s-k}.
\]

Then

\[
\int_0^1 |f(\alpha)|^s \, d\alpha = \int_{m_1/s} |f(\alpha)|^s \, d\alpha + \int_{\mathcal{M}_{1/8}} |f(\alpha)|^s \, d\alpha \ll X^{s-k}.
\]

\[
\square
\]

**Lemma 20.2.** Whenever $s > 2^k$, one has

\[
\int f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha = o(X^{s-k}).
\]

**Proof.** One has

\[
\int |f_1(\alpha) \ldots f_s(\alpha) K(\alpha)| \, d\alpha \ll \sum_{n=0}^\infty \int_{T(X)+n}^{T(X)+n+1} |f_1(\alpha) \ldots f_s(\alpha)| \frac{d\alpha}{|\alpha|^2} \\
\ll \sum_{n=0}^\infty \frac{1}{(T(X)+n)^2} \prod_{i=1}^s \left( \int_{T(X)+n}^{T(X)+n+1} |f_i(\alpha)|^s \, d\alpha \right)^{1/s}.
\]
But when $c \in \mathbb{R}_{\geq 0}$, one has
\[
\int_c^{c+1} |f_i(\alpha)|^s \, d\alpha = \frac{1}{|\lambda_i|} \int_{|\lambda_i|c}^{|\lambda_i||(c+1)} |f(\beta)|^s \, d\beta \\
\leq \frac{1}{|\lambda_i|} |\lambda_i| \int_{1}^{c} |f(\beta)|^s \, d\beta \\
\ll_X X^{s-k}.
\]
Hence
\[
\int_1 f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha \ll X^{s-k} \sum_{n=0}^{\infty} \frac{1}{(T(X) + n)^2} \ll X^{s-k} T(X)^{-1} = o(X^{s-k}).
\]
\[\blacksquare\]

21. The minor arcs

We subdivide the minor arcs into two subsets, namely
\[m_1 = \{ \alpha \in m : |\alpha| \geq S(X) \}\]
and
\[m_2 = \{ \alpha \in m : |\alpha| < S(X) \}.
\]
When $\alpha \in m_1$, one has $S(X) \leq |\alpha| \leq T(X)$, and thus it follows from Lemma 19.3 that
\[
\sup_{\alpha \in m_1} |f_1(\alpha) f_2(\alpha)| = o(X^2). \tag{21.1}
\]
We now aim to deliver a similar conclusion with $m_2$ in place of $m_1$.

Suppose that $\alpha \in m_2$, and suppose (if possible) that $|f_i(\alpha)| \geq \gamma X$ with $\gamma = S(X)^{1/(2^{k+1}+1)}$. We may suppose that $S(X) > X^{-1}$ without loss, for we have the liberty to take $S(X)$ decreasing as slowly as necessary to 0. Thus we have $\gamma \geq X^{-1/(2^{k+1}+1)}$, and Lemma 19.1 shows that there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(q, a) = 1$ and such that
\[1 \leq q \leq c_0 \gamma^{-2k} \quad \text{and} \quad |\lambda_1 q \alpha - a| \leq c_1 \gamma^{-2k} X^{-k}.
\]
Thus
\[
|a| \leq c_1 \gamma^{-2k} X^{-k} + |\lambda_1 a| q \ll c_1 \gamma^{-2k} X^{-k} + c_0 |\lambda_1| S(X) \gamma^{-2k} \\
\ll X^{-1} + S(X)^{1-2k/(2^{k+1}+1)} \ll S(X)^{1/5}.
\]
Since the right hand side here tends to 0 as $X \to +\infty$, we conclude that $a = 0$ whenever $X$ is sufficiently large. Hence
\[
|\alpha| \ll \gamma^{-2k} X^{-k} \ll S(X)^{-1} X^{-k} < \lambda^{-1} X^1 X^{-8-k},
\]
and from this it follows that $\alpha \in \mathfrak{M}$. This yields a contradiction from which we conclude that
\[
\sup_{\alpha \in m_2} |f_1(\alpha)| < \gamma X = S(X)^{1/(2^{k+1}+1)} X = o(X),
\]
and hence
\[
\sup_{\alpha \in m_2} |f_1(\alpha) f_2(\alpha)| = o(X^2). \tag{21.2}
\]
By combining the bounds (21.1) and (21.2), we find that
\[
\sup_{\alpha \in \mathbb{m}} |f_1(\alpha) f_2(\alpha)| = o(X^2). \tag{21.3}
\]

We now return to handle the minor arcs.

**Lemma 21.1.** Whenever \( s > 2^k \), one has
\[
\int_m f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha = o(X^{s-k}).
\]

**Proof.** We have
\[
\int_m f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha \ll \left( \sup_{\alpha \in \mathbb{m}} |f_1(\alpha) f_2(\alpha)| \right)^{1/4} \int_m \left| f_3(\alpha) \ldots f_s(\alpha) K(\alpha) \right| \, d\alpha. \tag{21.4}
\]

By Hölder’s inequality, the integral on the right hand side here is
\[
\ll \sum_{n=0}^{\infty} \prod_{i=1}^{s} \left( \int_{n+\lambda^{-1}X^{1/8-k}} |f_i(\alpha)|^{s-1/2} \min\{1, |\alpha|^{-2}\} \, d\alpha \right)^{1/s_i},
\]
in which we write
\[
S_i = \begin{cases} (4/3)(s - 1/2), & \text{when } i = 1, 2, \\ (s - 1/2), & \text{when } 3 \leq i \leq s. \end{cases}
\]

On recalling Lemma 20.1 and noting that \( s \geq 2^k + 1 \), we find that this integral is
\[
\ll \lambda \sum_{n=0}^{\infty} \min\{1, (n + \lambda^{-1}X^{1/8-k})^{-2}\} \int_{0}^{1} |f(\beta)|^{s-1/2} \, d\beta 
\ll X^{s-k-1/2}.
\]

Noting next from (21.3) that
\[
\left( \sup_{\alpha \in \mathbb{m}} |f_1(\alpha) f_2(\alpha)| \right)^{1/4} = o(X^{1/2}),
\]
we conclude via (21.4) that
\[
\int_m f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha = o(X^{s-k}).
\]

\( \square \)

**22. The major arc contribution**

We seek now to establish the lower bound
\[
\int_{\mathfrak{M}} f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha \gg X^{s-k}.
\]

This is fairly routine, and in some ways resembles our analysis of the singular integral in our discussion of the solubility of diagonal equations. Observe first that for \( 1 \leq j \leq s \), and for each \( a \in \mathbb{Z} \) and \( q \in \mathbb{N} \), one has
\[
f_j(\alpha) = f(\lambda_j \alpha) = q^{-1} S(q, a)v(\lambda_j \alpha - a/q) + O(q + X^k |q \lambda_j \alpha - a|).
\]
On putting \( a = 0 \) and \( q = 1 \), we find that whenever \( \alpha \in \mathfrak{M} \), so that \(|\alpha| \leq \lambda^{-1}X^{1/8-k} \)
, one has
\[
f_j(\alpha) = v(\lambda_j \alpha) + O(X^{1/8}).
\]
It is convenient temporarily to write \( v_j(\alpha) = v(\lambda_j \alpha) \). Thus, for \( \alpha \in \mathfrak{M} \), we have
\[
f_1(\alpha) \ldots f_s(\alpha) - v_1(\alpha) \ldots v_s(\alpha) \ll (X^{1/8})^s + X^{1/8} \max_{1 \leq i \leq s} |v_i(\alpha)|^{s-1}
\ll X^{s-7/8},
\]
whence
\[
\int_{\mathfrak{M}} (f_1(\alpha) \ldots f_s(\alpha) - v_1(\alpha) \ldots v_s(\alpha)) K(\alpha) \, d\alpha \ll X^{s-7/8} \int_0^{\lambda^{-1}X^{1/8-k}} \, d\alpha
\ll X^{s-k-3/4} = o(X^{s-k}).
\]

Employing the estimate \( v_j(\alpha) \ll \chi X(1 + X^k |\alpha|)^{-1/k} \), we find next that
\[
\int_{\mathbb{R}\setminus\mathfrak{M}} v_1(\alpha) \ldots v_s(\alpha) K(\alpha) \, d\alpha \ll X^s \int_{\lambda^{-1}X^{1/8-k}}^{\infty} (1 + X^k \alpha)^{-s/k} \, d\alpha \ll X^{s-k-1/(8k)},
\]
whence
\[
\int_{\mathfrak{M}} f_1(\alpha) \ldots f_s(\alpha) K(\alpha) \, d\alpha = J + o(X^{s-k}),
\]
where we write
\[
J = \int_{-\infty}^{\infty} v_1(\alpha) \ldots v_s(\alpha) K(\alpha) \, d\alpha.
\]

We now observe that
\[
J = \int_0^X \ldots \int_0^X \int_{-\infty}^{\infty} e((\lambda_1 \gamma_1^k + \ldots + \lambda_s \gamma_s^k) \alpha) K(\alpha) \, d\gamma
\]
\[
= \int_0^X \ldots \int_0^X \max\{0, 1 - |\lambda_1 \gamma_1^k + \ldots + \lambda_s \gamma_s^k|\} \, d\gamma
\]
\[
= X^s \int_0^1 \ldots \int_0^1 \max\{0, 1 - X^k |\lambda_1 \beta_1^k + \ldots + \lambda_s \beta_s^k|\} \, d\beta
\]
\[
= X^s \int_0^1 \ldots \int_0^1 k^{-s} (u_1 \ldots u_s)^{1/k-1} \max\{0, 1 - X^k |\lambda_1 u_1 + \ldots + \lambda_s u_s|\} \, du.
\]

But since \( \lambda_1 / \lambda_2 < 0 \), we can find \( \eta \in (0, 1)^s \) with
\[
\lambda_1 \eta_1 + \ldots + \lambda_s \eta_s = 0,
\]
and moreover we can find an \((s-1)\)-dimensional neighbourhood of \( \eta \), say
\[
U_\sigma = \{ z \in (0, 1)^s : |z_i - \eta_i| < \sigma (2 \leq i \leq s) \text{ and } z_1 = -\lambda_1^{-1}(\lambda_2 z_2 + \ldots + \lambda_s z_s) \},
\]
with \( \sigma \) sufficiently small, for which \( \lambda_1 z_1 + \ldots + \lambda_s z_s = 0 \) whenever \( z \in U_\sigma \). The measure of this neighbourhood is positive, and given any \( z \in U_\sigma \), one has
\[
|\lambda_1 z_1 + \lambda_2 z_2 + \ldots + \lambda_s z_s| < \frac{1}{2} X^{-k}
\]
for $|z'_1 - z_1| < \frac{1}{2}|\lambda_1|^{-1}X^{-k}$. Thus we find that since
\[
\int_0^1 \cdots \int_0^1 du \gg X^{-k},
\]
then $J \gg X^{s-k}$. On recalling (22.1), we may conclude as follows.

**Lemma 22.1.** Whenever $s \geq 2k + 1$, one has
\[
\int_{\mathfrak{m}} f_1(\alpha) \cdots f_s(\alpha) K(\alpha) d\alpha \gg X^{s-k}.
\]

By combining the conclusions of Lemmata 20.2, 21.1 and 22.1, we deduce that
\[
R_s(X) = \int_I f_1(\alpha) \cdots f_s(\alpha) K(\alpha) d\alpha + \int_{\mathfrak{m}} f_1(\alpha) \cdots f_s(\alpha) K(\alpha) d\alpha
\]
\[
\gg X^{s-k} + o(X^{s-k}).
\]

Thus $R_s(X) \to \infty$ as $X \to \infty$, and this establishes Theorem 18.2, and hence also Theorem 18.1.

Note that we can refine the condition on $s$ in Theorems 18.1 and 18.2 by using our work on Vinogradov’s mean value theorem. Thus, with little additional work one would be able to establish that $R_s(X) \gg X^{s-k}$ for $s \geq (4 + o(1))k^2 \log k$ for large values of $k$. The use of smooth numbers permits such a conclusion to be established under the condition $s \geq (1 + o(1))k \log k$ by employing the work of [30, 31].

23. Unrepresentation theory

We now consider the set of integers not represented in some prescribed manner, a subject to which I like to refer as unrepresentation theory. We illustrate ideas with two problems. Recall that we have established the upper bound $\widetilde{G}(k) \leq 2^k + 1$, so that when $s \geq 2k + 1$ and $n$ is large, one has the asymptotic formula
\[
R_{s,k}(n) = \text{card}\{x \in \mathbb{N}^s : n = x_1^k + \cdots + x_s^k\}
\]
\[
\sim \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n)n^{s/k-1}.
\]

We define $E_{s,k}(N)$ to be the number of integers $n$ with $1 \leq n \leq N$ for which
\[
\left| R_{s,k}(n) - \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n)n^{s/k-1} \right| > n^{s/k-1}(\log n)^{-1}, \tag{23.1}
\]
where, as usual, we write
\[
\mathfrak{S}_{s,k}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} (q^{-1}S(q,a))^s e(-na/q).
\]

Notice that when $s \geq 2^k + 1$, then in view of our earlier remarks, we have $E_{s,k}(N) \ll 1$. Moreover, if $E_{s,k}(N) = o(N)$, then almost all integers $n$ with $1 \leq n \leq N$ (in the sense
of natural density) are the sum of $s$ $k$-th powers of natural numbers, with the expected asymptotic formula holding for the number of associated representations.

**Theorem 23.1.** Whenever $s \geq 2k^{-1} + 1$, there is a positive number $\tau$ with the property that $E_{s,k}(N) \ll N^{1-\tau}$.

In fact, by utilising the results of Vaughan [23, 24] and Wooley [36], one can show that $E_{s,k}(N) \ll N(\log N)^{-A}$, for a suitable $A > 0$, for $s \geq 2k^{-1}$ when $k \geq 3$, and further $E_{s,k}(N) \ll N^{1-\tau}$ for $s \geq \frac{1}{2}(k^2 - k + \lceil \sqrt{2k+2} \rceil)$.

**The proof of Theorem 23.1.** Let $Z(N)$ denote the set of integers $n$ with $N/2 < n \leq N$ for which (23.1) holds. Write $X = N^{1/k}$ and

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

Also, define $\mathcal{M}_\delta$ and $\mathfrak{m}_\delta$ as we have throughout the course, and take $\delta = 1/8$. Then, for $N/2 < n \leq N$ and $s \geq 2k + 1$, we have

$$\int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha = \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} \mathcal{E}_{s,k}(n)n^{s/k-1} + o(n^{s/k-1}(\log n)^2).$$

This formula also remains true for $s \geq k + 2$, though the proof of this statement does not lie within the discussion of this course. Observe that if (23.1) holds for some integer $n$ with $N/2 < n \leq N$, we have

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

whence

$$\left| \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \right| \geq \frac{1}{2} n^{s/k-1}(\log n)^{-1}. \quad (23.2)$$

Define $\eta_n$ to be 0 when (23.1) does not hold, and otherwise (when (23.1) does hold) define $\eta_n$ via the relation

$$\left| \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \right| = \eta_n \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha.$$

Thus, we have

$$|\eta_n| = \begin{cases} 1, & \text{when } n \in Z(N), \\ 0, & \text{when } n \notin Z(N). \end{cases}$$

Furthermore, in view of the relation (23.2), it follows that

$$\sum_{n \in Z(N)} \eta_n \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha = \sum_{n \in Z(N)} \left| \int_{\mathfrak{m}_\delta} f(\alpha)^s e(-n\alpha) \, d\alpha \right| \gg \text{card}(Z(N)) N^{s/k-1}(\log N)^{-1}.$$

Then if we write

$$K(\alpha) = \sum_{N/2 < n \leq N} \eta_n e(n\alpha),$$

we deduce that

$$\int_{\mathfrak{m}_\delta} f(\alpha)^s K(-\alpha) \, d\alpha \gg \text{card}(Z(N)) N^{s/k-1}(\log N)^{-1}. $$
On recalling that $N \asymp X^k$, it follows by applying Schwarz’s inequality that
\[
\text{card}(\mathcal{Z}(N))X^{s-k}(\log X)^{-1} \ll \left( \int_{m_3} |f(\alpha)|^{2s} \, d\alpha \right)^{1/2} \left( \int_0^1 |K(\alpha)|^2 \, d\alpha \right)^{1/2}.
\]
By Parseval’s identity,
\[
\int_0^1 |K(\alpha)|^2 \, d\alpha = \sum_{N/2 < n \leq N} |\eta_n|^2 = \text{card}(\mathcal{Z}(N)).
\]
Meanwhile, by Hua’s lemma and Weyl’s inequality, provided that $2s > 2^k$, we have
\[
\int_{m_3} |f(\alpha)|^{2s} \, d\alpha \ll \left( \sup_{\alpha \in m_3} |f(\alpha)| \right)^{2s-2^k} \int_0^1 |f(\alpha)|^{2^k} \, d\alpha
\ll \left( X^{1-\delta 2^{-k}+\epsilon} \right)^{2s-2^k} X^{2^k-k+\epsilon}
\ll X^{2s-k-\delta 2^{-k}}.
\]
Thus,
\[
\text{card}(\mathcal{Z}(N))X^{s-k}(\log X)^{-1} \ll (\text{card}(\mathcal{Z}(N)))^{1/2} (X^{2s-k-\delta 2^{-k}})^{1/2},
\]
whence
\[
\text{card}(\mathcal{Z}(N)) \ll X^{k-\delta 2^{-k}}(\log X)^2 \ll N^{1-\tau},
\]
for any $\tau < \delta 2^{-k}/k$. The desired conclusion follows on noting that
\[
E_{s,k}(N) \leq \text{card}(\mathcal{Z}(N)) + \text{card}(\mathcal{Z}(N/2)) + \ldots + \text{card}(\mathcal{Z}(N/2^l)),
\]
with $l \asymp \log N$, so that $E_{s,k}(N) \ll N^{1-\tau}$. \hfill \Box

It is known that almost all positive integers are the sum of 4 cubes of natural numbers, or of at most 9 fifth powers of natural numbers, and so on. For large $k$, almost all positive integers are the sum of at most $\frac{1}{2}k(\log k + \log \log k + 2 + o(1))$ positive integral $k$-th powers.

We turn next to the tougher problem of representing almost all values of a thin subsequence of the natural numbers. In order to test ideas, consider a quadratic polynomial $\phi(t) \in \mathbb{Z}[t]$ having positive leading coefficient. We seek to represent its values in the shape
\[
x_1^k + \ldots + x_s^k = \phi(n),
\]
with $s$ as small as possible. Certainly, one could take $s = 2^k + 1$, because all sufficiently large values of $n$ would be represented in this way by virtue of our work on the classical Waring’s problem. If instead one were to take $s = 2^{k-1} + 1$, then the situation is far from clear, since we know only that the number of integers $m$ with $1 \leq m \leq \phi(N)$ with no representation as the sum of $s$ $k$-th powers of natural numbers is $o(\phi(N)) = o(N^2)$. So perhaps all integers with $1 \leq n \leq N$ might fail to have a representation in the shape $x_1^k + \ldots + x_s^k = \phi(n)$. One might fear that all of these $N$ integers might “fall between the cracks”, as it were.

In order to investigate this exceptional set problem, denote by $E_{s,k}(N; \phi)$ the number of integers $n$ with $1 \leq n \leq N$ for which
\[
\left| R_{s,k}(\phi(n)) - \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} \mathcal{E}_{s,k}(\phi(n))\phi(n)^{s/k-1} \right| > \phi(n)^{s/k-1}(\log \phi(n))^{-1}. \tag{23.3}
\]
Theorem 23.2. Whenever \( s \geq \frac{3}{4}2^k + 1 \), there is a positive number \( \tau \) with the property that \( E_{s,k}(N; \phi) \ll N^{1-\tau} \).

Proof. We proceed as in the proof of Theorem 23.1, but now define \( \mathcal{Z}(N) \) to be the set of integers \( n \) with \( N/2 < n \leq N \) for which (23.3) holds. Write now \( X = \phi(N)^{1/k} \asymp N^{2/k} \). Then, when \( N/2 < n \leq N \) and \( s \geq 2k + 1 \), one has

\[
\int_{\mathbb{R}^n} f(\alpha)^s e(-\phi(n)\alpha) \, d\alpha = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} E_{s,k}(\phi(n)) \phi(n)^{s/k-1} + o(\phi(n)^{s/k-1}(\log \phi(n))^{-2}).
\]

In a similar manner to our earlier treatment, if (23.3) holds for such an integer \( n \), we have

\[
\left| \int_{\mathbb{R}^n} f(\alpha)^s e(-\phi(n)\alpha) \, d\alpha \right| \geq \frac{1}{2} \phi(n)^{s/k-1}(\log \phi(n))^{-1}.
\]

Define now \( \eta_n \) to be 0 when (23.3) does not hold, and define \( \eta_n \) otherwise via

\[
\left| \int_{\mathbb{R}^n} f(\alpha)^s e(-\phi(n)\alpha) \, d\alpha \right| = \eta_n \int_{\mathbb{R}^n} f(\alpha)^s e(-\phi(n)\alpha) \, d\alpha.
\]

Then

\[
|\eta_n| = \begin{cases} 
1, & \text{when } n \in \mathcal{Z}(N), \\
0, & \text{when } n \notin \mathcal{Z}(N).
\end{cases}
\]

Also, adopting a similar line of reasoning as employed in the proof of Theorem 23.1, we obtain

\[
\text{card}(\mathcal{Z}(N)) \phi(N)^{s/k-1}(\log \phi(N))^{-1} \ll \int_{\mathbb{R}^n} f(\alpha)^s K(-\alpha) \, d\alpha,
\]

where we write

\[
K(\alpha) = \sum_{N/2 < n \leq N} \eta_n e(\phi(n)\alpha).
\]

On recalling that \( \phi(N) \asymp X^k \), we deduce via Hölder’s inequality that

\[
\text{card}(\mathcal{Z}(N)) X^{s-k}(\log X)^{-1} \ll \left( \int_{\mathbb{R}^n} |f(\alpha)|^{4s/3} \, d\alpha \right)^{3/4} \left( \int_0^1 |K(\alpha)|^4 \, d\alpha \right)^{1/4}.
\]

The first mean value on the right hand side may be estimated by means of Hua’s lemma and Weyl’s inequality provided that \( 4s/3 > 2^k \), much as in the corresponding treatment in the proof of Theorem 23.1. The second mean value, meanwhile, counts solutions of the equation

\[
\phi(n_1) - \phi(n_2) = \phi(n_3) - \phi(n_4),
\]

with \( n_i \in \mathcal{Z}(N) \), and with each solution counted with a unimodular weight. Here, by considering separately the diagonal solutions with \( n_1 = n_2 \) and \( n_3 = n_4 \), and those with \( n_3 \neq n_4 \) in which \( n_1 \) and \( n_2 \) may be determined via a divisor function estimate, we see that the second mean value is bounded above by a quantity

\[
\ll (\text{card}(\mathcal{Z}(N)))^2 + N^\varepsilon (\text{card}(\mathcal{Z}(N)))^2.
\]

Thus we deduce that

\[
\text{card}(\mathcal{Z}(N)) X^{s-k}(\log X)^{-1} \ll \left( X^{4s/3 - k - \tau} \right)^{3/4} \left( N^\varepsilon (\text{card}(\mathcal{Z}(N)))^2 \right)^{1/4}.
\]

We conclude from this discussion that

\[
\text{card}(\mathcal{Z}(N)) X^{s-k}(\log X)^{-1} \ll X^{s-\frac{3}{4}(k+\tau)} (N^\varepsilon \text{card}(\mathcal{Z}(N)))^{1/2},
\]

and this completes the proof of Theorem 23.2.
whence
\[ \text{card}(\mathcal{Z}(N)) \ll X^{k/2-\tau} \asymp N^{1-2\tau/k}. \]

The conclusion of the theorem consequently follows by summing over dyadic intervals. \(\square\)

The proof of Theorem 23.2 follows a line of reasoning that is fairly natural from the point of view of harmonic analysis. A strategy that is more convoluted, and contains more number theoretic input, would estimate the critical mean value via Cauchy-Schwarz in the form
\[ \int_{m_5} f(\alpha)^r K(-\alpha) \, d\alpha \leq \left( \int_0^1 |K(\alpha)^2 f(\alpha)^{2r}| \, d\alpha \right)^{1/2} \left( \int_{m_5} |f(\alpha)|^{2s-2r} \, d\alpha \right)^{1/2}, \]
where \( r = 2^{[k/2]-1} \) and \( 2s - 2r > 2^k \). The point here is that the first integral counts solutions of the equation
\[ \phi(n_1) - \phi(n_2) = \sum_{i=1}^r (x_i^k - y_i^k), \]
with \( 1 \leq x_i, y_i \leq X \) and \( n_j \in \mathcal{Z}(N) \). The solutions here are of two types. First, there are solutions with \( n_1 = n_2 \) and
\[ \sum_{i=1}^r (x_i^k - y_i^k) = 0. \]
By applying Hua’s lemma, one sees that these contribute at most
\[ \text{card}(\mathcal{Z}(N)) \int_0^1 |f(\alpha)|^{2r} \, d\alpha \ll \text{card}(\mathcal{Z}(N)) X^{2r-k/2}. \]

Consider next the contribution of the solutions with \( n_1 \neq n_2 \). There are at most \( X^{2r} \) possible choices for \( x_i \) and \( y_i \) (\( 1 \leq i \leq r \)), and for each of these possible choices a divisor function estimate reveals that there are at most \( O(N^\varepsilon) \) possible choices for \( n_1 \) and \( n_2 \). Hence we find that
\[ \int_0^1 |K(\alpha)^2 f(\alpha)^{2r}| \, d\alpha \ll \text{card}(\mathcal{Z}(N)) X^{2r-k/2} + N^\varepsilon X^{2r} \ll N^\varepsilon X^{2r}, \]
whilst the condition \( 2s - 2r > 2^k \) ensures that for a positive number \( \tau \) one has
\[ \int_{m_5} |f(\alpha)|^{2s-2r} \, d\alpha \ll X^{2s-2r-k-\tau}. \]

Utilising the estimates of the previous paragraph in the argument of the proof of Theorem 23.2, we perceive that
\[ \text{card}(\mathcal{Z}(N)) X^{s-k}(\log X)^{-1} \ll \left( X^{2r+\varepsilon} \right)^{1/2} \left( X^{2s-2r-k-\tau} \right)^{1/2} \ll X^{s-(k+\tau)/2}, \]
whence
\[ \text{card}(\mathcal{Z}(N)) \ll X^{k/2-\tau/4} \asymp N^{1-\tau/(2k)}. \]
Thus, we have \( E_{s,k}(N; \phi) = o(N) \) provided only that
\[ s \geq 2^{k-1} + 2^{[k/2]-1} + 1. \]
If one makes use of the (now proven) main conjecture in Vinogradov’s mean value theorem, then this line of reasoning yields the same conclusion concerning \( E_{s,k}(N; \phi) \) provided
only that $s \geq \frac{5}{8}k^2 + O(k)$. Both represent substantial improvements on the more pedestrian approach of Theorem 23.2.

©Trevor D. Wooley, Purdue University 2020. This material is copyright of Trevor D. Wooley at Purdue University unless explicitly stated otherwise. It is provided exclusively for educational purposes at Purdue University, and is to be downloaded or copied for your private study only.

References


Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA

E-mail address: twooley@purdue.edu