THE CIRCLE METHOD AND WARING'S PROBLEM

EDGAR ASSING

Abstract. We will give a introduction to the circle method based on its application to Waring's problem. These are some extended lecture notes for the course *Selected Topics in Number Theory - The Circle Method and Waring's problem* taught in the winter term 2021/22 at the University of Bonn. **Attention**: This manuscript probably contains many misprints and inaccuracies. For personal use only!

Contents

Date: 1^{st} April 2022.

1. Introduction

There is a variety of excellent articles and books on the topics treated in this lecture. We have used a combination of these in order to prepare these notes. A (non-exhaustive) selection of references can be found in the bibliography at the end.

Let us briefly introduce some general framework in which several interesting results and questions from number theory can be phrased.

Given k-sequences $A^{(1)}, \ldots, A^{(k)} \subset \mathbb{N} \cup \{0\}$ of non-negative integers we write

$$A = \sum_{i=1}^{k} A^{(i)} = \{n = a_1 + \ldots + a_k \colon a_i \in A^{(i)} \text{ for } i = 1, \ldots, k\}.$$

In general one is interested in the properties of A such as its density in $\mathbb{N}$. Let us give some examples.

- **(Fermat)** For $Q_2 = \{n^2 \colon n \in \mathbb{N} \cup \{0\}\}$ we have $\{p \text{ prime} \colon n \equiv 1 \bmod 4\} \subset Q_2 + Q_2$.
- **(Lagrange)** If Q_2 is again the set of squares, then we have $Q_2 + Q_2 + Q_2 + Q_2 = \mathbb{N} \cup \{0\}$.
- **(I. Vinogradov)** Let $P = \{p \text{ prime}\} \cup \{0\}$. There is a constant $C > 0$ such that $\{n \text{ odd} \colon n \geq C\} \subset P + P + P$. (Due to quite recent work of **H. Helfgott** it is now know that we can take $C = 5$. This is known as *Goldbach's weak conjecture.*)
- **(Hilbert)** Let $k \in N$ and set $Q_k = \{n^k \colon n \in \mathbb{N} \cup \{0\}\}$. Then there is an integer $g = g(k)$ such that $\mathbb{N} \cup \{0\} = \sum_{i=1}^{g} Q_k$. (This was conjectured by **Waring** and runs under the name *Waring's problem.* A more elegant proof was later supplied by **Hardy and Littlewood** as well as **I. Vinogradov**.)

We call a set $A \subset \mathbb{N} \cup \{0\}$ a *basis* (for $\mathbb{N}$) of *order* k if

$$\mathbb{N} \subset \underbrace{A + \ldots + A}_{k \text{ times}}.$$

Given a set A as above we define the *(Schnirelmann)-density* of A in $\mathbb{N}$ by

$$d(A) = \inf_n \frac{\sharp(A \cap [1, n])}{n}.$$

We will see the elementary but very interesting result which says that any set $\{0\} \subset A \subset \mathbb{N} \cup \{0\}$ with positive density is a basis for $\mathbb{N}$. This together with Vinogradov's result towards the weak Goldbach conjecture implies that $P \cup \{1\}$ is a basis for $\mathbb{N}$. (Note that the sequence P itself has density 0!) But it does not give us any useful information on the order of the basis. The (strong) Goldbach conjecture now asserts that

$$\{n \text{ even}\colon n \geq 4\} \subset P + P.$$

Thus, together with the weak Goldbach conjecture we would get that $P \cup \{1\}$ is a basis of order 3. Even though some results towards the (strong) Goldbach conjecture are available it is still open in general.

2. Notation

We summarize some standard notation:

- Let $f \in \mathbb{R}$ and $g \in \mathbb{R}_+$. We write $f \ll g$ if there exists a constant C such that $|f| \leq Cg$. If the constant C depends on some parameters we indicate this by adding a subscript. For example we could write $-x^n \ll_n e^x$ for all $x \in \mathbb{R}_{\geq 1}$.
- We write $e(x) = e^{2\pi i x}$.
- For $a, b \in \mathbb{Z}$ we write $(a, b) \in \mathbb{N}$ for the greatest common divisor (short gcd or ggT). This generalizes to $(a_1, \ldots, a_n)$, meaning the greatest common divisor of all the numbers in the bracket.
- If not explicitly stated otherwise the letter p is reserved for primes.
- Given an integer n we write $d(n)$ for the number of divisors. This is

$$d(n) = \sharp\{a \in \mathbb{N}\colon a \mid n\} = \sum_{a \mid n} 1.$$

- We write φ for Euler's totient function, μ for the Moebius function and Ω for the number or prime divisors.

3. Some Elementary considerations

We start this lecture by some elementary considerations following [**?**]. In particular, we will conclude this chapter by presenting an elementary solution of Waring's problem.

3.1. **On a theorem of Mann.** Let us start by listing some elementary properties of the (Schnirelmann)-density:

- Suppose $1 \notin A$, then $d(A) = 0$;
- Let $A = \{1 + r(n-1) \colon n \in \mathbb{N}\}$ be an arithmetic progression with difference r starting with 1, then $d(A) = \frac{1}{r}$;
- The density of any geometric progression is 0;
- We have $d(Q_k) = 0$ for all $k \geq 2$.
- $\mathbb{N} \subset A$ if and only if $d(A) = 1$;
- Suppose $d(A) = 0$ and $1 \in A$, then for every $\epsilon > 0$ there is $N = N(\epsilon)$ sufficiently large such that $\sharp(A \cap [1, n]) < \epsilon n$.

Next we try to understand how the density changes when adding two sets.

Lemma 3.1 (Schnirelmann). *For two sequences $\{0\} \subset A, B \subset \mathbb{N} \cup \{0\}$ we have*

$$d(A + B) \geq d(A) + d(B) - d(A)d(B).$$

Proof. We write $C = A + B$ and $d(A) = \alpha$, $d(B) = \beta$ and $d(C) = \gamma$. Further set

$$A(n) = \sharp([1, n] \cap A) \text{ and } B(n) = \sharp([1, n] \cap B). \tag{1}$$

We write $[1, n] \cap A = \{1 = a_1 < a_2 < \ldots < a_{A(n)}\}$. Since $0 \in B$ we have $A \subset C$. For $0 < k < A(n)$ we observe that there are exactly $l_k = a_{k+1} - a_k - 1$ numbers between a_k and a_{k+1} that don't belong to A. If $k = A(n)$, we simply have $l_k = n - a_k$. Note that $a_k + r \in C$ if and only if $r \in B$. Thus among the numbers $a_k + 1, \ldots, a_k + l_k$ at least $B(l_k)$ are in C. We deduce that

$$\sharp([1, n] \cap C) \geq A(n) + \sum_{k=1}^{A(n)} B(l_k).$$

Since $B(l_k) \geq \beta l_k$ and $A(n) \geq \alpha n$ we can estimate

$$\begin{aligned} \sharp([1, n] \cap C) &\geq A(n) + \beta \sum_{k=1}^{A(n)} l_k = A(n) + \beta \sum_{k=1}^{A(n)-1} (a_{k+1} - a_k - 1) + \beta(n - a_{A(n)}) \\ &= A(n) + \beta(n - A(n)) \geq \alpha n + \beta n - \alpha\beta n. \end{aligned}$$

This completes the proof. □

If we reformulate this inequality to

$$1 - d(A + B) \leq (1 - d(A))(1 - d(B)),$$

this easily generalises to

$$1 - d(A_1 + \ldots A_k) \leq \prod_{i=1}^{k} (1 - d(A_i)).$$

Theorem 3.2 (Schnirelmann). *If $\{0\} \subset A \subset \mathbb{N} \cup \{0\}$ has positive (Schnirelmann)-density, then it is a basis for $\mathbb{N}$.*

Proof. Write

$$A_k = \underbrace{A + \ldots + A}_{k \text{ times}}.$$

We need to show that there is $k = k(A)$ such that $\mathbb{N} \subset A_k$. Our previous result implies

$$d(A_k) \geq 1 - (1 - d(A))^k.$$

In particular, since $0 < d(A) \leq 1$, we can find sufficiently large k such that $d(A_k) \geq \frac{1}{2}$. But this implies that

$$2A_k(n) \geq 2d(A_k)n \geq n.$$

After noting that $A_{2k} = A_k + A_k$ the result follows from the following claim. Suppose $A(n) + B(n) > n - 1$, then $n \in A + B$. To see this we can assume without loss of generality that $n \notin A \cup B$. Thus, $r = A(n) = A(n-1)$ and $s = B(n) = B(n-1)$. If $A = \{0, 1 = a_1, a_2 \ldots\}$ and $B = \{0, 1 = b_1, b_2, \ldots\}$, then

$$a_1, a_2, \ldots, a_r \text{ and } n - b_1, n - b_2, \ldots, n - b_s$$

belong to $[1, n-1] \cap \mathbb{N}$. There are $r + s = A(n) + B(n) > n - 1$ numbers in this list, so that by the pigeon hole principle there must be $1 \leq i \leq r$ and $1 \leq j \leq r$ such that $a_i = n - b_j$. But then $n = a_i + b_j \in A + B$. □

Schnirelmann and Landau (fall 1931) conjectured that as long as $d(A)+d(B) \leq 1$ one actually has the stronger inequality

$$d(A + B) \geq d(A) + d(B).$$

This conjecture was finally resolved by Mann in 1942. Here we will present a proof given later (1943) by Artin and Scherk. This proof will occupy the remainder of this section. We write $\alpha = d(A)$ and $\beta = d(B)$ and assume $\alpha + \beta \leq 1$. As before we write use the notation from (1). Put $C = A + B$, $\gamma = d(C)$ and write $C(n) = \sharp([1, n] \cap C)$ as well as $C_n = [1, n] \cap C$. We set $C(0) = 0$.

Definition 3.1. We call C_n *normal* if $f, f' \notin C_n$ with $f \neq f'$ and $f, f' \in [0, n]$ implies $f + f' - n \notin C_n$.

Lemma 3.3 (and Definition)**.** *For fixed $n \notin C$, there are finitely many sets*

$$B = B^{(0)} \subset B^{(1)} \subset \ldots \subset B^{(h)} \text{ and } C = C^{(0)} \subset C^{(1)} \subset \ldots \subset C^{(h)} \qquad (2)$$

such that $A + B^{(i)} = C^{(i)}$ for $0 \leq i \leq h$ and $C_n^{(h)}$ is normal. We call $C^{(h)}$ the canonical extension of C. The numbers $\beta_0, \ldots, \beta_{h-1}$ are referred to as bases for the extensions.

Proof. Let $\beta_0 \in B$ be the smallest number such that there is $c, c' \in [0, n] \setminus C$ with $c \neq c'$ and $a \in A$ such that

$$c + c' - n = a + \beta_0 \in C_n.$$

The existence of β_0 is ensured since C_n is not normal. Define

$$C^* = \{c \in [0,n] \colon c \notin C,\, \exists c' \in [0,n] \setminus C \text{ and } a \in A \text{ such that } c + c' - n = a + \beta_0\}.$$

Obviously C and C^* are disjoint. We put $C^{(1)} = C \cup C^*$. Further set

$$B^* = \{\beta_0 + n - c \colon c \in C^*\}.$$

By construction every element $b^* \in B^*$ can be written as $b^* = c' - a$ for some $c' \in C^*$ and $a \in A$.

We claim that $B^* \subset [0,n]$ and $B \cap B^* = \emptyset$. To see this we note that $b^* \in B^*$ satisfies $0 \leq \beta_0 \leq b^* = c' - a \leq c' \leq n$. Further suppose $b^* \in B$. Then $c' = b^* + a \in B + A = C$, which is a contradiction. This establishes our claim.

Set $B^{(1)} = B \cup B^*$ and we have to show that $A + B^{(1)} = C^{(1)}$. Take $a \in A$ and $b_1 \in B^{(1)}$ We first show that $a + b_1 \in C^{(1)}$. If $a + b_1 \in C$ we are done. Therefore we assume $b_1 \in B^*$ and $a + b_1 \notin C$. We find

$$C \not\ni c = a + b_1 = a + \beta_0 + n - c' \text{ for } c' \notin C.$$

This implies $c + c' - n = a + \beta_0 \in A + B = C$. By construction of C^* we have $c \in C^* \subset C^{(1)}$. we have established that $A + B^{(1)} \subset C^{(1)}$. But the other inclusion follows directly from the construction so that we are done.

Suppose $n \in C^*$. Then we can take $c' = n \notin C$ and get $c = a + \beta_0 \in A + B = C$. This contradicts that $C \cap C^* = \emptyset$ by construction. Thus $n \notin C^{(1)}$.

If $C^{(1)}$ is still not normal, we can apply the same process to construct the extension $A + B^{(2)} = C^{(2)}$ with basis β_1. Since we are only adding integers from $[0, n-1]$ this procedure must terminate after finitely many steps and leave us with a normal set $A + B^{(\mu)} = C^{(\mu)}$. □

We now establish some properties of canonical extensions constructed above.

Lemma 3.4. *The base points of the canonical extensions satisfy*

$$\beta_0 < \beta_1 < \ldots < \beta_{h-1}.$$

Proof. Let $1 \leq i \leq \mu - 1$. Then $\beta_i \in B^{(i)} = B^{(i-1)} \cup (B^{(i-1)})^*$. Let us first assume that $\beta_i \in (B^{(i-1)})^*$. But then $\beta_i = \beta_{i-1} + n - c$ for some $c < n$ and we are done. Now we assume otherwise. By definition of β_i, we find $a \in A$, $c, c' \notin C^{(i)}$ such that

$$c + c' - n = a + \beta_i \in C^{(i)}.$$

But since we are assuming $\beta_i \in B^{(i-1)}$ we have $c + c' - n + \beta_i \in A + B^{(i-1)} = C^{(i-1)}$. But β_{i-1} is the minimal number with this property. Thus $\beta_{i-1} \leq \beta_i$. Finally assume $\beta_i = \beta_{i-1}$. This would imply $c, c' \in (C^{(i-1)})^* \subset C^{(i)}$, which is a contradiction. We conclude that $\beta_{i-1} < \beta_i$ must hold. □

Lemma 3.5. *For $0 \leq i \leq h$ let $m = \min\{k \in \mathbb{N} \colon k \notin C^{(h)}\}$. Suppose $c \in (C^i)^*$ for $0 \leq i \leq h-1$ and $n - m < c < n$. Then $c > n - m + \beta_i$.*

Proof. We reformulate the assumption to

$$0 < m + c - n < m.$$

By minimality of m this implies $m + c - n \in C^{(h)}$. We now write

$$C^{(h)} = C^{(i)} \cup (C^{(i+1)})^* \cup \ldots \cup (C^{(h-1)})^*$$

and consider two cases.

First, suppose $m + c - n \in C^{(i)}$. This implies

$$m + c - n = a + b_i \text{ for } a \in A \text{ and } b_i \in B^{(i)}.$$

Note that $m \notin C^{(i)}$ and $c \notin C^{(i)}$. Thus, by minimality of β_i, we get $b_i \geq \beta_i$. Equality would imply $m \in (C^{(i)})^* \subset C^{(h)}$, which is false. Thus $b_i > \beta_i$, so that

$$m + c - n = a + b_i \geq b_i > \beta_i$$

and we are done.

Second, suppose $c' = m + c - n \in (C^{(j)})^*$ for $i \leq j \leq h - 1$. But by construction of the set $(C^{(j)})^*$ we find that

$$c' - a = \beta_j + n - c'' \text{ for } a \in A \text{ and } c'' \in (C^{(j)})^*.$$

This yields the chain of inequalities:

$$c' \geq c' - a > \beta_j \geq \beta_i.$$

This concludes the proof. □

Lemma 3.6. *We have*

$$(C^{(i)})^*(n) - (C^{(i)})^*(n-m) = \sharp[(C^{(i)})^* \cap (n-m, n]] = \sharp[(B^{(i)})^* \cap [1, m)] = (B^{(i)})^*(m-1),$$

for $0 \leq i \leq h - 1$.

Proof. We look at the expression $b = \beta_i + n - c$. By definition of $(B^{(i)})^*$ and $(C^{(i)})^*$ we have the implications

$$c \in (C^{(i)})^* \Rightarrow b \in (B^{(i)})^* \text{ and } b \in (B^{(i)})^* \Rightarrow c \in (C^{(i)})^*.$$

If $n - m + \beta_i < c < n$, then $\beta_i < b < m$ and vice versa. Thus,

$$(C^{(i)})^*(n) - (C^{(i)})^*(n - m + \beta_i) = (B^{(i)})^*(m - 1) - (B^{(i)})^*(\beta_i).$$

By the previous lemma we have $(C^{(i)})^*(n - m + \beta_i) = (C^{(i)})^*(n - m)$. We are done if we can show $(B^{(i)})^*(\beta_i) = 0$, which follows directly from the construction. □

Lemma 3.7. *Let C be as above. For every $n \in \mathbb{N}$ there is $1 \leq m < n$ such that*

$$C(n) - C(n - m) \geq (\alpha + \beta)m.$$

Proof. We first treat a trivial case. Indeed, if $n \in C$, then

$$C(n) - C(n-1) = 1 \geq \alpha + \beta.$$

Thus we can assume without loss of generality that $n \notin C$.

Next we show the statement under the assumption that $C_n = C \cap [0, n]$ is normal. To do so let m be the smallest natural number that does not appear in C. Of course $m \leq n$. Let $n - n < s < n$ be an arbitrary integer. Suppose $s \notin C$. Then, by normality of C_n, $s + m - n \notin C$. But $0 < s + m - n < m$, so that we have a contradiction to minimality of m. We conclude that all integers between m and n are contained in C and get

$$C(n) - C(n-m) = m - 1.$$

Since $m \notin C$ the argument from the proof of Theorem 3.2 shows that $A(m) + B(m) \leq m - 1$. Combining these observations yields

$$C(n) - C(n-m) = m - 1 \geq A(m) + B(m) \geq (\alpha + \beta)m.$$

If C_n is not normal, then we take a canonical normal extenstion $C^{(h)}$ of C. For this extension we have the estimate

$$C^{(h)}(n) - C^{(h)}(n-m) \geq A(m) + B_h(m)$$

shown above, where m is the smallest number not appearing in $C^{(h)}$. We make the following two observations:

$$C^{(h)}(n) - C^{(h)}(n-m) = C(n) - C(n-m) + \sum_{i=0}^{h-1} ((C^{(i)})^*(n) - (C^{(i)})^*(n-m)) \text{ and}$$

$$B^{(h)}(m) = B^{(h)}(m-1) = B(m-1) + \sum_{i=0}^{h-1} (B^{(i)})^*(m-1).$$

Inserting this above yields

$$C(n) - C(n-m) + \sum_{i=0}^{h-1} ((C^{(i)})^*(n) - (C^{(i)})^*(n-m)) \geq A(m) + B(m-1) + \sum_{i=0}^{h-1} (B^{(i)})^*(m-1).$$

But we know from the previous lemma that $(C^{(i)})^*(n) - (C^{(i)})^*(n-m) = (B^{(i)})^*(m-1)$. But this gives

$$C(n) - C(n-m) \geq A(m) + B(m-1) = A(m) + B(m) \geq (\alpha + \beta)m.$$

This concludes the proof. □

With this *fundamental lemma* at hand we can proof the result promised earlier.

Theorem 3.8 (Mann). *Let* $\{0\} \subset A, B \subset \mathbb{N} \cup \{0\}$ *with* $d(A) + d(B) \leq 1$. *Then*

$$d(A + B) \geq d(A) + d(B).$$

Proof. Let $n \in \mathbb{N}$. We need to show that $C(n) \geq (\alpha+\beta)n$. We argue by induction. First, $C(1) = 1 \geq \alpha + \beta$ by assumption. Now assume that $C(k) \geq (\alpha+\beta)k$ for all $k < n$. To show the statement for n we apply Lemma 3.7 and find $1 \leq m \leq n$ such that $C(n) - C(n-m) \geq (\alpha+\beta)m$. By induction hypothesis we get

$$C(n) = C(n) - C(n-m) + C(n-m) \geq (\alpha+\beta)m + (\alpha+\beta)(n-m) = (\alpha+\beta)n.$$

□

3.2. **Waring's problem from an elementary point of view.** We will now present an elementary solution to Waring's problem following Linnik's argument.

We define

$$r_k(m) = \sharp\{(x_1, \ldots, x_k) \in \mathbb{Z}^k_{\geq 0} \colon x_1^n + \ldots + x_k^n = m\}.$$

Recall that we aim to show that for sufficiently large k (depending on the fixed n) we have $m \in \sum_{i=1}^k Q_n = A_n^{(k)}$ for all $m \in \mathbb{N}$. The latter obviously follows from $r_k(m) > 0$. Our solution of Waring's problem will rely on upper bounds for $r_k(m)$. We start by proving several Preliminary results.

Lemma 3.9. *Let $|a_2| \leq |a_1| \leq A$ be integers with $(a_1, a_2) = 1$. Then*

$$\sharp\{(z_1, z_2) \in \mathbb{Z}^2 \colon a_1 z_1 + y_2 z_2 = m \text{ and } |z_1|, |z_2| \leq A\} \leq \frac{3A}{|a_1|}.$$

Proof. Without loss of generality we assume $a_1 > 0$. Two solutions (z_1, z_2) and (z_1', z_2'). will satisfy

$$a_2(z_2' - z_2) = a_1(z_1 - z_1').$$

Since a_1 and a_2 are co-prime we find that $a_1 \mid (z_2' - z_2)$. But since we are assuming the solutions to be distinct we obtain $|z_2' - z_2| \geq a_1$.

Note that the solution (z_1, z_2) is uniquely determined by z_2, so that it is enough to count the number, t, of possible values for z_2 in $[-A, A]$. If z_2^- (resp. z_2^+) is the smallest (resp. biggest) possibility, then we must have

$$a_1(t-1) \leq z_2^+ - z_2^- \leq 2A.$$

But this yields

$$t \leq \frac{2A}{a_1} + 1 \leq \frac{3A}{a_1}.$$

□

Lemma 3.10. *Let $a_1, \ldots, a_l, m \in \mathbb{Z}$ such that $|a_i| \leq A$ for $i = 1, \ldots, l$ and $(a_1, \ldots, a_l) = 1$. Then*

$$\sharp[\{(z_1, \ldots, z_l) \in \mathbb{Z}^l \colon a_1 z_1 + \ldots + a_l z_l = m\} \cap B_A(0)] \ll_l \frac{A^{l-1}}{H},$$

for $H = \max_i |a_i|$. We will write $c(l)$ for the implicit constant.

Proof. We argue by induction over l. The previous result shows the claimed bound for $l = 2$. Thus we suppose that $l \geq 3$ and the statement of the lemma is true for $l - 1$. Without loss of generality we assume that $H = |a_l|$.

We start with the degenerate case $a_1 = \ldots = a_{l-1} = 0$. In this case the equation reads $\pm z_l = m$. It follows that z_l is uniquely determined and each z_i, for $i = 1, \ldots, l-1$ can be chosen arbitrarily in $[-A, A]$. Thus we have exactly $(2A+1)^{l-1} \leq 3^{l-1}A^{l-1}$ solutions. We are done since $H = 1$.

Suppose at least one of the numbers $a_1, \ldots, a_{l-1}$ is non-zero. Put $\delta = (a_1, \ldots, a_{l-1})$. Define $H' = \max_{i=1,\ldots,l-1} \frac{|a_i|}{\delta}$ and look at the equations

$$\frac{1}{\delta}(a_1z_1 + \ldots + a_{l-1}z_{l-1}) = m' \text{ and } \delta m' + a_lz_l = m.$$

We have the trivial estimate $m' \leq lH'A$. Further note that $\delta \leq |a_l|$ and $(\delta, a_l) = 1$. By the previous lemma we have

$$\sharp[\{(m', z_l) \in \mathbb{Z}^2 \colon \delta m' + a_lz_l = m\} \cap B_{lH'A}(0)] \ll_l \frac{H'A}{H}.$$

For each m' appearing as a solution we apply the induction hypothesis and count

$$\sharp[\{(z_1, \ldots, z_{l-1}) \in \mathbb{Z}^{l-1} \colon a_1z_1 + \ldots + a_lz_{l-1} = m'\} \cap B_A(0)] \ll_l \frac{A^{l-2}}{H'}.$$

Combining these two estimates gives the desired result. □

Before we state the next result we note that if $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{Z}^l \cap B_A(0)$ we can rewrite our equation as

$$\langle \mathbf{a}, \mathbf{z} \rangle = \mathbf{a} \cdot \mathbf{z}^t = m$$

for $\mathbf{z} = (z_1, \ldots, z_l) \in \mathbb{Z}^l \cap B_A(0)$.

Lemma 3.11. *Let $l > 2$ and $1 \leq A \leq B \ll_l A^{l-1}$. Then*

$$\sum_{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0)} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\} \ll_l (AB)^{l-1}.$$

Proof. We start by looking at the contribution of $\mathfrak{a} = (0, \ldots, 0)$. This case is trivially treated by observing that

$$\sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\} = (2B+1)^l \ll_l B^l \ll (AB)^{l-1}.$$

Next we consider the contribution of those $\mathfrak{a}$ such that $(a_1, \ldots, a_l) = 1$. Set

$$H = \|\mathbf{a}\|_\infty = \max_i |a_i|.$$

Of course there is m such that

$$\frac{A}{2^{m+1}} < H \leq \frac{A}{2^m}.$$

Given such an $\mathbf{a}$ we count solutions to the associated equation using the previous lemma and get

$$\sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\} \ll_l \frac{B^{l-1}}{H} \ll_l \frac{B^{l-1}2^m}{A}.$$

Note that the bound on H immediately implies $a_i \leq A2^{-m}$. Thus, the number of all $\mathbf{a}$ satisfying our current assumptions is bounded by

$$(2\frac{A}{2^m} + 1)^l \ll A^l 2^{-ml}.$$

Putting these observations together yields

$$\sum_{\substack{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0), \\ (a_1,\dots,a_l)=1}} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\} \ll_l \sum_{m \in \mathbb{Z}_{\geq 0}} \sum_{\substack{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0), \\ (a_1,\dots,a_l)=1, \\ A2^{-m-1} < \|\mathbf{a}\|_\infty \leq A2^{-m}}} \frac{B^{l-1}2^m}{A}$$
$$\ll_l (AB)^{l-1} \sum_{m \in \mathbb{Z}_{\geq 0}} 2^{-m(l-1)} \ll_l (AB)^{l-1}.$$

Finally we need to treat the contribution of those $\mathbf{a}$ with $(a_1, \dots, a_l) = \delta > 1$. These cases can be treated by reduction to the previous case. Indeed we can replace A by $\frac{1}{\delta}A$ and use the observation that the components of $\frac{1}{\delta}\mathbf{a}$ satisfy $(\frac{a_1}{\delta}, \dots, \frac{a_l}{\delta}) = 1$. We get

$$\sum_{\substack{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0), \\ (a_1,\dots,a_l)=\delta}} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\} = \sum_{\substack{\mathbf{a} \in \mathbb{Z}^l \cap B_{\frac{A}{\delta}}(0), \\ (a_1,\dots,a_l)=1}} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\}$$
$$\ll \frac{(AB)^{l-1}}{\delta^{l-1}}.$$

We have treated all cases and find that

$$\sum_{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0)} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\}$$
$$= \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle 0, \mathbf{z} \rangle = 0\} + \sum_{\delta \in \mathbb{N}} \sum_{\substack{\mathbf{a} \in \mathbb{Z}^l \cap B_A(0), \\ (a_1,\dots,a_l)=\delta}} \sharp\{\mathbf{z} \in \mathbb{Z}^l \cap B_B(0) \colon \langle \mathbf{a}, \mathbf{z} \rangle = 0\}$$
$$\ll_l B^l + (AB)^{l-1} \sum_{\delta \in \mathbb{N}} \delta^{1-l} \ll_l (AB)^{l-1}.$$

In the last step we used that $l \geq 3$, so that the δ-sum is finite. □

Lemma 3.12. *Let A, B be two finite multisets of numbers (i.e. elements can appear multiple times in A or B). We have*

$$\sharp\{(x,y) \in A \times B \colon c = x+y\} \leq \frac{1}{2}\sharp\{(x,y) \in A^2 \colon x-y = 0\} + \frac{1}{2}\sharp\{(x,y) \in B^2 \colon x-y = 0\}.$$

Note that is we apply this to $A = B$ we simply get

$$\sharp\{(x,y) \in A \times A \colon c = x + y\} \le \sharp\{(x,y) \in A^2 \colon x - y = 0\}.$$

Proof. We write

$$A = \{\underbrace{a_1, \dots, a_1}_{\lambda_1 \text{ times}}, \dots\dots\dots, \underbrace{a_r, \dots, a_r}_{\lambda_r \text{ times}}\} \text{ and } B = \{\underbrace{b_1, \dots, b_1}_{\mu_1 \text{ times}}, \dots\dots\dots, \underbrace{b_s, \dots, b_s}_{\mu_s \text{ times}}\}$$

for $a_1, \dots, a_r$ distinct and $b_1, \dots, b_s$ distinct. We now write

$$\sharp\{(x,y) \in A \times B \colon c = x + y\} = \sum_{\substack{1 \le i \le r, \\ 1 \le j \le s, \\ a_i + b_j = c}} \lambda_i \mu_j \le \frac{1}{2} \sum_{\substack{1 \le i \le r, \\ 1 \le j \le s, \\ a_i + b_j = c}} (\lambda_i^2 + \mu_j^2) \le \frac{1}{2} \sum_{1 \le i \le r} \lambda_i^2 + \frac{1}{2} \sum_{1 \le j \le s} \mu_j^2.$$

But the equation $x - y = 0$ for $x, y \in A$ holds exactly for $x = y = a_i$ for some $1 \le i \le r$. We get

$$\sharp\{(x,y) \in A^2 \colon x - y = 0\} = \sum_{1 \le i \le r} \lambda_i^2.$$

Running the same argument for B finishes the proof. □

We generalize this combinatorial result to more sets. This is the content of the next lemma.

Lemma 3.13. *Let $l = k2^s$ and $A_1, \dots, A_l$ be finite multisets of numbers. We have*

$$\sharp\{(x_1, \dots, x_l) \in A_1 \times \dots \times A_l \colon x_1 + \dots + x_l = c\}$$
$$\le \max_{0 \le m \le 2^s - 1} \sharp\{(y^{(1)}, \dots, y^{(2^s)}) \in [A^{(m)}]^{2^s} \colon y^{(1)} + \dots + y^{(2^{s-1})} - y^{(2^{s-1}+1)} - \dots - y^{(2^s)} = 0\},$$

where $A^{(m)} = \{x_1 + \dots + x_k \colon x_i \in A_{mk+i}\}$.

Proof. We set

$$A = \{x_1 + \dots + x_{\frac{l}{2}} \colon x_i \in A_i \text{ for } i = 1, \dots, \frac{l}{2}\}$$
$$\text{and } B = \{x_{\frac{l}{2}+1} + \dots + x_l \colon x_i \in A_i \text{ for } i = \frac{l}{2} + 1, \dots, l\}.$$

We obviously have

$$\sharp\{(x_1, \dots, x_l) \in A_1 \times \dots \times A_l \colon x_1 + \dots + x_l = c\} = \sharp\{(x,y) \in A \times B \colon c = x + y\}$$
$$\le \frac{1}{2} \sharp\{(x,y) \in A^2 \colon x - y = 0\} + \frac{1}{2} \sharp\{(x,y) \in B^2 \colon x - y = 0\},$$

where we applied the previous lemma. It remains to estimate the cardinalities on the right hand side. It is enough to treat the contribution from the set A, since the other one is analogous.

We bring the equation $x - y = 0$ for $x, y \in A$ in the form

$$0 = \sum_{i=1}^{\frac{l}{2}} (z_i - z_i') = \underbrace{\sum_{i=1}^{\frac{l}{4}} (z_i - z_i')}_{=x'} + \underbrace{\sum_{i=\frac{l}{4}+1}^{\frac{l}{2}} (z_i - z_i')}_{=y'} \text{ for } z_i, z_i' \in A_i.$$

As indicated in the previous equation we can continue to iterate this process s times to conclude the proof. □

Lemma 3.14. *There is $k = k(n) \in \mathbb{N}$ such that for an arbitrary $N \in \mathbb{N}$ we have*

$$r_k(m) \ll_n N^{\frac{k}{n}-1} \textit{ for all } 1 \leq m \leq N.$$

Proof. We will use induction to prove a slightly stronger statement. Indeed we will fix a polynomial

$$f(x) = a_0 x^n + \ldots + a_{n-1}x + a_n$$

and consider solutions to the equation

$$f(x_1) + \ldots + f(x_k) = m.$$

We define

$$r_{f,k}(m) = \{(x_1, \ldots, x_k) \in \mathbb{Z}^k \colon f(x_1) + \ldots + f(x_k) = m, \, |x_i| \leq N^{\frac{1}{n}} \text{ for } i = 1, \ldots, k\}.$$

Note that obviously $r_k(m) \leq r_{x^n,k}(m)$ as long as $m \leq N$.
Claim: Suppose the coefficients of f satisfy

$$|a_i| \ll_n N^{\frac{i}{n}},$$

then there is $k = k(n)$ such that

$$r_{f,k}(m) \ll N^{\frac{k}{n}-1} \text{ for all } 1 \leq m \leq N. \tag{3}$$

Let us first proof (3) for $n = 1$. In this case we must have $f(x) = a_0 \cdot x + a_1$. We take $k(1) = 2$, so that we need to solve

$$a_0(x_1 + x_2) = m - 2a_1.$$

But the condition $|x_1| \leq N$ implies that there are at most $2N + 1$ solutions. Since every x_1 uniquely determines x_2 we have

$$r_2(m) \leq 3N,$$

which concludes this case.

We now assume that (3) holds for $n' = n - 1$.[1] We put $k' = k(n')$ and choose

$$k = 2n \cdot 2^{s+1} \text{ for } s = \lfloor \log(k') \log(2)^{-1} \rfloor - 1.$$

[1] It may be instructive to work through the argument for $n = 2$ explicitly.

We now define the (multi)-set

$$A = \left\{ \sum_{i=1}^{\frac{k}{2}} f(x_i) \colon |x_i| \leq N^{\frac{1}{n}} \text{ for } i = 1, \dots, \frac{k}{2} \right\}.$$

Of course we have

$$r_{f,k}(m) = \sharp\{(x,y) \in A \times A \colon x + y = m\} \leq \sharp\{(x,y) \in A \times A \colon x - y = 0\},$$

where we applied Lemma 3.12. We can rewrite the equation $x-y=0$ with $x, y \in A$ as

$$\sum_{i=1}^{\frac{k}{2}} [f(x_i) - f(y_i)] = 0$$

for suitable x_i and y_i. We set $x_i - y_i = h_i$ and get

$$\sum_{i=1}^{\frac{k}{2}} [f(y_i + h_i) - f(y_i)] = 0$$

We allow $y_i, h_i \in [-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}}] \cap \mathbb{Z}$. By opening the powers $(y_i + h_i)^j$ we get

$$f(y_i + h_i) - f(y_i) = \sum_{v=0}^{n} a_v \sum_{t=1}^{n-v} \binom{n-v}{t} h_i^t y_i^{n-v-t}.$$

If we put $v + t = u$ we obtain

$$\begin{aligned} f(y_i + h_i) - f(y_i) &= h_i \sum_{v=0}^{n} a_v \sum_{u=v+1}^{n} \binom{n-v}{u-v} h_i^{u-v-1} y_i^{n-u} \\ &= h_i \sum_{u=1}^{n} y_i^{n-u} \sum_{v=0}^{u-1} a_v \binom{n-v}{u-v} h_i^{u-v-1} \\ &= h_i \sum_{u=1}^{n} a_{i,u} y_i^{n-u} = h_i \phi_i(y_i). \end{aligned}$$

Note that ϕ_i is a polynomial of degree $n-1$, but its coefficients depend on h_i. We get the new equation

$$h_1 \phi_1(y_1) + \ldots + h_{\frac{k}{2}} \phi_{\frac{k}{2}}(y_{\frac{k}{2}}) = 0.$$

Let us for now view the numbers h_i with $1 \leq i \leq \frac{k}{2}$ as fixed. Note that $\frac{k}{2} = 2n \cdot 2^s$, so that we set $k_0 = 2n$ and $l = k_0 2^s$. We define the sets

$$A_i = \{h_i \phi_i(y_i) \colon y_i \in [-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}}] \cap \mathbb{Z}\}.$$

Now we are trying to solve the equation

$$x_1 + \ldots + x_l = 0 \text{ with } x_i \in A_i.$$

But this is precisely the situation from Lemma 3.13. This leaves us to count solutions to

$$y^{(1)}+\ldots+y^{(2^{s-1})}-y^{(2^{s-1}+1)}-\ldots-y^{(2^s)}=0 \tag{4}$$

under the constraints

$$y^{(j)}=y_1^{(j)}+\ldots+y_{k_0}^{(j)} \text{ with } y_i^{(j)}\in A_{mk_0+i} \text{ for } 1\le j\le 2^s,$$

for some $0\le m\le 2^s-1$. We start by looking at $m=0$. Expanding our equation in this case, yields the expression

$$h_1\underbrace{\left[\phi_1(v_1^{(1)})+\ldots+\phi_1(v_1^{(2^{s-1})})-\phi_1(v_1^{(2^{s-1}+1)})-\ldots\phi_1(v_1^{(2^s)})\right]}_{z_1}$$
$$+\ldots+h_{k_0}\underbrace{\left[\phi_{k_0}(v_{k_0}^{(1)})+\ldots+\phi_{k_0}(v_{k_0}^{(2^{s-1})})-\phi_{k_0}(v_{k_0}^{(2^{s-1}+1)})-\ldots\phi_{k_0}(v_{k_0}^{(2^s)})\right]}_{z_{k_0}}=0.$$

Recall that the coefficients of $\phi_i(y)$ are given by the sums

$$a_{i,u}=\sum_{v=0}^{u-1}a_v\binom{n-v}{u-v}h_i^{u-v-1}.$$

If $h_i\in[-2N^{\frac{1}{n}},2N^{\frac{1}{n}}]\cap\mathbb{Z}$, then we can estimate

$$a_{i,u}\ll_n\sum_{v=0}^{u-1}a_v\binom{n-v}{u-v}N^{\frac{u-v-1}{n}}\ll_n\sum_{v=0}^{u-1}\binom{n-v}{u-v}N^{\frac{u-1}{n}}\ll_n N^{\frac{u-1}{n}}.$$

Further since we are assuming that $v_i^{(j)}\in[-2N^{\frac{1}{n}},2N^{\frac{1}{n}}]$ we get (estimating trivial) that

$$\phi_i(v_i^{(j)})\ll_n N^{\frac{n-1}{n}} \text{ for } 1\le i\le k_0 2^s \text{ and } 1\le j\le 2^s.$$

Thus we obtain the condition $z_i\ll_n N^{\frac{n-1}{n}}$.

The next step is to estimate the number of solutions to

$$z_i=m. \tag{5}$$

To do so we will use the induction hypothesis. Note that $1<k'=k(n-1)<2^{s-1}$ so that we can write (5) as

$$\phi_i(v_i^{(1)})+\ldots+\phi_i(v_i^{(k')})=m-\phi_i(v_i^{(k'+1)})-\ldots-\phi_i(v_i^{(2^{s-1})})+\phi_i(v_i^{(2^{s-1}+1)})+\ldots\phi_i(v_i^{(2^s)})$$
$$:=m'.$$

Before we apply the induction hypothesis we recall that

$$a_{i,u}\ll_n N^{\frac{u-1}{n}}=\left(N^{\frac{n-1}{n}}\right)^{\frac{u-1}{n-1}} \text{ and } m'\ll N^{\frac{n-1}{n}}.$$

Thus, applying (3) we find that

$$r_{\phi_i,k'}(m') \ll \left(N^{\frac{n-1}{n}}\right)^{\frac{k'}{n-1}-1} = N^{\frac{k'-n+1}{n}}.$$

This is for fixed $v_i^{(k'+1)},\ldots,v_i^{(2^s)}$. For these we have at most $(1+2N^{\frac{1}{n}})^{2^s-k'}$ choices. Thus we have

$$\sharp\{z_i = m\colon v_i^{(j)} \in [-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}}] \cap \mathbb{Z}\} \ll_n N^{\frac{2^s-n+1}{n}}.$$

The same argument can be made for all $m \neq 0$. We summarize our findings and translate them back to bounds on $r_{f,k}(m)$. We have

$$\begin{aligned} r_{f,k}(m) \ll_n \sharp\{|z_i| \ll N^{\frac{n-1}{n}} \text{ with multiplicity } \lambda_i \ll N^{\frac{2^s-n+1}{n}},\ h_i \in [-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}}]\cap\mathbb{Z}\colon \\ 0 = z_1h_1 + \ldots + z_{k_0}h_{k_0}\}. \end{aligned}$$

The right hand side is independent of f and we name it $A_{m,k}$. Note that only k_0 out of all the $\frac{k}{2}$ variables h_i (for $1 \leq i \leq \frac{k}{2}$). Thus we have to account for all the remaining choices trivially. This will account for a factor of

$$\ll_n \left(N^{\frac{1}{n}}\right)^{\frac{k}{2}-k_0} = N^{\frac{k}{2n}-2}.$$

We can write

$$\begin{aligned} A_{m,k} &\ll_n N^{\frac{k}{2n}-2} \cdot \left(N^{\frac{2^s-n+1}{n}}\right)^{k_0} \\ &\qquad \cdot \sum_{h_1,\ldots,h_{k_0}} \sharp\{(z_1,\ldots,z_k) \in \mathbb{Z}^{k_0}\colon h_1z_1 + \ldots + h_{k_0}z_{k_0} = 0 \text{ and } z_i \ll N^{\frac{n-1}{n}}\} \\ &\ll_n N^{2^{s+2}-2n} \sum_{\mathbf{h}\in\mathbb{Z}^{k_0}\cap B_{2N^{\frac{1}{n}}}(0)} \sharp\{\mathbf{z} \in \mathbb{Z}^{k_0} \cap B_{c_nN^{\frac{n-1}{n}}}(0)\colon \langle \mathbf{h}, \mathbf{z}\rangle = 0\}. \end{aligned}$$

Here the h_i-sums run over integers between $-2N^{\frac{1}{n}}$ and $2N^{\frac{1}{n}}$. Essentially we have simply removed the multiplicity of the z_i's by brute force and trivially dealt with the variables h_i that don't participate in the equation.

We have now arrived at the point where we can estmate $A_{m,k}$ and thus $r_{f,k}(m)$ using our results on linear equations. We get

$$r_{f,k}(m) \ll_n A_{m,k} \ll_n N^{2^{s+2}-2n}\left(N^{\frac{1}{n}} \cdot N^{\frac{n-1}{n}}\right)^{k_0-1} = N^{2^{s+2}-1} = N^{\frac{k}{n}-1}.$$

This concludes the induction step and thus the proof. □

Theorem 3.15 (Hilbert). *There is $k = k(n) \in \mathbb{N}$ such that $\mathbb{N} \subset A_n^{(k)}$. In other words, every natural number can be represented as a sum of at most k nth-powers of positive integers.*

Proof. Note that

$$R_k(N) = \sum_{m=0}^{N} r_k(m) = \sharp\{(x_1, \ldots, x_k) \in \mathbb{Z}_{\geq 0}^k \colon x_1^n + \ldots + x_k^n \leq N\}.$$

First we observe that $R_k(N) \geq \left(\frac{N}{k}\right)^{\frac{k}{n}}$. To see this we just take $(x_1, \ldots, x_k) \in ([0, \lfloor \sqrt[n]{\frac{N}{k}} \rfloor] \cap \mathbb{Z}_{\geq 0})^k$ arbitrary and observe that

$$x_1^n + \ldots + x_k^n \leq N.$$

We take $k = k(n)$ as in the previous Lemma and assume that $d(A_n^{(k)}) = 0$. Note that $1 \in A_n^{(k)}$. Thus, for $\epsilon > 0$ and sufficiently large N we have

$$A_n^{(k)}(N) \leq \epsilon N.$$

We apply the previous lemma to estimate

$$R_k(N) = r_k(0) + \sum_{m=1}^{N} r_k(m) \leq 1 + C_n N^{\frac{k}{m}-1} A_n^{(k)}(N) < 1 + C_n \epsilon N^{\frac{k}{n}},$$

for some positive constant C_n. If we take $2C_n\epsilon = k^{-\frac{k}{n}}$ this reads

$$R_k(N) < 1 + \frac{1}{2}\left(\frac{N}{k}\right)^{\frac{k}{n}}.$$

Note that for $N \leq k$ we have $A_n^{(k)}(N) = N$. Thus we have $N > k$. Thus $\left(\frac{N}{k}\right)^{\frac{k}{n}} > 1$. We conclude that $R_k(N) < \left(\frac{N}{k}\right)^{\frac{k}{n}}$. But this is a contradiction to the lower bound obtained above.

Therefore we have seen that $d(A_n^{(k)}) > 0$. But this proves the result, since according to Schnirelmann every set with positive density (containing 0) is a basis of $\mathbb{N}$. □

4. Partition numbers

A solution $(n_1, n_2, \ldots, n_l) \in \mathbb{N}^l$ to the equation

$$n = n_1 + n_2 + \ldots + n_l$$

is called a partition of n into l parts. Let $p_l(n)$ denote the total number of partitions of n into l parts. The total partition number $p(n)$ of n is then given by

$$p(n) = \sum_{l=1}^{n} p_l(n).$$

We introduce the associated power series

$$F(z) = \sum_{n=0}^{\infty} p(n) z^n = \prod_{m=1}^{\infty} \frac{1}{1 - z^m}.$$

The partition numbers have many reincarnations and are the star of many incredible formulae.

Studying the asymptotic behaviour of $p(n)$ is an important problem which ultimately led to the genesis of the circle method. We will derive a full asymptotic expansion of $p(n)$ following the work of H. Rademacher (1937). This is a direct extension of the work of G. H. Hardy and S. Ramanujan from 1918, where circle method was first introduced.

4.1. **The Dedekind eta function.** We first have to study the transformation behaviour of the Fourier Series

$$f(z) = \sum_{n=0}^{\infty} p(n)e(nz).$$

In view of the product expansion of $F(z)$ given above we put

$$\eta(z) = e\left(\frac{z}{24}\right)\prod_{m=1}^{\infty}(1-e(mz)) \text{ so that } f(z) = \frac{e(\frac{z}{24})}{\eta(z)}.$$

We will have to establish a certain transformation behaviour of $\eta(z)$. This will be nothing new for the modular form enthusiasts. Here we will only rigorously derive those facts that we need later on. A more exhaustive discussion can be found for example in [Ap].

Obviously $\eta(z+1) = e(\frac{1}{24})\eta(z)$. We also need the following result.

Lemma 4.1. *We have* $\eta(-\frac{1}{z}) = \sqrt{\frac{z}{i}}\eta(z)$.

This result can be established in a variety of ways. One very interesting option is to use Kronecker's limit formula or any other connection to Eisenstein series. Here we give an ad-hoc argument following C. L. Siegel.

Proof. Following standard terminology from the modular form world we put $q = e(z)$. We thus have

$$\frac{\pi i z}{12} - \log(\eta(z)) = -\sum_{m=1}^{\infty}\log(1-q^m) = \sum_{m=1}^{\infty}\sum_{k=1}^{\infty}\frac{1}{k}q^{km} = \sum_{k=1}^{\infty}\frac{1}{k}(q^{-k}-1)^{-1}.$$

Here we used the Taylor expansion of the logarithm and the geometric series.

We now consider the difference

$$\begin{aligned} f(z) = \pi i\frac{z+z^{-1}}{12} - \log(\eta(z)) + \log(\eta(-\frac{1}{z})) &= \sum_{k=1}^{\infty}\frac{1}{k}\left[\frac{1}{e^{-2\pi ikz}-1} - \frac{1}{e^{2\pi i/z}-1}\right] \\ &= \frac{i}{2}\sum_{k=1}^{\infty}\frac{1}{k}\left[\cot(\pi kz) + \cot(\pi k/z)\right]. \end{aligned}$$

In the last step we used that $\cot(x)$ is odd and the identity $\frac{1}{e^{-2\pi ix}-1} = \frac{i}{2}(\cot(\pi x)+i)$ Note that we have to show that

$$f(z) = \pi i \frac{z+z^{-1}}{12} + \frac{1}{2}\log(z/i).$$

We define

$$g_n(\tau) = \frac{1}{\tau}\cot(\pi(n+\frac{1}{2})\tau)\cot(\pi(n+\frac{1}{2})\tau/z)$$

for $n \in \mathbb{N}_0$. Of course g_n is a meromorphic function on $\mathbb{C}$ and we can analyse its pole structure. We claim that there is a pole of order 3 at 0 and we have

$$\mathrm{res}_{\tau=0} g_n(\tau) = -\frac{1}{3}(\tau+\tau^{-1}).$$

Further for $k \in \mathbb{N}$ we have

$$\mathrm{res}_{\tau=\pm\frac{k}{n+\frac{1}{2}}} = \frac{1}{\pi k}\cot(\pi k/z) \text{ and } \mathrm{res}_{\tau=\pm\frac{kz}{n+\frac{1}{2}}} = \frac{1}{\pi k}\cot(\pi kz).$$

We consider the path γ which goes around the rhombus with vertices at $1, z, -1, z$. By the residual theorem we compute

$$\frac{1}{8}\int_\gamma g_n(\tau)d\tau = -\frac{\pi i}{12}(\tau+\tau^{-1}) + \frac{i}{2}\sum_{k=1}^{n}\frac{1}{k}(\cot(\pi kz)+\cot(\pi k/z)).$$

Taking the limit $n \to \infty$ we get

$$\lim_{n\to\infty}\frac{1}{8}\int_\gamma g_n(\tau)d\tau = -\frac{\pi i}{12}(\tau+\tau^{-1}) + f(z).$$

Explicitly one can compute

$$\lim_{n\to\infty}\frac{1}{8}\int_\gamma g_n(\tau)d\tau = \left[\int_1^\tau - \int_\tau^{-1} + \int_{-1}^{-\tau} - \int_{-\tau}^{1}\right]\frac{1}{8\tau}d\tau = \frac{1}{2}\log(\frac{\tau}{i}).$$

This concludes the proof modulo some minor details. □

Given a matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ we define the action on the upper half plane $\mathbb{H}$ by

$$\gamma.z = \frac{az+b}{cz+d} \in \mathbb{H}.$$

This is a transitive group action. Further we write $j(\gamma, z) = cz+d$. We can now establish the following important transformation behaviour of η.

Theorem 4.2. *For $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ we have*

$$\eta(\gamma z) = \theta(\gamma)(-i\cdot j(\gamma,z))^{\frac{1}{2}}\eta(z),$$

where $\theta(-\gamma) = e(\frac{1}{4})\theta(z)$, $\theta(\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}) = e(\frac{b}{24})$ *and for* $c > 0$*:*

$$\theta(\gamma) = e\left(\frac{a+d}{24c} - \frac{1}{2}s(d,c)\right) \text{ where } s(d,c) = \sum_{0 \le x < c} \frac{x}{c}(\frac{dx}{c} - \left[\frac{dx}{c}\right] - \frac{1}{2}).$$

This turns η in a modular form of half-integral weight for the multiplier system θ. The function $s(d,c)$ is called Dedekind-sum. We have the important reciprocity formula

$$s(d,c) + s(c,d) = \frac{1}{12}\left(\frac{d}{c} + \frac{c}{d} + \frac{1}{cd} - 3\right).$$

Note that the matrices $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ generate $\mathrm{SL}_2(\mathbb{Z})$. Thus one we can apply the formulae for $\eta(z+1)$ and $\eta(-1/z)$ to derive the general formula. Note that there is a different approach based on an identity by Iseki. The latter seems conceptual more interesting but requires some non-trivial properties of the Hurwitz-zeta function which I would like to avoid.

Proof. We start with some preliminaries. First suppose $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c > 0$. We claim that

$$\theta(\gamma T^m) = e(\frac{m}{24})\theta(\gamma).$$

This can be derived from the observation that

$$\gamma T^m = \begin{pmatrix} a & am + b \\ c & cm + d \end{pmatrix}.$$

Thus by definition we have

$$\theta(\gamma T^m) = e(\frac{a + cm + d}{24} - s(cm + d, c)).$$

We are done once we have seen that $s(cm + d, c) = s(d,c)$. To verify this is left to the reader.

Next we claim that

$$\theta(\mathrm{sgn}(d)\gamma S) = \begin{cases} e(-\frac{1}{8})\theta(\gamma) & \text{if } d > 0, \\ e(\frac{1}{8})\theta(\gamma) & \text{if } d < 0. \end{cases}$$

To see this we first write

$$\gamma S = \begin{pmatrix} b & -a \\ d & -c \end{pmatrix}.$$

Consider the case $d > 0$. By definition of θ we have

$$\theta(\gamma S) = e(\frac{b-c}{24d} - \frac{1}{2}s(-c,d)) = e(\frac{b-c}{24d} + \frac{1}{2}s(c,d))$$

Here we used the identity $s(-c,d) = -s(c,d)$. Recall the reciprocity law for the Dedekind sum:

$$\begin{aligned} s(c,d) + s(d,c) &= \frac{c}{12d} + \frac{d}{12c} - \frac{1}{4} + \frac{1}{12cd} \\ &= \frac{c}{12d} + \frac{d}{12c} - \frac{1}{4} + \frac{ad-bc}{12cd}. \end{aligned}$$

We rearrange this to get

$$\frac{b-c}{12d} + s(c,d) = \frac{a+d}{12c} - s(d,c) - \frac{1}{4}.$$

Inserting this above concludes the treatment of the case $d > 0$. If $d < 0$ we compute

$$\theta(-\gamma S) = \theta(-\gamma S) = e(\frac{-b+c}{-24d} - \frac{1}{2}s(c,-d)).$$

Observe that $-d > 0$, so that we use the reciprocity law in the form

$$s(c,-d) + s(-d,c) = \frac{c}{-12d} - \frac{d}{12c} - \frac{1}{4} - \frac{ad-bc}{12cd}.$$

Rearranging terms and using $s(-d,c) = -s(d,c)$ yields

$$\frac{b-c}{12d} - s(c,-d) = \frac{1}{4} + \frac{a+d}{12c} - s(d,c).$$

The result follows immediately.

We come to the most important claim. Suppose the functional equation of η holds for γ. We claim that it then also holds for $\gamma' = \gamma T^m$ and $\gamma'' = \gamma S$. We start by computing

$$\eta(\gamma'.z) = \eta(\gamma.(T^m.z)) = \theta(\gamma)(-i(cT^m.z+d))^{\frac{1}{2}}\eta(T^m.z) = \theta(\gamma)e(\frac{m}{24})(-i(cz+cm+d))^{\frac{1}{2}}\eta(z).$$

By our first claim we have $\theta(\gamma)e(\frac{m}{24}) = \theta(\gamma T^m) = \theta(\gamma')$ and we are done. Similarly we compute

$$\eta(\gamma''.z) = \eta(\gamma.(S.z)) = \theta(\gamma)(-i(cS.z+d))^{\frac{1}{2}}\eta(Sz) = \theta(\gamma)(-i(cS.z+d))^{\frac{1}{2}}(-iz)^{\frac{1}{2}}\eta(z).$$

Here we used Lemma 4.1 to transform $\eta(S.z)$. We now treat $d > 0$ first. Write

$$cS.z + d = -\frac{c}{z} + d = \frac{dz-c}{z}.$$

We can thus rewrite

$$-i(cS.z+d) = \frac{-i(dz-c)}{-iz}e(-\frac{1}{4}).$$

Thus we have

$$\eta(\gamma''.z) = \eta(\gamma.(S.z)) = \theta(\gamma)(-i(cS.z+d))^{\frac{1}{2}}\eta(Sz) = \theta(\gamma)e^{-\frac{1}{8}}(-i(dz-c))^{\frac{1}{2}}\eta(z).$$

We are done with this case using our second claim. The case of negative d is similar but requires a little care with the branches of the square root.

With this final claim at hand we are done, since it is known that $\mathrm{SL}_2(\mathbb{Z})$ is generated by T and S. Furthermore we have seen above that the functional equation is true for $\gamma = S$.

□

We will need one particular instance of this transformation formula. Take $z'' = -\frac{d}{c} - \frac{1}{cz'}$. Now let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ with $c > 0$ and check

$$\gamma.z'' = \frac{a}{c} + \frac{z'}{c}.$$

Applying the transformation with $z' = \frac{iz}{c}$ we get

$$f(\frac{a}{c} + \frac{iz}{c^2}) = c^{-\frac{1}{2}} e(\frac{1}{2} s(d,c)) \underbrace{z^{\frac{1}{2}} \exp(\frac{\pi}{12}(\frac{1}{z} - \frac{z}{c^2}))}_{=E_c(z)} f(-\frac{d}{c} + \frac{i}{z}). \tag{6}$$

Note that we must restrict ourselves to those z such that $-\frac{d}{c} + \frac{i}{z}$ lies in the upper half plane. This is as far as our study of the eta function takes us.

4.2. **Farey fractions.** Given $C \in \mathbb{N}$ we define the Farey sequence of order C to be

$$\mathcal{F}(C) = \{\frac{a}{c} \colon 1 \leq c \leq C,\ (a,c) = 1\}.$$

We usually order this sequence to be increasing. Suppose we are looking at 3 consecutive Farey fractions

$$\ldots < \frac{a'}{c'} < \frac{a}{c} < \frac{a''}{c''} < \ldots.$$

We will first show that $ac' - a'c = 1$ (resp. $a''c - ac'' = 1$). This is done by induction the case of $C = 1$ being trivial. We need some simple observations which the reader will have to verify on its own. First note that the case $\frac{a}{c}, \frac{a'}{c'} \in \mathcal{F}(C) \setminus \mathcal{F}(C-1)$ can not occur. Thus we can assume that $\frac{a'}{c'} \in \mathcal{F}(C-1)$ (the opposite case being similar). Suppose $\frac{p}{q}$ is the (right) neighbour of $\frac{a'}{c'}$ in the Farey sequence $\mathcal{F}(C-1)$. If $\frac{a}{c} = \frac{q}{q}$ we are done according to the induction hypothesis. Otherwise $\frac{a}{c}$ is given as the mediant of $\frac{a'}{c'}$ and $\frac{p}{q}$. In symbols:

$$\frac{a'}{c'} < \frac{a}{c} = \frac{a'+p}{c'+q} < \frac{p}{q}.$$

Thus we conclude by

$$ac' - a'c = (a'+p)c' - a'(c'+q) = pc' - a'q = 1,$$

where we applied the induction hypothesis.

With this property established the numerators of consecutive elements in the Farey sequence can be related by

$$a' = \frac{ac'-1}{c} \text{ and } a'' = \frac{ac''+1}{c}.$$

Further, we have the following properties of the denominators:

$$C - c < c' \leq C \text{ and } ac' \equiv 1 \bmod c,$$
$$C - c < c'' \leq C \text{ and } ac'' \equiv -1 \bmod c.$$

With this at hand we can compute the mediants

$$\frac{a'+a}{c'+c} = \frac{a}{c} - \frac{1}{c(c+c')} = \frac{a'}{c'} + \frac{1}{c'(c+c')},$$
$$\frac{a+a''}{c+c''} = \frac{a}{c} + \frac{1}{c(c+c'')} = \frac{a''}{c''} - \frac{1}{c''(c+c'')}.$$

These mediants are reduced fractions but do not belong to the Farey sequence $\mathcal{F}(C)$. However, they will be very important to our applications.

Finally we introduce the so called Ford circles:[2]

$$\mathcal{C}(\frac{a}{c}) = \{z\colon |z - \frac{a}{c} - \frac{i}{2c^2}| = \frac{1}{2c^2}\}.$$

This is a circle in $\mathbb{H}$ which is tangent to the real line at $\frac{a}{c}$.

We will now verify the picture from Figure 1.

To do so we put

$$\alpha_{\frac{a}{c}} = \frac{a}{c} - \frac{1}{c(c'+ic)} \text{ and } \beta_{\frac{a}{c}} = \frac{a}{c} + \frac{1}{c(c''-ic)}.$$

It is a straight forward computation to check that this is the only point where the corresponding circles intersect (touch).

The segment of $\mathcal{C}(\frac{a}{c})$ connecting $\alpha_{\frac{a}{c}}$ and $\beta_{\frac{a}{c}}$ without passing through $\frac{a}{c}$ will be denoted by $\gamma_{\frac{a}{c}}$. We obtain a curve

$$\gamma_C = \bigcup_{\frac{a}{c}\in\mathcal{F}(C)} \gamma_{\frac{a}{c}}$$

which is one periodic.

Finally let us remark that in hyperbolic geometry (on $\mathbb{H}$) the Ford circles are certain horocylces. Indeed they are obtained from the horizontal line $\mathrm{Im}(z) = 1$ by applying a certain Möbius transform γ.

[2] Ford Circles can be nicely visualised with Mathematica. See https://demonstrations.wolfram.com/FordCircles/ for a demonstration.

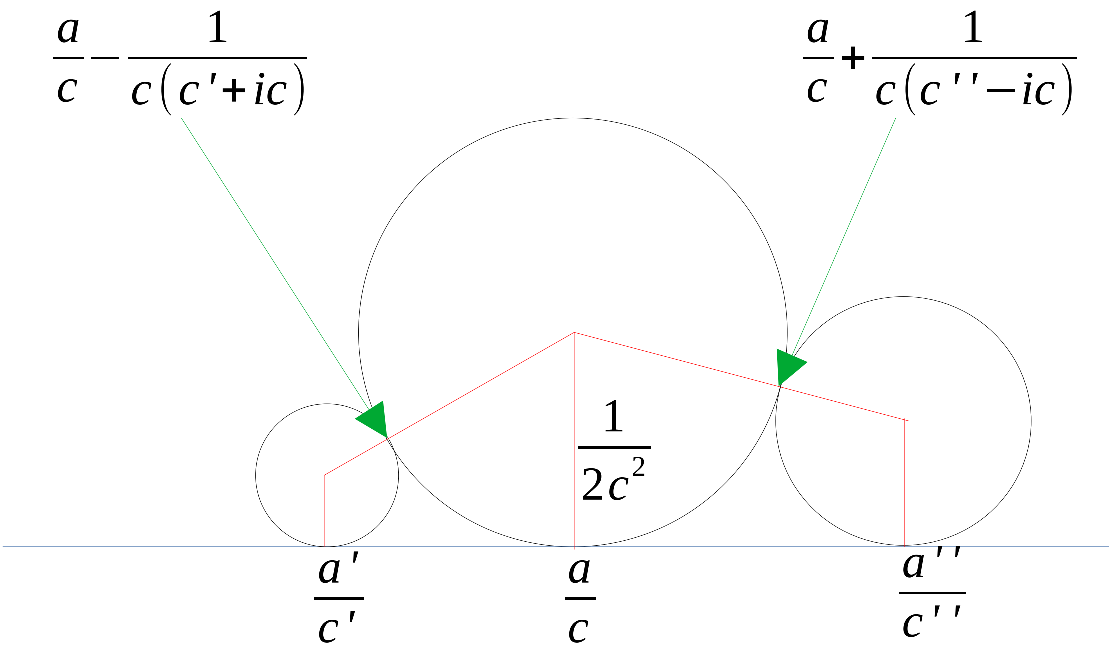
$\frac{a}{c}-\frac{1}{c(c'+ic)}$
$\frac{a}{c}+\frac{1}{c(c''-ic)}$
$\frac{1}{2c^2}$
$\frac{a'}{c'}$
$\frac{a}{c}$
$\frac{a''}{c''}$

FIGURE 1. Ford circles $\mathcal{C}(\frac{a'}{c'})$, $\mathcal{C}(\frac{a}{c})$ and $\mathcal{C}(\frac{a''}{c''})$

4.3. **Setting up the circle method.** Finally we will set up the circle method in order to produce an asymptotic formula for the partition numbers. Doing so we closely follow the exposition in [IK]. We can recover the coefficients $p(n)$ from the series $F(z)$ or $f(z)$ as follows:

$$p(n) = \frac{1}{2\pi i}\int_{\partial B_r(0)} F(z)z^{-n-1}dz = \int_w^{w+1} f(z)e(-nz)dz. \tag{7}$$

The first equality is simply Cauchy's integral, while the second follows from periodicity of the Fourier Series $f(z)$.

In the second integral of (7) we can take any continous path from w to $w+1$ in the upper half plane. Furthermore, $w \in \mathbb{H}$ can be chosen arbitrarily. We make our choices so that the integral is taken along the path

$$\gamma_C' = \bigcup_{\substack{1\le c\le C,\\ 1\le a\le C,\\ (a,c)=1}} \gamma_{\frac{a}{c}}.$$

We obtain the dissection

$$p(n) = \sum_{1\le c\le C}\ \sideset{}{^*}\sum_{a \bmod c} \underbrace{\int_{\gamma_{\frac{a}{c}}} f(z)e(-nz)dz}_{=H_{ac}(n)} \tag{8}$$

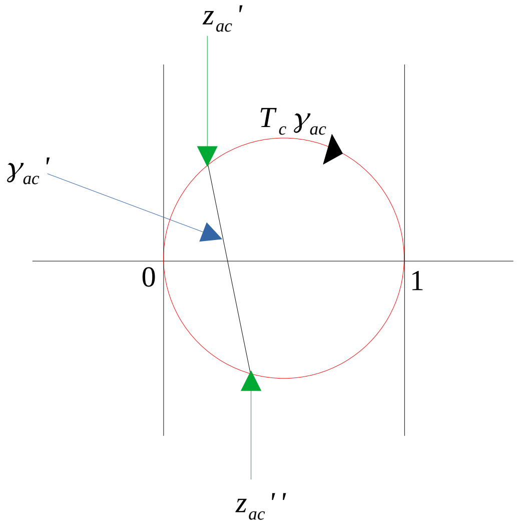

FIGURE 2. This is the translation of the Ford circle $\mathcal{C}(\frac{a}{c})$ by the change of variables T_c.

We start by making the change of variables (denoted by T_c^{-1})

$$z \mapsto \frac{a}{c} + \frac{iz}{c^2}.$$

This rotates, magnifies and translates the Ford circle $\mathcal{C}(\frac{a}{c})$ to the circle centered at $\frac{1}{2}$ with radius $\frac{1}{2}$. The end points of the piece we are integrating over become

$$z'_{ac} = \frac{ic}{c' + ic} \text{ and } z''_{ac} = \frac{-ic}{c'' - ic}.$$

See Figure 2 for illustration. This yields

$$H_{ac}(n) = \frac{i}{c^2} e(-\frac{an}{c}) \int_{z'_{ac}}^{z''_{ac}} f(\frac{a}{c} + \frac{iz}{c^2}) e(-\frac{inz}{c^2}) dz.$$

Applying (6) we get

$$H_{ac}(n) = c^{\frac{1}{2}} e(\frac{1}{2}s(a,c) - \frac{an}{c}) \underbrace{\frac{i}{c^2} \int_{z'_{ac}}^{z''_{ac}} f(-\frac{d}{z} + \frac{i}{z}) E_c(z) e^{2\pi nz/c^2} dz}_{=I_{a,c}(n)}.$$

Let us first compute the integral which provides the main therm and then justify why this is the case. This integral is

$$\begin{aligned}\tilde{I}_c(n) &= \frac{i}{c^2}\int_{\partial B_{\frac{1}{2}}(\frac{1}{2})} E_c(z)e^{2\pi nz/c^2}dz \\ &= \frac{i}{c^2}\int_{(1)} \exp(\frac{\pi z}{12} + \frac{2\pi}{c^2 z}(n - \frac{1}{24}))z^{-\frac{5}{2}}dz \\ &= \frac{2\pi}{c^2}\cdot\left(\frac{c}{12\lambda_n}\right)^{\frac{3}{2}} I_{\frac{3}{2}}\left(\frac{B}{c}\lambda_n\right).\end{aligned}$$

In the first step we changed $z \mapsto \frac{1}{z}$. The second step includes looking up the integral. Here $I_{\frac{3}{2}}$ is the I-Bessel function, $\lambda_n = \sqrt{n - \frac{1}{24}}$ and $B = \frac{2\pi}{\sqrt{6}}$. One could easily just work with the I-Bessel function, but in order to bring it in a more classical form we make the following observations:

$$\begin{aligned}I_{\frac{1}{2}}(x) &= \frac{\sqrt{2}\sinh(x)}{\sqrt{\pi x}} \text{ and} \\ I_{\frac{3}{2}}(x) &= \sqrt{x}\frac{d}{dx}\left[\frac{I_{\frac{1}{2}}(x)}{\sqrt{x}}\right].\end{aligned}$$

We obtain

$$\tilde{I}_c(n) = \frac{1}{\sqrt{2}\pi}\frac{d}{dn}\lambda_n^{-1}\sinh(\frac{B}{c}\lambda_n).$$

In particular $\tilde{I}_c(n)$ is independent of a. Thus we wish to replace $I_{a,c}$ by it with a controllable error. This is made precise in the following lemma.

Lemma 4.3. *We have*

$$H_{ac}(n) = \frac{1}{\sqrt{2}\pi}c^{\frac{1}{2}}e(\frac{1}{2}s(a,c) - \frac{an}{c})\frac{d}{dn}\lambda_n^{-1}\sinh(\frac{B}{c}\lambda_n) + O(c^{-1}C^{-\frac{3}{2}}\exp(2\pi nC^{-2})).$$

Proof. Since the strategy is clear we write

$$I_{a,c} - \tilde{I}_c(n) = I_1 - I_2 - I_3$$

for

$$\begin{aligned}I_1 &= \int_{z'_{ac}}^{z''_{ac}} \left(f\left(-\frac{d}{c} + \frac{i}{z}\right) - 1\right) E_c(z)e^{2\pi nz/c^2}dz, \\ I_2 &= \int_0^{z'_{ac}} E_c(z)e^{2\pi nz/c^2}dz \text{ and} \\ I_3 &= \int_{z''_{ac}}^1 E_c(z)e^{2\pi nz/c^2}dz.\end{aligned}$$

To estimate I_1 we need to move the path of integration. Indeed we move the path from the arc to the straight line connecting z'_{ac} and z''_{ac} (i.e. the chord). On this new path we have the estimate

$$|z| \leq \min(|z'_{ac}|, |z''_{ac}|) \leq 2cC^{-1} \text{ and } \operatorname{Re}(z) \leq \max(\operatorname{Re}(z'_{ac}), \operatorname{Re}(z''_{ac})) \leq c^2C^{-2}.$$

The length of the path can be bounded by $|z'_{ac}| + |z''_{ac}| \leq 4cC^{-1}$. We will use the estimate

$$\begin{aligned}
&\left(f\left(-\frac{d}{c} + \frac{i}{z}\right) - 1\right) E_c(z) e^{2\pi nz/c^2} \\
&= z^{\frac{1}{2}} \sum_{m=1}^{\infty} p(m) e(-\frac{dm}{c}) \exp\left(-\frac{2\pi}{z}(m - \frac{1}{24}) + \frac{2\pi z}{c^2}(n - \frac{1}{24})\right) \\
&\ll \left(\frac{c}{C}\right)^{\frac{1}{2}} \exp(\frac{2\pi n}{C^2}).
\end{aligned}$$

Here we used the Fourier expansion of f as well as the bounds

$$\operatorname{Re}(z^{-1}) = 1, \ \operatorname{Re}(z) \leq c^2C^{-2} \text{ and } p(m) \leq 2^m.$$

Estimating the integral trivially yields

$$I_1 \ll \left(\frac{c}{C}\right)^{\frac{3}{2}} \exp(\frac{2\pi n}{C^2}).$$

Since I_2 and I_3 are very similar we only show how to deal with I_2. First we observe that the length of the path around the relevant piece of the arc is bounded by $\frac{\pi}{2}|z'_{ac}| \leq \pi cC^{-1}$. Further, on this arc we have the estimate $|z| \leq cC^{-1}$. Bounding the integrand by

$$E_c(z) e^{2\pi nz/c^2} \ll \left(\frac{c}{C}\right)^{\frac{1}{2}} \exp(\frac{2\pi n}{C^2})$$

and estimating the integral trivially yields

$$I_2 \ll \left(\frac{c}{C}\right)^{\frac{3}{2}} \exp(\frac{2\pi n}{C^2}).$$

The same bound holds for I_3. The result follows directly. □

We are now ready to prove the following theorem.

Theorem 4.4. *For $n \geq 1$ we have*

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{c=1}^{\infty} c^{\frac{1}{2}} A_c(n) \frac{d}{dn} \frac{1}{\lambda_n} \sinh(\frac{B}{c}\lambda_n)$$

for $B = \frac{2}{\pi}\sqrt{6}$, $\lambda_n = \sqrt{n - \frac{1}{24}}$ and

$$A_c(n) = \sum_{\substack{a \bmod c, \\ (a,c)=1}} e(\frac{1}{2}s(a,c) - \frac{an}{c}).$$

Proof. Recall that by (8) we have

$$p(n) = \sum_{c=1}^{C} \sum_{\substack{1 \leq a \leq c, \\ (a,c)=1}} H_{ac}(n).$$

Inserting the expression from Lemma 4.3 and recognising the a-sum as $A_c(n)$ we get

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{c=1}^{C} c^{\frac{1}{2}} A_c(n) \frac{d}{dn} \frac{1}{\lambda_n} \sinh(\frac{B}{c}\lambda_n) + O\left(\sum_{c=1}^{C} \sum_{a=1}^{c} c^{-1} C^{-\frac{3}{2}} \exp(2\pi n C^{-2}) \right).$$

Of course we trivially have

$$\sum_{c=1}^{C} \sum_{a=1}^{c} c^{-1} C^{-\frac{3}{2}} \exp(2\pi n C^{-2}) \ll C^{-\frac{1}{2}} \exp(2\pi n C^{-2}).$$

Taking the limit $C \to \infty$ completes the proof. □

Remark 4.5. The asymptotic expansion of $p(n)$ as above is originally due to Hardy-Ramanujan and is essentially the birth of the circle method. The full asymptotic expansion was derived by H. Rademacher. Furthermore, A. Selberg succeeded in deriving the closed expression

$$A_c(n) = \left(\frac{c}{2}\right)^{\frac{1}{2}} \sum_{\substack{l \bmod 2c, \\ l(3l-1) \equiv -2n \bmod 2l}} (-1)^l \cos\left(\frac{\pi}{c} \left(l - \frac{1}{6} \right) \right).$$

This gives the bound $|A_c(n)| \leq 2\tau(c) c^{\frac{1}{2}}$.

5. Some results on exponential sums

Lemma 5.1 (Weyl's inequality). *Let $f(x)$ be a real polynomial of degree k with leading coefficient α (i.e. $f(x) = \alpha x^k + \alpha_1 x^{k-1} + \ldots$). Suppose that α has a rational approximation $\frac{a}{q}$ such that*

$$(a, q) = 1,\ q > 0,\ \text{and } |\alpha - \frac{a}{q}| \leq \frac{1}{q^2}.$$

Then, for any $\epsilon > 0$ and $K = 2^{k-1}$ we have

$$\sum_{x=1}^{P} e(f(x)) \ll_{\epsilon,k} P^{1+\epsilon} \left(P^{-\frac{1}{K}} + q^{-\frac{1}{K}} + \left(\frac{P^k}{q} \right)^{-\frac{1}{K}} \right).$$

Proof. We write $S_k(f) = \sum_{P_1 < x \le P_2} e(f(x))$ for $0 \le P_2 - P_1 \le P$. Considering the absolute value squared we find

$$\begin{aligned} |S_k(f)|^2 &= \sum_{P_1 < x_1, x_2 \le P_2} e(f(x_2) - f(x_1)) \\ &= P_2 - P_1 + 2\,\mathrm{Re}\left[\sum_{P_1 < x_1 < x_2 \le P_2} e(f(x_2) - f(x_1))\right]. \end{aligned}$$

We put $x_2 = x_1 + y$ and set

$$f(x_1 + y) - f(x_1) = \Delta_y f(x_1).$$

This yields

$$|S_k(f)|^2 = P_2 - P_1 + 2\,\mathrm{Re}\left[\sum_{y=1}^{P}\sum_{x \in I_y} e(\Delta_y f(x_1))\right].$$

Note that for some y the corresponding interval I_y is empty (i.e. the x_1-sum irrelevant). We obtain the estimate

$$|S_k(f)|^2 \le P + 2\sum_{y=1}^{P} |S_{k-1}(\Delta_y f)|.$$

Here the subscript indicates that $\Delta_y f(x)$ has degree (at most) $k-1$. This is the procedure we wish to iterate until we reach linear sums. The next iterate would look like

$$|S_{k-1}(\Delta_y f)|^2 \le P + 2\sum_{z=1}^{P} |S_{k-2}(\Delta_{y,z} f)|.$$

By Cauchy's inequality we get

$$\begin{aligned} |S_k(f)|^4 &\ll P^2 + P\sum_{y=1}^{P} |S_{k-1}(\Delta_y f)|^2 \\ &\ll P^3 + P\sum_{y=1}^{P}\sum_{z=1}^{P} |S_{k-2}(\Delta_{y,z} f)|. \end{aligned}$$

Continuing this process yields

$$|S_k(f)|^{2^v} \ll P^{2^v - 1} + P^{2^v - v - 1}\sum_{y_1=1}^{P} \dots \sum_{y_v=1}^{P} |S_{k-v}(\Delta_{y_1,\dots,y_v} f)|. \tag{9}$$

Taking $v = k-1$ we find

$$\Delta_{y_1,\dots,y_{k-1}} f(x) = k!\alpha y_1 \cdot \ldots \cdot y_{k-1} x + \beta.$$

Here β is a combination of all terms that are independent of x. Thus it is key to estimate

$$|S_1(\Delta_{y_1,\dots,y_{k-1}}f)| = |\sum_x e(k!\alpha y_1 \cdot \dots \cdot y_{k-1}x)|.$$

We note that we are summing part of a geometric series. In general this yields the estimate

$$|\sum_{x=x_1}^{x_2-1} e(\lambda x)| \leq \frac{2}{|1-e(\lambda)|} = \frac{1}{|\sin(\pi\lambda)|} \ll \frac{1}{\|\lambda\|}.$$

At this point we introduce the notation $\|\lambda\|$ to be the distance of λ to the nearest integer.

Inserting this above yields

$$|S_k(f)|^K \ll P^{K-1} + P^{K-k} \sum_{y_1=1}^{P} \dots \sum_{y_{k-1}=1}^{P} \min(P, \|k!\alpha y_1 \cdot \dots \cdot y_{k-1}\|^{-1}).$$

By the standard result $d(m) \ll_\epsilon m^\epsilon$ we find that

$$\sharp\{k!y_1 \cdot \dots \cdot y_{k-1} = m\} \ll_\epsilon m^\epsilon.$$

Arranging the y_i-sums appropriately yields

$$|S_k(f)|^K \ll P^{K-1} + P^{K-k+\epsilon} \sum_{m=1}^{k!P^{k-1}} \min(P, \|\alpha m\|^{-1}).$$

At this point we make use of the rational approximation $\frac{a}{q}$ to α. We divide the m-sum in blocks of length q. There are $\ll \frac{P^{k-1}}{q} + 1$ such blocks. The sum over such a block looks like

$$\sum_{m=0}^{q-1} \min(P, \|\alpha(m_1+m)\|^{-1}).$$

We observe that, since $|\alpha - \frac{a}{q}| \leq q^{-2}$ and $0 \leq m < q$, we have

$$\alpha(m_1+m) = \alpha m_1 + \frac{am}{q} + O(\frac{1}{q}).$$

Putting $am \equiv r \bmod q$ we observe that r runs through a complete sum of residues mod q when m runs through the full block. Thus the sum is

$$\sum_{r=0}^{q-1} \min(P, \frac{1}{\|(r+b)/q + O(1/q)\|}).$$

Here b is the integer closest to $q\alpha m_1$. We estimate

$$\sum_{r=0}^{q-1} \min(P, \frac{1}{\|(r+b)/q + O(1/q)\|}) \ll P + \sum_{s=1}^{q/2} \frac{q}{s} \ll P + q\log q.$$

Since we are doing this for every block we get

$$|S_k(f)|^K \ll P^{K-1} + P^{K-k+\epsilon}(\frac{P^{k-1}}{q} + 1)(P + q\log q) \ll P^{K+\epsilon}(P^{-1} + q^{-1} + P^{-k}q).$$

The result follows by taking the Kth root. □

Lemma 5.2 (Hua's inequality)**.** *Let $f_\alpha(x) = \alpha x^k$. We have*

$$\int_0^1 |\sum_{x=1}^{P} e(f_\alpha(x))|^{2^k} d\alpha \ll_{\epsilon,k} P^{2^k-k+\epsilon}.$$

Proof. Write

$$I_v = \int_0^1 |T(\alpha)|^{2^v} d\alpha, \text{ for } T(\alpha) = \sum_{x=1}^{P} e(f_\alpha(x)).$$

By induction on v we will show that

$$I_v \ll P^{2^v-v+\epsilon} \text{ for } v = 1, \ldots, k.$$

Of course $v = k$ is exactly what is claimed.

For $v = 1$ we have

$$I_1 = \int_0^1 \sum_{x_1,x_2} e(\alpha(x_1^k - x_2^k))d\alpha = P$$

by character orthogonality.

Suppose our claim holds for $1 \le v \le k-1$. As in the proof of Weyl's bound the differencing trick yields

$$|T(\alpha)|^{2^v} \ll P^{2^v-1} + P^{2^v-v-1}\,\mathrm{Re}\left[\sum_{y_1=1}^{P} \cdots \sum_{y_v=1}^{P} S_{k-v}(\alpha)\right],$$

where

$$S_{k-v}(\alpha) = \sum_{x \in I_{y_1,\ldots,y_v}} e(\alpha\Delta_{y_1,\ldots,y_v}(x^k)).$$

Note that by positivity this is true without absolute values inside the y-sums. From this we at once derive that

$$I_{v+1} \ll P^{2^v-1}I_v + P^{2^v-v-1} \sum_{y_1,\cdots,y_v} \mathrm{Re} \int_0^1 S_{k-v}(\alpha)|T(\alpha)|^{2^v} d\alpha.$$

We investigate the last integral:

$$\int_0^1 S_{k-v}(\alpha)|T(\alpha)|^{2^v} d\alpha = \int_0^1 \sum_x e(\alpha\Delta_{y_1,\ldots,y_v}(x^k)) \sum_{\substack{u_1,\ldots,u_{2^{v-1}},\\ v_1,\ldots,v_{2^{v-1}}}} e(\alpha u_1^k + \ldots - \alpha v_1^k - \ldots).$$

Including the y_i-sums we get the following counting problem:

$$N = \sharp\{\Delta_{y_1,\ldots,y_v}(x^k) + u_1^k + \ldots - v_1^k - \ldots = 0, 1 \le y_i, u_i, v_i, x \le P\}.$$

Thus we have

$$I_{v+1} \ll P^{2^v-1} I_v + P^{2^v-v-1} N.$$

The remainder of this proof we consider the counting problem to estimate N. Observe that $\Delta_{y_1,\ldots,y_v}(x^k)$ is positive and divisible by all $y_1, \ldots, y_v$. Fixing the u_i's and v_i's we have at most P^ϵ choices for each $1 \leq y_1, \ldots, y_v \leq P$. This is by the divisor bound. Given all y_i, u_i and v_i we note that x is uniquely determined. Thus, counting all choices for u_i and v_i, we get

$$N \ll P^{2^v+v\epsilon}.$$

Inserting this estimate above concludes the proof. □

Lemma 5.3 (van der Corput). *Let $f(x)$ be a real valued function, twice differentiable function with*

$$0 \leq f'(x) \leq \frac{1}{2} \text{ and } f''(x) > 0.$$

Then

$$\sum_{A \leq n \leq B} e(f(n)) = \int_A^B e(f(x))dx + O(1).$$

Proof. Without loss of generality we assume that $A < B$ are integers. We also count the end points with weight $\frac{1}{2}$. Further, by changing the constant term of f we assume that the difference between sum and integral is positive. The latter ensures that it is enough to consider $\cos(2\pi f(x))$ instead of $e(f(x))$.

We put $\Psi(x) = x - [x] - \frac{1}{2}$ and note that

$$\int_m^{m+1} \Psi(x)F'(x)dx = \frac{1}{2}(F(m+1) + F(m)) - \int_m^{m+1} F(x)dx.$$

Summing this over m yields

$$\sum_{m=A}^{B} F(m) = \int_A^B F(x)dx + \int_A^B \Psi(x)F'(x)dx.$$

Note that we stick to the convention that the end terms of the sum is weighted by $\frac{1}{2}$.

We are left with showing that

$$I = \int_A^B \Psi(x)(\cos(2\pi f(x)))'dx$$

is bounded.

Away from integers we develop $\Psi(x)$ into a Fourier series

$$\Psi(x) = -\sum_{v=1}^{\infty} \frac{\sin(2\pi vx)}{\pi v}.$$

Inserting this we get

$$\begin{aligned} I &= -\sum_{v=1}^{\infty} \frac{1}{v\pi} \int_A^B (\sin(2\pi v x))(\cos(2\pi f(x)))' dx \\ &= 2\sum_{v=1}^{\infty} \frac{1}{v} \int_A^B (\sin(2\pi v x))(\sin(2\pi f(x))) f'(x) dx \\ &= \sum_{v=1}^{\infty} \frac{1}{v} \int_A^B f'(x)[\cos(2\pi(vx - f(x))) - \cos(2\pi(vx + f(x)))] dx. \end{aligned}$$

Justifying the interchange of sum and integration is no issue here.

If we can show that

$$|I'| = |\int_A^B f'(x) \cos(2\pi(vx \pm f(x))) dx| < \frac{1}{\pi(2v-1)},$$

then we get

$$|I| < \frac{1}{\pi} \sum_{v=1}^{\infty} \frac{1}{v(2v-1)} < \frac{2}{\pi}$$

and this would concludes the proof. To show the desired estimate for $|I'|$ we write

$$I' = \frac{1}{2\pi} \int_A^B \frac{f'(x)}{v \pm f'(x)} \phi'(x) dx \text{ for } \phi(x) = \sin(2\pi(vx \pm f(x))).$$

The first factor is monotonic (for each positive integer v). The mean value theorem gives the desired bound. □

Remark 5.4. This can easily be generalised to the following situation. Suppose f' is monotonic on $[A, B]$ and suppose that $H_1 \leq f'(x) \leq H_2$ for $x \in [A, B]$. Then

$$\sum_{A \leq n \leq B} e(f(n)) = \sum_{h=H_1}^{H_2} \int_A^B e(f(x) - hx) dx + O(\log(2 + \max(|H_1|, |H_2|))).$$

6. An analytic approach to Waring's problem

We will now use the circle method to give another proof of Hilbert's theorem. The approach taken follows Vinogradov's modification of Hardy and Littlewood's original treatment. We closely follow the exposition from H. Davenport (see [Da2]). Recall that given $n \in \mathbb{N}$ we want to show that there is $k = k(n) \in \mathbb{N}$ such that

$$x_1^n + \ldots + x_k^n = m$$

has solutions in the non-negative integers for all $m \in \mathbb{N}$. We will do this by proving an asymptotic formula for

$$r_k(m) = \sharp\{(x_1, \ldots, x_k) \in (\mathbb{Z}_{\geq 0})^k \colon x_1^n + \ldots + x_k^n = m\}.$$

6.1. **An asymptotic formula.**

For a positive integer $P = \lceil m^{\frac{1}{n}} \rceil$ we define

$$T(\alpha) = \sum_{x=1}^{P} e(\alpha x^n).$$

Recall that $r_k(m)$ was the number of representations of m as sum of k nth powers of non-negative integers. Our starting point is the simple expression

$$r_k(m) = \int_0^1 T(\alpha)^k e(-m\alpha) d\alpha.$$

This is easily verified using character orthogonality.

For $\delta > 0$ consider the subset of the Farey sequence

$$\mathcal{F} = \mathcal{F}(P^\delta) \cap (0,1] = \{\frac{a}{q} \colon 1 \leq q \leq P^\delta,\ 1 \leq a \leq q,\ (a,q) = 1\}.$$

For each (reduced) fraction $\frac{a}{q} \in \mathcal{F}$ we consider the interval

$$\mathfrak{M}_{\frac{a}{q}} = \{\alpha \bmod 1 \colon |\alpha - \frac{a}{q}| \leq P^{-n+\delta}\}.$$

(Note that we understand $\mathfrak{M}_{\frac{1}{1}}$ in $(0,1]$, where it is essentially the union of two intervals!) These intervals do not overlap and we refer to them as *major arcs*. The *minor arcs* are given by

$$\mathfrak{m} = (0,1] \setminus \left(\bigcup_{\frac{a}{q} \in \mathcal{F}} \mathfrak{M}_{\frac{a}{q}} \right).$$

We can estimate the contribution of the minor arcs to $r_k(m)$ as follows.

Lemma 6.1. *If $k \geq 2^n + 1$, then there is $\delta' = \delta'(\delta) > 0$ such that*

$$\int_{\mathfrak{m}} |T(\alpha)|^k d\alpha \ll_{\delta,n,k} P^{k-n-\delta'}.$$

Proof. We apply Dirichlet's approximation theorem to $\alpha \in (0,1]$. This gives us a rational approximation $\frac{a}{q}$ such that

$$1 \leq q \leq P^{n-\delta} \text{ and } |\alpha - \frac{a}{q}| < \frac{1}{qP^{n-\delta}}.$$

Since $\alpha \in (0,1]$ we can achieve $1 \leq a \leq q$. From this we conclude that $\alpha \in \mathfrak{m}$ implies the important bound

$$q > P^\delta.$$

By Weyl's bound (Lemma 5.1) we get

$$|T(\alpha)| \ll_{\epsilon,n} P^{1-\frac{\delta}{K}+\epsilon} \text{ for } K = 2^{n-1}.$$

(We simplified the bound using $q > P^{\delta}$ and $P^n/q \geq P^{\delta}$. The latter follows from the size constraint on q in the approximation theorem.) We can now estimate

$$\int_{\mathfrak{m}} |T(\alpha)|^k d\alpha \ll_{\epsilon,n} P^{(k-2^n)(1-\frac{\delta}{K}+\epsilon)} \int_0^1 |T(\alpha)|^{2^n} d\alpha \ll_{\epsilon,n} P^{(k-2^n)(1-\frac{\delta}{K}+\epsilon)+2^n-n+\epsilon}.$$

Here we used Hua's inequality (Lemma 5.2). We conclude by using the assumption on k to see

$$\begin{aligned}(k-2^n)(1-\frac{\delta}{K}+\epsilon)+2^n-n+\epsilon &= k-n+(k+1-2^n)\epsilon+(2-\frac{k}{2^{n-1}})\delta \\ &\leq k-n+(k+1-2^n)\epsilon-\frac{\delta}{2^{n-1}} \leq k-n-\frac{\delta}{2^n}.\end{aligned}$$

In the last step we have chosen $\epsilon = \frac{\delta}{(k+1-2^n)2^n}$. The result follows with $\delta' = \delta 2^{-n}$. □

Remark 6.2. Generally speaking the treatment of the minor arcs is considered to be the challenging part of the circle method. Usually one requires some deep insights (such as Weyl's and Hua's inequality in this case) to make these estimates work. Once one has found a working argument to deal with minor arcs, one makes the minor arcs as large as the method permits. Then one hopes that the rest can be treated with the major arc machinery.

We turn towards the major arcs and define the integrals

$$S_{a,q} = \sum_{x=1}^{q} e(\frac{ax^n}{q}) \text{ and } I(\beta) = \int_0^P e(\beta\xi^n) d\xi.$$

Our first job is to approximate $T(\alpha)$ by these two objects.

Lemma 6.3. *For $\alpha \in \mathfrak{M}_{\frac{a}{q}}$ we have*

$$T(\alpha) = q^{-1} S_{a,q} I(\beta) + O(P^{2\delta}),$$

where $\beta = \alpha - \frac{a}{q}$.

Proof. We want to write $1 \leq x \leq P$ as $x = qy + z$ where $1 \leq z \leq q$. We get

$$T(\alpha) = \sum_{z=1}^{q} \sum_{y} e(\alpha(qy+z)^n) = \sum_{z=1}^{q} e(az^n/q) \sum_{y} e(\beta(qy+z)^n).$$

The next step is to write the y-sum as an integral. Here we will pick up some error. We make the following observations. For any differentiable function f the mean value theorem tells us

$$|f(t)-f(y)| \leq \frac{1}{2} \max|f'(t)| \text{ for } |t-y| \leq \frac{1}{2}.$$

Therefore we must have

$$\int_A^B f(x)dx - \sum_{A<x<B} f(x) \ll (B-A)\max|f(x)| + \max|f'(x)|.$$

We apply this to $f(y) = e(\beta(qy+z)^n)$. Of course $|f(x)| \leq 1$. Computing the derivative shows that

$$f'(y) = 2\pi inq\beta(qy+z)^{n-1}f(y) \ll_n q|\beta|P^{n-1}.$$

In our case we also have $B - A \ll P/q$. Thus we have

$$\begin{aligned}\sum_y e(\beta(qy+z)^n) &= \int_A^B e(\beta(qy+z)^n)dy + O_k(|\beta|P^n + 1)\\ &= q^{-1}\int_0^P e(\beta\xi^n)d\xi + O_n(|\beta|P^n+1).\end{aligned}$$

Altogether we obtain

$$T(\alpha) = q^{-1}S_{a,q}I(\beta) + O_n(q(|\beta|P^n + 1)).$$

Since we are currently dealing with major arcs we have $q \leq P^\delta$ and $|\beta| \leq P^{-n+\delta}$. Using these two bounds in the error completes the proof. □

Let $\mathfrak{M} = \bigcup_{\frac{a}{q}} \mathfrak{M}_{\frac{a}{q}}$ be the set of all major arcs.

Lemma 6.4. *There is $\delta' = \delta'(\delta) > 0$ such that*

$$\int_{\mathfrak{M}} T(\alpha)^k e(-m\alpha)d\alpha = P^{k-n}\mathfrak{S}(P^\delta, m)J(P^\delta) + O(P^{k-n-\delta'})$$

where

$$\mathfrak{S}(P^\delta, m) = \sum_{q<P^\delta}\sum_{\substack{1\leq a\leq q,\\ (a,q)=1}} q^{-k}S_{a,q}^k \cdot e(-ma/q)$$

and

$$J(P^\delta) = \int_{\gamma<P^\delta}\left(\int_0^1 e(\gamma\xi^n)d\xi\right)^k e(-\gamma)d\gamma.$$

Proof. First observe that we have the trivial bound

$$|q^{-1}S_{a,q}I(\beta)| \leq P.$$

Thus, using a binomial expansion we find

$$T(\alpha)^k = (q^{-1}S_{a,q})^k(I(\beta))^k + O(P^{k-1+2\delta}).$$

Thus each arc yields

$$\int_{\mathfrak{M}_{a,q}} T(\alpha)^k e(-m\alpha)d\alpha = (q^{-1}S_{a,q})^k e(-ma/q)\int_{|\beta|<P^{-n+\delta}} (I(\beta))^k e(-m\beta)d\beta + O(P^{k-n-1+3\delta}).$$

Summing over all admissible a and q we obtain

$$\int_{\mathfrak{M}} T(\alpha)^k e(-m\alpha)d\alpha = \mathfrak{S}(P^\delta, m) \int_{|\beta|<P^{-n+\delta}} (I(\beta))^k e(-m\beta)d\beta + O(P^{k-n-1+5\delta}),$$

where we bounded the number of tuples (a,q) trivially by $P^{2\delta}$. The same bound can be used to bound $\mathfrak{S}(P^\delta, m) \ll P^{2\delta}$.

We only have to manipulate the β-integral. To do so recall that $P = \lceil m^{\frac{1}{n}} \rceil$, so that $m - P^n \ll P^{n-1}$. Thus by the mean value theorem we have

$$|e(-\beta m) - e(-\beta P^n)| \ll |\beta| P^{n-1} \ll P^{-1+\delta}.$$

Thus the error obtained by replacing m with P^k in the integral can be observed in the big-O-term. We get

$$\int_{\mathfrak{M}} T(\alpha)^k e(-m\alpha)d\alpha = \mathfrak{S}(P^\delta, m) \int_{|\beta|<P^{-n+\delta}} (I(\beta))^k e(-P^n\beta)d\beta + O(P^{k-n-1+5\delta}),$$

Having a closer look at the remaining integral we obtain

$$\begin{aligned}\int_{|\beta|<P^{-n+\delta}} (I(\beta))^k e(-P^n\beta)d\beta &= \int_{|\beta|<P^{-n+\delta}} \left(\int_0^P e(\beta\xi^n)d\xi \right)^k e(-P^n\beta)d\beta \\ &= P^{k-n} J(P^\delta).\end{aligned}$$

This concludes the proof. □

We now define the singular series (for Waring's problem with k-variables and exponent n) by

$$\mathfrak{S}(m) = \sum_{q=1}^{\infty} \sum_{\substack{1\leq a\leq q,\\ (a,q)=1}} (q^{-1}S_{a,q})^k e(-ma/q).$$

Note that for $k \geq 2^n + 1$ this is absolutely convergent (as well as uniform in m). To see this we apply Weyl's boun (Lemma 5.1) to $S_{a,q}$ and get

$$\sum_{q=1}^{\infty} \sum_{\substack{1\leq a\leq q,\\ (a,q)=1}} (q^{-1}|S_{a,q}|)^k \ll_{k,n} \sum_{q=1}^{\infty} q^{1-\frac{k}{2^{n-1}}+\epsilon} \ll \sum_{q=1}^{\infty} q^{-1-2^{-n-1}+\epsilon} < \infty.$$

This suffices at the moment.

Theorem 6.5. *If $k \geq 2^n + 1$, the number $r_k(m)$ of representations of m by k positive integral nth powers satisfies*

$$r_k(m) = C_{n,k} m^{\frac{k}{n}-1} \mathfrak{S}(m) + O(m^{k/n-1-\delta'}),$$

for some fixed $\delta' > 0$ and

$$C_{k,s} = \frac{\Gamma(1+1/n)^k}{\Gamma(k/n)} > 0.$$

This gives an asymptotic formula for $r_k(m)$ as soon as we can show that the singular series $\mathfrak{S}(m)$ is bounded away from 0. We will leave this task for after the proof.

Proof. Combining all our results so far we obtain

$$r_k(m) = P^{k-n}\mathfrak{S}(P^\delta, m)J(P^\delta) + O(P^{k-n-\delta'}).$$

We first investigate the contribution from $J(P^\delta)$. To do so observe that

$$\int_0^1 e(\gamma\xi^n)d\xi = n^{-1}\int_0^1 \zeta^{-1+\frac{1}{n}}e(\gamma\zeta)d\zeta = n^{-1}\gamma^{-\frac{1}{n}}\int_0^\gamma \zeta^{-1+\frac{1}{n}}e(\zeta)d\zeta.$$

Since the remaining integral is bounded for all γ. (Dirichlet's convergence test for infinite integrals together with absolute convergence at 0), we obtain the estimate

$$\int_0^1 e(\gamma\xi^n)d\xi \ll_n \gamma^{-\frac{1}{n}}.$$

Thus we obtain

$$J(P^\delta) = \underbrace{\int_{\mathbb{R}}\left(n^{-1}\int_0^1 \zeta^{-1+\frac{1}{n}}e(\gamma\zeta)d\zeta\right)^k e(-\gamma)d\gamma}_{=C_{k,n}} + O(P^{-(k/n-1)+\delta})$$

At this point $C_{k,n}$ is simply a number depending on n and k. Thus, without making our error any large we can replace P by $m^{\frac{1}{n}}$ and $\mathfrak{S}(P^\delta, m)$ by $\mathfrak{S}(m)$. Thus we have seen

$$r_k(m) = C_{k,n}\mathfrak{S}(m)m^{\frac{k}{n}-1} + O(m^{\frac{k}{n}-1-\delta'}).$$

We are done as soon as we can evaluate $C_{k,n}$. To do so we first observe that

$$\int_{-\lambda}^{\lambda} e(\mu\gamma)d\gamma = \frac{\sin(2\pi\lambda\mu)}{\pi\mu}.$$

In the definition of $C_{k,n}$ we can replace the infinte γ-integral by a suitable limit and interchange integrals. We find

$$C_{k,n} = n^{-k}\lim_{\lambda\to\infty}\underbrace{\int_0^1\dots\int_0^1}_{k-\text{times}}(\zeta_1\dots\zeta_k)^{-1+\frac{1}{n}}\frac{\sin(2\pi\lambda(\zeta_1+\ldots+\zeta_k-1))}{\pi(\zeta_1+\ldots+\zeta_k-1)}d\zeta_1\dots\zeta_k.$$

We set

$$\phi(u) = \underset{u-1<\zeta_1+\ldots+\zeta_{k-1}<u}{\int_0^1\dots\int_0^1}[\zeta_1\cdot\ldots\cdot\zeta_{k-1}\cdot(u-\zeta_1-\ldots-\zeta_{k-1})]^{-1+\frac{1}{n}}d\zeta_1\dots d\zeta_{k-1}$$

By a change of variables we get

$$C_{k,n} = n^{-k}\lim_{\lambda\to\infty}\int_0^k \phi(u)\frac{\sin(2\pi\lambda(u-1))}{\pi(u-1)}du.$$

Note that $\phi(1)$ can be evaluated directly. Indeed, if $k = 2$, then we easily recognise Euler's integral representation for the Beta-function (i.e. $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$). The general case is a direct generalisation of this and we obtain

$$\phi(1) = \frac{\Gamma(1/n)^k}{\Gamma(k/n)}.$$

If we can show that $\phi(u)$ is of bounded variation, then we can use the Fourier integral theorem and obtain

$$\begin{aligned} C_{k,n} = n^{-k} \lim_{\lambda\to\infty} \int_0^k \phi(u) \frac{\sin(2\pi\lambda(u-1))}{\pi(u-1)} du &= n^{-k}\phi(1) \\ &= n^{-k}\frac{\Gamma(1/n)^k}{\Gamma(k/n)} = \frac{\Gamma(1+1/n)^k}{\Gamma(k/n)}. \end{aligned}$$

Thus we are done when we can argue that ϕ has bounded variation. To do so write

$$\phi(u) = u^{\frac{k}{n}-1} \int_0^{1/u} \dots \int_0^{1/u} [t_1 \cdot \ldots \cdot t_{k-1}(1 - t_1 - \ldots - t_{k-1})]^{-1+\frac{1}{n}} dt_1 \ldots dt_{k-1}.$$

Since the integrand is now independent of u and the range of integration contracts as u gets larger we are done. □

6.2. **The singular series.** We now turn towards the study of $\mathfrak{S}(m)$. To do so we set

$$A(q) = \sum_{\substack{a=1,\\(a,q)=1}}^{q} (q^{-1}S_{a,q})^k e(-am/q).$$

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

Lemma 6.6. *If* $(q_1, q_2) = 1$*, then*

$$A(q_1q_2) = A(q_1)A(q_2).$$

Proof. We start by writing $f(a, q) = S_{a,q}^k e(-am/q)$. If we have

$$\frac{a}{q} \equiv \frac{a_1}{q_1} + \frac{a_2}{q_2} (\text{mod } 1),\ q = q_1q_2,$$

then

$$f(a, q) = f(a_1, q_1)f(a_2, q_2).$$

To see this we compute

$$\frac{a}{q} q^k \left(\frac{z_1}{q_1} + \frac{z_2}{q_2}\right)^k \equiv \frac{a_1}{q_1}(q_2z_1)^k + \frac{a_2}{q_2}(q_1z_2)^k \text{ mod } 1.$$

With this at hand we can write

$$\begin{aligned} S_{a,q} &= \sum_{z=1}^{q} e(az^k/q) = \sum_{z_1=1}^{q_1}\sum_{z_2=1}^{q_2} e\left(\frac{a}{q}q^k\left(\frac{z_1}{q_1}+\frac{z_2}{q_2}\right)^k\right) \\ &= \sum_{z_1=1}^{q_1} e\left(\frac{a_1}{q_1}(q_2z_1)^k\right)\sum_{z_2=1}^{q_2} e\left(\frac{a_2}{q_2}(q_1z_2)^k\right) = S_{a_1,q_1}S_{a_2,q_2}. \end{aligned}$$

In the last step we simply re-ordered the sum after the change of variables $q_2z_1 \mapsto z_1$ (resp. $q_1z_2 \mapsto z_2$). The multiplicativity of f follows using

$$e(-\frac{a}{q}m) = e(-\frac{a_1}{q_1}m)e(-\frac{a_2}{q_2}m).$$

The statement of the lemma follows easily by observing

$$\sum_{\substack{a=1,\\(a,q)=1}}^{q} f(a,q) = \left(\sum_{\substack{a_1=1,\\(a_1,q_1)=1}}^{q_1} f(a_1,q_1)\right)\left(\sum_{\substack{a_2=1,\\(a_2,q_2)=1}}^{q_2} f(a_2,q_2)\right).$$

□

Lemma 6.7. *If $k \geq 2^n+1$, then we have*

$$\mathfrak{S}(m) = \prod_p \chi(p),$$

for $\chi(p) = 1 + \sum_{l=1}^{\infty} A(p^l)$. Further we have

$$\chi(p) = 1 + O(p^{-1-\delta}),\ \textit{for some } \delta > 0.$$

Proof. By the absolute convergence of $\mathfrak{S}(m)$ one can justify

$$\mathfrak{S}(m) = \sum_{q=1}^{\infty} A(q) = \prod_p \left(\sum_{v=0}^{\infty} A(p^v)\right) = \prod_p \chi(p)$$

using multiplicativity of $A(\cdot)$ and the fundamental theorem of arithmetic.

Recall that we already used the estimate

$$A(q) \ll_{n,k,\epsilon} q^{1-\frac{k}{2^n-1}+\epsilon} \ll q^{-1-\delta}$$

above. But this directly implies

$$\chi(p) - 1 = \sum_{l=1}^{\infty} A(p^l) \ll \sum_{l=1}^{\infty} p^{-l(1+\delta)} \ll p^{-1-\delta},$$

by a standard geometric series argument. □

Corollary 6.8. *If $k \geq 2^n+1$ there exists $p_0 = p_0(n)$ such that*

$$\frac{1}{2} \leq \prod_{p>p_0} \chi(p) \leq \frac{3}{2}.$$

This result will greatly improved by relating the numbers $\chi(p)$ to the number of solutions to a certain congruence.We now define

$$M(q) = \sharp\{0 < x_1, \dots, x_k \leq q \colon x_1^n + \dots + x_k^n \equiv m \bmod q\}.$$

Lemma 6.9. *We have*

$$1 + \sum_{v=1}^{N} A(p^v) = \frac{M(p^N)}{p^{N(k-1)}},$$

so that

$$\chi(p) = \lim_{N \to \infty} \frac{M(p^N)}{p^{N(k-1)}}.$$

Proof. For $q = p^N$ we write

$$\begin{aligned} M(q) &= q^{-1} \sum_{t=1}^{q} \sum_{x_1=1}^{q} \dots \sum_{x_k=1}^{q} e\left(\frac{t}{q}(x_1^n + \dots + x_k^n - m)\right) \\ &= q^{-1} \sum_{q_1 | q} \sum_{\substack{u=1, \\ (u,q_1)=1}}^{q_1} \sum_{x_1=1}^{q} \dots \sum_{x_k=1}^{q} e\left(\frac{u}{q_1}(x_1^n + \dots + x_k^n - m)\right) \end{aligned} \tag{10}$$

We also compute

$$\sum_{x=1}^{q} e(\frac{u}{q_1} x^n) = \frac{q}{q_1} \sum_{x=1}^{q_1} e(\frac{u}{q_1} x^n) = \frac{q}{q_1} S_{u,q_1}.$$

We now easily see

$$M(q) = q^{-1} \sum_{q_1 | q} \sum_{\substack{u=1, \\ (u,q_1)=1}} \left(\frac{q}{q_1}\right)^k (S_{u,q_1})^k e(-\frac{u}{q_1} m) = q^{k-1} \sum_{q_1 | q} A(q_1).$$

We conclude by setting $q = p^N$. □

It remains to thoroughly investigate the congruence at hand. For each p we set up the following notation. Let $\tau = \tau_p$ be such that $n = p^\tau n_0$ with $(n_0, p) = 1$. Further write

$$\gamma = \gamma_p = \gamma_{p,n} = \begin{cases} \tau + 1 & \text{if } p > 2, \\ \tau + 2 & \text{if } p = 2. \end{cases}$$

We will need the following result to lift certain congruences. This is a specific version of Hensel's Lemma taylor made for our purposes.

Lemma 6.10. *If the congruence $x^n \equiv m \bmod p^\gamma$ is soluble for $(m, p) = 1$, then $x^n \equiv m \bmod p^v$ is soluble for every $v > \gamma$.*

Proof. We first consider the case $p \neq 2$. Note that $(\mathbb{Z}/p^v\mathbb{Z})^\times$ is a cyclic group of order $\phi(p^v) = (p-1)p^{v-1}$. A generator g of this group is called a *primitive root to the modulus* p^v. If $v > \gamma$, then g is necessarily also a primitive root to the modulus p^γ. We write

$$m \equiv g^\mu,\ y \equiv g^\eta,\ x \equiv g^\xi \bmod p^v.$$

Looking at the exponents we find that the assumption $y^n \equiv m \bmod p^\gamma$ is equivalent to

$$n\eta \equiv \mu \bmod p^{\gamma-1}(p-1).$$

Inserting $n = p^\tau n_0 = p^{\gamma-1}n_0$ we get that μ is divisible by $p^{\gamma-1}$ and $(n_0, p-1)$. Now we find ξ such that

$$n\xi \equiv \mu \bmod p^{v-1}(p-1)$$

so that $x^k \equiv m \bmod p^v$.

We turn to the case $p = 2$. This case is slightly different due to the usual complications. But the idea is similar. Note that if n is odd (i.e. $\tau = 0$), then there is no problem, since every odd m is a nth power modulo 2^v.

Suppose that $\tau \geq 1$. In particular, since n is even we have $x^n \equiv 1 \bmod 4$ for all x. Further, 5 is a generating element (i.e. a primitive root) for the cyclic group of residue classes modulo 2^v with $\equiv 1 \bmod 4$. The order is 2^{v-2}. We proceed as earlier and write

$$m \equiv 5^\mu,\ y \equiv 5^\eta,\ x \equiv 5^\xi \bmod 2^v.$$

We get that $k\eta \equiv \mu \bmod 2^{\gamma-2}$. Again we see that μ is divisible by $2^\tau = 2^{\gamma-2}$ so that there is ξ with

$$k\xi \equiv \mu \bmod 2^{v-2}.$$

The corresponding x fulfills the desired congruence. □

Lemma 6.11. *If the congruence*

$$x_1^n + \ldots + x_k^n \equiv m \ \mathit{mod}\ p^\gamma$$

has a solution with $x_1, \ldots, x_k$ *not all divisible by* p*, then* $\chi(p) > 0$.

Proof. Suppose

$$a_1^n + \ldots + a_k^n \equiv m \bmod p^\gamma \text{ and } p \nmid a_1$$

Let $v > \gamma$ and observe that we can choose $x_2, \ldots, x_k$ in $p^{(v-\gamma)(k-1)}$ ways such that

$$x_j \equiv a_j \bmod p^\gamma,\ 0 < x_j \leq p^v \text{ for all } 2 \leq j \leq k.$$

The upshot is, that by the previous result we can choose $0 < x_1 \leq p^v$ such that

$$x_1^n \equiv m - x_2^n - \ldots - x_k^n \bmod p^v.$$

Since we have so many choices for $x_2, \ldots, x_k$ we get

$$M(p^v) \geq p^{(v-\gamma)(k-1)} = C_p p^{v(k-1)}.$$

But $C_p = p^{-\gamma(k-1)}$ is positive and independent of v. We use this to obtain

$$\chi(p) = \lim_{v\to\infty} M(p^v)p^{-v(k-1)} \geq C_p > 0.$$

□

Lemma 6.12. *If $k \geq 2n$ for n odd or $k \geq 4n$ for n even, then $\chi(p) > 0$ for all primes p and all m.*

Proof. We need to solve the congruence

$$x_1^n + \ldots + x_k^n \equiv m \bmod p^\gamma. \tag{11}$$

If $p \mid m$ we replace this congruence by

$$x_1^n + \ldots + x_{k-1}^n + 1^n \equiv m \bmod p^\gamma.$$

In this case we simply replace m by $m-1$ and k by $k-1$. We have reduced the problem to solve the congruence (11) for $k \geq 2n-1$ (resp. $k \geq 4n-1$ when n is even) and $(m,p) = 1$.

We start with the generic cases $p \neq 2$. There are $\phi(p^\gamma)$ congruences classes $0 < m < p^\gamma$ with $(p,m) = 1$. Let $k(m)$ denote the least k for which the congruence (11) is soluble. We observe that if $m \equiv z^n m' \bmod p^\gamma$, then $k(m) = k(m')$. We group the numbers m together by the value of $k(m)$:

$$M_i = \{0 < m < p^\gamma \colon (m,p) = 1,\, k(m) = i\}.$$

Note that if $M_i \neq \emptyset$, then we have $\sharp M_i \geq \sharp\{z^n \bmod p^\gamma \colon (z,p) = 1\}$. We can explicate this lower bound as follows. Put $z \equiv g^\zeta \bmod p^\gamma$ and $a \equiv g^\alpha \bmod p^\gamma$. Here g is again a primitive root. Observe that $z^n \equiv a \bmod p^\gamma$ is soluble if and only if α is divisible by $p^\tau(n, p-1)$. Taking all the distinct values for $\alpha \bmod p^\tau(p-1)$ into account we find that

$$\sharp\{z^n \bmod p^\gamma \colon (z,p) = 1\} = \frac{p^\tau(p-1)}{p^\tau(n,(p-1))} = \frac{p-1}{(n,p-1)} =: r.$$

We enumerate

$$M_1 = \{m_1^{(1)} < \ldots < m_{r_1}^{(1)}\} \text{ and } M_2 = \{m_1^{(2)} < \ldots < m_{r_2}^{(2)}\}.$$

Note that M_1 contains exactly all nth powers and is non-empty. Further $r_1, r_2 \geq r$.

We claim that M_j or M_{j+1} is non-empty. To see this we take the smallest m' with $(m',p) = 1$ that is not contained in $M_1 \cup \ldots \cup M_{j-1}$. Then $m'-1$ or $m'-2$ is not divisible by p. By minimality there is $1 \leq i \leq j-1$ such that $m'-1$ or $m'-2$ is in M_i. By writing

$$(m'-1) + 1^n \text{ or } (m'-2) + 1^n + 1^n$$

we see that $k(m') \leq j+1$. Again by assumption we see that M_j or M_{j+1} must be non-empty as claimed.

Suppose M_l is the last non-empty set. (of course only finitely many of these sets can be non-empty!) Then at least $\frac{1}{2}(l-1)$ of the first $l-1$ sets are non-empty. Including M_l this makes $\frac{1}{2}(l+1)$ non-emoty sets, each containing at least r elements. We get

$$\frac{1}{2}(l+1)r \leq \phi(p^\gamma) = p^\tau(p-1).$$

If we can show $l \leq 2n-1$, then we are obviously done. But this an easy task:

$$l+1 \leq \frac{2p^\tau(p-1)}{r} = 2p^\tau(n_0, p-1) \leq 2n.$$

We turn towards the exceptional case $p = 2$. Again we observe that for odd n there is really nothing to do. Thus we suppose $\tau \geq 1$. Without loss of generality we can assume that $0 < m < 2^\gamma$. Assuming $k \geq 2^\gamma - 1$ we can write down the explicit solution $x_i = 1$ for $1 \leq i \leq m$ and $x_i = 0$ for $i > m$. We are done since

$$2^\gamma - 1 = 2^{\tau+1} - 1 \leq 4n - 1.$$

□

We follow the notation of Hardy-Littlewood and set $\Gamma(n)$ (not to be confused with the Γ-function) to be the least number of variables k such that the congruence is soluble for all p and all m. We have seen so far, that

$$\Gamma(n) \leq \begin{cases} 2n & \text{if } n \text{ is odd,} \\ 4n & \text{if } n \text{ is even.} \end{cases} \tag{12}$$

We are now ready to establish the following important theorem.

Theorem 6.13. *If $k \geq 2^n + 1$, then*

$$\mathfrak{S}(m) \geq C_1(n,k) > 0 \textit{ for all } m.$$

Proof. We have seen by Corollary 6.8 that it suffices to show that $\chi(p) > 0$ for all $p \leq p_0$. But by (12) we have $2^n + 1 \geq \Gamma(n)$ for $n > 2$. However, according to Lemma 6.12 this suffices to conclude $\chi(p) > 0$ for all p. □

Note that our problem is, that we only established absolute convergence for $\mathfrak{S}(m)$ when $k \geq 2^n + 1$. We will spend the remainder of this section improving this. We will need a series of lemmata.

Lemma 6.14. *If $p \nmid a$ and $\delta = (n, p-1)$, then*

$$|S_{a,p}| \leq (\delta - 1)p^{\frac{1}{2}}.$$

Proof. Since $x^n \equiv m \bmod p$ has the same number of solutions as $x^\delta \equiv m \bmod p$. We get

$$S_{a,p} = \sum_{x \bmod p} e\left(\frac{a}{p}x^\delta\right).$$

Given a primitive character χ modulo p of order δ we observe that

$$\sharp\{x \bmod p \colon x^\delta \equiv t \bmod p\} = 1 + \chi(t) + \ldots + \chi^{\delta-1}(t).$$

This is a well known generalisation of the well known quadratic case. With this at hand we write

$$S_{a,p} = \sum_{l=0}^{\delta-1} \underbrace{\sum_{x \mod p} \chi^l(x) e(\frac{a}{p}x)}_{=G(a,\chi^l)}.$$

Here we interpret $\chi^0 = 1$ as the constant 1 function on $\mathbb{Z}$. (This is a slight abuse of notation!) Since $(a,p) = 1$, we have $G(a,1) = \sum_{x \bmod p} e(\frac{a}{p}x) = 0$ by character orthogonality. We get

$$S_{a,p} \leq (\delta - 1) \max_{1 \leq l \leq \delta-1} |G(a,\chi^l)|.$$

It is a classical result due to Gauß that $|G(a,\chi^l)| = \sqrt{p}$ as long as χ^l is non-principal. We repeat the standard argument for completeness.

We write ψ for any non-principal character modulo p. (The argument works for primitive characters of any modulus!) Consider

$$\begin{aligned} |G(a,\psi)|^2 &= \sum_{x,y} \psi(x)\overline{\psi(y)} e\left(\frac{a}{p}(x-y)\right) \\ &= \sum_x \sum_{0 \neq y} \psi(x) e\left(\frac{ay}{p}(x-1)\right). \end{aligned}$$

Note that

$$\sum_{0 \neq y} e\left(\frac{ay}{p}(x-1)\right) = \begin{cases} p-1 & \text{if } x = 1, \\ -1 & \text{else.} \end{cases}$$

This is seen by artificially inserting $y = 0$ and applying character orthogonality. We have

$$|G(a,\psi)|^2 = p\psi(1) - \sum_{x \neq 1} \psi(v) = p.$$

We are done after taking the square root. □

Lemma 6.15. *Suppose $p \nmid a$ and $p \nmid n$. Then for $1 < v \leq n$ we have*

$$S_{a,p^v} = p^{v-1},$$

and for $v > n$

$$S_{a,p^v} = p^{n-1} S_{a,p^{v-n}}.$$

Proof. In the definition of S_{a,p^v} we split the sum as follows:

$$S_{a,p^v} = \sum_{x=0}^{p^v-1} e(\frac{a}{p^v}x^n) = \sum_{0 \leq z < p^{v-1}} \sum_{0 \leq y < p} e(\frac{a}{p^v}(yp^{v-1} + z)^n).$$

Note that

$$(yp^{v-1}+z)^n \equiv z^n + np^{v-1}z^{n-1}y \bmod p^v.$$

We obtain

$$S_{a,p^v} = \sum_{0\leq z<p^{v-1}} e(\frac{az^n}{p^v}) \sum_{0\leq y<p} e(\frac{anz^{n-1}y}{p}).$$

Note that by assumption $p \nmid an$. Thus the inner sum vanishes unless $z \equiv 0 \bmod p$. Thus we write $z = pw$ and get

$$S_{a,p^v} = p\sum_{w=0}^{p^{v-2}-1} e(\frac{aw^n}{p^{v-n}}).$$

First, if $v \leq n$, then we are just summing ones and get $S_{a,p^v} = p^{v-1}$. Otherwise we observe that we are summing a function of period p^{v-k}. The claimed result follows at once. □

Lemma 6.16. *If $p \mid n$ we still have*

$$S_{a,p^v} = p^{n-1}S_{a,p^{v-n}}.$$

for $v > n$.

Proof. As earlier we write $n = p^\tau n_0$. We have

$$v \geq p^\tau n_0 + 1 \geq 2^\tau + 1 \geq \tau + 2.$$

We will follow the idea of the previous proof with slight modifications. In particular we write

$$x = p^{v-\tau-1}y + z \text{ for } 0 \leq y < p^{\tau+1},\ 0 \leq z < p^{v-\tau-1}.$$

Suppose we have

$$x^n \equiv z^n + np^{v-\tau-1}z^{n-1}y \bmod p^v. \tag{13}$$

Then we get

$$S_{a,p^v} = \sum_{z=0}^{p^{v-\tau-1}-1} e(\frac{az^n}{p^v}) \sum_{y=0}^{p^{\tau+1}-1} e(\frac{an_0z^{n-1}y}{p}).$$

Note that once again the inner sum vanishes unless $z \equiv 0$ mof p. Thus

$$S_{a,p^v} = p^{\tau+1}\sum_{z=0}^{p^{v-\tau-2}-1} e(\frac{az^n}{p^{v-n}}) = p^{n-1}S_{a,p^{v-n}}.$$

The latter is what we claimed, so that it suffices to verify (13). The critical case to consider is

$$(z+p^{v-\tau-1}y)^{p^\tau} \equiv z^{p^\tau} + p^{v-1}z^{p^\tau-1}y \bmod p^v.$$

Indeed raising this congruence to the power of n_0 is no problem.

Putting $\lambda \geq v-\tau-1$ we need to establish

$$(z+p^\lambda y)^{p^\tau} \equiv z^{p^\tau} + p^{\lambda+\tau}z^{p^\tau-1}y \bmod p^{\lambda+\tau+1}.$$

We will continue by induction on τ.

The starting point is $\tau = 1$. In this case we have $\lambda \geq 1 + \delta_{p=2}$. The critical term is the last term in the binomial expansion of $(z + p^\lambda y)^p$, which is $p^{\lambda p} y^p$. But we have $\lambda p \geq \lambda + 2$ under the hypothesis in place.

We continue the induction step. For $y_1 \equiv y \bmod p$ we have

$$\begin{aligned}(z + p^\lambda y)^{p^\tau} &= (z^{p^{\tau-1}} + p^{\lambda+\tau-1} z^{p^{\tau-1}-1} y_1)^p \\ &\equiv z^{p^\tau} + p^{\lambda+\tau} z^{p^\tau - 1} y_1 \bmod p^{\lambda+\tau+1} \\ &\equiv z^{p^\tau} + p^{\lambda+\tau} z^{p^\tau - 1} y \bmod p^{\lambda+\tau+1}.\end{aligned}$$

This holds true by assumptions on λ and we are done. □

Lemma 6.17. *For $(a, q) = 1$ we have*

$$S_{a,q} \ll_n q^{1-\frac{1}{n}}.$$

Proof. By multiplicativity established earlier it is enough to consider $q = p^v$. Write $T(a, p^v) = p^{-v+\frac{v}{n}} |S_{a,p^v}|$. We need to show that $T(a, p^v)$ is bounded independently of p^v. We have seen above that for $v > n$ we have

$$T(a, p^v) = T(a, p^{v-n}).$$

Applying this repeatedly allows us to assume that $v \leq n$. Also note that

$$T(a, p) \ll_n np^{\frac{1}{2}} p^{-(1-\frac{1}{n})} \leq np^{-\frac{1}{6}}.$$

Further, if $p \nmid n$ we get

$$T(a, p^v) \leq p^{v-1} p^{-v(1-\frac{1}{n})} \leq 1 \text{ for } 1 < v \leq n.$$

Thus $T(a, p^v) \leq 1$ except when $v = 1$ and $p \leq n^6$. But these are finitely many cases (their number only depending on n) and we can treat them trivially. □

Theorem 6.18. *The singular series $\mathfrak{S}(m)$ and the product $\prod_p \chi(p)$ are absolutely convergent if $k \geq 2n + 1$ and we have*

$$\mathfrak{S}(m) \geq C_1(n, k) > 0 \tag{14}$$

if $k \geq 2n + 1$ if n is odd or $k \geq 4n$ otherwise.

Proof. This is proved precisely as Theorem 6.13 using our improved bound on $S_{a,q}$. □

6.3. **Conclusion.** Combining the results from this section we obtain the following theorem.

Theorem 6.19. *Every sufficiently large number can be written as the sum of k positive integral nth powers for $k \geq 2^n + 1$.*

Proof. By Theorem 6.5 we have an asymptotic expansion for the representation numbers $r_k(m)$. Since by Theorem 6.13 the singular series appearing in the main term is positive. Thus, for sufficiently large m we have $r_k(m) \neq 0$, which proves the theorem. □

Thus we have established a new version of Hilbert's theorem using analytic tools. The upshot is, that for large m this approach even provides an asymptotic formula for the representation numbers and thus gives a lower bound for the number of variables necessary. (Hilbert's and probably also Linnik's approach can be made explicit to give a lower bound for the number of variables, but this lower bound will be much worse than what we obtained so far.)

The situation at hand is an adventure playground for analytic number theorists. The game being to improve the number of variables needed. To quantify progress we define the following numbers:

- We write $g(n)$ for the least positive integer k such that $r_k(m) > 0$ for all m;
- We write $G(n)$ for the least positive integer k such that $r_k(m) > 0$ for all sufficiently large m;
- We write $G_1(n)$ for the least positive integer k such that $r_k(m) > 0$ for $m \in \mathbb{N} \setminus E$, where E is an exceptional set of natural density 0. (I.e. $\sharp\{n \in E \colon n \leq N\} = o(N)$.)

Of course $G_1(n) \leq G(n) \leq g(n)$. In general $G(n) < g(n)$ as we will see later. A natural lower bounds for $G_1(n)$ is given by $\Gamma(n)$. This gives a general congruence obstruction, which makes the singular series vanish for a positive proportion of all m. Hardy-Littlewood classified all types of n for which $\Gamma(n) > n$.

7. Waring's problem for 3rd powers

As an example we consider the three number $g(3)$, $G(3)$ and $G_1(3)$. We are going to sketch the arguments to prove the following 3 results

$$g(3) = 9,\ G(3) \leq 7 \text{ and } G_1(3) = 4.$$

Theorem 7.1 (Wieferich, Kemper)**.** *We have $g(3) = 9$.*

All proofs I am aware of require some numerical computations as input. We will leave this to the reader.

Note that this result pre-dates the influential papers of Hardy-Littlewood. We follow the argument of Wieferich (1909). However it should be noted that Wieferich had to argue slightly more carefully since there were no powerful numerical devices at the time. They relied on tables of representations of numbers as cubes which were very limited. Indeed, Wieferich's proof had a small gap, which was later fixed by Kemper (1912).

Proof. First note that the number 23 (as well as 239) can not be written as a sum of less than 9 cubes. Thus $g(3) \geq 9$.

Suppose $z > 7,4 \cdot 5^{15}$. We assume that the claim can be checked for all smaller z using a computer.[3] Then there is $v \geq 5$ such that

$$7,4 \cdot 5^{3v} < z \leq 7,4 \cdot 5^{3v+3}.$$

We write $z_\alpha = z - \alpha^3$, for some positive integer α. Note that if $z_\alpha \geq 0$, then $\alpha \leq \sqrt[3]{7,4} \cdot 5^{v+1}$. Note that

$$z_{\alpha-1} - z_\alpha = 3\alpha^2 - 3\alpha + 1 \leq 3\alpha^2 \leq 3\sqrt[3]{7,4^2} \cdot 5^{2v+2}.$$

We define τ so that

$$3\sqrt[3]{7,4^2} \cdot 5^{2v+2} = \tau 5^{3v} \text{ i.e. } \tau = \frac{3\sqrt[3]{7,4^2}}{5^{v-2}} \asymp 284,8 \cdot 5^{-v}.$$

Since $v \geq 5 > 3$ we have $\tau < 2,3$. Thus we find α such that

$$7,4 \cdot 5^{3v} \leq z_\alpha < (7,4 + 2,3) \cdot 5^{3v} \text{ and } 7,4 \cdot 5^{3v} \leq z_\alpha < z_{\alpha-1} \leq 12 \cdot 5^{3v}.$$

We take $\alpha' \in \{\alpha, \alpha - 1\}$ such that $5 \nmid z_{\alpha'}$. (This can be achieved since for $5 \nmid \alpha$ we have $5 \nmid (3\alpha^2 - 3\alpha + 1)$, which can be seen by computing modulo 5.) Now we let β run through a reduced system of representatives modulo 5^v. Then there is β such that $\beta^3 \equiv z_{\alpha'} \bmod 5^v$. We rewrite this as

$$Z_\beta = z_{\alpha'} - \beta^3 = 5^v \cdot M.$$

We obviously $6,4 \cdot 5^{3v} < z_\alpha - \beta^3 < 12 \cdot 5^{3v}$. This yields $6,4 \cdot 5^{2v} < M < 12 \cdot 5^{2v}$. We put

$$M = 6 \cdot 5^{2v} + M_1 \text{ in particular } 0,4 \cdot 5^{2v} < M_1 < 6 \cdot 5^{2v}.$$

Now take $\epsilon \in \{0, 1, 2\}$ such that $v + \epsilon \equiv 0 \bmod 3$. We now claim that we can choose γ such that

$$M_1 = 5^\epsilon \cdot \gamma^3 + 6 \cdot M_3 \text{ for } M_3 \neq 4^\eta(8n + 7).$$

Here we cheat again. Indeed we observe that it is enough to consider $M_1 \equiv \theta \bmod 96$. Thus for the finitely many cases $\epsilon = 0, 1, 2$ and $\theta = 0, \ldots, 95$ one has to choose γ such that M_3 is not of the forbidden shape. This can again be done numerically.[4] One actually can do so for $\gamma \leq 22$. Thus we note that for $v \geq 5$ (this is not true for $v \leq 4$) that

$$0,4 \cdot 5^{2v} \geq 5^2 \cdot 22^3 \geq 5^\epsilon \gamma^3,$$

so that M_3 is positive. We have

$$0 < M_3 < 5^{2v}.$$

[3] I didn't check this myself and it was certainly not possible in 1909. Thus there were further reductions necessary for small z.

[4] Wieferich did all the cases by hand and includes corresponding tables in his paper.

We now observe the identity

$$\sum_{i=1}^{3}\left[(A+x_i)^3+(A-x_i)^3\right]=A\left[6A^2+6(x_1^2+x_2^3+x_3^2)\right].$$

We now Legendre's three-square theorem, which claims that M_3, since it is not of the form $4^{\eta}(8n+7)$, can be written as the sum of 3 squares:

$$M_3=x_1^2+x_2^2+x_3^2.$$

Thus, applying our identity with $A=5^v$ we find that

$$6\cdot 5^{3v}+6\cdot 5^v M_3=y_1^3+\ldots+y_6^3.$$

But now we insert everything and obtain

$$\begin{aligned} z &= \alpha^3+\beta^3+5^v\cdot M \\ &= \alpha^3+\beta^3+5^{v+\epsilon}\gamma^3+6\cdot 5^{3v}+6\cdot 5^v M_3 \\ &= \alpha^3+\beta^3+(\gamma 5^{\frac{v+\epsilon}{3}})^3+y_1^3+\ldots+y_6^3. \end{aligned}$$

We count 9 cubes and (up to our omissions) the proof is complete. □

Theorem 7.2 (Linnik). *We have $G(3)\leq 7$.*

Even before the arrival of the circle method Landau had shown $G(3)\leq 8$. (Showing that there are only finitely many numbers which need 9 cubes.) Linnik improved this to $G(3)\leq 7$ but one might even conjecture $G(3)=4$.

Linnik's original argument was quite complicated using some deep result on representation numbers of quadratic forms. We instead follow the easier proof of Watson. This proof is more closely related to Landau's earlier argument and uses some prime number theory, which we will use as black box.

Proof. We will use the following claim without proof: *If X is sufficiently large and $k<\log(X)^{100}$ with $(k,l)=1$, then there is a prime $p\equiv l$ mod k with $X<p<1,01\cdot X$.*[5]

Before we start the main argument we prove an elementary claim.

Claim: *Let N be a positive integer and assume that there exist distinct primes p,q,r such that*

- $p\equiv q\equiv r\equiv -1$ *mod* 6;
- $r<q<1,01\cdot r$;
- $\frac{3}{4}q^{18}p^3<N<q^{18}p^3$;
- $N\equiv 3p$ *mod* $6p$;
- $4N\equiv r^{18}p^3$ *mod* q^6;
- $2N\equiv q^{18}p^3$ *mod* r^6;

[5] This can be deduced form a suitable version of the Siegel-Walfisz theorem for primes in arithmetic progressions.

then N is representable as the sum of six positive integral cubes.

To see this we first observe that the conditions ensure

$$(4q^{18}+2r^{18})p^3 < 8N < (4q^{18}+8r^{18})p^3.$$

Further we have the congruence

$$8N \equiv (4q^{18}+2r^{18})p^3 + 18q^6r^6p \bmod p^6r^6p.$$

One can check that both sides of this congruence are congruent 24 mod 48. Thus we can strengthen the modulus to $48 \cdot p^6r^6p$. We therefore write

$$8N = (4q^{18}+2r^{18})p^3 + 18q^6r^6p + 48q^6r^6p \cdot u \text{ for } 0 < 8u+3 < q^{-6}r^{12}p^2.$$

Now write $8u+3 = x_1^2+x_2^2+x_3^2$ by Legendre's three square theorem. We have the obvious bounds $x_i \leq q^{-3}r^6p$. We conclude by writing

$$\begin{aligned} 8N &= (4q^{18}+2r^{18})p^3 + 6q^6r^6p(x_1^2+x_2^2+x_3^2) \\ &= \sum_{i=1}^{2} \left[(q^6p+r^3x_i)^3+(q^6p-r^3x_i)^3\right] + [(r^6p+q^3x_3)^3+(r^6p-q^3x_3)^3]. \end{aligned}$$

We are done since $8 = 2^3$ and all terms on the right hand side are even and positive. (Note that this is a similar idea we have seen earlier!)

We turn towards the proof of the theorem. Let n be sufficiently large. We claim that there are primes r, q less $\log(n)^2$ such that $q \equiv r \equiv -1 \bmod 6$, $r < q < 1,01 \cdot r$ and $(qr, n) = 1$. This can also be shown using standard results from prime number theory. (In particular a suitable Siegel-Walfisz theorem and a standard bound for the number of prime factors of n.)

Now we write $X = n^{\frac{1}{3}}q^{-6}$. Every number co-prime with $q \cdot r$ is congruent to a cube modulo q^6 and r^6, we can find a number l such that

$$4n \equiv r^{18}l^3 \bmod q^6 \text{ and } 2n \equiv q^{18}l^3 \bmod r^6.$$

Now we apply our first claim (note that if n is large enough so is X) to find a prime p with

$$p \equiv -aq^6r^6 + lb6r^6 + lc6q^6 \bmod 6q^6r^6 \text{ and } X < p < 1,01 \cdot X.$$

for suitable a, b, c. This constructions ensures that

$$4n \equiv r^{18}p^3 \bmod q^6,\ 2n \equiv q^{18}p^3 \bmod r^6 \text{ and } p \equiv -1 \bmod 6.$$

Finally we choose an integer t so that

$$t^3 \equiv n - 3p \bmod 6p,\ t \equiv 0 \bmod q^2r^2 \text{ and } 0 < t \leq 6pq^2r^2.$$

This is certainly possible, since every number is congruent to a cube modulo $6p$. But now we are done, since $N = n - t^3$ satisfies all the assumptions from the claim above and can thus be written as a sum of 6 (integral positive) cubes.

□

Theorem 7.3 (Davenport). *We have $G_1(3) = 4$.*

Here we follow Davenport's original treatment (see [Da1]), which uses a variation ot the circle method. The proof will take us a while and we first establish several lemmata.

However, before we start lets observe that every cube is congruent 0, 1 or -1 modulo 9. Thus the sum of 3 cubes can not be congruent 4 or 5 modulo 9. This already shows $G_1(3) > 3$. To derive upper bounds is the hard part.

We introduce the following notation:

$$T(\alpha) = \sum_{P \le x \le 2P} e(\alpha x^3),$$
$$T_1(\alpha) = \sum_{P^{\frac{4}{5}} \le x \le 2P^{\frac{4}{5}}} e(\alpha x^3) \text{ and}$$
$$V(\alpha) = T(\alpha)^2 T_1(\alpha)^2 = \sum_n \rho(n) e(n\alpha).$$

Note that here

$$\rho(n) = \sharp\{w^3 + x^3 + y^3 + z^3 = n \colon P \le w, x \le 2P \text{ and } P^{\frac{4}{5}} \le y, z \le 2P^{\frac{4}{5}}\}.$$

We need some further notions, some of which we have seen earlier:

$$I(\beta) = \frac{1}{3} \sum_{P^3 \le n \le (2P)^3} n^{-\frac{2}{3}} e(\beta n),\ I_1(\beta) = \frac{1}{3} \sum_{(P^{\frac{4}{5}})^3 \le n \le (2P^{\frac{4}{5}})^3} n^{-\frac{2}{3}} e(\beta n),$$

$$S_{a,q} = \sum_{x=1}^{q} e(\frac{a}{q} x^3),\ A(n,q) = q^{-4} \sum_{a=1,(a,q)=1}^{q} S_{a,q}^4 e(-\frac{an}{q}),$$

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

$$T^*(\alpha,a,q) = q^{-1} S_{a,q} I(\alpha - \frac{a}{q}),\ T_1^*(\alpha,a,q) = q^{-1} S_{a,q} I_1(\alpha - \frac{a}{q}) \text{ and}$$

$$V^*(\alpha,a,q) = (T^*(\alpha,a,q) T_1^*(\alpha,a,q))^2.$$

Recall that by Lemma 6.17 we had

$$S_{a,q} \ll q^{\frac{2}{3}}.$$

Further we have seen in an exercise that, for $n \neq 0$ we have

$$S_{a,n,q} = \sum_{x=1}^{q} e(\frac{a}{q} x^3 + \frac{n}{q} x) \ll q^{\frac{2}{3}+\epsilon}(n,q).$$

Lemma 7.4. *If $|\beta| \le \frac{1}{2}$, then*

$$I(\beta) \ll \min(P, P^{-2}|\beta|^{-1}) \text{ and } I_1(\beta) \ll \min(P^{\frac{4}{5}}, P^{-\frac{8}{5}}|\beta|^{-1}).$$

Proof. The first choice in the minimum is simply the trivial estimate. The second one follows from

$$\sum_{n_1<x\leq n_2} e(\beta n) \ll |\beta|^{-1}$$

and partial summation. □

A direct consequence of this lemma are the estimates

$$T^*(\alpha,a,q) \ll q^{-\frac{1}{3}}\min(P,P^{-2}|\beta|^{-1}) \text{ and } T_1^*(\alpha,a,q) \ll q^{-\frac{1}{3}}\min(P^{\frac{4}{5}},P^{-\frac{8}{5}}|\beta|^{-1}), \tag{15}$$

as long as $\alpha=\frac{a}{q}+\beta$, where $|\beta|\leq\frac{1}{2}$.

Lemma 7.5. *For $H\geq 1$, $q\leq H^{1-\delta}$, $\beta\ll q^{-1}H^{-2-\delta}$ and $n\neq 0$, we have*

$$\int_0^H e(\beta x^3-\frac{nx}{q})dx = -\frac{q}{2\pi in}\left(e\left(\beta H^3-\frac{nH}{q}\right)-1\right)+O(qn^{-2}H^{-\delta}).$$

Proof. We integrate by parts l-times. This yields

$$\int_0^H e(\beta x^3-\frac{nx}{q})dx = -\frac{q}{2\pi in}\left(e(\beta H^3-\frac{nH}{q})-1\right)-\sum_{h=1}^{l-1}\left(\frac{q}{2\pi in}\right)^{h+1}\left[e(-\frac{nx}{q})D^h(e(\beta x^3))\right]_0^H$$
$$+\left(\frac{q}{2\pi in}\right)^l\int_0^H e(-\frac{nx}{q})D^l(e(\beta x^3))dx.$$

We write

$$D^h(e(\beta x^3)) = \sum_{\frac{1}{3}h\leq r\leq h} c(r,h)\beta^r x^{3r-h}e(\beta x^3).$$

Using the assumptions we can estimate $\beta^r x^{3r-h}\ll q^{-r}H^{-r(2+\delta)}H^{3r-h}\ll q^{-h}H^{-h\delta}$. Therefore we can estimate

$$D^h(e(\beta x^3))\ll_h q^{-h}H^{-h\delta}.$$

Using these estimates the error term can be bounded by

$$\sum_{h=1}^{l-1}\left(\frac{q}{|n|}\right)^{h+1}c_1(h)q^{-h}H^{-h\delta}+\left(\frac{q}{|n|}\right)^l c_1(h)q^{-l}H^{1-l\delta}.$$

By choosing l large enough the result follows. □

Lemma 7.6. *If $\alpha=\frac{a}{q}+\beta$ with $q\leq P^{1-\delta}$ and $|\beta|\leq q^{-1}P^{-2-\delta}$, the*

$$T(\alpha)=T^*(\alpha,a,q)+O(q^{\frac{2}{3}+\epsilon}).$$

Further we have the estimate

$$T(\alpha)\ll q^{-\frac{1}{3}}\min(P,P^{-2}|\beta|^{-1}).$$

Proof. Splitting the summation in the definition of $T(\alpha)$ in congruence classes modulo q yields

$$T(\alpha)=\sum_{h=1}^{q}\sum_{\frac{P-h}{q}\le m\le\frac{2P-h}{q}} e\left(\left(\frac{a}{q}+\beta\right)(mq+h)^3\right)=\sum_{h=1}^{q}e(\frac{a}{q}h^3)\sum_{\frac{P-h}{q}\le m\le\frac{2P-h}{q}} e\left(\beta(mq+h)^3\right).$$

We will apply Poisson summation to the inner sum. For convenience we modify the m-sum by counting the boundary terms with weight $\frac{1}{2}$. For this sum we get

$$\begin{aligned}\sum_{\frac{P-h}{q}\le m\le\frac{2P-h}{q}} e\left(\beta(mq+h)^3\right)&=\int_{\frac{P-h}{q}}^{\frac{2P-h}{q}} e(\beta(xq+h)^3)dx+\sum_{n\in\mathbb{Z}\setminus\{0\}}\int_{\frac{P-h}{q}}^{\frac{2P-h}{q}} e(\beta(xq+h)^3-nx)dx\\&=q^{-1}\int_P^{2P}e(\beta x^3)dx+q^{-1}\sum_{n\in\mathbb{Z}\setminus\{0\}}e(\frac{nh}{q})\int_P^{2P}e(\beta x^3-nx)dx.\end{aligned}$$

We will use the approximation $y^{-\frac{2}{3}}e(\beta y)=n^{-\frac{2}{3}}e(\beta n)+O(P^{-3})$ for $n\le y\le n+1$ and $P^3\le n\le(2P)^3$. We obtain

$$\int_P^{2P}e(\beta x^3)dx=\frac{1}{3}\int_{P^3}^{(2P)^3}y^{-\frac{2}{3}}e(\beta y)dy=I(\beta)+O(1).$$

So far we have obtained

$$T(\alpha)=T^*(\alpha,a,q)+\sum_{0\ne n\in\mathbb{Z}}q^{-1}S_{a,n,q}\int_P^{2P}e(\beta x^3-\frac{nx}{q})dx+O(1).$$

Applying the previous result to the integral in the n-sum we get

$$\begin{aligned}T(\alpha)=T^*(\alpha,a,q)-\frac{1}{2\pi i}\sum_{0\ne n\in\mathbb{Z}}n^{-1}S_{a,n,q}\left(e(8\beta P^3-\frac{2nP}{q})-e(\beta p^3-\frac{nP}{q})\right)\\+O(\sum_{n\ne0}|S_{a,q,n}|n^{-2}P^{-\delta}).\end{aligned}$$

We can control the error term trivially:

$$\sum_{n\ne0}|S_{a,q,n}|n^{-2}P^{-\delta}\ll q^{\frac{2}{3}+\epsilon}P^{-\delta}\sum_{n=1}^{\infty}n^{-2}(n,q)\ll q^{\frac{2}{3}+\epsilon}.$$

The n-sum is treated differently. The first part of the sum (small n) is treated trivially:

$$\ll\sum_{n\le q^2}n^{-1}|S_{a,q,n}|\ll\sum_{n\le q^2}n^{-1}q^{\frac{2}{3}+\epsilon}(n,q)\ll q^{\frac{2}{3}+\epsilon}. \tag{16}$$

Finally we treat the remaining part:

$$\begin{aligned}
&e(8\beta P^3)\sum_{|n|>q^2} n^{-1}S_{a,q,n}e(-\frac{2n}{q}P) - e(\beta P^3)\sum_{|n|>q^2} n^{-1}S_{a,q,n}e(-\frac{n}{q}P)\\
&= e(8\beta P^3)\sum_{h=1}^{q} e(\frac{ah^3}{q})\sum_{n=q^2+1}^{\infty}\frac{2i}{n}\sin(2\pi\frac{n(h-2P)}{q})\\
&\qquad - e(\beta P^3)\sum_{h=1}^{q} e(\frac{ah^3}{q})\sum_{n=q^2+1}^{\infty}\frac{2i}{n}\sin(2\pi\frac{n(h-P)}{q})\\
&\ll \sum_{h=1}^{q}\Big(\min(1,q^{-2}\|\frac{h-2P}{q}\|^{-1}) + \min(1,q^{-2}\|\frac{h-P}{q}\|^{-1})\Big)\\
&\ll q^{\frac{2}{3}+\epsilon}.
\end{aligned}$$

The claimed upper bound follows directly after inserting a previous estimate for $T^*(\alpha,a,q)$ and observing that $q^{1+\epsilon} \le \min(P, P^{-2}|\beta|^{-1})$. □

As a direct consequence (replacing P by $P^{\frac{4}{5}}$) we obtain the following result concerning $T_1(\alpha)$.

Lemma 7.7. *If $\alpha = \frac{a}{q} + \beta$ with $q \le P^{\frac{4}{5}(1-\delta)}$ and $|\beta| \le q^{-1}P^{-\frac{4}{5}(2+\delta)}$, then*

$$T_1(\alpha) = T_1^*(\alpha,a,q) + O(q^{\frac{2}{3}+\epsilon}).$$

Further we have the estimate

$$T_1(\alpha) \ll q^{-\frac{1}{3}}\min(P^{\frac{4}{5}}, P^{-\frac{8}{5}}|\beta|^{-1}).$$

Recalling our original treatment of Waring's problem it should appear natural that we will use the Farey sequence together with the rational approximations to $T(\alpha)$. In our case at hand the distinction between major and minor arcs will be different.

Consider the full Farey sequence $\mathcal{F} = \{\frac{a}{q} \colon 0 < q \le P^{2+\delta}, (a,q) = 1\} \cap [0,1]$. The Farey arc surrounding $\frac{a}{q} \in \mathcal{F}$ is given by

$$A_{a,q} = (\frac{a+a_-}{q+q_-}, \frac{a+a_+}{q+q_+}].$$

Here $\frac{a_-}{q_-} < \frac{a}{q} < \frac{a_+}{q_+}$ are 3 neighbouring Farey fractions in our sequence. These arcs cover the interval $[0,1]$ (modulo 1). For $\alpha \in A_{a,q}$ we have

$$\alpha = \frac{a}{q} + \beta \text{ for } -\theta_1 q^{-1}P^{-2-\delta} \le \beta \le \theta_2 q^{-1}P^{-2-\delta},$$

where $\theta_1, \theta_2 \in [\frac{1}{2}, 1]$. We define

$$\mathfrak{M} = \bigcup_{\underbrace{q \le P^{\frac{4}{5}(1-\delta)}, (a,q)=1}} A_{a,q} \tag{17}$$

to be the major arcs and rename $\mathfrak{M}_{a,q} = A_{a,q}$ in this case. The minor arcs are

$$\mathfrak{m} = \bigcup_{\substack{P^{\frac{4}{5}(1-\delta)} < q \le P^{2+\delta}, \\ (a,q)=1}} A_{a,q} \tag{18}$$

A single minor arc is relabeled to be $\mathfrak{m}_{a,q} = A_{a,q}$.

We need the following lemma for the major arcs.

Lemma 7.8. *Under our current hypothesis we have*

$$\sum_{a,q} \int_{\mathfrak{M}_{a,q}} |V(\alpha) - V^*(\alpha, a, q)|^2 d\alpha \ll P^{3+\frac{2}{3}}.$$

Proof. Take $\alpha \in \mathfrak{M}_{a,q}$. By the lemmata above we have

$$T(\alpha) - T^*(\alpha, a, q) \ll q^{\frac{2}{3}+\epsilon} \text{ and } T(\alpha), T^*(\alpha, a, q) \ll q^{-\frac{1}{3}} \min(P, P^{-2}|\beta|^{-1}).$$

Similarly we have

$$T_1(\alpha) - T_1^*(\alpha, a, q) \ll q^{\frac{2}{3}+\epsilon} \text{ and } T_1(\alpha), T_1^*(\alpha, a, q) \ll q^{-\frac{1}{3}} P^{\frac{4}{5}}.$$

Using this we conclude that

$$\begin{aligned} &T(\alpha)^2 T_1(\alpha)^2 - T^*(\alpha, a, q)^2 T_1^*(\alpha, a, q)^2 \\ &\qquad \ll q^{-\frac{1}{3}+\epsilon} P^{\frac{8}{5}} \min(P, P^{-2}|\beta|^{-1}) + q^{-\frac{1}{3}+\epsilon} P^{\frac{4}{5}} \min(P^2, P^{-4}|\beta|^{-1}). \end{aligned}$$

Thus we can estimate

$$\begin{aligned} \int_{\mathfrak{M}_{a,q}} |V(\alpha) - V^*(\alpha, a, q)|^2 d\alpha &\ll q^{-\frac{2}{3}+\epsilon} P^{\frac{16}{5}} \int_0^\infty \min(P^2, P^{-4}\beta^{-2}) d\beta \\ &\quad + q^{-\frac{2}{3}+\epsilon} P^{\frac{8}{5}} \int_0^\infty \min(P^4, P^{-8}\beta^{-4}) d\beta \\ &\ll q^{-\frac{2}{3}+\epsilon} P^{\frac{11}{5}} + q^{-\frac{2}{3}+\epsilon} P^{\frac{13}{5}} \ll q^{-\frac{2}{3}+\epsilon} P^{\frac{13}{5}}. \end{aligned}$$

Summing this over all the arcs in question we get

$$\sum_{a,q} \int_{\mathfrak{M}_{a,q}} |V(\alpha) - V^*(\alpha, a, q)|^2 d\alpha \ll p^{\frac{13}{5}} \sum_{q \le P^{\frac{4}{5}(1-\delta)}} q^{\frac{1}{3}+\epsilon} \ll P^{3+\frac{7}{15}}.$$

The result follows at once. □

Before we can estimate the minor arc contribution we need another preliminary result.

Lemma 7.9. *If $P^{1-\delta} < q \leq P^{2+\delta}$ and $|\beta| \leq q^{-1}P^{-2-\delta}$, then*

$$T(\alpha) \ll P^{\frac{3}{4}+\delta}.$$

Further, if $P^{\frac{4}{5}(1-\delta)} < q \leq P^{\frac{4}{5}(2+\delta)}$ and $|\beta| \leq q^{-1}P^{-\frac{4}{5}(2+\delta)}$, then

$$T_1(\alpha) \ll P^{\frac{4}{5}(\frac{3}{4}+\delta)}.$$

Proof. We recall that Weyl's bound reads

$$\sum_{x=1}^{m} e(\frac{ax^3}{q}) \ll (mq)^{\epsilon}(m^{\frac{3}{4}} + mq^{-\frac{1}{2}} + m^{\frac{1}{4}}q^{\frac{1}{4}}).$$

We set $S_m = \sum_{1 \leq x \leq m^{\frac{1}{3}}} e(\frac{ax^3}{q})$. For $m \leq 8P^3$ we get

$$S_m \ll (Pq)^{\epsilon}(P^{\frac{3}{4}} + Pq^{-\frac{1}{2}} + P^{\frac{1}{4}}q^{\frac{1}{4}}) \ll P^{\frac{3}{4}+\frac{1}{4}\delta+\epsilon}.$$

With this at hand we can conclude the proof using partial summation as follows:

$$\begin{aligned} T(\alpha) &= \sum_{n=P_1}^{P_2} (S_n - S_{n-1})e(\beta n) \\ &= \sum_{n=P_1}^{P_2} S_n(e(n\beta) - e((n+1)\beta)) - S_{P_1-1}e(\beta(P_1-1)) + S_{P_2}e(\beta P_2) \\ &\ll P^{\frac{3}{4}+\frac{1}{4}\delta+\epsilon}(P^3|\beta| + 1). \end{aligned}$$

The claim follows since $P^3|\beta|^{-1} \leq q^{-1}P^{1-\delta} \leq q$. □

We will also use the following counting result.

Lemma 7.10. *The number of integral solutions of*

$$x_1^3 + y_1^3 + z_1^3 = x_2^3 + y_2^3 + z_2^3$$

with

$$P \leq x_1, x_2 \leq 2P \text{ and } P^{\frac{4}{5}} \leq y_1, z_1, y_2, z_2 \leq 2P^{\frac{4}{5}}.$$

Is bounded by $O(P^{\frac{13}{5}+\epsilon})$.

Proof. We first count solutions of the special shape $x_1 = x_2$. Thus we can choose x_1, y_1, z_1 freely in $O(P^{\frac{15}{3}})$ ways. Since there are at most $O(P^{\epsilon})$ solutions to $m = y_2^3 + z_2^3$ we have in total $O(P^{\frac{15}{3}+\epsilon})$ solutions of this form.

For the rest of the argument we assume that $x_1 > x_2$ and write $x_2 = x$ and $x_1 = x + t$. Inserting this parametrisation yields

$$3tx^2 + 3t^2x + t^3 + y_1^3 + z_1^3 = y_2^3 + z_2^3.$$

Thee right hand side is not larger than $16P^{\frac{12}{5}}$, but the left hand side is greater than $3P^2t$. Thus we obtain the range $0 < t < 6P^{\frac{2}{5}}$ for t.

We write

$$r(m) = \sharp\{P^{\frac{4}{5}} \leq y_2, z_2 \leq 2P^{\frac{4}{5}} \colon m = y_2^3 + z_2^3\}$$

and

$$r(t,m) = \sharp\{P^{\frac{4}{5}} \leq y_1, z_1 \leq 2P^{\frac{4}{5}}, P \leq x \leq 2P \colon 3tx^2 + 3t^2x + t^3 + y_1^3 + z_1^3 = m\}.$$

We then are left to estimate

$$\sum_{0<t<6P^{\frac{2}{5}}} \sum_m r(m)r(t,m) \leq \underbrace{\left(\sum_t \sum_m r(m)^2\right)^{\frac{1}{2}}}_{\ll P^{1+\epsilon}} \cdot \left(\underbrace{\sum_t \sum_m r(m,t)^2}_{\leq M}\right)^{\frac{1}{2}} \ll P^{1+\epsilon} M^{\frac{1}{2}}.$$

A key observation is that we can take M above to be the number of solutions to

$$3tx_1^2 + 3t^2x_1 + t^3 + y_1^3 + z_1^3 = 3tx_2^2 + 3t^2x_2 + t^3 + y_2^3 + z_2^3,$$

where we impose the usual restrictions upon t, x_i, y_i, z_i.

To estimate M we first look again at the quasi-diagonal contribution $y_1^3 + z_1^3 = y_2^3 + z_2^3$. Choosing t, x_1, y_1, z_1 freely and using the same trick as above shows that this contribution is bounded by $O(P^{3+\epsilon})$.

We turn to the solutions with $y_1^3 + z_1^3 - y_2^3 - z_2^3 \neq 0$. Given any such y_i, z_i by the divisor bound we see that there are only up to $O(P^\epsilon)$ choice for t and $x_1 - x_2$. But having made these choices $x_1 + x_2$ and so x_1 and x_2 ae ultimately determined as well. Thus the contribution of these solutions is at most $O(P^{\frac{16}{5}+\epsilon})$ and we find that $M = O(P^{\frac{16}{5}+\epsilon})$. Inserting this estimate for M above completes the proof. □

We can now estimate the minor arcs.

Lemma 7.11. *Under our current assumptions we have*

$$\int_{\mathfrak{m}} |V(\alpha)|^2 d\alpha \ll P^{4+\frac{1}{10}+3\delta}.$$

Proof. We start by estimating the contribution of a minor arc $\mathfrak{m}_{a,q}$ with $P^{\frac{4}{5}(1-\delta)} < q \leq P^{1-\delta}$. Here we use

$$T(\alpha) \leq q^{-\frac{1}{3}} \min(P, P^{-2}|\beta|^{-1}) \text{ and } T_1(\alpha) \ll P^{\frac{3}{5}+\delta}.$$

We get

$$\int_{\mathfrak{m}_{a,q}} |V(\alpha)|^2 d\alpha \ll q^{\frac{-4}{3}} P^{4(\frac{3}{5}+\delta)} \int_0^\infty \min(P^4, P^{-8}\beta^{-4}) d\beta \ll q^{-\frac{4}{3}} P^{\frac{12}{5}+4\delta+1}.$$

Summing over all such minor arcs gives a contribution $\ll P^{4+\frac{1}{15}+4\delta}$. Now we consider the remaining denominators $P^{1-\delta} < q \leq P^{2+\delta}$. On these arcs we use $T(\alpha) \ll P^{\frac{3}{4}+\delta}$. We can bound

$$\int_{\mathfrak{m}} |V(\alpha)|^2 d\alpha \ll P^{\frac{3}{2}+2\delta} \int_0^1 |T(\alpha)T_1(\alpha)^2|^2 d\alpha + P^{4+\frac{1}{15}+4\delta}.$$

The key observation is now that we can reverse the starting point of circle method and recognise that the remaining α-integral counts the same solutions as does Lemma 7.10. Inserting this result concludes the proof. □

We can now attack the analytic main part of the proof. We will set $R = [P^{\frac{4}{5}(1-\delta)}]$. In the following we reserve the capital letter A, Q for positive integers with $Q \leq R$, $A \leq Q$ and $(A, Q) = 1$.

Lemma 7.12. *With our current hypothesis we have*

$$\int_0^1 |\sum_Q {\sum_A}' V^*(\alpha, A, Q)|^2 d\alpha \ll P^{4-\frac{8}{15}}.$$

Here the $\sum'$ indicates that if α is on the major arc $\mathfrak{M}_{a,q}$, then we omit the term $Q = q$ and $A = a$ from the sum.

Proof. First we use (15) to estimate

$$V^*(\alpha, A, Q) \ll Q^{-\frac{4}{3}} P^{-4} \|\alpha - \frac{A}{Q}\|^{-2} P^{\frac{8}{5}}.$$

Inserting this estimate above yields

$$\int_0^1 |\sum_Q \sum_A V^*(\alpha, A, Q)|^2 d\alpha \ll \sum_Q \sum_A P^{-\frac{24}{5}} Q^{-\frac{8}{3}} \int \|\alpha - \frac{A}{Q}\|^{-4} d\alpha$$
$$+ \sum_{\substack{Q_1, Q_2 \\ (A_1,Q_1) \neq (A_2,Q_2)}} \sum_{A_1, A_2} P^{-\frac{24}{5}} Q_1^{-\frac{4}{3}} Q_2^{-\frac{4}{3}} \int \|\alpha - \frac{A_1}{Q_1}\|^{-2} \|\alpha - \frac{A_2}{Q_2}\|^{-2} d\alpha.$$

We turn to the first sum, the diagonal contribution. Note that the integral is taken over $[0, 1)$ omitting the major arc $\mathfrak{M}_{A,Q}$. Thus we can estimate

$$\begin{aligned}\sum_Q \sum_A P^{-\frac{24}{5}} Q^{-\frac{8}{3}} \int \|\alpha - \frac{A}{Q}\|^{-4} d\alpha &\ll P^{-\frac{24}{5}} \sum_A \sum_Q Q^{-\frac{8}{3}} \int_{\frac{1}{2}Q^{-1}P^{-2-\delta}}^{\infty} \beta^{-4} d\beta \\ &\ll P^{-\frac{24}{5}} \sum_Q Q \cdot Q^{-\frac{8}{3}} Q^3 P^{3(2+\delta)} \\ &\ll P^{3+\frac{1}{15}+3\delta}.\end{aligned}$$

We turn to the off-diagonal. Here both major arcs $\mathfrak{M}_{A_1,Q_1}$ and $\mathfrak{M}_{A_2,Q_2}$ are removed from the domain of integration. For any α we have

$$\|\alpha - \frac{A_1}{Q_1}\| \geq \frac{1}{2}\|\frac{A_1}{Q_1} - \frac{A_2}{Q_2}\| \text{ or } \|\alpha - \frac{A_2}{Q_2}\| \geq \frac{1}{2}\|\frac{A_1}{Q_1} - \frac{A_2}{Q_2}\|.$$

Without loss of generality we can assume that the first is the case. Since α is not in the major arc $\mathfrak{M}_{A_2,Q_2}$ we have

$$\|\alpha - \frac{A_2}{Q_2}\| \geq \frac{1}{2}Q_2^{-1}P^{-2-\delta}.$$

With this at hand the integral can be estimated by

$$\int \|\alpha - \frac{A_1}{Q_1}\|^{-2}\|\alpha - \frac{A_2}{Q_2}\|^{-2}d\alpha \ll P^{4+2\delta}Q_2^2\|\frac{A_1}{Q_1} - \frac{A_2}{Q_2}\|^{-2}.$$

Inserting this in the sum we get

$$\sum_{\substack{Q_1,Q_2 \\ (A_1,Q_1)\neq(A_2,Q_2)}}\sum_{A_1,A_2} P^{-\frac{24}{5}}Q_1^{-\frac{4}{3}}Q_2^{-\frac{4}{3}}\int \|\alpha - \frac{A_1}{Q_1}\|^{-2}\|\alpha - \frac{A_2}{Q_2}\|^{-2}d\alpha$$

$$\ll P^{-\frac{24}{5}+4+2\delta}\sum_{\substack{Q_1,Q_2 \\ (A_1,Q_1)\neq(A_2,Q_2)}}\sum_{A_1,A_2}\frac{Q_1^{\frac{2}{3}}Q_2^{\frac{8}{3}}}{\langle A_1Q_2 - A_2Q_1\rangle^2}.$$

Here $\langle A_1Q_2 - A_2Q_1\rangle$ denotes the absolutely least residue of $A_1Q_2 - A_2Q_1$ modulo Q_1Q_2. Now if Q_1, Q_2 are fixed and we have $\langle A_1Q_2 - A_2Q_1\rangle = n$, then A_1, A_2 are determined uniquely. Thus we can finish our estimate as follows:

$$\sum_{\substack{Q_1,Q_2 \\ (A_1,Q_1)\neq(A_2,Q_2)}}\sum_{A_1,A_2} P^{-\frac{24}{5}}Q_1^{-\frac{4}{3}}Q_2^{-\frac{4}{3}}\int \|\alpha - \frac{A_1}{Q_1}\|^{-2}\|\alpha - \frac{A_2}{Q_2}\|^{-2}d\alpha$$

$$\ll P^{-\frac{24}{5}+4+2\delta}\sum_{Q_1}Q_1^{\frac{2}{3}}\sum_{Q_2}Q_2^{\frac{8}{3}}\sum_{n=1}^{\infty}n^{-2} \ll P^{4-\frac{8}{15}}.$$

□

Lemma 7.13. *We have*

$$\int_0^1 |V(\alpha) - \sum_Q\sum_A V^*(\alpha, A, Q)|^2 d\alpha \ll P^{4+\frac{1}{10}+3\delta}.$$

Proof. When α lies on the major arc $\mathfrak{M}_{a,q}$ we estimate

$$|V(\alpha) - \sum_Q\sum_A V^*(\alpha, A, Q)| \leq |V(\alpha) - V^*(\alpha, a, q)| + |\sum_Q{\sum_A}' V^*(\alpha, A, Q)|.$$

If α is on a minor arc we simply use

$$|V(\alpha) - \sum_Q\sum_A V^*(\alpha, A, Q)| \leq |V(\alpha)| + |\sum_Q{\sum_A}' V^*(\alpha, A, Q)|.$$

All the so obtained pieces can be estimated using previous results. Namely Lemma 7.8, 7.11 and 7.12. □

Lemma 7.14. *We have*

$$\sum_n (\rho(n) - \psi(n)\mathfrak{S}(P^{2+\delta}, n))^2 \ll P^{4+\frac{1}{10}+3\delta},$$

where $\psi(n) \asymp P^{\frac{3}{5}}$ *for* $3P^3 \leq n \leq 15P^3$.

Proof. We set

$$\psi(n) = \frac{1}{81} \sum_{\substack{P^3 \leq n_1, n_2 \leq (2P)^3 \\ n = n_1+n_2+n_3+n_4}} \sum_{(P^{\frac{4}{5}})^3 \leq n_3, n_4 \leq (2P^{\frac{4}{5}})^3} (n_1 n_2 n_3 n_4)^{-\frac{2}{3}}.$$

According to our definitions we have

$$I(\alpha - \frac{A}{Q})^2 I_1(\alpha - \frac{A}{Q})^2 = \sum_n \psi(n) e((\alpha - \frac{A}{Q})n).$$

Further unraveling the definitions yields

$$V^*(\alpha, A, Q) = \sum_n Q^{-4}(S_{A,Q})^4 \psi(n) e((\alpha - \frac{A}{Q})n) \tag{19}$$

and

$$\sum_Q \sum_A V^*(\alpha, A, Q) = \sum_n \psi(n) \mathfrak{S}(P^{2+\delta}, n) e(n\alpha).$$

This yields

$$\begin{aligned}\sum_n (\rho(n) - \psi(n)\mathfrak{S}(P^{2+\delta}, n))^2 &= \int_0^1 \sum_{n_1, n_2} (\rho(n_1) - \psi(n_1)\mathfrak{S}(P^{2+\delta}, n_1)) \\ &\qquad \cdot (\rho(n_2) - \psi(n_2)\mathfrak{S}(R, n_2)) e(\alpha(n_1 - n_2)) d\alpha \\ &\leq \int_0^1 |V(\alpha) - \sum_Q \sum_A V^*(\alpha, A, Q)|^2 d\alpha.\end{aligned}$$

This can be estimated using the previous lemma and yields the first part of the assertion.

We still need to consider the size of $\psi(n)$. Suppose $3P^3 \leq n \leq 15P^3$, then we can choose n_3, n_4 arbitrarily. Further we note that $n - n_3 - n_4 \sim P^3$. Thus we can also choose n_2 in at least $\gg P^3$ ways. But then n_1 is uniquely determined. Thus we have

$$\psi(n) \gg P^3 (P^{\frac{12}{5}})^2 \left(P^3 \cdot P^3 \cdot P^{\frac{12}{5}} \cdot P^{\frac{12}{5}}\right)^{-\frac{2}{3}} \gg P^{\frac{3}{5}}.$$

The upper bound is seen similarly. □

We now have to recall some facts about the singular series. Note that we already know that $A(n,q) \ll q^{-\frac{1}{3}}$. Further we have the multiplicativity $A(n,q_1q_2) = A(n,q_1)A(n,q_2)$ for $(q_1,q_2)=1$. Recall that we already investigated the numbers

$$M(q) = \sharp\{0 \le x_1, \ldots, x_4 < q \colon x_1^3 + \ldots + x_4^3 \equiv n \bmod q\}.$$

Note that from Lemma 6.9 it follows that

$$A(p^l) = \frac{M(p^l)}{p^{3l}} - \frac{M(p^{l-1})}{p^{3(l-1)}}.$$

We will now require the corresponding primitive count:

$$N(p^l) = \sharp\{0 \le x_1, \ldots, x_4 < p^l,\ (x_1,x_2,x_3,x_4,p)=1,\ x_1^3 + \ldots + x_4^3 \equiv n \bmod q\}.$$

We need the following result, refining some of our earlier investigations. The upshot of this modification is that one can show

$$N(p^l) = p^{(l-\gamma)3} N(p^\gamma) \text{ for } l \ge \gamma, \tag{20}$$

where $\gamma = 2$ if $p = 2,3$ and $\gamma = 1$ otherwise.

Lemma 7.15. *Write $p^{3\rho+\sigma} \| n$ where $0 \le \sigma \le 2$. Set $l_0 = \max(3\rho+\sigma+1, 3\rho+\gamma)$. Then $A(n,p^l) = 0$ if $l > l_0$ and*

$$\chi_p(n) = \sum_{v=0}^{\infty} A(n,p^v) = p^{-3\gamma} N(p^\gamma, 0) \sum_{v=0}^{\rho-1} p^{-v} + p^{-\rho-3\gamma} N(p^\gamma, np^{-3\rho}).$$

In particular, if $p \nmid 6n$, then $A(n,p^l) = 0$ for $l > 1$.

A similar results holds for any s and any k but we will stick to the special case at hand.

Proof. It suffices to show that $p^{-3l}M(p^l,n)$ equals the claimed expression. Thus given $l > l_0$ we have to reduce the count of solutions to the congruence in question to primitive solutions.

Write $h_1^3 + \ldots + h_4^3 \equiv n \bmod p^l$. First note that not all h_i are divisible by $p^{\rho+1}$. Indeed if this were the case, then $p^{3\rho+\sigma+1}$ would divide $h_1^3 + \ldots + h_4^3$. (This is since $3\rho+\sigma+1 \le 3\rho+3$.) But $l \ge l_0 \ge 3\rho+\sigma+1$, so that $p^{3\rho+\sigma+1} \mid n$. This is a contradiction.

We divide the $M(p^l,n)$ solutions to the congruence in $\rho+1$ classes as follows. For $0 \le i \le \rho$ write $M_i(p^l,n)$ for the number of solutions $(h_1,\ldots,h_4)$ for which $h_1^3 + \ldots + h_4^3 \equiv n \bmod p^l$ and $p^i \mid\mid (h_1,\ldots,h_4)$. Given a solution contributing to $M_i(p^l,n)$. Then we can write $h_v = p^i y_v$. Since $0 < l_0 - 3\rho \le l - 3i$ we find that

$$y_1^3 + \ldots + y_4^3 \equiv p^{-3i} n \bmod p^{l-3i},\ 0 \le y_v < p^{l-i} \text{ and } p \nmid (y_1,\ldots,y_4).$$

The correspondence $(h_1,\ldots,h_4) \to (y_1,\ldots,y_4)$ is actually one to one and by counting the possible quadruples $(y_1,\ldots,y_4)$ we get

$$M_i(p^l,n) = p^{8i} N(p^{l-3i}, p^{-3i}n).$$

We combine this to

$$M(p^l,\alpha) = \sum_{i=0}^{\rho} p^{8i} N(p^{l-3i}, p^{-3i}n) = p^{3(l-\gamma)} \sum_{i=0}^{\rho} p^{-i} N(p^{\gamma}, p^{-3i}n).$$

In the last step we inserted (20) which is possible since $l - 3i \geq l_0 - 3\rho \geq \gamma$.

Since $\gamma \leq 3$ we have $\gamma + 3i \leq 3 + 3i \leq 3\rho$ for $i < \rho$. In particular $p^{\gamma} \mid p^{-3i}n$. So that in this case $N(p^{\gamma}, p^{-3i}n) = N(p^{\gamma}, 0)$. Inserting this completes the proof. □

Lemma 7.16. *We have $N(p^{\gamma}, n) > 0$ for all n.*

Proof. We first deal with the exceptional case $p = 3$. In particular $3^{\gamma} = 9$ and one can exhibit solutions to the congruence explicitly. For example $1^3 + 8^3 + 0^3 + 0^3 \equiv 0 \bmod 9$ and more generally

$$\underbrace{1^3 + \ldots + 1^3}_{m \text{ times}} + \underbrace{0^3 + \ldots + 0^3}_{(4-m) \text{ times}} \equiv m \bmod 9 \tag{21}$$

for $1 \leq m \leq 4$. The remaining m can be dealt with by replacing 1 by 8 on the left hand side and m by $-m$ on the right hand side.

The other exceptional case $p = 2$ is even easier, so that we can turn towards $p > 3$, where $\gamma = 1$. We write $r = (3, p-1) \leq 3$. For $p \nmid n$ and we claim that r-variables suffice to solve the congruence. In particular, by setting the remaining variables we see that 3 variables suffice in this case.

We call $n_1, n_2 \in \mathbb{Z}/p\mathbb{Z}$ equivalent if there x with is $(x,p) = 1$ such that $x^3 n_1 \equiv n_2 \bmod p$. This is indeed an equivalence relation and we claim that there are $1 + r$ equivalence classes. The trivial class is obviously just 0. Thus we need to see that there are r classes remaining. To see this we choose a primitive root of unity g, so that

$$g^0, g, \ldots, g^{p-2}$$

ganz $(\mathbb{Z}/p\mathbb{Z})^{\times}$ representieren. Note that $g^{m_1} \equiv g^{m_2} \bmod p$ if and only if $(p-1) \mid (m_1 - m_2)$. We will compute how many of the number g^{3m} with $0 \leq m \leq p-2$ are distinct modulo p. But as remarked above this is detected by $(p-1) \mid 3(m_1 - m_2)$. But the latter is equivalent to $m_1 \equiv m_2 \bmod \frac{p-1}{(3,p/1)}$. Thus each non-trivial equivalence class has $\frac{p-1}{(3,p-1)}$ elements. Therefore there are $(3, p-1)$ distinct non-trivial classes as claimed.

Let $0 < n_1 < \ldots < n_r < p$ be the first representative in each of the r non-trivial equivalence classes. Obviously it is enough to show that the congruence $h_1^3 + \ldots + h_i^3 \equiv n_i \bmod p$ has a solution for all $i = 1, \ldots, r$. We continue by induction. Of course $n_1 = 1 = 1^3$. Now $n_{i+1} - 1$ is (by minimality) in a class $[n_j]$ with $1 \leq j \leq i$. Thus there is x such that $(n_{i+1} - 1) = x^3 n_j$. By induction hypothesis we can write

$$n_{i+1} \equiv 1^3 + x^3(h_1^3 + \ldots + h_j^3) \bmod p$$

and we are done.

Finally, if $p \mid n$, we can write $n - 1 \equiv y_1^3 + \ldots + y_3^3 \bmod p$. But then

$$n \equiv y_1^3 + \ldots + y_3^3 + 1^3 \bmod p.$$

□

Lemma 7.17. *For any p and any n we have*

$$\chi_p(n) \geq p^{-6}.$$

Proof. If $p^3 \nmid n$, then we have seen earlier that

$$\chi_p(n) = p^{-3\gamma} N(p^\gamma, n) \geq p^{-3\gamma} \geq p^{-6}.$$

Similarly, if $p^3 \mid n$, we get the lower bound

$$\chi_p(n) \geq p^{-3\gamma} N(p^\gamma, 0) \geq p^{-3\gamma} \geq p^{-6}.$$

□

Lemma 7.18. *For any prime p we have the estimate*

$$A(n,p) \ll \begin{cases} p^{-\frac{3}{2}} & \text{if } p \nmid n, \\ p^{-1} & \text{if } p \mid n. \end{cases}$$

Furthermore,

$$\chi_p(n) - 1 \ll p^{-\frac{3}{2}} \text{ for } p \nmid n$$

and $\chi_p(n) > 1 - Cp^{-1}$ if $p \mid n$.

Proof. The bounds on $\chi_p(n)$ follow from the bounds on $A(n,p)$ by taking into account when the sum defining $\chi_p(n)$ terminates. (Some cases need to be distinguished but this is not difficult.)

If $p \mid n$ we simply conclude using the previously established bound $|S_{a,q}| \leq (\delta - 1)\sqrt{p} \leq 2\sqrt{p}$ and trivial estimates.

Finally, if $p \nmid n$, we recall that we have seen before that

$$A(p,n) = p^{-4} \sum_{r=1}^{p-1} e(-\frac{rn}{p})(\sum_{\psi} \overline{\psi}(r)\tau(\psi))^4,$$

where the ψ-sum runs over non-principal characters whose 3rd power is principal. The bound follows by opening the 4th power, taking the r-sum inside to find another Gauß sum and using standard estimates for Gauß sums. □

Lemma 7.19. *The series $\mathfrak{S}(n)$ is absolutely convergent and satisfies*

$$\mathfrak{S}(n) \gg \log\log(n)^{-O(1)}.$$

Furthermore,

$$\sum_{q \geq \eta} A(n,q) \ll \eta^{-\frac{1}{6}} n^{\epsilon}.$$

Proof. Absolute convergence follows directly from the estimates given in the previous lemma. To obtain the lower bound we compute

$$\mathfrak{S}(n) = \prod_p \chi_p(n) > \left(\prod_{p\leq 2C} p^{-6}\right)\left(\prod_{\substack{p>2C,\\ p\nmid n}} (1-Cp^{-\frac{3}{2}})\right)\left(\prod_{\substack{p>2C,\\ p|n}} (1-Cp^{-1})\right)$$
$$\gg \prod_{p|n}(1-p^{-1})^{O(1)} \gg \log\log(n)^{-O(1)}.$$

We still have to show the upper bound on the tails of the singular series. Given q we write $q = q'q_3$ for $(q',6)=1$ and $q_3 \mid 6^\infty$. We further write $q' = q_1q_2$ by defining $q_2 = q'/q_1$ and $p \mid q_1$ if and only if $p \| q'$. Of course $q = q_1q_2q_3$ and the q_i's are pairwise co-prime.

If $p^l \| q_2$, then $p > 3$ and $l > 1$. Thus $A(n,p^l) = 0$ if $l > l_0 = 3\rho+\sigma+1$. Therefore $A(n,p^l) \neq 0$ implies $l \leq 3\rho+\sigma+1$, so that $p^{l-1} \mid n$. This in turn implies $p^l \mid n^2$ and $q_2 \mid n^2$.

We will use the bounds

$$A(n,q_i) \ll q_i^{-\frac{1}{3}} \text{ for } i = 2,3.$$

The q_1 part can be estimated differently:

$$A(n,q_1) \ll q_1^{-\frac{3}{2}+\epsilon}(n,q_1)^{\frac{1}{2}}.$$

With this at hand we are ready to estimate

$$\sum_{q\geq\eta} A(n,q) \ll \sum_{\substack{q_1q_2q_3\geq\eta,\\ q_2|n^2}} q_1^{-\frac{3}{2}+\epsilon}(n,q_1)^{\frac{1}{2}}(q_2q_3)^{-\frac{1}{3}} \ll \eta^{-\frac{1}{6}} \sum_{\substack{q_1,q_2,q_3,\\ q_2|n^2}} q_1^{-\frac{4}{3}+\epsilon}(n,q_1)^{\frac{1}{2}}(q_2q_3)^{-\frac{1}{6}}.$$

We are done after inserting the estimates

$$\sum_{q_2|n^2} q_2^{-\frac{1}{6}} \leq d(n^2) \ll n^\epsilon,$$
$$\sum_{q_3} q_3^{-\frac{1}{6}} \leq (1-2^{-\frac{1}{6}})^{-1}(1-3^{-\frac{1}{6}})^{-1} \ll 1 \text{ and}$$
$$\sum_{q_1} q_1^{-\frac{4}{3}+\epsilon}(n,q_1)^{\frac{1}{2}} \ll \sum_{d|n} d^{\frac{1}{2}} \sum_{r=1}^{\infty} (rd)^{-\frac{4}{3}+\epsilon} \ll d(n) \ll n^\epsilon.$$

□

Lemma 7.20. *We have*

$$\sum_{3P^3\leq n\leq 15P^3} (\psi(n)\mathfrak{S}(P^{2+\delta},n) - \psi(n)\mathfrak{S}(n))^2 \ll P^4.$$

Proof. Note that $\psi(n) \ll P^{\frac{3}{5}}$. The result follows directly from

$$\mathfrak{S}(n) - \mathfrak{S}(P^{2+\delta}, n) = \sum_{q > P^{2+\delta}} A(n,q) \ll P^{-\frac{1}{6}(2+\delta)} P^{\epsilon}.$$

□

We are now finally ready to proof the main result of this section.

Proof of Theorem 7.3. Let $E(N)$ denote the number of positive integers less than N that are not representable as sum of four positive integral cubes. We aim to show that

$$E(N) \ll N^{1-\frac{1}{30}+4\delta}.$$

Set $P = (\frac{1}{5}N)^{\frac{1}{3}}$. This choice is made so that $3P^3 < N < 2N < 15P^3$. Thus we get

$$\sum_{N<n\leq 2N} (\rho(n) - \psi(n)\mathfrak{S}(n))^2 \ll P^{4+\frac{1}{10}+3\delta}.$$

Of course we have $\rho(n) = 0$, when n contributes to the count of $E(2N)$. For such an n with $N < n \leq 2N$ we have

$$(\rho(n) - \psi(n)\mathfrak{S}(n))^2 = \psi(n)^2\mathfrak{S}(n)^2 \gg P^{\frac{6}{5}} \log\log(P^3)^{-O(1)} \gg P^{\frac{6}{5}-\epsilon}.$$

By assuming P to be large and making ϵ smaller if necessary we can assume that the implicit constant is 1. We can now conclude that

$$E(2N) - E(N) \leq P^{-\frac{6}{5}+\epsilon} \sum_{N<n\leq 2N} (\rho(n) - \psi(n)\mathfrak{S}(n))^2 \ll P^{3-\frac{1}{10}+3\delta+\epsilon} \ll N^{1-\frac{1}{30}+4\delta},$$

for $N > N_0$. We now fix r_0 such that $2^{r_0+1} < \frac{N}{N_0}$. A standard dyadic sum argument yields

$$E(N) \ll N2^{-r_0-1} + \sum_{r=0}^{r_0} \left(\frac{N}{2^{r+1}}\right)^{1-\frac{1}{30}+\delta} \ll N^{1-\frac{1}{30}+\delta}, \tag{22}$$

for $2^{r_0} \leq N^{\frac{1}{30}} < 2^{r_0+1}$ and N large enough. This concludes the proof. □

8. Decoupling and Vinogradov's mean value theorem

We now change gear a little and discuss some very recent developments in the intersection of harmonic analysis and number theory.

8.1. **Harmonic Analysis and decoupling.**

We will discuss some ideas that ultimately lead to the infamous decoupling theorems that brought much progress to harmonic analysis and number theory recently. Due to time and place constraints we can not give full details. Most of the ideas are taken from a survey article of L. Pierce on the topic (see [Pi]).

8.1.1. *From orthogonality, Littlewood-Paley theory and restriction problems.* Let $\mathcal{H}$ be a Hilbert space and $(f_i)_{i\in I}$ a (finite) family of elements. By Cauchy-Schwarz we have the estimate

$$\|\sum_{i\in I} f_i\|_{\mathcal{H}} \leq \sqrt{\sharp I}\left(\sum_{i\in I}\|f_i\|_{\mathcal{H}}^2\right)^{\frac{1}{2}}.$$

We all know, that if the f_i's are orthogonal (i.e. $\langle f_i, f_j\rangle_{\mathcal{H}} = 0$ for $i \neq j$), then this improves to

$$\|\sum_{i\in I} f_i\|_{\mathcal{H}} = \left(\sum_{i\in I}\|f_i\|_{\mathcal{H}}^2\right)^{\frac{1}{2}}.$$

Essentially we have already used this observation in the setting of $\mathcal{H} = L^2([0,1])$, where the set $f_n(x) = e^{2\pi i n x}$ is a orthogonal sequence.

Example 8.1. For $\mathcal{H} = L^2(\mathbb{R}^n)$ we can choose (nice) disjoint sets $(U_i)_{i\in I}$. If we pick f_i such that $\operatorname{supp}(\hat{f}_i) \subset U_i$, then the functions f_i are orthogonal. This follows from the well known Plancherel theorem:

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx = \int_{\mathbb{R}^n}\hat{f}(x)\overline{\hat{g}(x)}dx.$$

Building on this example we can take a function and dissect its Fourier transform into pieces. This technique (Littlewood-Paley theory) yields a sequence of orthogonal elements of $L^2(\mathbb{R}^n)$. To be more precise we put

$$\Delta_j(\xi) = \mathbb{1}_{U_j} \text{ for } U_j = B_0(2^{j+1}) \setminus B_0(2^j).$$

Now we obtain the following operators

$$[P_j f](x) = (\Delta_j\hat{f})^{\vee}(x).$$

(Here $f^{\vee}$ denotes the Fourier Inversion.) The argument from the example shows that the functions $f_j = P_j f$ are pairwise orthogonal. One sees that

$$\|(\sum_j |P_j f|^2)^{\frac{1}{2}}\|_{L^2} = \|f\|_{L^2}.$$

It is a key inside that (at least for $n = 1$) we can salvage this equality to some extend. Indeed we get

$$\|(\sum_j |P_j f|^2)^{\frac{1}{2}}\|_{L^p(\mathbb{R})} \asymp_p \|f\|_{L^p(\mathbb{R})} \tag{23}$$

The deep *Ball Multiplier Theorem* due to C. Fefferman shows that we can not use sharp cut-offs such as the functions Δ_j in higher dimension. I.e. the upper and lower bounds (23) fail (for general f) in $L^p(\mathbb{R}^n)$ with $p \neq 2$ and $n > 1$.

In the setting above we call the f_i's *quasi orthogonal* if (23) holds for all $1 < p < \infty$. Quasi orthogonal families can for example be constructed (in all dimensions) by replacing the Δ_i's with smooth bumb function that are essentially concentrated in the annulus U_i and are of rapid decay outside.

The construction above also makes it natural to study so called square-functions. Say we have a family of operators T_i then the *square function* is the operator

$$f \mapsto \left(\sum_i |T_i f|^2\right)^{\frac{1}{2}}.$$

Half of the content of (23) says that the square function (associated to the U_j) is bounded in L^p.

At this point we will not define what we mean by a decoupling inequality but we give some examples that feature de-coupling.

Example 8.2. Let $(f_i)_{i\in I}$ be a family of functions such that there is $N \in \mathbb{N}$ such that for each ξ:

$$\sharp\{i \in I \colon \hat{f}_i(\xi) \neq 0\} \leq N.$$

We write $w_i = \operatorname{supp}(\hat{f}_i)$. Using Plancherel and Cauchy-Schwarz yields

$$\|\sum_i f_i\|_{L^2} = \|\sum_j \hat{f}_j\|_{L^2} \leq \|\left(\sum_i |\hat{f}_i|^2\right)^{\frac{1}{2}} \underbrace{\left(\sum_i \mathbb{1}_{w_i}\right)^{\frac{1}{2}}}_{\leq N} \|_{L^2}.$$

We obtain the following l^2 *decoupling inequality for* L^2:

$$\|\sum_i f_i\|_{L^2} \leq \sqrt{N}\|\left(\sum_i |\hat{f}_i|^2\right)^{\frac{1}{2}}\|_{L^2} = \sqrt{N}\left(\sum_i \|\hat{f}_i\|_{L^2}^2\right)^{\frac{1}{2}} = \sqrt{N}\left(\sum_i \|f_i\|_{L^2}^2\right)^{\frac{1}{2}}.$$

Example 8.3. Let $f_j(x) = e(j^2 x)$ for $j \in \{1, \ldots, N\}$. We claim that these functions satisfy the following l^2 decoupling inequality for $L^4([0,1])$:

$$\|\sum_{j=1}^N e(j^2 x)\|_{L^4} \ll N^\epsilon \left(\sum_{j=1}^N \|e(j^2 x)\|_{L^4}^2\right)^{\frac{1}{2}}.$$

To see this we argue as follows. First observe that

$$\|\sum_{j=1}^N e(j^2 x)\|_{L^4}^4 = \sharp\{1 \leq x_1, \ldots, x_4 \leq N \colon x_1^2 + x_2^2 = x_3^2 + x_4^2\}$$

by character orthogonality. If we choose x_1, x_2 freely, then there are up to N^ϵ possibilities for x_3, x_4. Thus

$$\|\sum_{j=1}^N e(j^2 x)\|_{L^4}^4 \ll N^{2+\epsilon}.$$

But the right hand side can be computed to be

$$\left(\sum_{j=1}^{N} \|e(j^2x)\|_{L^4}^2\right)^{\frac{1}{2}} = N^{\frac{1}{2}}.$$

This is because

$$\int_0^1 |e(j^2x)|^4 dx = \int_0^1 1dx = 1.$$

Let $S \subset \mathbb{R}^n$ be a compact smooth submanifold with induced Lebesgue measure $d\sigma$. We say that (L^p, L^q) *restriction holds for S* if

$$\|\hat{f}|_S\|_{L^q(S,d\sigma)} \ll_{p,q,S} \|f\|_{L^p(\mathbb{R}^n)}$$

holds for every Schwartz-function f. There are many interesting phenomena related to these restriction problems, which we unfortunately don't discuss here.

Let us define the *restriction operator*

$$R_S f = \hat{f}|_S.$$

We have the closely related *extension operator*

$$[E_S g](x) = \int_S g(\xi)e(x\cdot\xi)d\sigma(\xi).$$

(This is the inverse Fourier transform $(gd\sigma)^{\vee}$ along S) The point of these operators is that one can formulate the restriction problem by saying

$$R_S\colon L^p(\mathbb{R}^n) \to L^q(S,d\sigma) \text{ or } E_S\colon L^{q'}(S,d\sigma) \to L^{p'}(\mathbb{R}^n)$$

is bounded for $1/p + 1/p' = 1 = 1/q + 1/q'$. The main conjecture among these lines is the so called *restriction conjecture.* We briefly state the adjoint form:

Conjecture 8.1. *Let S be a compact C^2 hypersurface in $\mathbb{R}^n$ with nonvanishing Gaussian curvature at every point. Then for $p' > \frac{2n}{n-1}$ and $q \leq \left(\frac{n-1}{n+1}\right) p'$ we have*

$$\|E_S g\|_{L^{p'}(\mathbb{R}^n)} \ll_{p,q,S} \|g\|_{L^{q'}(S,d\sigma)}.$$

The number $\frac{2n}{n-1}$ is called the restriction exponent.

The Tomas-Stein Restriction Theorem makes progress towards this conjecture. (Instead of the restriction exponent only the Tomas-Stein exponent $\frac{2(n+1)}{n-1}$ is reached.)

We now state a quite general framework in which one can consider decoupling estimates. Later we specialise to more restricted, but still important cases.

Consider a compact smooth manifold S in $\mathbb{R}^n$ with associated measure σ. We partition (or cover) S by caps τ of size δ. Given $g\colon S \to \mathbb{C}$ we write $g_\tau = \mathbb{1}_\tau \cdot b$ for the restriction of g to τ. An l^2 decoupling result for L^p states that there exists a critical index $p_c > 2$ and some $\kappa \geq 2$ such that

$$\|E_S g\|_{L^p(B)} \ll_\epsilon \delta^{-\epsilon} \left(\sum_\tau \|E_S g_\tau\|_{L^p(B)}^2\right)^{\frac{1}{2}}$$

for each ball $B \subset \mathbb{R}^n$ with radius $\delta^{-\kappa}$ and for each $2 \leq p \leq p_c$.

The first (sharp) l^2 decoupling result was shown for compact C^2 hypersurfaces $S \subset \mathbb{R}^n$ with appropriate curvature by Bourgain and Demeter. We are more interested for a version concerning curves which will apply to Vinogradov's mean value theorem.

8.1.2. *Decoupling for the moment curve.* For $n \geq 2$ and any interval $J \subset [0,1]$ we define the moment curve

$$\Gamma_J = \{(t, t^2, \ldots, t^n) \colon t \in J\} \subset \mathbb{R}^n.$$

Given an integrable function $g \colon [0,1] \to \mathbb{C}$ we define the extension operator

$$[E_J g](x_1, \ldots, x_n) = \int_J g(t) e(tx_1 + t^2 x_2 + \ldots + t^n x_n) dt.$$

Let B be a ball of radius $1 \leq \delta^{-n}$ centered at x_0 in $\mathbb{R}^n$. We define the weight

$$w_B(x) = (1 + |x - x_0|\delta^n)^{-E} \text{ for } E \gg n \text{ sufficiently large.}$$

We work with the weighted L^p-norms:

$$\|f\|_{L^p(w_B)} = \left(\int_{\mathbb{R}^n} |f(x)|^p w_B(x) \right)^{\frac{1}{p}}.$$

We use these weighted version as replacement for $L^p(B)$. Indeed thinking of E sufficiently large, there is essentially no weight outside of B.

Theorem 8.4 (l^2 decoupling for $L^{n(n+1)}$ for the moment curve in $\mathbb{R}^n$)**.** *In the notation above we have*

$$\|E_{[0,1]} g\|_{L^{n(n+1)}(w_B)} \ll_{\epsilon, n} \delta^{-\epsilon} \left(\sum_{\substack{J \subset [0,1], \\ |J| = \delta}} \|E_J g\|^2_{L^{n(n+1)}(w_B)} \right)^{\frac{1}{2}}$$

for all integrable $g \colon [0,1] \to \mathbb{C}$. (Note that the implied constant is independent of δ, B and g.)

We outline the originally proof by Bourgain-Demeter-Guth from the harmonic analysis perspective. Note that there is a simplified argument due to Guo, Li, Yung and Zorin-Kranich, which was inspired by the (nested) efficient congruencing approach due to Wooley.

Proof (Outline): We write

$$\|E_{[0,1]} g\|_{L^p(w_B)} \leq V_{p,n}(\delta) \left(\sum_{\substack{J \subset [0,1], \\ |J| = \delta}} \|E_J g\|^2_{L^p(w_B)} \right)^{\frac{1}{2}}, \tag{24}$$

so that we need to show $V_{p,n}(\delta) \ll_{\epsilon,n} \delta^{-\epsilon}$ for $p = p_n = n(n+1)$.

Step 1 (Initial reductions). We start with some basic reduction steps and easy cases.

Claim 1.a *It is enough to show $V_{p,n}(\delta) \ll_{n,\epsilon} \delta^{-\epsilon}$ for all $p < p_n$ sufficiently close to the critical case.* To see this one first shows

$$\|E_J g\|_{L^{p_n}(w_B)} \ll \|E_J g\|_{L^p(w_B)} \text{ for every } 1 \le p < p_n.$$

This is non-trivial and it uses that $E_J g$ is frequency localised. (Ideas related to a Bernstein inequality are necessary.) With this at hand we assume (24) is known for p. We apply Hölder with $q = \frac{p_n}{p} > 1$ to see

$$\left(\sum_J \|E_J g\|^2_{L^p(w_B)}\right)^{\frac{1}{2}} \ll \|\mathbb{1}_{[0,q]}\|_{L^{pq'}(w_B)} \left(\sum_J \|E_J g\|^2_{L^{p_n}(w_B)}\right)^{\frac{1}{2}}.$$

The idea is that $pq' = \frac{pp_n}{p_n - p}$ goes to infinity as p approaches p_n. Thus by taking p close enough to p_n we can easily control the contribution of $\|\mathbb{1}_{[0,q]}\|_{L^{pq'}(w_B)}$.

Claim 1.b *Decoupling is translation and dilation invariant.* This relies on a tricky affine change of variables and instead of giving any details we only state this more precisely. Suppose (24) is known. Then we have the following. For any $0 < \delta \le 1$, any $0 < \rho \le 1$ and every interval I of length δ^ρ we have

$$\|E_I g\|_{L^p(w_B)} \le V_{p,n}(\delta^{1-\rho}) \left(\sum_{\substack{J \subset I, \\ |J| = \delta}} \|E_J g\|^2_{L^p(w_B)}\right)^{\frac{1}{2}}.$$

Note that the ball B still has radius δ^{-n}.

Claim 1.c *For every $0 < \delta \le 1$ and for any interval I, and for any ball B of radius δ^{-1} in $\mathbb{R}^n$ we have*

$$\|E_I g\|_{L^2(w_B)} \ll \left(\sum_{\substack{J \subset I, \\ |J| = \delta}} \|E_J g\|^2_{L^2(w_B)}\right)^{\frac{1}{2}}$$

for a dissection of I into subintervals J of length δ. (In other words, $V_{2,n}(\delta) \ll 1$.)

This can be reduced to the case where we replace the weight w_B on the left hand side by the sharp cut-off $\mathbb{1}_B$ and the weight w_B on the left hand side by $\eta_B(x) = \eta((x - x_0)R^{-1})$. Here η is chosen such that the Fourier transform of $\sqrt{\eta}$ is supported in a small neighborhood of the origin. We also assume that $\eta(x) \ge 1$ on the unit ball centered at 0. With this choice made we observe that

$$\|E_I g\|^2_{L^2(B)} \ll \|E_I g\|^2_{L^2(\eta_B)} = \|\sqrt{\eta_B} E_I g\|^2_{L^2(\mathbb{R})} = \|\sum_J \sqrt{\eta_B} E_J g\|^2_{L^2(\mathbb{R}^n)}.$$

As in our example at the beginning we now want to show that the Fourier transforms of $\sqrt{\eta_B}E_{J_i}g$, $i = 1, 2$ have disjoint support as soon as J_1 and J_2 are distinct and non-adjacent. If we know this we obtain the result by an almost orthogonality result that we have already seen. But this support property can be shown using the convolution theorem (for Fourier transforms) and the construction of η (and of course the extension operator $E_J g$).

Step 2 (Controlling linear decoupling by multilinear objects). We first need to define the multilinear decoupling parameter, which will play a key role here.

Let M and K be sufficiently large (in terms of n) parameters. Further take $n \geq 2$, $2 \leq p < p_n$ and $0 < \delta \leq 1$ as usual. Define $V_{p,n}(\delta, K)$ to be the smallest positive real number such that for every collection $I_1, \dots, I_M$ of pairwise non-adjacent intervals of the form $[i/K, (i+1)/K]$ and all balls B of radius δ^{-n} we have

$$\| \left(\prod_{j=1}^{M} E_{I_i} g \right)^{\frac{1}{M}} \|_{L^p(w_B)} \leq V_{p,n}(\delta, K) \left(\prod_{j=1}^{M} \left(\sum_{J \subset I_j, |J| = \delta} \|E_J g\|^2_{L^p(w_B)} \right)^{\frac{1}{2}} \right)^{\frac{1}{M}},$$

for every $g \colon [0,1] \to \mathbb{C}$.

Claim 2.a *Linear decoupling implies multilinear decoupling. In other words, $V_{p,n}(\delta, K) \leq V_{p,n}(\delta)$.* This claim relies on two simple facts. First, by Hölder we have

$$\| \left(\prod_{j=1}^{M} E_{I_i} g \right)^{\frac{1}{M}} \|_{L^p(w_B)} \leq \prod_{j=1}^{M} \left(\|E_{i_j} g\|_{L^p(w_B)} \right)^{\frac{1}{M}} \tag{25}$$

Second, we put $g = \sum_{j=1}^{M} g_j$, where g_j is supported on I_j. Since the intervals are non-adjacent and distinct we have

$$E_{I_j} g = E_{[0,1]} g_j.$$

Thus we can apply linear decoupling to each factor, which gives the claim directly.

Claim 2.b *Multilinear decoupling implies linear decoupling.* To be more precise one can show that there exist constants $C_{K,p}$ and $\epsilon_(K)$ with $\lim_{K \to \infty} \epsilon_p(K) = 0$ such that for $0 < \delta \leq 1$ we have

$$V_{p,n}(\delta) \leq C_{K,p} \delta^{-\epsilon_p(K)} \sup_{\delta \leq \delta' < 1} V_{p,n}(\delta', K).$$

Compared to the other implication discussed in Claim 2.a this is highly non-trivial. The proof proceeds via induction on scales. In particular one combines clever partition arguments, the arithmetic-geometric mean inequality and the rescaling principle. We omit the details.

Key Definition: We come to the definition of a key quantity. We start by rescaling our weighted L^p-norm to:

$$\|F\|_{L^p_\sharp(w_B)} = \left(\frac{1}{\mathrm{Vol}(B)}\int |F|^p w_B dx\right)^{\frac{1}{p}}.$$

It can be seen from the definitions, that this re-normalisation works well with multilinear decoupling. For $1 \leq t < \infty$ and $q, r > 0$ we set

$$D_t(q, B^r, g) = \left(\prod_{j=1}^{M}\left(\sum_{J_j \subset I_j, |J_j| = \delta^1} \|E_{J_j} g\|^2_{L^t_\sharp(w_{B^r})}\right)^{\frac{1}{2}}\right)^{\frac{1}{M}},$$

where B^u is a ball of radius δ^{-u}. In this notation the definition of $V_{p,n}(\delta, K)$ reads:

$$\|\left(\prod_{j=1}^{M} E_{I_j} g\right)^{\frac{1}{M}} \|_{L^p_\sharp(w_{B^n})} \leq V_{p,n}(\delta, K) D_p(1, P^n, g).$$

The definition we want to make concerns an average of these quantities. For any ball B^r (of radius δ^{-r}) and a finitely overlapping cover $\mathcal{B}_s(B^r)$ of B^r by balls B^s (of radius δ^{-s}) we define

$$A_p(q, B^r, s, g) = \left(\frac{1}{|\mathcal{B}_s(B^r)|}\sum_{B^s \in \mathcal{B}_s(B^r)} D_2(q, B^s, g)^p\right)^{\frac{1}{p}}.$$

Claim 2.c *We have the estimate:*

$$\|\left(\prod_{j=1}^{M} E_{I_j} g\right)^{\frac{1}{M}} \|_{L^p_\sharp(w_{B^n})} \ll \delta^{-\frac{u}{2}} A_p(u, B^n, u, g).$$

This follows from standard inequalities (Cauchy-Schwarz, Minkowski), a Bernstein type property and some tricks to convert weights. We omit the details.

Claim 2.d *If we can control the quantity $A_p(q, B^r, s, g)$, then decoupling follows.* We will first explain what we mean by controlling $A_p(\ldots)$. To state this precisely we define $\eta_p \geq 0$ to be the unique real number such that

$$\lim_{\delta \to 0} V_{p,n}(\delta)\delta^{\eta_p + \sigma} = 0 \text{ and } \lim_{\delta \to 0} V_{p,n}(\delta)\delta^{\eta_p - \sigma} = \infty$$

for every $\sigma > 0$. We assume that the following holds:

Let $n \geq 3$, $2 \leq p < p_n$. Suppose that Theorem 8.4 is known for all dimensions $k \leq n-1$. Then for every $W > 0$ and for every sufficiently small $u > 0$, we have for every $g\colon [0,1] \to \mathbb{C}^n$, every $0 < \delta \leq 1$ and every ball $B^n \subset \mathbb{R}^n$ of radius δ^{-n} we have

$$A_p(u, B^n, u, g) \ll_{\sigma,\epsilon,K,W} \delta^{-\epsilon}\delta^{-(\eta_p+\sigma)(1-uW)} D_p(1, B^n, g), \tag{26}$$

for every $\epsilon, \sigma > 0$.

Using this result, taking the supremum over g, $(I_j)_j$ and balls B^n in Claim 2.c, and unraveling the definitions yields

$$V_{p,n}(\delta, K) \ll_{\sigma,\epsilon,K,W} \delta^{-\frac{u}{2}}\delta^{-\epsilon}\delta^{-(\eta_p+\sigma)(1-uW)}.$$

According to Claim 2.b the linear decoupling constant $V_{p,n}(\delta)$ is controlled by the multilinear version. Thus, for a sequence $\delta \to 0$ we get

$$\delta^{-(\eta_p-\sigma)} \ll_{\sigma,\epsilon,K,W} \delta^{-\epsilon_p(K)}\delta^{-\frac{u}{2}}\delta^{-\epsilon}\delta^{-(\eta_p+\sigma)(1-uW)}.$$

Looking at the exponents we find that this implies

$$\eta_p \leq \frac{1}{2W} + \frac{\epsilon + \epsilon_p(K) + \sigma(2-uW)}{uW}.$$

Note that this holds for all $\epsilon, \sigma > 0$ arbitrarily small. Further we can take K arbitrarily large. Finally recall that $\epsilon_p(K) \to 0$ as $K \to \infty$. Thus $\eta_p \leq \frac{1}{2W}$. Since we can take W as large as necessary we get $\eta_p = 0$. By definition of η_p this implies the decoupling theorem.

Step 3 (Tools to control $A_p(\ldots)$). It remains to derive (26). This is an iterative process and in this step we provide some key tools one needs to proceed.
Claim 3.a *Multilinear l^2 decoupling for L^2.* We have already seen that l^2 decoupling for L^2 is special, see Claim 1.c. This holds true for multilinear decoupling as well. For every $0 < \delta \leq 1$ let $\mathcal{I}_1, \ldots, \mathcal{I}_M$ be collections of intervals of length (a multiple of) δ. Assume that the elements of $\mathcal{I}_i$ are pairwise disjoint intervals. For every ball B of radius δ^{-1}, we have

$$\left(\prod_{j=1}^{M}\left(\sum_{I\subset\mathcal{I}_j}\|E_I g\|^2_{L^2_\sharp(w_B)}\right)^{\frac{1}{2}}\right)^{\frac{1}{M}} \ll \left(\prod_{j=1}^{M}\left(\sum_{J\subset I,|J|=\delta}\|E_J g\|^2_{L^2_\sharp(w_B)}\right)^{\frac{1}{2}}\right)^{\frac{1}{M}}.$$

Note that this is a very strong form of decoupling since it works down to the scale δ. We skip the proof.
Claim 3.c *Lower dimensional decoupling.* This tools is derived from the assumption that the decoupling theorem is known for dimensions $2 \leq k < n$. From this one can derive the following useful result. *Let $n \geq 3$ be fixed. For every $0 < \delta \leq 1$ and every $3 \leq k \leq n$, for any interval $I \subset [0,1]$ of length (a multiple of) $\delta^{\frac{n}{k-1}}$, for every ball $B \in \mathbb{R}^n$ of radius δ^{-n} and every $2 \leq p \leq p_n$ we have*

$$\|E_I g\|_{L^p_\sharp(w_B)} \ll V_{p,k-1}(\delta^{\frac{n}{k-1}})\left(\sum_{J\subset I,|J|=\delta^{\frac{n}{k-1}}}\|E_J g\|^2_{L^p_\sharp(w_B)}\right)^{\frac{1}{2}}.$$

Somehow the idea here is that at an appropriate scale we can approximate parts of the moment curve by its lower dimensional versions.

Claim 3.c *The key ball inflation statement.* This seems to be the most sophisticated tool we are going to name here.

Fix $1 \le k \le n-1$, $2n \le p \le p_n$, and take $M = n!$. For any ball B^{k+1} of radius δ^{-k+1} and a cover $\mathcal{B}_k(B^{k+1})$ and for every g we have

$$\frac{1}{|\mathcal{B}_k(B^{k+1})|}\sum_{B^k\in\mathcal{B}_k(B^{k-1})}\left(\prod_{j=1}^{M}\left(\sum_{J_j\subset I_j,|J_j|=\delta}\|E_{J_j}g\|^2_{L_\sharp^{\frac{pk}{n}}(w_{B^k})}\right)^{\frac{1}{2}}\right)^{\frac{p}{M}}$$
$$\ll_{\epsilon,K}\delta^{-\epsilon}\left(\prod_{j=1}^{M}\left(\sum_{J_j\subset I_j,|J_j|=\delta}\|E_{J_j}g\|^2_{L_\sharp^{\frac{pk}{n}}(w_{B^{k+1}})}\right)^{\frac{1}{2}}\right)^{\frac{p}{M}},$$

with an implicit constant independent of δ, g and the ball B^{k+1}.

The point is that the size of the balls changes, while the intervals stay the same! The proof includes some intricate reduction steps which finally make it possible to apply some multilinear Kakey result, which is itself a deep theorem from the realm of (multilinear) restriction problems.

Step 4 (The iteration process). We will not say to much about this very complicated step. Note that the key inequality reads

$$A_p(u,B^n,u)\ll\delta^{-\epsilon}V_{p,n}(\delta)^{1-\sum_{j=0}^{r}\gamma_j}\cdot D_p(1,B^n)^{1-\sum_{j=0}^{r}\gamma_j}\cdot\prod_{j=0}^{r}A_p(b_ju,B^n,b_ju)^{\gamma_j},\tag{27}$$

for r big, u small, $p < p_n$ close to p_n and suitable number γ_j, b_j. The proof of this inequality uses several applications of the tools provided in Step 3. In particular $n-1$ instances of ball inflation. The constants γ_j, b_j arise through Hölder and interpolation arguments.

Now we iterate (27) and get

$$A_p(u,B^n,u)\ll\delta^{-\epsilon}V_{p,n}(\delta)^{1-\sum_{\mathbf{j}=0}^{r}\gamma_{\mathbf{j}}}\cdot D_p(1,B^n)^{1-\sum_{\mathbf{j}=0}^{r}\gamma_{\mathbf{j}}}\cdot\prod_{j_1=0}^{r}\cdots\prod_{j_L=0}^{r}A_p(\beta_{\mathbf{j}}u,B^n,\beta_{\mathbf{j}}u)^{\gamma_j},\tag{28}$$

for $\mathbf{j} = (j_1,\dots,j_L) \in [0,r]^L$, $\beta_{\mathbf{j}} = b_{j_1}\cdot\ldots\cdot b_{j_L}$ and $\gamma_{\mathbf{j}} = \gamma_{j_1}+\ldots+\gamma_{j_L}$.

We need one last estimate which is easily derived from Höelder and rescaling:

$$A_p(\beta u,B^n,\beta u)\ll V_{p,n}(\delta^{1-u\beta})D_p(1,B^n).$$

Inserting this in (28) we find

$$A_p(u,B^n,u)\ll\delta^{-\epsilon}\delta^{\eta_p(1-u\sum_{\mathbf{j}}\beta_{\mathbf{j}}\gamma_{\mathbf{j}})}D_p(1,B^n).$$

Here we win since we can make

$$\sum_{\mathbf{j}} \beta_{\mathbf{j}}\gamma_{\mathbf{j}} = \left(\sum_{j=0}^{r} b_j\gamma_j\right)^{L}$$

sufficiently large by first taking r large and then L! (Note that if we would not iterate (26), then we could not ensure that $\sum_j b_j\gamma_j$ is arbitrarily large.) □

A direct consequence of the decoupling theorem is the following discrete version, which is what we actually need.

Corollary 8.5. *Let $n \geq 2$ and $p \geq 2$ be fixed. For every $\epsilon > 0$ there is a constant $C_\epsilon = C(\epsilon, n, p)$ such that the following holds: For every $N \geq 1$, and for each choice of a fixed set of points $\{t_1, \dots, t_N\}$ with $t_i \in (\frac{i-1}{N}, \frac{i}{N}]$, for each ball B_R of radius $R \geq N^n$ in $\mathbb{R}^n$, and every set of coefficients $\{a_i\}_{1\leq i\leq N}$ with $a_i \in \mathbb{C}$ we have*

$$\left(\frac{1}{|B_R|}\int_{\mathbb{R}^n} |\sum_{i=1}^{N} a_i e(t_i x_1 + \dots t_i^n x_n)|^p w_{B_R}(x)dx_1 \dots dx_n\right)^{\frac{1}{p}}$$
$$\leq C_\epsilon N^\epsilon \left(1 + N^{\frac{1}{2}(1-\frac{n(n+1)}{p})}\right)\left(\sum_{i=1}^{N}|a_i|^2\right)^{\frac{1}{2}}.$$

We only indicate how this is proved from the general decoupling inequality.

Proof. First, note that the statement of Theorem 8.4 holds for any ball B of radius $R \gg \delta^{-n}$ (where $0 < \delta \leq 1$). Second,, typical reductions show that it is sufficient to deduce the critical case $p = n(n+1)$.

Formally we want to apply the general decoupling estimate to $g = \sum_{i=1}^{N} a_i\delta_{t=t_i}$. With this choice we of course have

$$E_{[0,1]}g(x_1, \dots, x_n) = \sum_{i=1}^{N} a_i e(t_i x_1 + \dots + t_i^n x_n).$$

Also for each interval $J = (\frac{i-1}{,}\frac{i}{N}]$ we get

$$E_J g(x_1, \dots, x_n) = a_i e(t_i x_1 + \dots + t_i^n x_n),$$

for the unique $t_i \in J$. The L^p-norm of this is trivially computed and we get

$$\|E_J g\|_{L^p(w_{B_R})} = \left(\int_{\mathbb{R}^n} |a_i|^p w_{B_R}(x)dx\right)^{\frac{1}{p}} \ll_p |B_R|^{\frac{1}{p}}|a_i|.$$

A direct application of Theorem 8.4 with $\delta = \frac{1}{N}$ (for larger radii) in this setting yields the desired inequality. To make this rigorous we simply have to choose suitable approximations for the δ-function in our choice of g. □

Along these lines one can prove decoupling estimates for other convenient curves replacing the moment curve. We mention the following example, which we need later to improve Hua's inequality. The curve in question is

$$\Gamma = \{(t^k, t^{s-1}, \ldots, t) \colon 1 \leq t \leq 2\}.$$

We restrict ourselves to state the discrete version of the corresponding decoupling inequality.

Theorem 8.6. *The following discrete decoupling inequality holds:*

$$N^{-s^2} \int_{[-N,N]^s} |\sum_{n=N}^{2N} e\left((\frac{n}{N})^k x_0 + (\frac{n}{N})^{s-1} x_1 + \ldots + \frac{n}{N} x_{s-1}\right)|^{s(s+1)} d\mathbf{x} \ll N^{\frac{1}{2}s(s+1)+\epsilon},$$

for $\epsilon > 0$ *and* $N \geq 1$.

8.2. **Vinogradov's mean value Theorem.** We consider

$$f_k(x, N) = \sum_{1 \leq n \leq N} e(nx_1 + n^2 x_2 + \ldots + n^k x_k).$$

In this section we are interested in the mean value

$$J_{s,k}(N) = \int_{[0,1]^k} |f_k(x, N)|^{2s} dx_1 \ldots dx_k.$$

Reversing the circle method approach (i.e. orthogonality of characters) we can interpret this as

$$J_{s,k}(N) = \sharp\{\mathbf{x} \in ([0, N] \cap \mathbb{N})^{2s} \colon x_1^j + \ldots + x_s^j = x_{s+1}^j + \ldots + x_{2s}^j \text{ for all } 1 \leq j \leq k\}.$$

Note that this system of equation is translation invariant.

The bound

$$J_{s,k}(N) \ll_\epsilon N^\epsilon (N^s + N^{2s - \frac{1}{2}k(k+1)}) \text{ for all } \epsilon > 0 \tag{29}$$

is the *Main Conjecture* and is now a theorem due to Bourgain-Demeter-Guth (and Wooley). The *critical exponent* is $s = \frac{1}{2}k(k+1)$.

Note that we have the trivial lower bound $J_{s,k}(N) \gg N^s$ coming from the solutions $x_1, \ldots, x_s \in [1, N]$ and $x_i = x_{i+s}$. We can give an alternative lower bound as follows. Observe that for $1 \leq x_i \leq N$ we have

$$|(x_1^j - x_{s+1}^j) + \ldots + (x_s^j - x_{2s}^j)| \leq sX^j \text{ for } 1 \leq j \leq k.$$

By arranging all $\asymp N^{2s}$ possible tuples $(x_1, \ldots, x_{2s}) \in [0, N]^s$ according to the values estimated in the previous equation we get

$$\begin{aligned} N^{2s} &\ll \sum_{|h_1| \leq sN} \cdots \sum_{|h_k| \leq sN^k} \int_{(0,1]^k} |f_k(\alpha; N)|^{2s} e(\alpha \cdot \mathbf{h}) d\alpha \\ &\ll X \cdot \ldots \cdot X^k J_{s,k}(N) = N^{\frac{1}{2}k(k+1)} J_{s,k}(N). \end{aligned}$$

Thus we have $N^{2s-\frac{1}{2}k(k+1)} \ll J_{s,k}(N)$. In particular the upper bound in the main conjecture is sharp up to N^ϵ.

Theorem 8.7 (Bourgain-Demeter-Guth, Wooley)**.** *For all $s, k \geq 1$ we have*

$$J_{s,k}(N) \ll N^\epsilon (N^s + N^{2s-\frac{1}{2}k(k+1)}),$$

for every $N \geq 1$ and all $\epsilon > 0$.

Note that the cases $k = 1, 2$ are relatively easy. For general k and $s \leq k$ it can be shown that $J_{s,k}(N) = s!N^s + O(N^{s-1})$, so that we essentially only have the diagonal contribution and its permutations. There has been a lot of progress through the years showing that the estimate holds for $s \geq s_0(k)$. It is relatively easy to see that once the conjecture is resolved for the critical exponent $s = \frac{1}{2}k(k+1)$, then it holds for all s. For $k = 3$ the critical case and thus the full conjecture was first resolved by Wooley using his very successful efficient congruencing method. This method has been extended to handle all k. However, here we will only indicate how this theorem is derived from the decoupling result discussed above.

Proof. First we reduce to the critical exponent $s_c = \frac{1}{2}k(k+1)$. To do so assume we know the theorem for s_c. Suppose $s > s_c$. Then we trivially have

$$J_{s,k}(N) \leq \sup_\alpha |f_k(\alpha; N)|^{2s-2s_c} \int_{(0,1]^k} |f_k(\alpha; N)|^{2s_c} d\alpha.$$

Applying the theorem for s_c and the trivial estimate $|f_k(\alpha, N)| \leq N$ yields the result. The opposite situation is $s < s_c$. Here we apply Hölder with $q = \frac{s_c}{s}$. This yields

$$J_{s,k}(N) \leq \left(\int_{(0,1]^k} 1 \right)^{\frac{1}{p}} (J_{s_c,k}(N))^{\frac{1}{q}} \ll \left(N^{s_c+\epsilon} \right)^{\frac{1}{q}} \ll N^{s+\epsilon}.$$

Here we used again the theorem in the critical case.

Now the case $n = 1$ is trivial. Here we have $s_c = 1$. Thus we are counting $x_1, x_2 \in [0, N]$ with $x_1 - x_2 = 0$. Therefore $J_{1,1}(N) \ll N$ and we are done with this case.

We are left with $n \geq 2$ and $s = s_c$ critical. To treat this situation we start with some test function yoga. We choose a Schwartz function $\phi \in \mathcal{S}(\mathbb{R}^k)$ with $\phi(x), \hat{\phi}(\xi) \geq 0$ and $\hat{\phi}(\xi) \geq 1$ for $|\xi| \leq 1$. We rescale this as follows:

$$\phi_M(x) = \phi(\frac{x}{M}) \text{ so that } \hat{\phi_M}(x) = M^k \phi(Mx).$$

By varying the radius $R \asymp N^k$ we can majorise $\phi_{N^k}(x)$ by w_{B_R}.

We apply Corollary 8.5 with $p = 2s$, $n = k$ and $t_i = \frac{i}{N}$ to see

$$N^{-k^2} \int_{\mathbb{R}^k} |\sum_{i=1}^N e(\frac{i}{N}x_1 + \ldots + \frac{i^k}{N^k}x_k)|^{2s} \phi_{N^k}(x_1, \ldots, x_k) dx_1 \ldots dx_k \ll N^{s+\epsilon}.$$

Changing variables and inserting the definition of ϕ_{N^k} yields

$$N^{\frac{1}{2}k(k+1)-k^2}\int_{\mathbb{R}^k}|\sum_{i=1}^{N} e(ix_1+\ldots+i^kx_k)|^{2s}\phi(\frac{x_1}{N^{k-1}},\ldots,x_k)dx_1\ldots dx_k \ll N^{s+\epsilon}.$$

Now we open up the $2s$-moment to get

$$N^{\frac{1}{2}k(k+1)-k^2}\sum_{i_1,\ldots,i_{2s}}\int_{\mathbb{R}^k} e(\theta_i(\mathbf{i})x_1+\ldots+\theta_k(\mathbf{i})x_k)\phi(\frac{x_1}{N^{k-1}},\ldots,x_k)dx_1\ldots dx_k \ll N^{s+\epsilon},$$

where

$$\theta_j(\mathbf{i}) = i_1^j+\ldots+i_s^j-i_{s+1}^j-\ldots-i_{2s}^j \text{ for } 1\leq j\leq k.$$

We now recognise that the remaining integral is nothing but the Fourier transform. Indeed

$$\int_{\mathbb{R}^k} e(\theta_i(\mathbf{i})x_1+\ldots+\theta_k(\mathbf{i})x_k)\phi(\frac{x_1}{N^{k-1}},\ldots,x_k)dx_1\ldots dx_k = N^{\frac{1}{2}(k-1)k}\cdot\hat{\phi}(N^{k-1}\theta_1(\mathbf{i}),\ldots,\theta_k(\mathbf{i})).$$

Inserting this above yields

$$\sum_{i_1,\ldots,i_{2s}}\hat{\phi}(N^{k-1}\theta_1(\mathbf{i}),\ldots,\theta_k(\mathbf{i})) \ll N^{s+\epsilon}.$$

By our choice of ϕ we get

$$J_{s,k}(N) = \sum_{\substack{i_1,\ldots,i_{2s},\\ |\theta_j(\mathbf{i})|<N^{j-k}}} 1 \leq \sum_{i_1,\ldots,i_{2s}}\hat{\phi}(N^{k-1}\theta_1(\mathbf{i}),\ldots,\theta_k(\mathbf{i})) \ll N^{s+\epsilon}.$$

Indeed, since $\theta_j(\mathbf{i})$ is an integer and $N^{j-k}\leq 1$ for all $1\leq j\leq k$ we must have $\theta_j(\mathbf{i})=0$ for all j. □

8.3. **Applications to Exponential sums.** The results in the previous sections can be used to strengthen certain estimates of exponential sums. We start by looking at a strengthening of Weyl's bound in a setting critical to Waring's problem.

Theorem 8.8. *Suppose $k\geq 3$. For $2\leq j\leq k$ assume*

$$|x_j-\frac{a}{q}|\leq\frac{1}{q^2} \textit{ with } (a,q)=1.$$

We have the estimate

$$f_k(x,N) \ll N^{1+\epsilon}(q^{-1}+N^{-1}+qN^{-j})^{\frac{1}{K}} \textit{ with } K=k(k-1).$$

Proof. Let $c_1,\ldots,c_N\in\mathbb{C}$ and put $S=\sum_{n=1}^N c_n$. We obtain

$$S^b = \sum_{1\leq n_1,\ldots,n_b\leq N} c_{n_1}c_{n_2}\cdots c_{n_b}.$$

We will write $\mathbf{n}$ for a b-tuple of integers lying in $[1,N]^b$. Further put

$$s_j(\mathbf{n}) = n_1^j+\ldots+n_b^j.$$

Thus given $\mathbf{n}$ we can associate

$$\mathbf{s} = (s_1(\mathbf{n}), \ldots, s_{k-2}(\mathbf{n})).$$

Note that $\mathbf{s} \in \mathcal{S}$ for

$$\mathcal{S} = \mathbb{N}^{k-2} \cap [1, bN] \times \ldots \times [1, bN^{k-2}].$$

Turning tables we put

$$\mathcal{N}(\mathbf{s}) = \{\mathbf{n} \colon s_j(\mathbf{n}) = s_j \text{ for } 1 \leq j \leq k-2\}$$

for $\mathbf{s} = (s_j)_{1 \leq j \leq k-2} \in \mathcal{S}$.

With this notation set up we get

$$S^b = \sum_{\mathbf{s} \in \mathcal{S}} \sum_{\mathbf{n} \in \mathcal{N}(\mathbf{s})} c_{n_1} \cdots c_{n_b}.$$

Note that $\sharp\mathcal{S} = b^{k-2} N^{(k-1)(k-2)/2}$. Applying Cauchy-Schwarz we get

$$\begin{aligned} |S|^{2b} &\leq b^{k-2} N^{(k-1)(k-2)/2} \sum_{\mathbf{s} \in \mathcal{S}} | \sum_{\mathbf{n} \in \mathcal{N}(\mathbf{s})} c_{n_1} \cdots c_{n_b}|^2 \\ &= b^{k-2} N^{(k-1)(k-2)/2} \sum_{\substack{\mathbf{n}, \mathbf{m}, \\ s_j(\mathbf{n}) = s_j(\mathbf{m})}} c_{m_1} \cdots c_{m_b} \overline{c_{n_1}} \cdots \overline{c_{n_b}}. \end{aligned}$$

We now specialise to $c_n = e(P(n))$ for $P(x) = \sum_{j=1}^k \alpha_j x^j$. We obtain

$$c_{m_1} \cdots c_{m_b} \overline{c_{n_1}} \cdots \overline{c_{n_b}} = e((s_k(\mathbf{m}) - s_k(\mathbf{n}))\alpha_k + (s_{k-1}(\mathbf{m}) - s_{k-1}(\mathbf{n}))\alpha_{k-1})$$

if $\mathbf{n}, \mathbf{m}$ satisfy the constraints of the sum.

We put $m = m_1$ and $m_i = m + u_i$. Further write $n_i = m + v_i$. Then binomial expansion implies

$$\begin{aligned} s_j(\mathbf{u}) &= \sum_{i=1}^b (m_i - m)^j = \sum_{r=0}^j \binom{j}{r} s_r(\mathbf{m})(-m)^{j-r} \text{ and} \\ s_j(\mathbf{v}) &= \sum_{i=1}^b (n_i - m)^j = \sum_{r=0}^j \binom{j}{r} s_r(\mathbf{n})(-m)^{j-r}. \end{aligned}$$

In particular, $s_j(\mathbf{m}) = s_j(\mathbf{n})$ for $1 \leq j \leq k-2$ implies that $s_j(\mathbf{u}) = s_j(\mathbf{v})$ for the same indices. Further we get

$$s_{k-1}(\mathbf{u}) - s_{k-1}(\mathbf{v}) = s_{k-1}(\mathbf{m}) - s_{k-1}(\mathbf{n}).$$

Similarly we see

$$s_k(\mathbf{u}) - s_k(\mathbf{v}) = s_k(\mathbf{m}) - s_k(\mathbf{n}) - km(s_{k-1}(\mathbf{m}) - s_{k-1}(\mathbf{n})).$$

To save chalk we put $d_j = d_j(\mathbf{u}, \mathbf{v}) = s_j(\mathbf{u}) - s_j(\mathbf{n})$. These considerations lead us to

$$|S|^{2b} \leq b^{k-2} N^{(k-1)(k-2)/2} \sum_{\substack{\mathbf{u},\mathbf{v} \\ d_j=0 \\ j=1,\ldots,k-2}} e(d_k\alpha_k + d_{k-1}\alpha_k - 1) \sum_m e(kd_{k-1}m\alpha_k).$$

Let us have a closer look at the ranges of the summation parameters. First, u_i and v_i lie in $\{-N, \ldots, N\}$. Further, m must satisfy

$$\begin{aligned} &1 \leq m \leq N, \\ &1 - u_i \leq m \leq N - u_i \text{ for } 2 \leq i \leq b, \\ &1 - v_i \leq m \leq N - v_i \text{ for } 1 \leq i \leq b. \end{aligned}$$

So the inner sum is linear and runs over some interval of length $\leq N$. Thus we have the familiar estimate

$$\sum_m e(kd_{k-1}m\alpha_k) \ll \min(N, \|kd_{k-1}\alpha_k\|^{-1}).$$

Let $R_1(h)$ denote the number of solutions to the system of equations

$$\begin{aligned} &u_2^j + \ldots + u_b^j = v_1^j + \ldots + v_b^j \text{ for } 1 \leq j \leq k-2, \\ &u_2^{k-1} + \ldots + u_b^{k-1} = h + v_1^{k-1} + \ldots + v_b^{k-1} \end{aligned}$$

with $|ui|, |v_i| \leq N$. We can estimate

$$\sum_{\substack{\mathbf{u},\mathbf{v} \\ d_j=0 \\ j=1,\ldots,k-2}} e(d_k\alpha_k + d_{k-1}\alpha_k - 1) \sum_m e(kd_{k-1}m\alpha_k) \ll \sum_h R_1(h) \min(N, \|kh\alpha_k\|^{-1}).$$

Next we shift the variables ($m_i = m + u_i$ and $n_i + m + v_i$ for $N + 1 \leq m \leq 2N$) in the equations underlying $R_1(h)$. This way we find that $R_1(h)$ counts solutions to

$$\begin{aligned} &s_j(\mathbf{m}) = s_j(\mathbf{n}) \text{ for } 1 \leq j \leq k-2, \\ &s_{k-1}(\mathbf{m}) = h + s_{k_j}(\mathbf{n}) \end{aligned}$$

Within the constraints

$m_1 = m$, $m - N \leq m_i \leq m + N$ for $2 \leq i \leq b$ and $m - N \leq n_i \leq m + N$ for $1 \leq i \leq b$.

We relax these constraints to $1 \leq m_i, n_i \leq 3N$ for $1 \leq i \leq b$ and write $R_2(h)$ for the new count of solutions. Note that since we allow m_1 to vary we have $R_2(h) \geq NR_1(h)$.

By character orthogonality we can write

$$R_2(h) = \int_{[0,1]^{k-1}} |\sum_{n=1}^{3N} e(P(n, \boldsymbol{\alpha}))|^{2b} e(-h\alpha_{k-1}) d\boldsymbol{\alpha},$$

for

$$P(x, \boldsymbol{\alpha}) = \sum_{j=1}^{k-1} \alpha_j x^j.$$

Estimating the integral trivially yields $F_2(h) \leq R_2(0)$. We recognise $R_2(0) = J_{k-1}(3N; b)$ as the Vinogradov mean value.

We obtain the estimate

$$\sum_h R_1(h) \min(N, \|kh\alpha_k\|^{-1}) \ll N^{-1} J_{k-1}(3N; b) \left(N + \sum_{h=1}^{2bkN^{k-1}} \min(N, \|h\alpha_k\|^{-1}) \right)$$
$$\ll bkN^{k-1} J_{k-1}(3N; b) \left(\frac{1}{q} + \frac{\log(q)}{N} + \frac{q\log(q)}{N^k} \right).$$

In the last step the h-sum was estimated as in the proof of Weyl's inequality.

Backtracking everything we have done we see that we have obtained the inequality

$$\sum_{n=1}^N e(P(n)) \ll (b^k k)^{\frac{1}{2b}} N \left(\frac{J_{k-1}(3N; 2b)}{N^{2b-k(k-1)/2}} \right)^{\frac{1}{2b}} \left(\frac{1}{q} + \frac{\log(q)}{N} + \frac{q\log(q)}{N^k} \right)^{\frac{1}{2b}}.$$

We now take $b = k(k-1)/2$ and estimate $J_{k-1}(3N, 2b)$ using the mean value estimate. This yields the claim. □

Theorem 8.9. *For $0 < s \leq k$ we have*

$$\int_0^1 |T(\alpha)|^{s(s+1)} d\alpha \ll N^{s^2+\epsilon}.$$

Recall that $T(\alpha) = \sum_{n=1}^N e(\alpha n^k)$.

Proof. We deduce this from Theorem 8.6. First, we rescale the discrete decoupling inequality and use periodicity to get

$$\int_{[-1,1]} \int_{[0,1]^{s-1}} |\sum_{n=N}^{2N} e\left(\frac{n^k}{N^{k-s}} x + n^{s-1} x_{s-1} + \ldots + n x_1 \right)|^{s(s+1)} dx_1 \ldots dx_{s-1} dx \ll N^{\frac{1}{2}s(s+1)+\epsilon}.$$

We now write $K_r = K_r(t)$ for the kernel on $\mathbb{R}/\mathbb{Z}$ whose Fourier transform $\hat{K}_r$ is trapezoidal and satisfies $\hat{K}_r(n) = 1$ for $|n| \leq r$ and $\operatorname{supp} \hat{K}_r \subset [-2r, 2r]$. Note that

$$K(\mathbf{x}) = K_{2N}(x_1) \cdot \ldots \cdot K_{2N^{s-1}}(x_{s-1}) \ll N^{\frac{1}{2}s(s-1)}.$$

Similarly to the derivation of Vinogradov's mean value theorem we multiply the integrand by $K(\mathbf{x})$, open the $s(s+1)$-power and realise the Fourier transform. This yields

$$\int_{[-1,1]} |\sum_{n=N}^{2N} e(\frac{n^k}{N^{k-s}} x)|^{s(s+1)} dx \ll N^{s^2+\epsilon}.$$

We continue from here using another test function. Pick $0 \leq \varphi \leq 1$ with $\operatorname{supp}\varphi \subset [-1,1]$ and $\hat{\varphi}(0) = 1$, $\hat{\varphi} \geq 0$. Then a by now familiar argument shows

$$\sum_{N \leq n_1,\ldots,n_{s(s+1)} \leq 2N} \mathbb{1}_{n_1^k+\ldots-n_{s(s+1)}^k} \leq \sum_{N \leq n_1,\ldots,n_{s(s+1)} \leq 2N} \hat{\varphi}(N^{s-k}(n_1^k+\ldots-n_{s(s+1)}^k)) \ll N^{s^2+\epsilon}.$$

This implies

$$\int_0^1 |\sum_{n=N}^{2N} e(n^k x)|^{s(s+1)} dx \ll N^{s^2+\epsilon}$$

and the assertion follows by a typical dyadic dissection. □

9. Refined analysis of the asymptotic formula for Waring's problem

Using the improved versions of Weyl's estimate and Hua's inequality given in the last section we now revisit the circle method. This will lead us to highly non-trivial estimates for the number $G(n)$ as n gets large. We will prove the following theorem.

Theorem 9.1 (Bourgain). *We have*

$$G(k) \leq k^2 + 1 - \max_{s \leq k} \left\lceil s\frac{k-s-1}{k-s+1} \right\rceil.$$

For large k this implies $G(k) < k^2 - k + O(\sqrt{k})$. To the best of my knowledge this is still the record for large k. Note that for small k there are different individual records as we can see from the case $k = 3$ considered earlier. The following argument is essentially taken from [Wo].

9.1. **The minor arc estimate.** We introduce some notation (some of which should be familiar):

$$f_k(\mathbf{x}; P) = \sum_{1 \leq n \leq P} e(x_1 n + \ldots + x_k n^k),$$

$$F_k(\mathbf{y}, \theta; P) = \sum_{1 \leq n \leq P} e(y_1 n + \ldots + y_{k-2} n^{k-2} + \theta n^k),$$

$$g_k(\alpha; P) = \sum_{1 \leq n \leq P} e(\alpha n^k) \text{ and } \sigma_{s,j}(\mathbf{x}) = \sum_{i=1}^{s} (x_i^j - x_{s+i}^j) \text{ for } 1 \leq j \leq k.$$

The set of minor arcs $\mathfrak{m}$ to be the set of $\alpha \in [0,1]$ such that whenever

$$|\alpha - \frac{a}{q}| \leq (2kq)^{-1} P^{1-k} \text{ and } (a,q) = 1 \tag{30}$$

then $q > (2k)^{-1}P$. (Note that here we are taking smaller minor arcs than in our original circle method application to Waring's problem.)

Proposition 9.2. *One has*

$$\int_{\mathfrak{m}} |g_k(\alpha; P)|^{2s} d\alpha \ll P^{\frac{1}{2}k(k-1)-1} \log(P)^{2s+1} J_{s,k}(2P).$$

Proof. First we take $\mathbf{h} \in \mathbb{Z}^{k-2}$ and observe that

$$\int_{\mathfrak{m}} \int_{[0,1)^{k-2}} |F_k(\boldsymbol{\beta}, \theta; P)|^{2s} e(-\beta_1 h_1 - \ldots - \beta_{k-2} h_{k-2}) d\boldsymbol{\beta} d\theta = \sum_{1 \le \mathbf{x} \le P} \delta(\mathbf{x}, \mathbf{h}) \int_{\mathfrak{m}} e(\theta \sigma_{s,k}(\mathbf{x})) d\theta,$$

for

$$\delta(\mathbf{x}, \mathbf{h}) = \prod_{j=1}^{k-2} \left(\int_0^1 e(\beta_j (\sigma_{s,j}(\mathbf{x}) - h_j)) d\beta_j \right).$$

Character orthogonality tells us that

$$\int_0^1 e(\beta_j (\sigma_{s,j}(\mathbf{x}) - h_j)) d\beta_j = \delta_{\sigma_{s,j}(\mathbf{x}) = h_j}.$$

This justifies the notation $\delta(\mathbf{x}, \mathbf{h})$. Since $|\sigma_{s,j}(\mathbf{x})| \le sP^j$ we have

$$\sum_{|h_1| \le sP} \cdots \sum_{|h_{k-2}| \le sP^{k-2}} \delta(\mathbf{x}, \mathbf{h}) = 1.$$

Also note that

$$\sum_{1 \le \mathbf{x} \le P} e(\theta \sigma_{s,k}(\mathbf{x})) = |g_k(\theta)|^{2s}.$$

With this at hand we can deduce

$$\sum_{|h_1| \le sP} \cdots \sum_{|h_{k-2}| \le sP^{k-2}} \int_{\mathfrak{m}} \int_{[0,1)^{k-2}} |F_k(\boldsymbol{\beta}, \theta; P)| e(-\beta_1 h_1 - \ldots - \beta_{k-2} h_{k-2}) d\boldsymbol{\beta} d\theta$$

$$= \int_{\mathfrak{m}} \sum_{1 \le \mathbf{x} \le P} \left(\sum_{\mathbf{h}} \delta(\mathbf{x}, \mathbf{h}) \right) e(\theta \sigma_{s,k}(\mathbf{x})) d\theta = \int_{\mathfrak{m}} |g_k(\theta)|^{2s} d\theta.$$

By estimating the left hand side of this inequality trivially we get

$$\int_{\mathfrak{m}} |g_k(\theta)|^{2s} d\theta \ll P^{\frac{1}{2}(k-1)(k-2)} \int_{\mathfrak{m}} \int_{[0,1)^{k-2}} |F_k(\boldsymbol{\beta}, \theta; P)|^{2s} d\boldsymbol{\beta} d\theta. \tag{31}$$

Character orthogonality applied essentially as above yields

$$\int_{\mathfrak{m}} \int_{[0,1)^{k-2}} |F_k(\boldsymbol{\beta}, \theta; P)|^{2s} d\boldsymbol{\beta} d\theta = \sum_{|h| \le sP^{k-1}} \int_{\mathfrak{m}} \int_{[0,1)^{k-1}} |f(\boldsymbol{\alpha}, \theta)|^{2s} e(-\alpha_{k-1} h) d\boldsymbol{\alpha} d\theta, \tag{32}$$

with $f(\boldsymbol{\alpha}, \theta) = f_k((\alpha_1, \ldots, \alpha_{k-1}, \theta); P)$. Next we write

$$f(\boldsymbol{\alpha}) = \sum_{1+y \le x \le P+y} e(\psi(x-y, \boldsymbol{\alpha})) \text{ with } \psi(z, \boldsymbol{\alpha}) = \alpha_1 z + \ldots + \alpha_k z^k.$$

Using the binomial theorem we can write

$$\psi(x-y;\boldsymbol{\alpha}) = \sum_{i=0}^{k} \tilde{\alpha}_i x^i \text{ where } \tilde{\alpha}_i = \sum_{j=i}^{k} \binom{j}{i} (-y)^{j-i}\alpha_j.$$

Now we define

$$K_y(\gamma) = \sum_{1+y\le z\le P+y} e(-\gamma z) \text{ and } f_y(\boldsymbol{\alpha},\gamma) = \sum_{1\le x\le 2P} e(\psi(x-y;\boldsymbol{\alpha})+\gamma x).$$

This definitions allow us to write

$$f(\boldsymbol{\alpha}) = \int_0^1 f_y(\boldsymbol{\alpha};\gamma)K_y(\gamma)d\gamma. \tag{33}$$

Next define

$$\mathcal{F}_y(\boldsymbol{\alpha},\theta,\boldsymbol{\gamma}) = \prod_{i=1}^{s} f_y(\boldsymbol{\alpha},\theta;\gamma_i) f_y(-\boldsymbol{\alpha},-\theta;-\gamma_{s+i}),$$

$$I_h(\gamma,y) = \int_{\mathfrak{m}}\int_{[0,1)^{k-1}} \mathcal{F}_y(\boldsymbol{\alpha},\theta,\boldsymbol{\gamma})e(-\alpha_{k-1}h)d\boldsymbol{\alpha}d\theta \text{ and}$$

$$\tilde{K}(\boldsymbol{\gamma}) = \prod_{i=1}^{s} K_y(\gamma_i)K_y(-\gamma_{s+i}).$$

This is set up so that by inserting (33) in (32) we get

$$\int_{\mathfrak{m}}\int_{[0,1)^{k-2}} |F_k(\boldsymbol{\beta},\theta;P)|^{2s} d\boldsymbol{\beta}d\theta = \sum_{|h|\le sX^{k-1}} \int_{[0,1)^{2s}} I_h(\boldsymbol{\gamma},y)\tilde{K}(\boldsymbol{\gamma})d\boldsymbol{\gamma}.$$

We will use character orthogonality to treat the $\boldsymbol{\alpha}$-integral in the definition of I_h. We set $\Delta(\theta,\boldsymbol{\gamma},h,y;\mathbf{x})$ to be

$$e(\theta\sigma_{s,k}(\mathbf{x}-y) + \gamma_1x_1 + \ldots + \gamma_sx_s - \gamma_{s+1}x_{s+1} - \ldots - \gamma_{2s}x_{2s})$$

when

$$\sum_{i=1}^{s}((x_i-y)^j - (x_{s+i}-y)^j) = 0 \text{ for } 1 \le j \le k-2 \text{ and}$$

$$\sum_{i=1}^{s}((x_i-y)^{k-1} - (x_{s+i}-y)^{k-1}) = h \tag{34}$$

and zero otherwise. This yields

$$\int_{[0,1)^{k-1}} \mathcal{F}_y(\boldsymbol{\alpha},\theta,\boldsymbol{\gamma})e(-\alpha_{k-1}h)d\boldsymbol{\alpha} = \sum_{1\le\mathbf{x}\le 2P} \Delta(\theta,\boldsymbol{\gamma},h,y;\mathbf{x}).$$

Using the binomial theorem we observe that if $\mathbf{x}$ satisfies (34), then

$$\sum_{i=1}^{s}(x_i^j - x_{s+1}^j) = 0 \text{ for } 1 \leq j \leq k-2 \text{ and } \sum_{i=1}^{s}(x_i^{k-1} - x_{s+i}^{k-1}) = h.$$

This directly implies

$$\sigma_{s,k}(\mathbf{x} - y) = \sum_{i=1}^{s}((x_i - y)^k - (x_{s+i} - y)^k) = \sigma_{s,k}(\mathbf{x}) - khy.$$

Reversing character orthogonality yields

$$\int_{[0,1)^{k-1}} \mathcal{F}_y(\boldsymbol{\alpha}, \theta, \boldsymbol{\gamma}) e(-\alpha_{k-1}h) d\boldsymbol{\alpha} = \int_{[0,1)^{k-1}} \mathcal{F}_0(\boldsymbol{\alpha}, \theta, \boldsymbol{\gamma}) e(-khy\theta - h\alpha_{k-1}) d\boldsymbol{\alpha}.$$

We use this to obtain

$$\sum_{|h| \leq sP^{k-1}} I_h(\boldsymbol{\gamma}, y) = \int_{\mathfrak{m}} \int_{[0,1)^{k-1}} \mathcal{F}_0(\boldsymbol{\alpha}, \theta, \boldsymbol{\gamma}) \sum_{|h| \leq sP^{k-1}} e(-khy\theta - h\alpha_{k-1}) d\boldsymbol{\alpha} d\theta$$

The linear sum can be estimated by

$$\sum_{|h| \leq sP^{k-1}} e(-khy\theta - h\alpha_{k-1}) \ll \min(P^{k-1}, \|ky\theta + \alpha_{k+1}\|^{-1})$$

We set

$$\psi(\theta, \alpha_{k-1}) = P^{-1} \sum_{1 \leq y \leq P} \min(P^{k-1}, \|ky\theta + \alpha_{k+1}\|^{-1})$$

and deduce the estimate

$$P^{-1} \sum_{1 \leq y \leq P} \sum_{|h| \leq sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \ll \int_{\mathfrak{m}} \int_{[0,1)^{k-1}} |\mathcal{F}_0(\boldsymbol{\alpha}, \theta, \boldsymbol{\gamma})| \psi(\theta, \alpha_{k-1}) d\boldsymbol{\alpha} d\theta.$$

We have seen how to treat functions such as $\psi(\theta, \alpha_{k-1})$ several times. Suppose $\theta \in \mathbb{R}$ is such that $|\theta - \frac{b}{r}| \leq r^{-2}$ with $b, r \in \mathbb{Z}$ and $(b, r) = 1$. Then we have

$$\psi(\theta, \alpha_{k-1}) \ll P^{k-1}(P^{-1} + r^{-1} + rP^{-k}) \log(2r).$$

By Dirichlet approximation we can assume that $r \leq 2kP^{k-1}$. But since $\theta \in \mathfrak{m}$ we have $r > (2k)^{-1}P$ by default. Thus we get

$$\sup_{\theta \in \mathfrak{m}} \psi(\theta, \alpha_{k-1}) \ll P^{k-2} \log(P).$$

With this at hand we can estimate

$$P^{-1} \sum_{1 \leq y \leq P} \sum_{|h| \leq sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \ll P^{k-2} \log(P) \int_{\mathfrak{m}} \int_{[0,1)^{k-1}} |\mathcal{F}_0(\boldsymbol{\alpha}, \theta, \boldsymbol{\gamma})| d\boldsymbol{\alpha} d\theta.$$

Inserting the definition of $\mathcal{F}_0$ and applying Hölder we obtain

$$\begin{aligned}P^{-1}\sum_{1\le y\le P}\sum_{|h|\le sP^{k-1}} I_h(\boldsymbol{\gamma},y) &\ll P^{k-2}\log(P)\prod_{i=1}^{2s}\left(\int_{\mathfrak{m}}\int_{[0,1)^{k-1}}|f_0(\boldsymbol{\alpha},\theta;\gamma_i)|^{2s}d\boldsymbol{\alpha}d\theta\right)^{\frac{1}{2s}}\\ &\le P^{k-2}\log(P)\sup_{\gamma\in[0,1]}\int_0^1\int_{[0,1)^{k-1}}|f_0(\boldsymbol{\alpha},\theta;\gamma)|^{2s}d\boldsymbol{\alpha}d\theta\\ &= P^{k-2}\log(P)\int_{[0,1)^k}|f_k(\boldsymbol{\alpha},2P)|^{2s}d\boldsymbol{\alpha}.\end{aligned}$$

The remaining integral is nothing but the Vinogradov mean value.

Inserting this estimate above yields

$$\begin{aligned}\int_{\mathfrak{m}}\int_{[0,1)^{k-1}}|F(\boldsymbol{\beta},\theta)|^{2s}d\boldsymbol{\beta}d\theta &= P^{-1}\sum_{1\le y\le P}\int_{\mathfrak{m}}\int_{[0,1)^{k-1}}|F(\boldsymbol{\beta},\theta)|^{2s}d\boldsymbol{\beta}d\theta\\ &= \int_{[0,1)^{2s}}\left(P^{-1}\sum_{1\le y\le P}\sum_{|h|\le sP^{k-1}}I_h(\boldsymbol{\gamma},y)\right)\tilde{K}(\boldsymbol{\gamma})d\boldsymbol{\gamma}\\ &\ll P^{k-2}\log(P)J_{s,k}(2P)\int_{[0,1)^{2s}}|\tilde{K}(\boldsymbol{\gamma})|d\boldsymbol{\gamma}.\end{aligned}$$

Since $\tilde{K}$ factors in a product of linear sums $K_y(\gamma)$. Each of these can be estimated by

$$\int_0^1|K_y(\gamma)|d\gamma \le \int_0^1\min(P,\|\gamma\|^{-1})d\gamma \ll \log(P).$$

This implies

$$\int_{[0,1)^{2s}}|\tilde{K}(\boldsymbol{\gamma})|d\boldsymbol{\gamma} \ll \log(P)^{2s},$$

which in turn gives the bound

$$\int_{\mathfrak{m}}\int_{[0,1)^{k-1}}|F_k(\boldsymbol{\beta},\theta;P)|^{2s}d\boldsymbol{\beta}d\theta \ll P^{k-2}\log(P)^{2s+1}J_{s,k}(2P).$$

We conclude the proof by inserting this in (31). □

The proof of the above estimate goes back to Wooley, but injecting the full strength of the main conjecture yields the following useful estimate.

Corollary 9.3. *We have*

$$\int_{\mathfrak{m}}|g_k(\alpha;P)|^{k(k+1)}d\alpha \ll P^{k^2-1+\epsilon}.$$

Proof. We simply apply the above proposition with $s=\frac{1}{2}k(k+1)$ and use (29) to estimate the mean value. □

This suffices to give a good estimate for the minor arcs.

Theorem 9.4. *Let $s < k$ and*

$$k^2 - s\frac{k-s-1}{k+1-s} < s_0 \le k(k+1).$$

Then we have

$$\int_{\mathfrak{m}} |T(\alpha)|^{s_0} d\alpha \ll N^{s_0-k-\delta'}.$$

This improves considerably the number of variables for which we can control the minor arcs for Waring's problem.

Proof. We start with a general consideration involving the Hölder inequality (sometimes called interpolation). Suppose $0 < a \le s_0 \le b$. And assume that we have good bounds

$$\|f\|_a^a \le A \text{ and } \|f\|_b^b \le B.$$

We now write $s_0 = \theta s_0 + (1-\theta)s_0$ for $\theta \in [0,1]$. We want to choose p and θ such that $p\theta s_0 = b$ and $q(1-\theta)s_0 = a$ (where $q = p/(p-1)$ as usual). This is satisfied with the choice

$$p = \frac{b-a}{s_0-a} \text{ and } \theta = \frac{b(s_0-a)}{s_0(b-a)}.$$

Applying Hölder and inserting out bounds thus yields

$$\|f\|_{s_0}^{s_0} \le A^{\frac{1}{q}} B^{\frac{1}{p}} = A^{\frac{b-s_0}{b-a}} B^{\frac{s_0-a}{b-a}}.$$

We set $b = k(k+1)$, $a = s(s+1)$ and $f = T$ considered as a function on the minor arcs $\mathfrak{m}$. By Theorem 8.9 we have $A = P^{s^2+\epsilon}$. Further our corollary above yields $B = P^{k^2-1+\epsilon}$. Thus we get

$$\int_{\mathfrak{m}} |T(\alpha)|^{s_0} d\alpha \ll P^{(k^2-1)\frac{s_0-s(s+1)}{k(k+1)-s(s+1)}+s^2\frac{k(k+1)-s_0}{k(k+1)-s(s+1)}+\epsilon}.$$

If we write $a = \frac{k(k+1)-s_0}{k(k+1)-s(s+1)}$ and $\eta = (1-a)(k+1)+as$ we can rewrite the exponent as

$$\begin{aligned}(k^2-1)(1-a)+s^2a+\epsilon &= (k+1)(k-1)(1-a)+s(s+1)a-as+\epsilon\\ &= k(k+1)(1-a)+s(s+1)a-\eta+\epsilon\\ &= k(k+1)-[k(k+1)-s(s+1)]a-\eta+\epsilon\\ &= s_0-\eta+\epsilon.\end{aligned}$$

We are done as soon as we can verify that $\eta > k$. A moderately annoying shows that this is satisfied for

$$s_0 > k^2 - s\frac{k-s-1}{k+1-s}.$$

□

9.2. **The major arcs revisited.** We consider large n and put $P = \lfloor n^{\frac{1}{k}} \rfloor$. The complement of the minor arcs in $[0,1)$, the major arcs, are as usual denoted by $\mathfrak{M}$. We can decompose them in the pieces

$$\mathfrak{M} = \bigcup_{\substack{0<q\le(2k)^{-1}P,\\ 0\le a\le q,\\ (a,q)=1}} \underbrace{\{\alpha \in [0,1) \colon |q\alpha - a| \le (2k)^{-1}P^{1-k}\}}_{=\mathfrak{M}(q,a)}.$$

We claim that for

$$k^2 - s_0\frac{k-s_0-1}{k+1-s_0} < s \le k(k+1)$$

we have

$$\int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha = \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(s/k)}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1} + o(n^{\frac{s}{k}-1}).$$

To derive this we have to be slightly more carefully than in our first analysis of the major arcs.

We put

$$v(\beta) = \sum_{x\le n} \frac{1}{k}x^{\frac{1}{k}-1}e(\beta x)$$

and

$$V(\alpha,q,a) = q^{-1}S_{a,q}v(\alpha - \frac{a}{q}).$$

Note that by our bounds on $S_{a,q}$ (see Lemma 6.17) and standard bounds on $v(\beta)$ (using partial summation and bounds on linear exponential sums) we have

$$V(\frac{a}{q}+\beta,q,a) \ll (q^{-1}\min(n,\|\beta\|^{-1}))^{\frac{1}{k}}.$$

First we need the following approximation.

Lemma 9.5. *Let $(a,q)=1$ and $\alpha = \frac{a}{q}+\beta$. We have*

$$T(\alpha) - V(\alpha,q,a) \ll q^{\frac{1}{2}+\epsilon}(1+n|\beta|)^{\frac{1}{2}}.$$

Furthermore, if $|\beta| \le (2kq)^{-1}n^{\frac{1}{k}-1}$, then the bound simplifies to

$$T(\alpha) - V(\alpha,q,a) \ll q^{\frac{1}{2}+\epsilon}.$$

The same result holds if, in the definition of $V(\alpha,a,q)$, we replace $v(\beta)$ by

$$v_1(\beta) = \int_0^{n^{\frac{1}{k}}} e(\beta x^k)dx.$$

Proof. By splitting the sum defining f into congruence classes modulo q, and detecting these by character orthogonality yields

$$T(\alpha) = q^{-1}\sum_{-\frac{q}{2}<b\le\frac{q}{2}} S_{a,b,q}F(b)$$

for

$$F(b) = \sum_{x \le P} e(\beta x^k - \frac{bx}{q}).$$

By Remark 5.4 we have

$$F(b) = \sum_{h=-H}^{H} I(b + hq) + O(\log(2 + H)),$$

where $-H_1 = H_2 = H = \lfloor |\beta| k P^{k-1} + \frac{3}{2} \rfloor$ and

$$I(c) = \int_0^X e(\beta y^k - cy/q) dy.$$

Using our estimate for $S_{a,b,q}$ and

$$q^{-1} \sum_{b=1}^{q} (q, b) \le d(q) \ll q^\epsilon$$

we obtain

$$T(\alpha) - q^{-1} S_{a,q} v_1(\beta) = q^{-1} \sum_{\substack{-B < b \le B, \\ b \ne 0}} S_{a,b,q} I(b) + O(q^{\frac{1}{2}+\epsilon} \log(2 + H)).$$

Starting from here we can first deal with the case $|\beta| \le (2kq)^{-1} P$. In this case $0 < H \le 2$. Suppose $b \ne 0$ and $0 \le y \le P$, then

$$|\beta k y^{k-1} - \frac{b}{q}| \ge \frac{1}{2} |\frac{b}{q}|.$$

Thus, in this range the integral $I(b)$ can be treated simply using partial integration. This yields

$$I(b) \ll |\frac{b}{q}|^{-1}.$$

We get the estimate

$$f(\alpha) - q^{-1} S_{a,q} v_1(\beta) \ll q^{-1} \sum_{1 \le b \ll q} q^{\frac{1}{2}+\epsilon} (b, q) q b^{-1} \ll q^{\frac{1}{2}+\epsilon}.$$

Next we will argue that we can replace $v_1(\beta)$ by $v(\beta)$. This is seen as follows. To do so consider the sum

$$G(Y) = \sum_{m \le Y} \frac{1}{k} m^{\frac{1}{k}-1} = Y^{\frac{1}{k}} + C_k + O(Y^{\frac{1}{k}-1}).$$

The last asymptotic can for example be derived using Euler-Maclaurin summation. By partial summation we get

$$\begin{aligned} v(\beta) &= G(P^k)e(\beta P^k) - 2\pi i\beta \int_1^{P^k} G(y)e(\beta y)dy \\ &= (P+C_k)e(\beta P^k) - 2\pi i\beta \int_1^{P^k} (y^{\frac{1}{k}} + C_k)e(\beta y)dy + O(P^{1-k} + |\beta|P). \end{aligned}$$

Changing variables and integrating by parts yields

$$v(\beta) = v_1(\beta) + O(1 + |\beta|P).$$

Thus we are free to pass between v_1 and v at the cost of an negligible error.

To treat the case of general β we argue as follows. We split the integrals $I(b) = I_1(b) + I_2(b)$ in two parts. The part of $I(b)$ on which y satisfies

$$|k\beta y^{k-1} - \frac{b}{q}| \geq \frac{|b|}{2q} \tag{35}$$

can be estimated as before by $I_1(b) \ll q/|b|$. The contribution of these integrals to the sum is bounded as above and we get

$$f(\alpha) - q^{-1}S_{a,q}v_1(\beta) = q^{-1} \sum_{\substack{-B<b\leq B,\\ b\neq 0}} S_k(a,b;q)I_2(b) + O(q^{\frac{1}{2}+\epsilon}\log(2+H) + q^{\frac{1}{2}+\epsilon}\log(2B)).$$

On the domain of integration of $I_2(b)$ we have $|k\beta y^{k-1} - \frac{b}{q}| \leq \frac{|b|}{2q}$. One notes that $I_2(b)$ vanishes unless

$$\frac{|b|}{2q} \leq k|\beta|y^{k-1} \leq \frac{3|b|}{2q},$$

so that $|b| \leq 2kq|\beta|P^{k-1}$. For such b set

$$\delta_b = |\beta|^{\frac{1}{2k-2}}(|b|/q)^{\frac{k-2}{2k-2}}.$$

Parts of the integral where $|k\beta y^{k-1} - \frac{b}{q}| \geq \delta_b$ are again estimated by integration by parts obtaining the bound $\ll \delta^{-1}$. The remaining part of the integral is over an interval $[\gamma_1, \gamma_2]$ of length δ^{-1}. Thus this remaining part can be estimated trivially by $\ll \delta^{-1}$. (The choice of δ is precisely made to balance the length of the interval where the phase is small with the upper bound coming from integration by parts on the rest!) We can now estimate

$$q^{-1} \sum_{\substack{-B<b\leq B,\\ b\neq 0}} S_{a,b,q}I_2(b) \ll q^{-1} \sum_{0<b\leq 2kq|\beta|X^{k-1}} q^{\frac{1}{2}-\epsilon}\delta_b^{-1} \ll q^{\frac{1}{2}+\epsilon}d(q)|\beta|^{\frac{1}{2}}P^{\frac{k}{2}}.$$

Inserting this above completes the proof. □

Next we need to revisit the singular series. Recall

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} A(q) \text{ for } A(q) = \sum_{\substack{a=1,\\ (a,q)=1}}^{q} [q^{-1}S_{a,q}]^s e(-\frac{an}{q}).$$

Lemma 9.6. *Suppose $s \geq 1$ and $l = uk + v$ with $1 \leq v \leq k$. Then*

$$p^{us}A(p^l) \ll \begin{cases} p^{-\frac{s}{2}}(p^{\frac{1}{2}}(p^{l-1},n) + (p^l,n)) & \textit{if } l \equiv 1 \textit{ mod } k, \\ p^{-s}(p^l,n) & \textit{else.} \end{cases}$$

Furthermore, if $\lambda = l - \max(k,\gamma) > 0$ and $p^\lambda \nmid h$, then

$$A(p^l) = 0.$$

Here $p^\tau \| k$ and $\gamma = \tau + 1$ if $p > 2$ or $p = 2$ and $\tau = 0$ or $\gamma = \tau + 2$ otherwise.

Proof. We start with the case $p > k$, so that $\gamma = 1$. By Lemma 6.15 we have

$$p^{ls}A(p^l) = (p^{u(k-1)})^s \sum_{a=1,\, p\nmid a}^{p^l} S^s_{p^v,a} e(-anp^{-l}). \tag{36}$$

We write $a = xp^v + y$ with $0 \leq x < p^{l-v}$ and $1 \leq y < p^v$. Note that we can execute the x-sum by character orthogonality picking up the condition $p^{l-v} \mid n$. Using Lemma 6.15 again we get

$$\sum_{y=1,\, p\nmid y}^{p^v} S^s_{p^v,y} e(-ynp^{-l}) = p^{s(v-1)} \sum_{y=1,\, p\nmid y}^{p^v} e(-ynp^{-l})$$

for $v > 1$. We can bound this sum by $p^{s(v-1)}(p^v, np^{v-l})$, which leads to

$$|A(p^l)| \leq p^{-us-s}(p^l,n).$$

For $v = 1$ we evaluate the y-sum above differently. Indeed we can write

$$\sum_{y=1,\, p\nmid y}^{p^v} (S_{y,p^v})^s e(-yhp^{-l}) = \sum_{\substack{\chi_i \in \mathcal{A},\\ i=1,\dots,s}} \tau(\chi_1)\dots\tau(\chi_s) \sum_{y=1}^{p} \chi_1^{-1}(y)\dots\chi_s^{-1}(y)e(-ynp^{-l}).$$

Here $\mathcal{A}$ is the set of non-principal characters χ mod p such that χ^k is principal. Note that $\mathcal{A} = (k, p-1) - 1$. The y-sum can be identified as another Gauß sum. Estimating the Gauß sums by $|\tau(\chi)| = \sqrt{p}$ we get

$$A(q) \ll p^{-us-\frac{s}{2}}(p^{\frac{1}{2}}(p^{l-1},n) + (p^l,n)).$$

Note that we had to take the exceptional cases into account for which the product $\chi_1 \dots \chi_s$ is the principal character.

We turn towards $p \leq k$. When $l \leq \max(\gamma, k)$ the conclusion is trivial. We assume $l > \max(\gamma, k)$. But in this situation (36) holds as well. As above we obtain the vanishing condition for $A(p^l)$. In the non-vanishing situation (i.e. $p^{l-v} \mid n$) we have

$$p^{ls}A(p^l) = p^{us(k-1)}p^{l-v} \sum_{y=1,\, p\nmid y}^{p^v} (S_{y,p^v})^s e(-ynp^{-l}) \ll p^{us(k-1)}(p^l, h).$$

This leads the desired bound immediately. □

With this at hand we can derive finer properties of the singular series.

Lemma 9.7. *Suppose* $s \geq 4$. *Then* $\mathfrak{S}(n)$ *converges absolutely and is non-negative. Further, if* $s \geq \max(5, k+2)$*, then* $\mathfrak{S}(n) \ll 1$. *If* $\max(4, k) \leq s < \max(5, k+2)$ *one still has* $\mathfrak{S}(n) \ll n^\epsilon$.

Proof. First note that the previous lemma implies

$$\sum_{l=1}^{\infty} A(p^l) \ll np^{-(s-1)/2} \ll np^{-\frac{3}{2}}. \tag{37}$$

Thus, exploiting multiplicativity, we get

$$\sum_{q\leq Q} |A(q)| \leq \prod_{p\leq Q} (1 + Cp^{-\frac{3}{2}})^n \leq (C')^n.$$

This yields absolute convergence and non-negativity is clear from the product representation in terms of local densities.

We turn towards the claimed upper bounds. Let $p^\theta \| n$ and write $l = uk + v$ with $1 \leq v \leq k$. Further define w by

$$w + us - \min(l, \theta) = \begin{cases} -\frac{s}{2} & \text{if } l \leq \theta \text{ and } v = 1, \\ -\frac{s-1}{2} & \text{if } l > \theta \text{ and } v = 1, \\ -s & \text{if } v \neq 1. \end{cases}$$

We have

$$A(p^l) \ll \begin{cases} p^w & \text{if } l \leq \theta + \max(k, \gamma), \\ 0 & \text{else.} \end{cases}$$

With this we can estimate

$$\sum_{l=1}^{\infty} |A(p^l)| \ll \begin{cases} p^{-\frac{3}{2}} & \text{if } \theta = 0, \text{ or } \theta \geq 1 \text{ and } s \geq \max(5, k+2), \\ \theta & \text{if } \theta \geq 1 \text{ and } s \geq \max(4, k). \end{cases}$$

The claimed estimates for the singular series follow directly from the Euler product

$$\mathfrak{S}(n) = \prod_p (1 + \sum_{l=1}^{\infty} A(p^l)).$$

□

The final preliminary lemma is the following estimate.

Lemma 9.8. *Suppose* $s \geq \max(4, k+1)$. *Then*

$$\sum_{q \leq Q} q^{\frac{1}{k}} A(q) \ll (nQ)^{\epsilon}.$$

Proof. Using the notation from the previous proof we obtain

$$p^{\frac{l}{k}} A(p^l) \ll \begin{cases} \theta & \text{if } l \leq \theta, \\ p^{-1} & \text{if } \theta < l \leq \theta + \max(k, \gamma), \\ 0 & \text{if } l > \theta + \max(k, \gamma). \end{cases}$$

Thus we obtain

$$\sum_{q \leq Q} q^{\frac{1}{k}} |A(q)| \leq \prod_{p \leq Q} (1 + \sum_{l=1}^{\infty} p^{\frac{l}{k}} |A(p^l)|) \leq d(n)^C \prod_{p \leq Q} (1 + C/p).$$

This already completes the proof. □

Completely analogously we can show that

$$\sum_{q \leq Q} q^{-t\lambda} \sum_{\substack{a=1, \\ (a,q)=1}}^{q} |S_{a,q}|^t \ll Q^{\epsilon}$$

whenever $t \geq \max(4, k)$. Here we take $\lambda = 0$ if $t \geq k+1$ and $\lambda = \frac{1}{k}$ if $t = k$.

We can now prove the following asymptotic evaluation of the major arc integral.

Theorem 9.9. *Suppose that* $s \geq \max(5, k+1)$. *Then there is* $\delta > 0$ *such that*

$$\int_{\mathfrak{M}} T(\alpha)^s e(-\alpha n) d\alpha = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} n^{\frac{s}{k}-1} \mathfrak{S}(n) + O(n^{\frac{s}{k}-1-\delta}).$$

Proof. Take $\alpha \in \mathfrak{M}_{a,q}$. We start by collecting some estimates. From $T(\alpha) = V(\alpha, a, q) + O(q^{\frac{1}{2}+\epsilon})$ we obtain

$$T(\alpha)^s - V(\alpha, a, q)^s \ll q^{\frac{s}{2}+\epsilon} + q^{\frac{1}{2}+\epsilon} |V(\alpha, a, q)|^{s-1}.$$

Inserting this into the q-part of the major arc integral yields

$$\sum_{\substack{a=1, \\ (a,q)=1}}^{q} \int_{\mathfrak{M}_{a,q}} T(\alpha)^s e(-\alpha n) d\alpha = \sum_{\substack{a=1, \\ (a,q)=1}}^{q} \int_{\mathfrak{M}_{a,q}} V(\alpha, a, q)^s e(-\alpha n) d\alpha$$
$$+ O(P^{1-k} q^{\frac{s}{2}+\epsilon} + q^{\frac{3}{2}-s+\epsilon} \sum_{\substack{a=1, \\ (a,q)=1}}^{q} |S_{a,q}|^{s-1} \int_{-\frac{1}{2}}^{\frac{1}{2}} |v(\beta)|^{s-1} d\beta).$$

Summing this over $q \ll P$ yields

$$\int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha = \sum_{0<q\ll P} \sum_{\substack{a=1,\\(a,q)=1}}^{q} \int_{\mathfrak{M}_{a,q}} V(\alpha,a,q)^s e(-\alpha n)d\alpha + \mathcal{E},$$

where we can estimate the error by

$$\mathcal{E} \ll P^{2+\frac{s}{2}-k+\epsilon} + P^{\frac{3}{4}+\epsilon} \sum_{q\le P} q^{\frac{3}{4}-s+\epsilon} \sum_{\substack{a=1,\\(a,q)=1}}^{q} |S_{q,a}|^{s-1} n^{\frac{s-1}{k}-1+\epsilon} \ll n^{\frac{s}{k}-1-\delta}.$$

Writing $\mathfrak{N}_{a,q} = [-\frac{1}{2},\frac{1}{2}] \setminus (\mathfrak{M}_{a,q} - \frac{a}{q})$ and estimating

$$\sum_{\substack{a=1,\\(a,q)=1}}^{q} \int_{\mathfrak{N}_{a,q}} V(\alpha,a,q)^s e(-\alpha n)d\alpha \ll |A(q)| \int_{q^{-1}P^{1-k}}^{\infty} \beta^{-\frac{s}{k}} d\beta \ll n^{\frac{s}{k}-1-\delta} \tag{38}$$

allows us to enlarge the remaining major arc integral to $[-\frac{1}{2},\frac{1}{2}]$. Thus we have seen that

$$\int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha = \mathfrak{S}(n,P)I(n) + O(n^{\frac{s}{k}-1-\delta}),$$

for

$$\mathfrak{S}(m,P) = \sum_{q\le P} A(q) \text{ and } I(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} v(\beta)^s e(-\beta m)d\beta.$$

It is clear that we can replace the partial singular series $\mathfrak{S}(m;P)$ by the full one $\mathfrak{S}(m)$ by making a negligible error. The result follows by asymptotically evaluating $I(m)$ as we have seen it done before. □

9.3. **Completion of the Proof of Theorem 9.1.** We are now essentially done. We get the asymptotic formula

$$\begin{aligned} R_{s,k}(n) &= \int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha + \int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha \\ &= \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(s/k)} \mathfrak{S}_{s,k}(n) n^{\frac{s}{k}-1} + o(n^{\frac{s}{k}-1}), \end{aligned}$$

for

$$k^2 - s\frac{k-s-1}{k+1-s} < s_0 \le k(k+1).$$

The smallest possible choice for s within these constraints is

$$s_0 = k^2 + 1 - \max_{s\le k} \left\lceil s\frac{k-s-1}{k-s+1} \right\rceil.$$

Since our earlier analysis showed that the singular series can be bounded from below in this situation. Thus we get that every sufficiently large number is represented by a sum s_0 positive kth powers. Therefore we must have $G(k) \leq s_0$ as desired.

10. Exercises

Exercise Block 1. The goal of this exercise is recreate an (elementary) argument due to M. A. Lukomskaja and proof the following theorem due to van der Waerden.

Theorem 10.1 (Van der Waerden)**.** *Let k, l be two arbitrary natural numbers. Then there is an integer $n = n(k, l)$ such that the following holds. If any piece of the natural numbers of length n (i.e. $[a, a+n] \cap \mathbb{N}$ for $a \in \mathbb{N}$) is decomposed into k arbitrary (sub)-sets, then at least one of these k sets contains an arithmetic progression of length l.*

The proof proceeds by induction on l and it is an exercise to fill in the details of the following steps.

a) Show the statement for $l = 2$.

From now on we assume that the statement folds for some $l \geq 2$ and put $l' = l+1$. We also set $q_0 = 1$ and $n_0 = n(k, l)$. Now we make the following inductive construction:

$$q_s = 2n_{s-1}q_{s-1}, \text{ and } n_s = n(k^{q_s}, l)$$

for $s \geq 1$. To conclude the induction step we will establish that $n(k, l+1) = q_k$.

b) Put $\Delta = [a, a+q_k] \cap \mathbb{N}$ for some $a \in \mathbb{N}$ and decompose it in k disjoint sets

$$\Delta = S_1 \cup \ldots \cup S_k.$$

we can interpret this as equivalence relation and write $a \sim b$ if and only if $a, b \in S_i$ for some $1 \leq i \leq k$. Two intervals $I, J \subset \Delta$ of successive integers with the same cardinality (i.e $I = \{b, b+1, \ldots, b+r\}$ and $J = \{c, c+1, \ldots, c+r\}$) are called equivalent if $a+i \sim c+i$ for all $i = 0, 1, \ldots, r$. Construct l' sets $\Delta_1, \ldots, \Delta_{l'}$ of successive integers with cardinality q_{k-1} such that

   i. The minimal elements $a_i \in \Delta_i$ form an arithmetic progression of length l'. (For later reference we set $d_1 = a_i - a_{i-1}$.)
   ii. The sets $\Delta_1, \ldots, \Delta_l$ are all equivalent.

c) Continue this construction to find integers $\Delta_{i_1,\ldots,i_k} \in \Delta$ with indices $1 \leq i_1, \ldots, i_k \leq l'$ with the following properties

   i. $\Delta_{i_1,\ldots,i_k} \sim \Delta_{j_1,\ldots,j_k}$ for $1 \leq i_1, \ldots, i_k, j_1, \ldots, j_k \leq l$.
   ii. For $s < k$, $1 \leq i_1, \ldots, i_s, j_1, \ldots, j_s \leq l$ and $1 \leq i_{s+1}, \ldots, i_k \leq l'$ we have $\Delta_{i_1,\ldots,i_s,i_{s+1},\ldots,i_k} \sim \Delta_{j_1,\ldots,j_s,i_{s+1},\ldots,i_k}$.
   iii. For $s < k$, $1 \leq i_1, \ldots, i_k \leq l'$ and $i'_s = i_s+1 \leq l'$ we have $\Delta_{i_1,\ldots,i'_s,i_{s+1},\ldots,i_k} - \Delta_{i_1,\ldots,i_s,i_{s+1},\ldots,i_k} = d_s$. Where d_s is some natural number. (Recall that d_1 is fixed above.)

d) Use the numbers constructed above to find an arithmetic progression of length l' which lies completely on one of the sets S_i (i.e. all terms of the arithmetic progression are equivalent). *Hint:* Consider the numbers

$$a_0 = \Delta_{l',l',\ldots,l'},\ a_1 = \Delta_{1,l',\ldots,l'}, \ldots, a_k = \Delta_{1,1,\ldots,1}.$$

Solution: To see *a.*) we put $n(k,2) = k+1$. Then we observe that if we decompose $k+1$ integers in k sets there is at least one set with 2 elements. We are done since any 2 integers form an arithmetic progression of length 2.

We turn to *b.*). First note that there are at most k^m equivalence classes of intervals of length m in Δ. The trick is to decompose the first half of Δ in at most $k^{q_{k-1}}$ equivalence classes of intervals of length q_{k-1}. Note that the first half of Δ is an interval of length $n_{k-1} = n(k^{q_{k-1}}, l)$, so that we can apply the induction hypothesis to find an arithmetic progression of length l which are the first elements of intervals $\Delta_1, \ldots, \Delta_l$ with q_{k-1} elements of the same type. We add the last interval $\Delta_{l'}$ artificially.

The idea behind *c.*) is pretty clear. We apply the same argumentation as in *b.*) to each of the intervals Δ_{i_1} producing intervals Δ_{i_1,i_2} for $1 \leq i_2 \leq l$. We add the final interval $\Delta_{i_1,l'}$ of length q_{k-1} artificially. We continue this procedure till we have $\sharp\Delta_{i_1,\ldots,i_k} = q_0 = 1$ and we identify these sets of one elements with the element contained. To check the properties *i*, *ii* and *iii* is no only a matter of unraveling the notation.

Finally we come to *d.*). As the hint suggests we look at the $k+1$ numbers $a_0, \ldots, a_k$. The pigeon hole principle tells us that there is $r < s$ such that a_r and a_s are in the same set (i.e. equivalent). We now define

$$c_i = \Delta_{\underbrace{1,\ldots 1}_{r \text{ times}},\underbrace{i,\ldots,i}_{s-r \text{ times}},\underbrace{l',\ldots,l'}_{k-s}}.$$

for $1 \leq i \leq l'$. We claim that this is the desired arithmetic progression of length l' of the same type. Indeed the first $c_1 \sim c_2 \sim \ldots \sim c_l$ by construction. But $c_{l'} = a_r \sim a_s = c_1$. So that they are all equivalent. We still have to check that the distance between two consecutive elements c_i and C_{i+1} remains constant as i varies. To do so put

$$c_{i,m} = \Delta_{\underbrace{1,\ldots 1}_{r \text{ times}},\underbrace{i',\ldots i'}_{m \text{ times}},\underbrace{i,\ldots,i}_{s-r-m \text{ times}},\underbrace{l',\ldots,l'}_{k-s}} \tag{39}$$

for $0 \leq m \leq s-r$. Of course $c_{i,0} = c_i$ and $c_{i,s-r} = c_{i'} = c_{i+1}$. Note that by construction we have $c_{i,m} - c_{i,m+1} = d_{r+m}$. By the telescoping trick we have

$$c_{i+1} - c_i = \sum_{m=1}^{s-r} (c_{i,m} - c_{i,m-1}) = d_{r+1} + \ldots + d_s.$$

Since the latter is independent of i the proof is concluded.

Exercise Block 2. It should be well known (for example from algebraic number theory) when the equation

$$n = x^2 + y^2$$

is soluble over the integers. Furthermore, if n can be expressed as a sum of two squares, the number of ways in which this can be achieved is well understood. Our goal is to give another proof of the following theorem due to Lagrange.

Theorem 10.2 (Lagrange). *For every $n \in \mathbb{N}$ we can solve*

$$n = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

with $(x_1, x_2, x_3, x_4) \in \mathbb{Z}_{\geq 0}^4$.

It is an exercise to complete the proof using the following steps.

a) Show the following identity due to Euler:

$$\begin{aligned}(x_1^2 + x_2^2 + x_3^2 + x_4^2)&(y_1^2 + y_2^2 + y_3^2 + y_4^2)\\ &= (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2\\ &\quad + (x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4)^2 + (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2.\end{aligned}$$

b) Show that for every $p > 2$ there is $1 \leq m < p$ such that

$$mp = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

is soluble over the (non-negative) integers. (**Hint:** Find $0 \leq x, y \leq \frac{p-1}{2}$ such that $x^2 \equiv -1 - y^2 \bmod p$.)

c) Show that for every p the equation

$$p = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

is soluble over the (non-negative) integers. (**Hint:** Take m to be the minimal natural number satisfying *b.*). Then argue by contradiction.)

d) Finish the proof of Lagrange's theorem.

Note that the theorem of Lagrange has a strengthening which is due to Jacobi. Indeed one can actually prove the explicit expression for the representation numbers:

$$r(n) = \sharp\{(x_1, \ldots, x_4) \in \mathbb{Z}^4 \colon n = x_1^2 + \ldots + x_4^2\} = 8 \sum_{\substack{m|n,\\ 4\nmid m}} m.$$

This can be proved elementary but also has a very nice proof using modular forms.

Solution: Part *a.*) can be seen by brute force.

To see part *b.*) we proceed as follows. We first claim that there is a pair (x, y) such that

$$x^2 \equiv -1 - y^2 \bmod p \text{ and } |x|, |y| < \frac{p}{2}.$$

To see this we look at the sets

$$A = \{x^2 \colon 0 \leq x \leq \frac{p-1}{2}\} \text{ and } B = \{-1-y^2 \colon 0 \leq y \leq \frac{p-1}{2}\}.$$

The elements of set A (and similarly of set B) are pairwise incongruent modulo p. (Indeed $x_1^2 \equiv x_2^2 \bmod p$ would imply $p \mid (x_1 - x_2)(x_1 + x_2)$, which is impossible.) Since $\sharp(A \cup B) = p + 1$, the claim follows from the pigeon hole principle. But this implies $x^2 + y^2 + 1 \equiv 0 \bmod p$. The latter can be rewritten as

$$x^2 + y^2 + 1^1 + 0^2 = mp$$

for $m \in \mathbb{Z}$. We complete this part of the exercise by observing

$$0 < mp < \frac{p^2}{4} + \frac{p^2}{4} + 1 < p^2.$$

Turning to part *c*.) we observe that the statement is obviously true for $p = 2$. For $p > 2$ let $m = m(p)$ be the smallest (positive) number such that

$$m \cdot p = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

Suppose $m > 1$. We first show that m must be odd. Note that $mp \equiv 0 \bmod 2$ implies (after renaming the x_i's if necessary) that $x_1 + x_2 \equiv 0 \bmod 2$ and $x_3 + x_4 \equiv 0 \bmod 2$. But then we can write

$$\frac{m}{2} \cdot p = \left(\frac{x_1 + x_2}{2}\right)^2 + \left(\frac{x_1 - x_2}{2}\right)^2 + \left(\frac{x_3 + x_4}{2}\right)^2 + \left(\frac{x_3 - x_4}{2}\right)^2,$$

which contradicts minimality of m. Thus m must be odd. Take $-\frac{m-1}{2} \leq y_i \leq \frac{m-1}{2}$ such that $y_i \equiv x_i \bmod m$ for $i = 1, 2, 3, 4$. We obtain the congruence

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 \equiv 0 \bmod m.$$

If $m \mid x_i$ for $i = 1, 2, 3, 4$, then $m \mid p$ which implies $m = 1$ and we would be done. Thus, not all y_i, $i = 1, 2, 3, 4$ can be 0 and we find $n > 0$ such that

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = m \cdot n.$$

A trivial estimate shows that $mn < m^2$ such that $m < n$. By Euler's identity we get

$$m^2 np = (mn) \cdot (mp) = (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) z_1^2 + z_2^2 + z_3^2 + z_4^2,$$

for

$$\begin{aligned} z_1 &= x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4, \\ z_2 &= x_1 y_2 - x_2 y_1 + x_3 y_4 - x_4 y_3, \\ z_3 &= x_1 y_3 - x_3 y_1 + x_4 y_2 - x_2 y_4 \text{ and} \\ z_4 &= x_1 y_4 - x_4 y_1 + x_2 y_3 - x_3 y_2. \end{aligned}$$

One checks that $m \mid z_i$ for $i = 1, 2, 3, 4$. Thus we get

$$np = \left(\frac{z_1}{m}\right)^2 + \left(\frac{z_2}{m}\right)^2 + \left(\frac{z_3}{m}\right)^2 + \left(\frac{z_4}{m}\right)^2,$$

which contradicts the minimality of m.

Finally $d.$) follows easily from Euler's identity (see $a.$)) and the fundamental theorem of arithmetic.

Exercise Block 3. We still need to fill in some details left out in our proof of the transformation behaviour of the Dedekind-eta-function. Recall that we had defined the Dedekind sum:

$$s(d,c) = \sum_{o \leq x < c} \frac{x}{c}(\frac{dx}{c} - \left[\frac{dx}{c}\right] - \frac{1}{2}).$$

Prove the following facts about the group $\mathrm{SL}_2(\mathbb{Z})$ as well as the sum $s(d,c)$:

a) The group $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices T, S.

b) For $(c,d) = 1$ we have

$$s(d,c) = \sum_{x \bmod c} ((\frac{x}{c}))((\frac{dx}{c})),$$

with

$$((x)) = \delta_{x \notin \mathbb{Z}} \cdot (x - [x] - \frac{1}{2}).$$

c) $s(d,c)$ satisfies the following properties:
- $s(\pm d + mc, c) = \pm s(d,c)$;
- If $\overline{d}$ is the inverse of d modulo c, then $s(\overline{d},c) = s(d,c)$.
- If $d^2 \equiv -1 \bmod c$, then $s(d,k) = 0$.

d) Show the reciprocity formula in the form

$$12cd(s(d,c) + s(c,d)) = c^2 + d^2 - 3cd + 1.$$

for $c, d > 0$.

Solution: We start with a). Note that $S^2 = -1_2 \in \mathrm{SL}_2(\mathbb{Z})$ and $T^m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$. Now suppose we have $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Without loss of generality we can assume $c > 0$. Indeed if $c = 0$, then $\gamma = \pm T^m$ for some m and we are done. If $c < 0$ we can replace γ by $-\gamma$. Further, we set

$$\gamma' = T^m \gamma = \begin{pmatrix} a + cm & b + dm \\ c & d \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

By choosing m appropriately we can arrange that $-c < a' \leq 0$. We then observe that

$$\gamma'' = S\gamma' = \begin{pmatrix} c' & d' \\ -a' & -d' \end{pmatrix} = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}.$$

We have set things up so that $0 \leq c'' = -a' < c$. Thus we have reduced the size of the lower left entry. Continuing this procedure allows us to reach the situation

$c'' = 0$ after finitely many steps.

We turn towards $b)$. We first observe that $((x))$ is odd and of period 1. Thus we have

$$\sum_{x \mathrm{mod} c} ((\frac{x}{c})) = 0.$$

By a change of variables in the sum this leads to

$$\sum_{x \mathrm{mod} c} ((\frac{dx}{c})) = 0 \text{ for } (d,c) = 1.$$

We compute

$$\sum_{x \bmod c} ((\frac{x}{c}))((\frac{dx}{c})) \sum_{x \bmod c} (\frac{x}{c} - \frac{1}{2})((\frac{dx}{c})) = \sum_{x=1}^{c-1} \frac{x}{c}((\frac{dx}{c})) = s(d,c).$$

Starting from $b)$ it is not hard to derive $c)$. Indeed the first two bullet points follow directly from the periodicity and oddness of $((\cdot))$. The final bullet point is not much harder since $d^2 \equiv -1 \bmod c$ implies $d \equiv -\overline{d} \bmod c$. We can thus simply apply the first two points.

The most difficult part is $d)$, which we treat now. Note that this reciprocity formula was first derived from the functional equation of the Dedekind eta function (or more precisely $\log(\eta(z))$). However, since we use the reciprocity result to deduce the transformation behaviour we will discuss a purely arithmetic proof here. We first compute

$$\sum_{x=1}^{c-1}((\frac{dx}{c}))^2 = \sum_{x \bmod c} ((\frac{dx}{c}))^2 \sum_{x \bmod c} ((\frac{x}{c}))^2 = \sum_{x=1}^{c-1}(\frac{x}{c}-\frac{1}{2})^2 = \frac{1}{c^2}\sum_{x=1}^{c-1} x^2 - \frac{1}{c}\sum_{x=1}^{c-1} x + \frac{1}{4}\sum_{x=1}^{c-1} 1.$$

On the other hand we have

$$\begin{aligned}\sum_{x=1}^{c-1}((\frac{dx}{c}))^2 &= \sum_{x=1}^{c-1}(\frac{dx}{c} - \left[\frac{dx}{c}\right] - \frac{1}{2})^2 \\ &= \sum_{x=1}^{c-1}(\frac{d^2x^2}{c^2} + \left[\frac{dx}{c}\right]^2 + \frac{1}{4} - \frac{dx}{c} + \left[\frac{dx}{c}\right] - \frac{2dx}{c}\left[\frac{dx}{c}\right]) \\ &= 2d\sum_{x=1}^{c-1}\frac{x}{c}\left(\frac{dx}{c} - \left[\frac{dx}{c}\right] - \frac{1}{2}\right) + \sum_{x=1}^{c-1}\left[\frac{dx}{c}\right]\left(\left[\frac{dx}{c}\right]+1\right) - \frac{d^2}{c^2}\sum_{x=1}^{c-1}x^2 + \frac{1}{4}\sum_{x=1}^{c-1}1.\end{aligned}$$

Combining these two identities we obtain

$$2d \cdot s(d,c) + \sum_{x=1}^{c-1}\left[\frac{dx}{c}\right]\left(\left[\frac{dx}{c}\right]+1\right) = \frac{d^2+1}{c^2}\sum_{x=1}^{c-1}x^2 - \frac{1}{c}\sum_{x=1}^{c-1}x. \qquad (40)$$

We set $\left[\frac{dx}{c}\right] = v-1$ and observe that possible values for v are $v = 1, 2, \dots, d$. Further put

$$N(v) = \sharp\{x\colon \left[\frac{dx}{c}\right] = v-1\}.$$

Thus we can write

$$\sum_{x=1}^{c-1} \left[\frac{dx}{c}\right]\left(\left[\frac{dx}{c}\right]+1\right) = \sum_{v=1}^{d}(v-1)v \cdot N(v).$$

We investigate $N(v)$. Note that x contributes to $N(v)$ if $v-1 < \frac{dx}{c} < v$. We rewrite this as

$$\frac{c(v-1)}{d} < x < \frac{cx}{d}.$$

We conclude that

$$N(v) = \begin{cases} \left[\frac{cv}{d}\right] - \left[\frac{c(v-1)}{d}\right] & \text{if } 1 \le v \le d-1, \\ c-1-\left[\frac{c(d-1)}{d}\right] & \text{if } v = d. \end{cases}$$

With this at hand we compute

$$\begin{aligned} &\sum_{x=1}^{c-1} \left[\frac{dx}{c}\right]\left(\left[\frac{dx}{c}\right]+1\right) \\ &\quad = \sum_{v=1}^{d}(v-1)v\left(\left[\frac{cv}{d}\right] - \left[\frac{c(v-1)}{d}\right]\right) - d(d-1) \\ &\quad = \sum_{v=1}^{d-1}\left[\frac{cv}{d}\right]((v-1)v - v(v+1)) + cd(d-1) - d(d-1) \\ &\quad = -2\sum_{v=1}^{d-1} v\left[\frac{cv}{d}\right] + d(d-1)(c-1). \end{aligned}$$

Finally we compute

$$2ds(c,d) = 2\sum_{v=1}^{d-1} v\left(\frac{cv}{d} - \left[\frac{cv}{d}\right] - \frac{1}{2}\right) = -2\sum_{v=1}^{d-1} v\left[\frac{cv}{d}\right] + \frac{2c}{d}\sum_{v=1}^{d-1} v^2 - \sum_{v=1}^{d-1} v.$$

Inserting this above yields

$$\sum_{x=1}^{c-1} \left[\frac{dx}{c}\right]\left(\left[\frac{dx}{c}\right]+1\right) = 2d \cdot s(c,d) - \frac{2c}{d}\sum_{v=1}^{d-1} v^2 + \sum_{v=1}^{d-1} v + d(d-1)(c-1).$$

In view of (40) and multiplying with $6c$ we get

$$13cd\cdot s(d,c) + 12cd\cdot s(c,d) = 6\frac{d^2+1}{c}\sum_{x=1}^{c-1} x^2 - 6\sum_{x=1}^{c-1} x + \frac{12c^2}{d}\sum_{v=1}^{d-1} v^2 - 6c\sum_{v=1}^{d-1} v - 6cd(d-1)(c-1).$$

Recall the formulae

$$\sum_{x=1}^{c-1} x = \frac{(c-1)c}{2} \text{ and } \sum_{x=1}^{c-1} x^2 = \frac{(c-1)c(2c-1)}{6}.$$

Inserting these evaluations we get

$$\begin{aligned} &13cd \cdot s(d,c) + 12cd \cdot s(c,d) \\ &= (d^2+1)(c-1)(2c-1) - 3(c-1)c + 2c^2(d-1)(2d-1) - 3cd(d-1) - 6cd(d-1)(c-1). \end{aligned}$$

We are done after multiplying out the right hand side.

Exercise Block 4. We consider some standard techniques that are frequently used in number theory. Show the following:

a) (**Dirichlet Approximation**) For every real number α and every $N \in \mathbb{N}$, there is $1 \leq q \leq N$ and $a \in \mathbb{Z}$ such that $|\alpha - \frac{a}{q}| < \frac{1}{qN}$.

b) (**Divisor bound**) Show that $d(n) = \sum_{d|n} 1 \ll_\epsilon n^\epsilon$ for all $\epsilon > 0$.

c) (**Partial summation**) Let $y \in \mathbb{N}$ and $x \in \mathbb{R}$ with $y < x$. For $g \in \mathcal{C}^1([y,x])$ we have

$$\sum_{y \leq n \leq x} f(n)g(n) = S_f(y,x)g(x) - \int_y^x S_f(y,z)g'(z)dz.$$

d) (**Useful Fourier series**) For $\alpha \notin \mathbb{Z}$ we have

$$\{\alpha\} - \frac{1}{2} = \sum_{0 \neq |m| \leq M} \frac{e(-m\alpha)}{2\pi i m} + O(\frac{1}{M\|\alpha\|}).$$

Solution: To prove *a*) we consider the $N+1$ real numbers

$$0, 1, \{\alpha\}, \{2\alpha\}, \ldots, \{(N-1)\alpha\}$$

all lying in the interval $[0,1]$. Now we divide the $[0,1]$ in N disjoint sub-intervals of length N^{-1}. For example the intervals $[\frac{i-1}{N}, \frac{i}{N})$ for $i = 1, \ldots, N_1$ and $[\frac{N-1}{N}, 1]$. By the pigeon hole principle two of the numbers considered above must lie in the same interval. But then by definition of the bracket $\{x\} = x - \lfloor x \rfloor$ the difference of these two numbers is of the shape $q\theta - p$, where $q \leq N$. Since the difference is obviously bounded by $\frac{1}{N}$ we are done after dividing by q.

We turn to *b*) and write $n = \prod_p p^{l_p}$. Then $d(n) = \prod_p (l_p + 1)$. We get

$$\frac{d(n)}{n^\epsilon} = \prod_{p|n} \frac{l_p + 1}{p^{\epsilon l_p}} \leq \prod_{\substack{p|n, \\ p < 2^{\frac{1}{\epsilon}}}} \frac{l_p + 1}{p^{\epsilon l_p}}.$$

Indeed this follows since $p \geq 2^{\frac{1}{\epsilon}}$ implies

$$p^{\epsilon l_p} \geq 2^{l_p} = (1+1)^{l_p} \geq l_p + 1.$$

We continue our estimate as follows

$$\frac{d(n)}{n^{\epsilon}} \leq \prod_{\substack{p|n,\\ p<2^{\frac{1}{\epsilon}}}} \frac{l_p+1}{2^{\epsilon l_p}} \leq \prod_{\substack{p|n,\\ p<2^{\frac{1}{\epsilon}}}} \frac{l_p+1}{\epsilon l_p \log(2)} \leq \prod_{\substack{p|n,\\ p<2^{\frac{1}{\epsilon}}}} \frac{2}{\epsilon \log(2)} \ll 1.$$

To prove *c*) we use the fundamental theorem of calculus and get

$$\begin{aligned} S_f(y,x)g(x) - \sum_{y\leq n\leq x} f(n)g(n) &= \sum_{y\leq n\leq x} f(n)(g(x)-g(n)) \\ &= \sum_{y\leq n\leq x} f(n)\int_n^x g'(\xi)d\xi = \int_y^x g'(\xi)\sum_{y\leq n\leq \xi} f(n)d\xi. \end{aligned}$$

Finally we turn to *d*). Without loss of generality we assume $0<\alpha\leq\frac{1}{2}$. We observe that, for $m\neq 0$, we have

$$\int_{\alpha}^{\frac{1}{2}} e(-mt)dt = \frac{(-1)^{m+1}}{2\pi i m} + \frac{e(-\alpha m)}{2\pi i m}.$$

Summing up this identity and completing the geometric series in the integral yields

$$\sum_{0\neq|m|\leq M} \frac{e(-m\alpha)}{2\pi i m} - \alpha + \frac{1}{2} = \int_{\alpha}^{\frac{1}{2}} \sum_{|m|\leq M} e(mt)dt = \int_{\alpha}^{\frac{1}{2}} \frac{\sin((2M+1)\pi t)}{\sin(\pi t)}dt.$$

By the mean value theorem we get

$$\sum_{0\neq|m|\leq M} \frac{e(-m\alpha)}{2\pi i m} - \alpha + \frac{1}{2} = \int_{\alpha}^{\xi} \frac{\sin((2M+1)\pi t)}{\sin(\pi\alpha)}dt.$$

This implies the result by estimating the integral trivially and using the bound $\sin(\pi\alpha)^{-1}\leq\|\alpha\|^{-1}$.

Exercise Block 5. In the elementary proof of Hilbert's theorem actually a slightly more general problem was considered. Indeed, fix a (monic) polynomial f of degree n. Write

$$r_{k,f}(m) = \{\mathbf{x}\in(\mathbb{N}_0)^k\colon f(x_1)+\ldots+f(x_k)=m\}.$$

Use the circle method to obtain an asymptotic formula for this generalised representation number.

Solution: We write $F(\mathbf{x}) = f(x_1)+\ldots+f(x_k)-m$. Note that large bits of the argument will work for a general polynomial of degree n in k-variables as soon as k is sufficiently large compared to n.

Consider the box $\mathcal{B} = [0, X]^k$ for $X = \lceil m^{\frac{1}{n}} \rceil$. As before we make the Ansatz

$$r_{k,f}(m) = \int_0^1 \underbrace{\left(\sum_{\mathbf{x} \in \mathcal{B} \cap \mathbb{N}_0^k} e(\alpha F(\mathbf{x})) \right)}_{=S_F(\alpha)} d\alpha.$$

For $P \geq 2Q \geq 2$ we consider the major arcs

$$\mathfrak{M} = \{\alpha \colon |\alpha - \frac{a}{q}| \leq \frac{1}{qP}, \, q \leq Q \text{ and } (a, q) = 1\}.$$

First we observe that the intervals of $\mathfrak{M}$ centred at different rational points $\frac{a}{q}$ do not overlap. The minor arcs are as usual the compliment of the major arcs $\mathfrak{m} = [0, 1] \setminus \mathfrak{M}$.

We start by treating the exponential sum on the minor arcs. For $\alpha \in [0, 1]$ we set

$$S_f(\alpha) = \sum_{0 \leq n \leq X} e(\alpha f(x)).$$

Note that by the construction of F we have $S_F(\alpha) = S_f(\alpha)^k e(-\alpha m)$. Our observation is, that Weyl's bound for exponential sums applies to this setting as before and we can bound

$$S_f(\alpha) \ll X^{1+\epsilon} \left(\frac{1}{Q} + \frac{1}{X} + \frac{P}{X^n} \right)^{2^{1-n}}.$$

This yields the minor arc bound

$$|\int_{\mathfrak{m}} S_F(\alpha) d\alpha| \leq \max_{\alpha \in \mathfrak{m}} |S_F(\alpha)^k| \ll X^{n+\epsilon} \left(\frac{1}{Q} + \frac{1}{X} + \frac{P}{X^n} \right)^{k 2^{1-n}}.$$

We choose $Q = X^{\frac{1}{4}}$ and $P = X^{n-\frac{1}{3}}$, so that for $k \geq n 2^{n+1}$ have

$$\int_{\mathfrak{m}} S_F(\alpha) d\alpha \ll X^{k-n-\delta} \ll m^{\frac{k}{n}-1-\delta'}.$$

for some small $\delta > 0$.

We turn to the major arc case. For $\alpha \in \mathfrak{M}$ near $\frac{a}{q}$ we put $\beta = \alpha - \frac{a}{q}$. One writes

$$S_F(\alpha) = \sum_{\mathbf{u} \bmod q} e(\frac{a}{q} F(\mathbf{u})) \sum_{\substack{\mathbf{x} \in \mathcal{B} \cap \mathbb{N}_0^k, \\ \mathbf{x} \equiv \mathbf{u} \bmod q}} e(\beta F(\mathbf{x})).$$

We want to replace the $\mathbf{x}$-sum by the integral

$$\mathcal{B}_f(\beta) = \int_{\mathcal{B}} e(\beta F(\mathbf{x})) d\mathbf{x}.$$

To do so we note that for our choice of P (actually $P \gg_f X^{n-1}$ would do the job) one has

$$|\frac{\partial}{\partial x_v} f(\mathbf{x})| \leq \frac{P}{2} \text{ for } 1 \leq v \leq k \text{ and } \mathbf{x} \in \mathcal{B}.$$

Thus applying Lemma 5.3 in each variable we get

$$\sum_{\substack{\mathbf{x}\in\mathcal{B}\cap\mathbb{N}_0^k,\\ \mathbf{x}\equiv\mathbf{u} \bmod q}} e(\beta F(\mathbf{x})) = q^{-k}\mathcal{B}_F(\beta) + O((1+X/q)^{k-1}).$$

Since the integral is independent of u we get

$$S_F(\alpha) = \mathcal{B}_F(\beta)C_F(\frac{a}{q}) + O(q(q+X)^{k-1}),$$

for the complete sums[6]

$$C_F(\frac{a}{q}) = q^{-k}\sum_{\mathbf{u} \bmod q} e(\frac{a}{q}F(\mathbf{u})).$$

Note the trivial bound $C_F(\frac{a}{q}) \leq 1$.

Integrating this expression over the major arcs yields

$$\int_{\mathfrak{M}} S_F(\alpha)d\alpha = \sum_{q\leq Q} c_F(q)\int_{|\beta|\leq(qP)^{-1}} \mathcal{B}_F(\beta)d\beta + O(P^{-1}Q^2(Q+X)^{k-1}).$$

Here

$$c_F(q) = q^{-k}\sideset{}{^\star}\sum_{a \bmod q}\sum_{\mathbf{u} \bmod q} e(\frac{a}{q}F(\mathbf{u})).$$

We will now replace the integral by

$$V_F(\mathcal{B}) = \int_{\mathbb{R}} \mathcal{B}_F(\beta)d\beta.$$

Put $\gamma = \frac{k}{n} - 1$. Using the special shape of F one obtains the standard bound

$$\mathcal{B}_F(\beta) \ll (\beta X^n)^{-1-\gamma}X^k.$$

This immediately provides the estimate

$$\int_{|\beta|\leq(qP)^{-1}} \mathcal{B}_F(\beta)d\beta = V_F(\mathcal{B}) + O((qPX^{-n})^{\gamma}X^{k-n}).$$

Further, using Weyl's bound (Lemma 5.1) we get

$$C_F(\frac{a}{q}) \ll q^{-k2^{1-n}+\epsilon}.$$

[6] Note that in our special case this k-dimensional sum factors in 1 dimensional ones. More precisely $C_F(\frac{a}{q}) = C_f(\frac{a}{q})^k$ in the obvious notation.

In particular $c_F(q) \ll q^{1-k2^{1-n}+\epsilon}$. For k large, say $k > n2^n$, we have $c_F(q) \ll q^{-1-\eta}$ for some $\eta > 0$. Setting $\mathfrak{S}_F = \sum_{q=1}^{\infty} c_F(q)$ we get

$$\sum_{q\leq Q} c_F(q) = \mathfrak{S}_F + O(Q^{-\eta}).$$

Thus we have seen that

$$\int_{\mathfrak{M}} S_F(\alpha)d\alpha = \mathfrak{S}_F V_F(\mathcal{B}) + \underbrace{O((Q^{-\eta} + Q^2(QPX^{-n})^{\gamma})X^{k-n})}_{=O(X^{k-n-\delta})}.$$

Recall our choice of Q and P from the minor arc estimate to control the error term.

It remains to treat singular series and singular integral appropriately. We start with the singular series. First observe that

$$\sum_{a \bmod q}{}^{\star} e(\frac{a}{q}F(\mathbf{u})) = \sum_{d|q,\, d|F(\mathbf{q})} \mu(\frac{q}{d})d.$$

Inserting this in the definition of $c_F(q)$ we get

$$c_F(q) = q^{-k}\sum_{d|q}\mu(\frac{q}{d})d\cdot(\sharp\{\mathbf{u} \bmod q\colon F(\mathbf{u}) \equiv 0 \bmod d\}).$$

If $d^{k-1}w_F(d)$ is the number of solutions to the congruence $F(\mathbf{x}) \equiv 0 \bmod d$, then we have found

$$c_F(q) = \mu(\frac{q}{d})w_F(d) = [\mu \star w_F](q).$$

In particular μ and w_F are multiplicative, so that c_F is multiplicative. This allows us to write

$$c_F(q) = \prod_{p^{\alpha}\|q}(w_F(p^{\alpha}) - w_F(p^{\alpha-1})).$$

Inserting this in the singular series yields

$$\mathfrak{S}_F = \prod_p(1 + \sum_{\alpha=1}^{\infty}(w_F(p^{\alpha}) - w_F(p^{\alpha-1})) = \prod_p \delta_F(p).$$

This can be seen as a definition of $\delta_F(p)$, but we see straight away that

$$\delta_F(p) = \lim_{\alpha\to\infty} w_F(p^{\alpha}).$$

We turn towards the singular integral. We will do so using the formula

$$\int_{\mathbb{R}} \frac{\sin(2\pi\beta V)}{\pi\beta}e(\beta y)d\beta = \delta_{|y|<V} \text{ for } y \neq V.$$

With this at hand we can compute

$$\begin{aligned}V_F(\mathcal{B}) &= \int_{\mathbb{R}} \mathcal{B}_F(\beta)d\beta \\ &= \int_R \frac{\sin(2\pi\beta V)}{2\pi\beta V}\mathcal{B}_F(\beta)d\beta + O(\int_0^T \beta V|\mathcal{B}_F(\beta)|d\beta + \int_T^\infty (\beta V)^{-1}|\mathcal{B}_F(\beta)|d\beta) \\ &= \frac{1}{2V}\sharp\{\mathbf{x}\in\mathcal{B}\colon |F(\mathbf{x})| \leq V\} + O((TV+(TV)^{-1})(TX^n)^{-\gamma}X^{k-n}).\end{aligned}$$

Here we used our bound on $\mathcal{B}_F(\beta)$ as well as the estimates

$$\sin(2\pi\theta V) = 2\pi\theta V + O(\theta^2V^2) \text{ and } \sin(2\pi\theta V) \ll 1.$$

We choose $T = V^{-1}$ and recalling that $\gamma = \frac{k}{n} - 1$ we get

$$V_F(\mathcal{B}) = \frac{1}{2V}\sharp\{\mathbf{x}\in\mathcal{B}\colon |F(\mathbf{x})| \leq V\} + O(V^\gamma).$$

In particular one could take the limit $V \to 0$ to get a clean formula for the singular integral. For our specific case we can take this computation one step further. Combining the above formula for $V_F(\mathcal{B})$ with the easy bound $V_F(\mathcal{B}) \ll X^{k-n}$ we get

$$\sharp\{\mathbf{x}\in\mathcal{B}\colon |f(x_1)+\ldots+f(x_k)-m| \leq U\} \ll UX^{k-n} \text{ for any } U>0.$$

Since f is monic we have the approximation $f(x) = x^n + O(X^{n-1})$. This gives

$$V_F(\mathcal{B}) = \frac{1}{2U}\sharp\{\mathbf{x}\in\mathcal{B}\colon |x_1^n+\ldots+x_k^n-m| \leq U\} + O(U^{-1}X^{k-1} + (UX^{-n})^\gamma X^{k-n}).$$

Choose $U = X^{k-\frac{1}{\gamma+1}}$ and find

$$V_F(\mathcal{B}) = V(m) + O(m^{\frac{k}{n}-1-\delta}).$$

But $V(m)$ is the same singular integral we evaluated for the classical Waring Problem.

Combining everything we have seen that

$$r_{k,f}(m) = C_{n,k}\mathfrak{S}_F(m)m^{\frac{k}{n}-1} + O(m^{\frac{k}{n}-1-\delta})$$

for some $\delta > 0$ and $k > n2^n$. Here we have $C_{n,k} = \Gamma(1+\frac{1}{n})^k\Gamma(\frac{k}{n})^{-1}$ and

$$\mathfrak{S}_F(m) = \prod_p \delta_F(p).$$

This asymptotic formula becomes only useful if enough information (such as good lower bounds) on the singular series is available. This can be very hard in general. For the special F under consideration here this is doable but we omit the details.

Exercise Block 6. So far we have estimated exponential sums (for example using Weyl's inequality) and we have evaluated/estimated certain complete character sums (for example Gauß sums). Throughout this exercise we will focus a little more on complete character sums.

Let us start by looking at the general sum

$$S(f;q) = \sum_{x \bmod q} e\left(f(x)q^{-1}\right),$$

for $q \in \mathbb{N}$ and $f \in \mathbb{Z}[X]$. Note that one could include a Dirichlet character modulo q as well as rational functions $f \in \mathbb{Z}(X)$, but for our purposes this level of generality suffices.

a) Suppose $q = rs$ with $(r,s) = 1$. Show that

$$S(f;q) = S(\overline{s}f;r)S(\overline{r}f;s).$$

Here $\overline{s}$ is the inverse of s modulo r and similarly $\overline{r}$ is the inverse of r modulo s.

This reduces the problem of understanding sums $S(f;q)$ for prime powers $q = p^l$. Now there are two situations which turn out to be very different in nature. First, if $l = 1$, then we are dealing with sums over finite fields. In this case Weil, using methods from algebraic geometry, and later (Bombieri)-Stephanov, using elementary methods involving auxiliary polynomials, showed the incredible bound

$$S(f;p) \leq (\deg(f) - 1)\sqrt{p}.$$

This shows square root cancellation in the complete sum and is essentially the best one can hope for. We will take this estimate for granted in what follows.

On the other hand, if $l > 1$, there are elementary techniques to handle the sums in question. These resemble the method of stationary phase as known from analysis.

b) Write $l = 2k + \rho > 1$. Define the Gauß sum

$$G_p(a,b) = \sum_{x \bmod p} e((ax^2 + bx)p^{-1}).$$

Show that

$$S(f;p^l) = p^k \sum_{\substack{y \bmod p^k, \\ f'(y) \equiv 0 \bmod p^k}} e(f(y)p^{-2k-\rho}) \cdot \begin{cases} 1 & \text{if } \rho = 0, \\ G_p(\frac{1}{2}f''(y), \frac{f'(y)}{p^{-k}}) & \text{if } \rho = 1. \end{cases}$$

For $(2ab, p) = 1$ we can reduce $G_p(a, b)$ to the classical Gauß sum by completing the square.[7] One then obtains the evaluation

$$G_p(a, b) = \epsilon_p \sqrt{p} \left(\frac{a}{p}\right) e(-\overline{4b} \left(\frac{f'(y)}{p^k}\right)^2 p^{-1}).$$

The upshot of this evaluation is that if the polynomial equation $f'(y) \equiv 0 \bmod p^k$ is well behaved, then we get essentially square root cancellation.

To end this exercise block we will put these tools to good use.

c) Suppose $(a, q) = 1$, then

$$S_k(a, b; q) = \sum_{x=1}^{q} e((ax^k + bx)q^{-1}) \ll_{k,\epsilon} q^{\frac{1}{2}+\epsilon}(q, b).$$

Note that this relies on Weil's bound for exponential sums claimed above. One can proof weaker bounds without it. For example, if $k = 3$, Davenport proved $S_3(a, b; q) \ll q^{\frac{2}{3}}$ in a more elementary fashion.

Solution: To see a) we write

$$\frac{1}{q} \equiv \frac{\overline{s}}{r} + \frac{\overline{r}}{s} \bmod 1.$$

Since f has integer coefficients we end up with

$$S(f; q) = \sum_{x \bmod q} e\left(\overline{s} f(x) r^{-1}\right) e\left(\overline{r} f(x) s^{-1}\right).$$

The claim follows from the Chinese Remainder Theorem.

We turn to b). We write the summation variable x as

$$x = y + zp^k \text{ for } y \in \mathbb{Z}/p^k\mathbb{Z} \text{ and } z \in \mathbb{Z}/p^{k+\rho}\mathbb{Z}.$$

Taylor expanding f at y yields

$$f(x) \equiv f(y) + f'(y)zp^k + \frac{1}{2} f''(y)z^2 p^{2k} \bmod p^{2k+1}.$$

This leads to

$$S(f; q) = \sum_{y \bmod p^k} e(f(y)p^{-2k-\rho}) \sum_{z \bmod p^{k+\rho}} e(f'(y)zp^{-k-\rho} + \frac{1}{2} f''(y)z^2 p^{-\rho}).$$

Here one takes a closer sum at the inner sum. First, if $\rho = 0$, Then we get

$$\sum_{z \bmod p^{k+\rho}} e(f'(y)zp^{-k-\rho} + \frac{1}{2} f''(y)z^2 p^{-\rho}) = \sum_{z \bmod p^k} e(f'(y)zp^{-k-\rho}) = p^k \delta_{f(y) \equiv 0 \bmod p^k}$$

[7] If $p \mid b$, completing the square is not necessary and one can simply drop the exponential in the evaluation below.

by character orthogonality. On the other hand, if $\rho = 1$, then we proceed as follows

$$\begin{aligned}&\sum_{z \bmod p^{k+\rho}} e(f'(y)zp^{-k-\rho} + \frac{1}{2}f''(y)z^2p^{-\rho})\\ &= \sum_{z' \bmod p} e(f'(y)zp^{-k-1} + \frac{1}{2}f''(y)z^2p^{-1}) \sum_{w \bmod p^k} e(f'(y)wp^{-k})\\ &= p^k\delta_{f'(y\equiv 0 \bmod p^k)} \sum_{z' \bmod p} e(f'(y)zp^{-k-1} + \frac{1}{2}f''(y)z^2p^{-1}).\end{aligned}$$

We recognise the remaining z'-sum as a Gaußsum and the argument is complete.

Finally we turn towards c). According to a) it is enough to study $S_k(a, b; p^l)$. Now if $l = 1$, we get

$$S_k(a, b; p) \leq (k-1)\sqrt{p} \ll_k p^{\frac{1}{2}+\epsilon}(p, b)$$

from Weil's bound. Thus we can assume $l > 1$. We write p^θ for the largest power of p dividing b and set $l = 2k+\rho$. Note that if $\theta \geq k+\rho$ (including the case $b = 0$), then the estimate is trivial. Since the implicit constant may depend on k we can also assume that $\tau < \frac{l}{2}$ for $p^\tau \| k$.

We first consider even l (i.e. $\rho = 0$). According to 2) we need to solve the congruence

$$kay^{k-1} + b \equiv 0 \bmod p^k.$$

Let N denote the number of solutions (mod p^k) to this congruence. Then we have

$$S_k(a, b; p^l) \ll Np^k.$$

If the congruence is insoluble, then we are done. This happens unless $\theta \geq \tau$ and $(k-1) \mid \theta - \tau$, which we assume from now on. We put $\lambda = (\theta - \tau)/(k-1)$ and $y = p^\lambda w$ so that the new congruence to consider is

$$(kp^{-\tau})aw^{k-1} + (bp^{-\theta}) \equiv 0 \bmod p^{k-\theta}$$

for $1 \leq w \leq p^{k-\lambda}$. If $k - \lambda \leq k - \theta$ we are done since $N \ll 1$. Otherwise we have $N \ll p^{k-\lambda-(k-\theta)} = p^{\theta-\lambda}$ and we are done.

We turn towards the situation of odd l (i.e. $\rho = 1$). Note that for $f(y) = ax^k+bx$ we have $\frac{1}{2}f''(y) = a\binom{k}{2}$. Thus, if $p \mid \binom{k}{2}$, then the Gauss sum reads

$$G_p(\frac{1}{2}f''(y), \frac{f'(y)}{p^{-k}}) = \sum_{x \bmod p} e(\frac{f'(y)}{p^{-k}}p^{-1}) = p \cdot \delta_{p|\frac{f'(y)}{p^{-k}}}.$$

We obtain

$$S_k(a, b; p^l) = p^{k+1} \sum_{\substack{y \bmod p^k,\\ f'(y)\equiv 0 \bmod p^{k+1}}} e(f(y)p^{-2k-\rho}) \tag{41}$$

One can complete the proof similar as above by counting solutions to the congruence condition. (Here we use that the implicit constant depends on k to absorb the additional factor of p.)

Finally look at $p \nmid \binom{k}{2}$. Here we still have to consider two cases. First we look at those y with $p \mid y$. In these cases the Gauß sum again reduces to a linear sum and we can argue as earlier. Thus it remains to look at those y with $p \nmid y$. For those we have square root cancellation in the Gauß sum so that

$$S_k(a, b; p^l) \leq p^{k+\frac{\varrho}{2}} \sharp\{y \bmod p^k \colon aky^{k-1} + b \equiv 0 \bmod p^k\}.$$

Note that this has only solutions with $p \nmid y$ if $\tau = \theta$. We can count solutions as usual to conclude the proof and the exercise.

Exercise Block 7. We have tried to sketch the derivation of Vinogradov's main conjecture from certain decoupling inequalities. Another approach mainly developed by T. Wooley is based on the so called *efficient congruencing method.* This approach is more number theoretic in nature and has many nice features as well. Nowadays its understood that efficient congruencing can be seen as p-adic version of decoupling. The case $k = 3$ case of the main conjecture was actually established by T. Wooley via his method before the seminal work of Bourgain-Demeter-Guth saw the light of day. In this exercise we explore the efficient congruencing method following the excellent exposition of Heath-Brown (see [He]).

Recall that $J_{s,k}(X)$ is defined to count the number of solutions to

$$x_1^j + \ldots + x_s^j = x_{s+1}^j + \ldots + x_{2s}^j \text{ for } 1 \leq j \leq k \tag{42}$$

with $1 \leq x_1, \ldots, x_{2s} \leq X$. We want to prove the following critical case of Vinogradov's main conjecture for $k = 3$:

$$J_{6,3}(X) \ll X^{6+\epsilon}.$$

We start by introducing some notation. As usual we write

$$f(\alpha) = \sum_{x \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k) \text{ so that } J_{s,k}(X) = \int_{(0,1]^k} |f(\alpha)|^{2s} d\alpha.$$

For a fixed prime $p \geq 5$ and a positive exponent a we write

$$f_a(\alpha, \xi) = \sum_{\substack{x \leq X, \\ x \equiv \xi \bmod p^a}} e(\alpha_1 x + \ldots + \alpha_k x^k).$$

We put

$$I_m(X; \xi, \eta; a, b) = \int_{(0,1]^k} |f_a(\alpha, \xi)|^{2m} |f_b(\alpha, \eta)|^{2(s-m)} d\alpha$$

It will be helpful to observe that this counts solutions to (42) with the additional conditions

$$x_i \equiv \xi \bmod p^a \text{ for } 1 \leq i \leq m \text{ and } s+1 \leq i \leq s+m \text{ and}$$
$$x_i \equiv \eta \bmod p^b \text{ for } m+1 \leq i \leq s \text{ and } s+m+1 \leq i \leq 2s.$$

Note that $I_0(X;\xi,\eta;a,b)$ is obviously independent of ξ and a. We define

$$I_m(X;a,b) = \max_{\eta \not\equiv \xi \bmod p} I_m(X;\xi,\eta;a,b) \text{ and } I_0(X;a,b) = \max_{\eta \bmod p} I_0(X;\xi,\eta;a,b).$$

We also shorthe notation and write $J(X) = J_{6,3}(X)$. It is now an exercise to prove the following preliminary lemmata:

a) If $p^b \leq X$ we have
$$I_0(X;a,b) \leq J(2Xp^{-b}).$$

b) If $p \leq X$ we have
$$J(X) \ll pJ(2Xp^{-1}) + p^{12} I_2(X;1,1).$$

c) We have
$$I_2(X;a,b) \leq I_2(X,b,a)^{\frac{1}{3}} I_1(X;a,b)^{\frac{2}{3}}.$$

d) If $p^b \leq X$ we have
$$I_1(X;a,b) \leq I_2(X;b,a)^{\frac{1}{4}} J(2Xp^{-b})^{\frac{3}{4}}.$$

e) We have
$$I_1(X;a,b) \leq p^{3b-a} I_1(X;3b,b)$$
if $1 \leq a \leq 3b$.

f) If $1 \leq a \leq b$ we have
$$I_2(X;a,b) \leq 2bp^{4(b-a)} I_2(X;2b-a,b).$$

g) If $1 \leq a \leq b$ and $p^b \leq X$ we have
$$I_2(X;a,b) \leq 2bp^{-10a/3+14b/3} I_2(X;b,2b-a)^{\frac{1}{3}} I_2(X;b,3b)^{\frac{1}{3}} J(2Xp^{-b})^{\frac{1}{2}}.$$

With these results at hand we can continue the proof. The lower bound from the diagonal as well as the trivial upper bound tell us

$$X^6 \ll J(X) \ll X^{12}.$$

We define

$$\Delta = \inf\{\delta \in \mathbb{R} \colon J(X) \ll 6 + \delta \text{ for } X \geq 1\}.$$

Of course if we show $\Delta = 0$, then the proof is complete. We note that

$$I_2(X;a,b) \leq J(X) \ll X^{6+\epsilon+\Delta} \leq X^{6+\epsilon+\Delta} p^{-2a-4b} p^{3(3b-a)}.$$

The latter holds for $a \leq b$. Starting from this we will argue by induction and show that

$$I_2(X;a,b) \ll_{\epsilon,a,b,n} X^{6+\Delta+\epsilon} p^{-2a-4b} p^{(3-n\Delta/6)(3b-a)}, \tag{43}$$

for $1 \le a \le b$ and $p^{3^n b} \le X$. Suppose this is done. Then we apply it with $a = b = 1$ and choose p in

$$\frac{1}{2} X^{\frac{1}{3^n}} \le p \le X^{\frac{1}{3^n}}.$$

Note that if we want $p \ge 5$, then this is possible for $X \ge 10^{3^n}$ by Bertrand's Postulate. From b) we get

$$J(X) \ll pJ(2X/p) + p^{12} I_2(X;1,1) \ll p(X/p)^{6+\Delta+\epsilon} + X^{6+\Delta+\epsilon} p^{12-n\Delta/3}.$$

Suppose $\Delta > 0$, then we find n such that $n\Delta \ge 39$. Thus we get

$$J(X) \ll X^{6+\Delta+\epsilon} p^{-1} \ll X^{6+\Delta-3^{-n}+\epsilon}.$$

This contradicts the definition of Δ and we must have $\Delta = 0$.

We are left with showing (43). Since the case $n = 0$ is settled we continue inductively. We assume $p^{3^{n+1}b} \le X$ and use (43) to see

$$I_2(X;b,2b-a) \ll X^{6+\Delta+\epsilon} p^{-2b-4(2b-a)} p^{(3-n\Delta/6)(3(2b-a)-b)} = X^{6+\Delta+\epsilon} p^{4a-10b} p^{(3-n\Delta/6)(5b-3a)}.$$

(Whenever $1 \le a \le b$ also $a \le b \le 2b-a$ and $p^{3^n}(2b-a) \le p^{3^{n+1}b} \le X$.) Similarly (43) we get

$$I_2(X;b,3b) \ll X^{6+\Delta+\epsilon} p^{-14b} p^{(3-n\Delta/6)(8b)}.$$

Finally note that

$$J(2Xp^{b-1}) \ll X^{6+\Delta+\epsilon} p^{-6b-\Delta b},$$

for $p^v \le X$. Using these estimates with g) yields

$$\begin{aligned} I_2(X;a,b) &\le 2bp^{-10a/3+14b/3} (X^{6+\Delta+\epsilon} p^{4a-10b} p^{(3-n\Delta/6)(5b-3a)})^{\frac{1}{3}} \\ &\quad \cdot (X^{6+\Delta+\epsilon} p^{-14b} p^{(3-n\Delta/6)(8b)})^{\frac{1}{3}} (X^{6+\Delta+\epsilon} p^{-6b-\Delta b})^{\frac{1}{2}} \\ &= X^{6+\Delta+\epsilon} p^{-2a-4b} p^{(3-n\Delta/6)(3b-a)} p^{-\Delta b/2} \\ &\le X^{6+\Delta+\epsilon} p^{-2a-4b} p^{(3-n\Delta/6)(3b-a)}. \end{aligned}$$

This is exactly what we claimed and the proof is complete. Admittedly this seems very ad-hoc but there are some nice concepts behind.

Solution: We start with a). Note that $I_0(X;a,b)$ counts those solutions of the Vinogradov-system which are of the shape $x_i = \eta + p^b y_i$. Obviously we have $0 \le y_i \le Xp^{-b}$. Put $z_i = y_i + 1$. Because the system of equations is translation and dilation invariant we see that z_i also solves the equation. Further $1 \le z_1 \le 2Xp^{-b}$ trivially. Thus z_i is counted by $J(2Xp^{-b})$. This directly implies the claim.

We turn to b). We split the solutions to (42) in congruence classes $\xi \equiv \xi_i \bmod p$ for $1 \le i \le 12$. We bound the number of solutions with $\xi_i = \xi_j$ for all $1 \le i, j \le 12$ by

$$\sum_{\eta \bmod p} I_0(X;0,\eta;1,1) \le pI_o(X;1,1) \le pJ(2Xp^{-1}),$$

where we used a) and the definitions. For the remaining number of variables we always have i, j with $\xi_i \neq \xi_j$. This yields

$$J(X) \leq pJ(2Xp^{-1}) + \binom{12}{2} p(p-1) \int_{(0,1]^3} |f_1(\alpha, \xi) f_1(\alpha, \eta) f(\alpha)^{10}| d\alpha.$$

Here we have chose ξ and η to maximise the right hand side. Applying Hölder we get

$$\int_{(0,1]^3} |f_1(\alpha, \xi) f_1(\alpha, \eta) f(\alpha)^{10}| d\alpha \leq \left(\int_{(0,1]^3} |f_1(\alpha, \xi)|^4 |f_1(\alpha, \eta)|^8 d\alpha \right)^{\frac{1}{12}} \\ \cdot \left(\int_{(0,1]^3} |f_1(\alpha, \xi)|^8 |f_1(\alpha, \eta)|^4 d\alpha \right)^{\frac{1}{12}} \left(\int_{(0,1]^3} |f(\alpha)|^{12} d\alpha \right)^{\frac{5}{6}}.$$

We obtain the estimate

$$J(X) \ll pJ(2X/p) + p^2 I_2(X; 1, 1)^{\frac{2}{12}} J(X)^{\frac{5}{6}}.$$

This implies the result by dividing both sides by $J(X)^{\frac{5}{6}}$ and taking 6th powers.

Part c) is a simple application of Hölder:

$$\begin{aligned} I_2(X; \xi, \eta; a, b) &= \int_{(0,1]^3} |f_a(\alpha; \xi)|^4 |f_b(\alpha, \eta)|^8 d\alpha \\ &\leq \left(\int_{(0,1]^3} |f_a(\alpha; \xi)|^8 |f_b(\alpha, \eta)|^4 d\alpha \right)^{\frac{1}{3}} \left(\int_{(0,1]^3} |f_a(\alpha; \xi)|^2 |f_b(\alpha, \eta)|^{10} d\alpha \right)^{\frac{2}{3}} \\ &\leq I_2(X; b, a)^{\frac{1}{3}} I_1(X; a, b)^{\frac{2}{3}}. \end{aligned}$$

Taking the supremum over $\xi \not\equiv \eta \bmod p$ gives the result.

Part d) works similar:

$$\begin{aligned} I_1(X; \xi, \eta; a, b) &= \int_{(0,1]^3} |f_a(\alpha; \xi)|^2 |f_b(\alpha, \eta)|^1 0 d\alpha \\ &\leq \left(\int_{(0,1]^3} |f_a(\alpha; \xi)|^4 |f_b(\alpha, \eta)|^8 d\alpha \right)^{\frac{1}{4}} \left(\int_{(0,1]^3} |f_b(\alpha, \eta)|^{12} d\alpha \right)^{\frac{3}{4}} \\ &\leq I_2(X; b, a)^{\frac{1}{4}} I_0(X; b, b)^{\frac{3}{4}} \leq I_2(X; b, a)^{\frac{1}{4}} J(2Xp^{-b})^{\frac{3}{2}}. \end{aligned}$$

This is as desired.

Let us treat e). Recall that $I_1(X; \xi, \eta; a, b)$ counts solutions to (42) with

$$x_i = \begin{cases} \xi + p^a y_i & \text{if } i = 1, 7, \\ \eta + p^b y_i & \text{else.} \end{cases}$$

We put $\nu = \xi - \eta$ and deduce by translation invariance that the new variables

$$z_i = \begin{cases} \nu + p^a y_i & \text{if } i = 1, 7, \\ p^b y_i & \text{else.} \end{cases}$$

Considering the top equation (degree 3) yields the congruence

$$(\nu + p^a y_1)^3 \equiv (\nu + p^a y_7)^3 \bmod p^{3b}.$$

We now recall that $\xi \not\equiv \eta \bmod p$. Thus $p \nmid \nu$, so that we get $\nu + p^a y_1 \equiv \nu + p^a y_7 \bmod p^{3b}$. At then end this yields $y_1 \equiv y_7 \bmod p^{3b-a}$. We conclude that there is a unique ξ' out of p^{3b-a} possibilities, so that $x_1 \equiv x_7 \equiv \xi' \bmod p^{3b}$. We obtain the upper bound

$$I_1(X; \xi, \eta, a, b) \leq p^{3b-a} I_1(X; 3b, b).$$

The claimed inequality follows by taking the maximum.

The hardest part will now be to deduce f). Again we look at the solutions to (42) counted by $I_2(X; \xi, \eta; a, b)$. These satisfy

$$x_i = \begin{cases} \xi + p^a y_i \text{ for } i = 1, 2, 7, 8, \\ \eta + p^b y_i \text{ else.} \end{cases}$$

Again we write $p \nmid \nu = \xi - \eta$ and find new solutions $z_i = x_i - \eta$. This directly yields the congruence

$$(\eta + p^a y_1)^j + (\eta + p^a y_2)^j \equiv (\eta + p^a y_7)^j + (\eta + p^a y_8)^j \bmod p^{jb},$$

for $j = 1, 2, 3$. We put $S_j = y_1^j + y_2 j - y_7^j - y_8^j$. Using the congruence above for $j = 2, 3$ we find that

$$\begin{aligned} &2\nu S_1 + paS_2 \equiv 0 \bmod p^{2b-a} \text{ and} \\ &3\nu^2 S_1 + 3\nu p^a S_2 + p^{2a} S_3 \equiv 0 \bmod p^{3b-a}. \end{aligned}$$

Combining these in order to eliminate S_1 yields

$$3\nu p^a S_2 + 2p^{2a} S_3 \equiv 0 \bmod p^{2b-a}.$$

We divide out p^a and are left with the two congruences

$$3\nu S_2 + 2p^a S_3 \equiv 0 \bmod p^{2b-2a} \text{ and } 2\nu S_1 + 2p^a S_2 \equiv 0 \bmod p^{2b-2a}.$$

To continue we need the following result, which we prove at the end:

Let $N(p; a, c)$ denote the number of solutions (y_1, y_2, y_7, y_8) to

$$3\nu S_2 + 2p^a S_3 \equiv 2\nu S_1 + 2p^a S_2 \equiv 0 \textit{ mod } p^c. \tag{44}$$

Then if $a \geq 1$ and $c \geq 0$ we have $N(p; a, c) \leq (c+1)p^{2c}$.

With this at hand we continue as follows. Suppose $y_i \equiv y_{i0} \bmod p^{2b-2a}$ for $i = 1, 2, 7, 8$ then $x_i \equiv \xi_i \bmod p^{2b-a}$ for $\xi_i = \xi + p^a y_{i,0}$. The key observation is that the contribution of those y_i's is given by

$$\begin{aligned}
&\int_{(0,1]^3} f_{2b-a}(\alpha, \xi_1) f_{2b-a}(\alpha, \xi_2) \overline{f_{2b-a}(\alpha, \xi_7) f_{2b-a}(\alpha, \xi_8)} |f_b(\alpha, \eta)|^8 d\alpha \\
&\qquad \leq \int_{(0,1]^3} | \prod_{i=1,2,7,8} f_{2b-a}(\alpha, \xi_i) | |f_b(\alpha; \eta)|^8 d\alpha \\
&\qquad \leq \prod_{i=1,2,7,8} \left(\int_{(0,1]^3} |f_{2b-a}(\alpha, \xi_i)|^4 |f_b(\alpha, \eta)|^8 \right)^{\frac{1}{4}} \\
&\qquad \leq \prod_{i=1,2,7,8} I_2(X, \xi_i, \eta; 2b-a, a)^{\frac{1}{4}} \\
&\qquad \leq I_2(X; 2b-a, a).
\end{aligned}$$

Counting all possible y_{i0} we use the claim above. Collecting everything together yields

$$I_2(X; a, b) \leq N(p; a, 2(b-a)) I_2(X; 2b-a, b) \leq 2bp^{4(b-a)} I_2(X; 2b-a, a)$$

as desired.

It remains to proof (44). This is done by induction on c. Note that the case $c = 0$ is trivial. If $c = 1$ we have $p \mid S_1$ and $p \mid S_2$. We get $2p^2 - p$ solutions and are done. We turn towards the general case. We call a solution (y_1, y_2, y_7, y_8) singular if $y_1 \equiv y_2 \equiv y_7 \equiv y_8 \bmod p$. The remaining solutions are obviously called non-singular. We observe that for a non-singular solution the vectors

$$\nabla(2\nu S_1 + p^a S_2) \text{ and } \nabla(3\nu S_2 + 2p^a S_3)$$

are not proportional modulo p. The upshot is that a non-singular solution modulo p^c will lift to precisely p^2 solutions modulo p^{c+1}. We therefore write $N_0(p; a, c)$ for the number of non-singular solutions modulo p^c. By induction we easily see that $N_0(p; a, c) \leq 2p^{2c}$. It remains to estimate the non-singular solutions. We set

$$y_1 \equiv y_2 \equiv y_7 \equiv y_8 \equiv \beta \bmod p.$$

We set $y_i = \beta + pu_i$ and $S'_j = u_1^j + u_2^j - u_7^j - u_8^j$. Of course we have

$$2\nu S_1 + p^a S_2 = 2(\nu + \beta p^a) p S'_1 + p^{a+2} S'_2.$$

Similarly we get

$$3\nu S_2 + 2p^a S_3 = 6\beta(\nu + \beta p^a) p S'_1 + 3(\nu + 2\beta p^a) p^2 S'_2 + 2p^{a+3} S'_3.$$

We set $\nu' = \nu + \beta p^a$. Eliminating S'_1 from the second equation we get the two congruences

$$2\nu' S'_1 + p^{a+1} S'_2 \equiv 0 \bmod p^{c-1} \text{ and } 3\nu' S'_2 + 2p^{a+1} S'_3 \equiv 0 \bmod p^{c-2}.$$

We have to count solutions u_i modulo p^{c-1}. Thus we first view the two congruences above as a system of congruences modulo p^{c-2}. The solutions of this system are precisely counted by $N(p;a+1,c-2)$. We have to investigate how such a solution lifts to solutions modulo p^{c-1}. To do so we note that

$$\nabla(2\nu' S_1' + p^{a+1}S_2') \equiv 2\nu'(1,1,-1,-1) \not\equiv 0 \bmod p.$$

Thus each solution lifts to p^3 solutions of the two congruences modulo p^{c-1}. Taking all possibilities for β into account we get at must $p^4N(p;a+1,c-2)$ singular solutions. Summing things up we combine singular and non-singular solutions to get

$$N(p;a,c) \leq N_0(p;a,c) + p^4N(p;a+1,c-2) \leq 2p^{2c} + p^4(c-1)p^{2c-4} = (c+1)p^{2c}.$$

This completes the proof of (44) and the proof of f).

To see part g) we use f), c), e) and d) to obtain the desired inequality:

$$\begin{aligned} I_2(X;a,b) &\leq 2bp^{4(b-a)}I_2(X;2b-a,b)\\ &\leq 2bp^{4(b-a)}I_2(X,b,2b-a)^{\frac{1}{3}}I_1(X;2b-a,b)^{\frac{2}{3}}\\ &\leq 2bp^{4(b-a)}I_2(X,b,2b-a)^{\frac{1}{3}}\left(p^{3b-(2b-a)}I_1(X;3b,b)\right)^{\frac{2}{3}}\\ &\leq 2bp^{4(b-a)+2(a+b)/3}I_2(X,b,2b-a)^{\frac{1}{3}}\left(I_2(X;b,3b)^{\frac{1}{4}}J(2Xp^{-b})^{\frac{3}{4}}\right)^{\frac{2}{3}}\\ &= 2bp^{-10a/3+14b/3}I_2(X,b,2b-a)^{\frac{1}{3}}I_2(X;b,3b)^{\frac{1}{6}}J(2Xp^{-b})^{\frac{1}{2}}. \end{aligned}$$

Exercise Block 8. What are the best bounds for $G_1(k)$ we can prove with the tools from these lectures?

Solution: We consider the following situation. Take $k \geq 2$ and $s > k+1$. Suppose that N is sufficiently large and let ψ be a function that is monotonically increasing to infinity but satisfies $\psi(t) \ll t^{\delta}$ for δ sufficiently small.

We write $\mathcal{Z}_{s,k}(N)$ for the set of positive integers $\frac{N}{2} < n \leq N$ with

$$|r_s(n) - \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(\frac{s}{k})}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1}| > n^{\frac{s}{k}-1}\psi(n)^{-1}.$$

We write $Z = \sharp\mathcal{Z}_{s,k}(N)$ and $X = N^{\frac{1}{k}}$.

We claim that there are numbers $\eta_n(s,k) \in S^1$ such that

$$\int_{\mathfrak{m}} |T(\alpha)^s K(\alpha)|d\alpha \gg N^{\frac{s}{k}-1}\psi(N)^{-1}Z \tag{45}$$

for

$$K(\alpha) = \sum_{n\in\mathcal{Z}_{s,k}(N)} \eta_n(s,k)e(n\alpha).$$

Here we use the same minor arcs as defined around (30).

If we take the major arcs $\mathfrak{M}$ in accordance with the minor arcs above we have seen that

$$\int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1} + O(n^{\frac{s}{k}-1-2\delta}).$$

Note that we have

$$\begin{aligned}
n^{\frac{s}{k}-1}\psi(n)^{-1} &< |r_s(n) - \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(\frac{s}{k})}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1}| \\
&= |\int_0^1 f(\alpha)^s e(-n\alpha)d\alpha - \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(\frac{s}{k})}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1}| \\
&\leq |\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha| + |\int_{\mathfrak{M}} T(\alpha)^s e(-n\alpha)d\alpha - \frac{\Gamma(1+\frac{1}{k})^s}{\Gamma(\frac{s}{k})}\mathfrak{S}_{s,k}(n)n^{\frac{s}{k}-1}| \\
&= |\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha| + O(n^{\frac{s}{k}-1-2\delta}).
\end{aligned}$$

Thus, for $\frac{N}{2} < n \leq N$ and N sufficiently large we have

$$|\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha| \geq \frac{1}{2}n^{\frac{s}{k}-1}\psi(n)^{-1}.$$

We define the desired numbers $\eta_n(s,k)$ by requiring

$$|\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha| = \eta_n(s,k)\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha \tag{46}$$

when $n \in \mathcal{Z}_{s,k}(N)$. We may assume that that $\eta_{s,k}(n)$ vanishes when n is not in the exceptional set $\mathcal{Z}_{s,k}(N)$. With choice we get

$$N^{\frac{s}{k}-1}\psi(N)^{-1}\cdot Z \ll \sum_{\frac{N}{2}<n\leq N} \eta_{s,k}(n)\int_{\mathfrak{m}} T(\alpha)^s e(-n\alpha)d\alpha = \int_{\mathfrak{m}} T(\alpha)^s \underbrace{\left(\sum_{\frac{N}{2}<n\leq N} \eta_{s,k}(n)e(-n\alpha)\right)}_{=K(\alpha)} d\alpha$$

which is what we claimed.

Using Cauchy-Schwarz in (45) yields

$$N^{\frac{s}{k}-1}\psi(N)^{-1}Z \ll \left(\int_{\mathfrak{m}} |T(\alpha)|^{2s}d\alpha\right)^{\frac{1}{2}}\left(\int_0^1 |K(\alpha)|^2 d\alpha\right)^{\frac{1}{2}}.$$

By Parseval we have

$$\int_0^1 |K(\alpha)|^2 d\alpha = \sum_{n\in\mathcal{Z}_{s,k}(N)} 1 = Z.$$

From our minor arc estimate in Theorem 9.4 we get

$$\int_{\mathfrak{m}} |T(\alpha)|^{2s} \ll X^{2s-k-\delta},$$

for

$$2s \geq k^2 + 1 - \max_{j \leq k} \left\lceil j \frac{k-j-1}{k-j+1} \right\rceil .$$

Combining both yields

$$Z^{\frac{1}{2}} \ll N^{1-\frac{s}{k}} \psi(N) X^{s-\frac{k}{2}-\frac{\delta}{2}} = (N^{1-\frac{\delta}{k}} \psi(N)^2)^{\frac{1}{2}}.$$

Thus, if we assume that ψ grows slowly enough we obtain

$$Z \ll N^{1-\delta'} \text{ for some small } \delta' > 0.$$

From this we obtain directly that for such s the number or n which are not representable by a sum of s positive kth powers has density 0 (i.e. there are $o(N)$ such exceptional integers smaller than N). We get

$$G_1(k) \leq \frac{k^2+1}{2} - \frac{1}{2} \max_{j \leq k} \left\lceil j \frac{k-j-1}{k-j+1} \right\rceil .$$

References

[Ap] Apostol, Tom M. Modular functions and Dirichlet series in number theory. Second edition. Graduate Texts in Mathematics, 41. Springer-Verlag, New York, 1990.

[BGD] Bourgain, Jean; Demeter, Ciprian; Guth, Larry. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. Ann. of Math. (2) 184 (2016), no. 2, 633–682.

[Ch] Chintschin, A. J. Drei Perlen der Zahlentheorie. (German) [[Three pearls of number theory]] Reprint of the 1951 translation from the Russian. With a foreword by Helmut Koch. Akademie-Verlag, Berlin, 1984.

[Da1] Davenport, H. On Waring's problem for cubes. Acta Math. 71 (1939), 123–143.

[Da2] Davenport, H. Analytic methods for Diophantine equations and Diophantine inequalities. Second edition. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman. Edited and prepared for publication by T. D. Browning. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2005.

[He] Heath-Brown, D. R. The cubic case of Vinogradov's mean value theorem. Essent. Number Theory 1 (2022), no. 1, 1–12.

[IK] Iwaniec, Henryk; Kowalski, Emmanuel. Analytic number theory. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[Pi] Pierce, Lillian B. The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth]. Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135. Astérisque 2019, no. 407, Exp. No. 1134, 479–564.

[Va] Vaughan, R. C. The Hardy-Littlewood method. Second edition. Cambridge Tracts in Mathematics, 125. Cambridge University Press, Cambridge, 1997.

[Wo] Wooley, Trevor D. The asymptotic formula in Waring's problem. Int. Math. Res. Not. IMRN 2012, no. 7, 1485–1504.

Email address: `assing@math.uni-bonn.de`