A density version of Waring’s problem

by

JUHO SALMENSUU (Turku)

1. Introduction

1.1. Statements of results. In this paper, we investigate when a positive density subset of $k$th powers forms an asymptotic additive basis. This problem is motivated by similar results related to Goldbach’s problem [LP10], [Sha14]. For example Shao [Sha14] proved that if $A$ is a subset of the primes, and the lower density of $A$ in the primes is larger than $5/8$, then all sufficiently large odd positive integers can be written as the sum of three primes in $A$. The key to studying this kind of problems is the transference principle introduced by Green [Gre05].

Let $k \geq 2$ be an integer. Set

$$\mathbb{N}(k) := \{t^k : t \in \mathbb{N}\} \quad \text{and} \quad \mathbb{Z}_m(k) := \{t^k : t \in \mathbb{Z}_m\},$$

where $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$. Let $A \subseteq \mathbb{N}(k)$. Define

$$\delta_A = \delta(A) := \liminf_{N \to \infty} \frac{|A \cap [N]|}{|\mathbb{N}(k) \cap [N]|},$$

where $[N] := \{1, \ldots, N\}$. For $m \in \mathbb{N}$ let $P(m) := \prod_{p \leq m} p^k$ and

$$Z_k := \lim_{m \to \infty} \frac{|\mathbb{Z}_P(m)(k)|}{|\{a \in \mathbb{Z}_P(m)(k) : (a, P(m)) = 1\}|}. \quad (1)$$

We prove later that $\lim_{k \to \infty} Z_k = 1$.

For each prime $p$ and $k \in \mathbb{N}$, define $\tau(k, p)$ so that $p^{\tau(k, p)} \parallel k$, where $p^h \parallel k$ means that $p^h \mid k$ and $p^{h+1} \nmid k$. Let

$$R_k := \prod_{(p-1) \mid k} p^{\eta(k, p)}, \quad (2)$$

2020 Mathematics Subject Classification: Primary 11P05; Secondary 11B13.
Key words and phrases: Waring’s problem, transference principle.
Received 1 June 2020; revised 12 January 2021.
Published online 31 May 2021.

DOI: 10.4064/aa200601-1-2
where
\[
\eta(k, p) := \begin{cases} 
\tau(k, p) + 2 & \text{if } p = 2 \text{ and } \tau(k, p) > 0, \\
\tau(k, p) + 1 & \text{otherwise.}
\end{cases}
\]

For \( n \in \mathbb{N} \) let \( \omega(n) \) denote the number of distinct prime divisors of \( n \).

Our main result is the following.

**Theorem 1.1.** Let \( s, k \in \mathbb{N}, k \geq 2, s > \max(16k\omega(k) + 4k + 3, k^2 + k) \) and let \( A \subseteq \mathbb{N}^{(k)} \) be such that \( \delta(A) > (1 - Z^{-1/2}k^{-1/2})^{1/k} \). Then, for all sufficiently large integers \( n \equiv s \pmod{R_k} \), we have \( n \in sA \).

In the proof of the last theorem, due to some technical difficulties, we need to restrict to the elements of \( A \) which do not have small prime factors. This leads to the congruence condition in the theorem. We expect that this condition can be removed. In the following corollary, we have done so at the cost of some extra summands.

**Corollary 1.2.** Let \( s, k \in \mathbb{N}, k \geq 2, s > \max(16k\omega(k) + 4k + 3, k^2 + k) \) and \( s \geq \max p \in P \frac{p-1}{p} \) for all \( p \in P \). Hence \( sA = \mathbb{Z^k} \) for all \( s \geq \max p \in P \frac{p-1}{p} \). The rest now follows from Theorem 1.1.

The density condition in Theorem 1.1 is not optimal. We expect the result to hold as long as \( A \) is not a subset of a non-trivial arithmetic progression. The set \( A \subseteq \mathbb{N}^{(k)} \) can be contained in a non-trivial arithmetic progression if and only if
\[
\delta(A) \leq \max_{p} \frac{\max_a |\{b \in [p] : b^k \equiv a \pmod{p}\}|}{p} = \max_{p} \frac{p}{p - 1}.
\]

The following theorem shows that the condition \( \delta(A) > \max p \in P \frac{p-1}{p} \) can be guaranteed if we assume that the number of summands is very large depending on \( k \).

**Theorem 1.3.** Let \( k \geq 2 \) and \( \delta > 0 \). Let \( A \subseteq \mathbb{N}^{(k)} \) be such that \( A \) is not a subset of any non-trivial arithmetic progression and \( \delta(A) > \delta \). There

\( ^{(1)} \) By non-trivial we mean that the common difference of the arithmetic progression is not 1.
exists \( s = s(k, \delta) \in \mathbb{N} \) such that all sufficiently large natural numbers belong to the set \( sA \).

### 1.2. Outline of the proof of Theorem 1.1

We prove Theorem 1.1 using the transference principle, which we introduce in Section 3.

Let \( f \) be, roughly speaking, the characteristic function of the set \( A \) in Theorem 1.1. In order for the transference principle to work we need \( f \) to satisfy three conditions: 1) \( f \) satisfies a sufficient mean condition. 2) \( f \) has a pseudorandom majorant function. 3) \( f \) satisfies a suitable restriction estimate. We establish these conditions in Sections 6, 7 and 8 respectively.

Both the pseudorandomness condition and the restriction estimate can be proved with the standard circle method machinery with minor alterations. The mean condition also follows from simple calculations.

Another main ingredient in the proof of Theorem 1.1 is solving the local density version of Waring’s problem. Essentially we want to prove that if \( A \subseteq \mathbb{Z} := \{ a \in \mathbb{Z}^{(k)} : (a, P(w)) = 1 \} \) and \( |A| > \frac{1}{2}|\mathbb{Z}| \), then \( sA = \mathbb{Z}_{P(w)} \) for some suitably large \( s \) depending on \( k \), where \( P(w) = \prod_{p \leq w} p^k \) and \( w \in \mathbb{N} \). We prove this in Section 5. This is done using the Chinese remainder theorem, Hensel’s lemma and the Cauchy–Davenport theorem.

**Remark 1.** The transference lemma (Proposition 3.9) gives us the limitation \( \delta_A > 2^{-1/k} \). Our result (Theorem 1.1) comes close to this when \( k \) is sufficiently large. In particular, for small \( k \), we have some density loss because it is not possible to prove the pseudorandomness condition for \( f_b \) (for the definition of \( f_b \) see (9)), when \( (W, b) > 1 \): If \( (W, b) > 1 \), we will eventually lose \( w \)-smoothness of \( W \) in calculations, which is crucial for proving pseudorandomness. There is a way to define \( f_b \) so that it satisfies the pseudorandomness condition for all \( b \in \mathbb{Z}^{(k)}_W \), but this leads to a significantly more difficult local problem, which we have not been able to solve.

### 2. Notation

For the rest of the paper we assume that \( k \geq 2 \) is a fixed integer.

Let \( s \in \mathbb{N} \) and \( s \geq 2 \). For a set \( A \subseteq \mathbb{N} \) we define the sumset by

\[
sA = \{a_1 + \cdots + a_s : a_1, \ldots, a_s \in A\}.
\]

For any integers \( q, b \), we define

\[
b + A = \{b\} + A \quad \text{and} \quad q \cdot A = \{qa : a \in A\}.
\]

For finitely supported functions \( f, g : \mathbb{Z} \to \mathbb{C} \), we define their convolution \( f \ast g \) by

\[
f \ast g(n) = \sum_{a+b=n} f(a)g(b).
\]
For a set $A$, write $1_A(x)$ for its characteristic function. Let $A, B \subseteq [N]$ and $\eta > 0$. We define $S_\eta(A, B)$ by

$$S_\eta(A, B) = \{ n : 1_A * 1_B(n) \geq \eta N \}.$$ 

The Fourier transform of a finitely supported function $f : \mathbb{Z} \to \mathbb{C}$ is defined by

$$\hat{f}(\alpha) = \sum_{n \in \mathbb{Z}} f(n) e(-n\alpha)$$

where $e(x) = e^{2\pi ix}$. We will also use the notation $e_W(n)$ for $e(n/W)$.

Let $f : \mathbb{R} \to \mathbb{C}$ and $g : \mathbb{R} \to \mathbb{R}_+$. We write $f = O(g)$ or $f \ll g$ if there exists a constant $C > 0$ such that $|f(x)| \leq Cg(x)$ for all $x$ in the domain of $f$. If $f$ takes only positive values then similarly $f \gg g$ means that there exists a constant $C > 0$ such that $f(x) \geq Cg(x)$ for all $x$ in the domain of $f$. If the implied constant $C$ depends on some constant $\epsilon$ we use the notations $O_\epsilon, \ll_\epsilon, \gg_\epsilon$. If $f \ll g$ and $f \gg g$ we write $f \asymp g$. We also write $f = o(g)$ if

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

The function $f$ is asymptotic to $g$, denoted $f \sim g$, if

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1.$$ 

We will write $\mathbb{T}$ for $\mathbb{R}/\mathbb{Z}$. We also define the $L^p$-norm

$$\|f\|_p = \left( \int_\mathbb{T} |f(\alpha)|^p \, d\alpha \right)^{1/p}$$

for a function $f : \mathbb{T} \to \mathbb{C}$.

3. Transference principle. In this section, we apply the transference principle to prove the transference lemma (Proposition 3.9 below), which we use to prove our main theorem. The idea is to transfer an additive combinatorial result from the integers to a sparse subset of the integers. In particular, these sparse subsets should be pseudorandom.

3.1. The sumset problem in dense settings. In this subsection, we prove a sumset result, where the sets involved are positive density sets of natural numbers. We later transfer the solution of this problem, using the transference principle, to the solution of our sparse problem (the density version of Waring’s problem).

We need the following lemma from [GR05, Corollary 6.2], which is a quantitative version of the Cauchy–Davenport theorem.
Lemma 3.1. Let $\eta > 0$ and $p$ be a prime. Let $A, B \subseteq \mathbb{Z}_p$ and $|A|, |B| \geq \sqrt{\eta}p$. Then
\[ |S_{\eta}(A, B)| \geq \min(p, |A| + |B| - 1) - 3\sqrt{\eta}p. \]

Using the previous lemma inductively we will prove the following result.

Lemma 3.2. Let $p$ be a prime, $s \in \mathbb{N}$, $s \geq 2, \epsilon > 2s/p$ and let $B_1, \ldots, B_s \subseteq \mathbb{Z}_p$ be such that $\sum_i |B_i| > (1 + \epsilon)p$ and $|B_i| > (\epsilon/s)p$ for all $i \in \{1, \ldots, s\}$. Then, for all $n \in \mathbb{Z}_p$,
\[ 1_{B_1} \cdots 1_{B_s}(n) \geq \epsilon/s p^{s-1}. \]

Proof. Let $\eta = \epsilon/(6s^2)$ and
\[ R_1 := B_1, \quad R_{i+1} := S_{\eta^2}(R_i, B_{i+1}) \quad \text{for} \quad i \in \{1, \ldots, s - 1\}. \]

It follows from Lemma 3.1 that
\[ |R_2| = |S_{\eta^2}(B_1, B_2)| \geq \min(p, |B_1| + |B_2| - 1) - 3\eta p. \]

Similarly, \[ |R_3| = |S_{\eta^2}(R_2, B_3)| = \min(p, |R_2| + |B_3| - 1) - 3\eta p \]
\[ \geq \min(p, |B_1| + |B_2| + |B_3| - 2) - 6\eta p. \]

Repeating this argument inductively, for each $i \in \{1, \ldots, s - 2\}$, we get
\[ |R_{s-1}| \geq \min\left(p, \sum_{1 \leq i \leq s - 1} |B_i| - (s - 2)\right) - 3(s - 2)\eta p. \]

For $n_0 \in \mathbb{N}$ let $N(n_0) := |\{(a, b) \in R_{s-1} \times B_s : a + b \equiv n_0 \pmod{p}\}|$. We see that
\[ N(n_0) = |B_s \cap (n_0 - R_{s-1})| = |B_s \setminus (\mathbb{Z}_p \setminus (n_0 - R_{s-1}))|. \]

Hence
\[ N(n_0) \geq |R_{s-1}| - (p - |B_s|) \]
\[ \geq \min\left(p + |B_s|, \sum_{i=1}^{s} |B_i| - (s - 2)\right) - 3(s - 2)\eta p - p \]
\[ > \min(p + (\epsilon/s)p, (1 + \epsilon)p - s) - 3s\eta p - p = \epsilon' p, \]

where $\epsilon' = \min(\epsilon/s - 3s\eta, \epsilon - (3s\eta + s/p)) > \epsilon/4$. Now, for all $n \in \mathbb{Z}_p$,
\[ 1_{B_1} \cdots 1_{B_s}(n) \geq \sum_{a+b=n} 1_{B_1} \cdots 1_{B_s-1}(a)1_{B_s}(b) \]
\[ \geq \epsilon' p \min_{b \in B_s} 1_{B_1} \cdots 1_{B_{s-1}}(b) \]
\[ \geq \epsilon' p \min_{b \in R_{s-1}} \sum_{i+j=b} 1_{B_1} \cdots 1_{B_{s-2}}(i) \geq \epsilon' \eta^2 p \min_{i \in R_{s-2}} 1_{B_1} \cdots 1_{B_{s-2}}(i). \]
Repeating the last two steps in the above argument $s - 3$ times yields

$$1_{B_1} \ast \cdots \ast 1_{B_s}(n) \geq \epsilon' \eta^{2(s-2)} p^{s-1}.$$  

Now we are ready to prove the following sumset lemma.

**Lemma 3.3.** Let $\epsilon > 0$, $s \geq 2$ and let $A_1, \ldots, A_s \subseteq [N]$ be such that $\sum_i |A_i| > (s(1+\epsilon)/2)N$ and $|A_i| > (\epsilon/2)N$ for all $i \in \{1, \ldots, s\}$. Then there exists $c(\epsilon, s) > 0$ such that, for all $n \in ((1 - \frac{\epsilon^2}{16}) \frac{sN}{2}, (1 + \frac{\epsilon}{4}) \frac{sN}{2})$,

$$1_{A_1} \ast \cdots \ast 1_{A_s}(n) \geq c(\epsilon, s)N^{s-1}$$

provided that $N$ is sufficiently large depending on $\epsilon$.

**Proof.** Let $p$ be a prime such that $p \in ((1+\kappa)sN/2, (1+2\kappa)sN/2)$, where $\kappa = \epsilon/4$. Such a prime exists by the prime number theorem provided that $N$ is large enough depending on $\epsilon$. For $i \in \{1, \ldots, s\}$ define $B_i \subseteq \mathbb{Z}_p$ by $B_i := \{a \pmod{p}: a \in A_i\}$. We see that

$$\sum_{i=1}^s |B_i| = \sum_{i=1}^s |A_i| > \frac{s(1+\epsilon)}{2}N = \frac{1+4\kappa}{2} sN > (1+\kappa')p,$$

where $\kappa' = \frac{2\kappa}{1+2\kappa}$. Similarly $|B_i| > (\kappa'/s)p$ for all $i \in \{1, \ldots, s\}$. Assuming that $N$ is sufficiently large depending on $\epsilon$, we find that $\kappa' > 2s/p$. Hence it follows from Lemma 3.3 that, for any $n \in \mathbb{Z}_p$,

$$1_{B_1} \ast \cdots \ast 1_{B_s}(n) \gg_{\epsilon,s} p^{s-1} \gg_{\epsilon,s} N^{s-1}.$$

For each $n \in A_1 + \cdots + A_s$ we have $n \leq sN < \frac{2}{1+\kappa}p$. On the other hand, for $n \in (\frac{1-\kappa}{1+\kappa}p, p)$, we have $p + n > \frac{2}{1+\kappa}p$. Thus, for $n \in (\frac{1-\kappa^2}{2} sN, \frac{1+\kappa}{2} sN)$,

$$1_{B_1} \ast \cdots \ast 1_{B_s}(n) = 1_{A_1} \ast \cdots \ast 1_{A_s}(n)$$

and the claim follows.  

**3.2. Transference.** In this subsection, we establish the transference lemma, which we will use to prove our main theorem. But first we introduce some definitions.

**Definition 3.4.** Let $\eta > 0$ and $N \in \mathbb{N}$. We say that a function $f : [N] \to \mathbb{R}_{\geq 0}$ is $\eta$-pseudorandom if there exists a majorant function $\nu_f$ such that $f \leq \nu_f$ pointwise and $\|\hat{\nu}_f - \hat{1}_{[N]}\|_\infty \leq \eta N$.

**Definition 3.5.** Let $q > 1$, $N \in \mathbb{N}$ and $K \geq 1$. We say that a function $f : [N] \to \mathbb{R}_{\geq 0}$ is $q$-restricted with constant $K$ if $\|\hat{f}\|_q \leq KN^{1-1/q}$.

**Definition 3.6.** Let $\delta > 0$ and $N \in \mathbb{N}$. We say that $f : [N] \to \mathbb{R}$ is $\delta$-uniform if $\|\hat{f}\|_\infty \leq \delta N$.

Let $N \in \mathbb{N}$, $\delta > 0$ and $f : [N] \to \mathbb{R}_{\geq 0}$ be a function. Let $T$ be the set of large frequencies of $f$:

$$T := \{\gamma \in \mathbb{T} : |\hat{f}(\gamma)| \geq \delta N\}.$$
We also define a Bohr set using these frequencies:

\[ B(\delta, N) = \{ 1 \leq b \leq \delta N : \| b\gamma \|_T < \delta/2\pi, \forall \gamma \in T \}. \]

For \( N, \delta, f \) as above we define

\[ f_{\delta,N}^*(n) := \mathbb{E}_{a,b \in B} f(n + a - b) \quad \text{and} \quad f_{\delta,N}^{\text{unf}} := f - f_{\delta,N}^*. \]

Now we can state the following lemma that is the core of the transference principle.

**Lemma 3.7.** Let \( \delta > 0, N, \eta \in \mathbb{N} \) and \( K \geq 1 \). Let \( f : [N] \to \mathbb{R}_{\geq 0} \) be \( \eta \)-pseudorandom and \( q \)-restricted with constant \( K \). Then

(i) \( 0 \leq f_{\delta,N}^*(n) \leq 1 + O_\delta(\eta) \) for all \( n \in [N] \),

(ii) \( f_{\delta,N}^{\text{unf}} \) is \( \delta \)-uniform,

(iii) \( f_{\delta,N}^* \) and \( f_{\delta,N}^{\text{unf}} \) are \( q \)-restricted with constant \( K \).

**Proof.** See [MMS17, proof of Lemma 4.3].

Next, we prove that the functions \( f_1 \cdots f_s \) and \( (f_1)^*_{\delta,N} \cdots (f_s)^*_{\delta,N} \) are in a certain sense close to each other.

**Lemma 3.8.** Let \( \delta > 0, \eta > 0, N \in \mathbb{N} \) and \( K \geq 1 \). Let also \( s \in \mathbb{N}, q \in (s - 1, s) \) and, for each \( i \in \{ 1, \ldots, s \} \), let \( f_i : [N] \to \mathbb{R}_{\geq 0} \) be \( \eta \)-pseudorandom and \( q \)-restricted with constant \( K \). Then, for all \( n \in [N] \),

\[ |f_1 \cdots f_s(n) - (f_1)^*_{\delta,N} \cdots (f_s)^*_{\delta,N}(n)| \leq 2^{s-\delta q} N^{1-q}. \]

**Proof.** Denote \( f_i^{\text{unf}} = (f_i)^*_{\delta,N} \) and \( f_i^* = (f_i)^*_{\delta,N} \) for all \( i \in \{ 1, \ldots, s \} \). We see that

\[ f_1 \cdots f_s(n) = f_1^* \cdots f_s^*(n) + \sum_{g_i \in \{ f_i^*, f_i^{\text{unf}} \} \atop \exists i, g_i = f_i^{\text{unf}}} g_1 \cdots g_s(n). \]

Now choose \( a = q - s + 1 \in (0, 1) \). Let \( i \in \{ 1, \ldots, s \} \) be such that \( g_i = f_i^{\text{unf}} \); we can assume that \( i = 1 \). By Hölder’s inequality and Lemma 3.7, we have

\[ |g_1 \cdots g_s(n)| \leq \int_\mathbb{T} |g_1(\gamma) \cdots g_s(\gamma)| d\gamma \]

\[ \leq \| f_1^{\text{unf}} \|_1^{1-a} \int_\mathbb{T} |f_1^{\text{unf}}(\gamma)|^{a/q} \| \hat{g}_2(\gamma) \cdots \hat{g}_s(\gamma) \| d\gamma \]

\[ \leq \| f_1^{\text{unf}} \|_1^{1-a} \| (f_1^{\text{unf}}(\gamma))^a \|_q \| \hat{g}_2 \|_q \cdots \| \hat{g}_s \|_q \]

\[ = (\delta N)^{1-a} K^a N^{(1-1/q)} K^{s-1} N^{(s-1)(1-1/q)} \]

\[ = \delta^{1-a} K^{s-1+a} N^{s-1}. \]
Thus
\[ |\sum_{g_i \in \{f^*_i, f^\text{unf}_i\}} g_1 \cdots g_s(n)| \leq 2^s \delta^{s-q} K^q N^{s-1}. \]

**Proposition 3.9** (Transference lemma). Let \( s \geq 2, s-1 < q < s, K \geq 1 \) and \( \epsilon, \eta \in (0, 1) \). Let \( N \in \mathbb{N} \) and, for each \( i \in \{1, \ldots, s\} \) let \( f_i : [N] \to \mathbb{R}_{\geq 0} \) be \( \eta \)-pseudorandom and \( q \)-restricted with constant \( K \). Assume also that
\[ (5) \quad \mathbb{E}_{n \in [N]} f_1(n) + \cdots + f_s(n) > s(1 + \epsilon)/2 \]
and
\[ (6) \quad \mathbb{E}_{n \in [N]} f_i(n) > \epsilon/2 \]
for all \( i \in \{1, \ldots, s\} \). Write \( \kappa := \epsilon/32 \). Assume that \( \eta \) is sufficiently small depending on \( \epsilon, K, q \) and \( s \). Then, for all \( n \in \left((1 - \kappa^2)\frac{sN}{2}, (1 + \kappa)\frac{sN}{2}\right) \),
\[ f_1 \cdots f_s(n) \geq c(\epsilon, s) N^{s-1}, \]
where \( c(\epsilon, s) > 0 \) depends only on \( \epsilon \) and \( s \).

**Proof.** Let \( \delta \in (0, \epsilon/8) \) to be chosen later depending on \( \epsilon, s, K \) and \( q \). Denote \( f^\text{unf}_i = (f_i)_{\delta,N}^\text{unf} \) and \( f^*_i = (f_i)_{\delta,N}^* \) for all \( i \in \{1, \ldots, s\} \). Write \( \lambda := \epsilon/8 \) and let \( A_i = \{n : f^*_i(n) > \lambda \} \) for \( i \in \{1, \ldots, s\} \). By Lemma 3.8,
\[ (7) \quad f_1 \cdots f_s(n) \geq f_1^* \cdots f_s^*(n) - 2^s \delta^{s-q} K^q N^{s-1} \]
\[ = \sum_{\substack{a_1 + \cdots + a_s = n \\text{a}_i \in [N]}} f_1^*(a_1) \cdots f_s^*(a_s) - 2^s \delta^{s-q} K^q N^{s-1} \]
\[ \geq \sum_{\substack{a_1 + \cdots + a_s = n \\text{a}_i \in A_i}} f_1^*(a_1) \cdots f_s^*(a_s) - 2^s \delta^{s-q} K^q N^{s-1} \]
\[ \geq \lambda^s \sum_{a_1 + \cdots + a_s = n} 1_{A_1}(a_1) \cdots 1_{A_s}(a_s) - 2^s \delta^{s-q} K^q N^{s-1} \]
\[ \geq \lambda^s 1_{A_1} \cdots 1_{A_s}(n) - 2^s \delta^{s-q} K^q N^{s-1}. \]
For all \( i \in \{1, \ldots, s\} \), by the definition of \( f^*_i \) and Lemma 3.7(ii), we get
\[ \mathbb{E}_{n \in [N]} f^*_i(n) = \mathbb{E}_{n \in [N]} f_i(n) - \mathbb{E}_{n \in [N]} f^\text{unf}_i(n) \]
\[ \geq \mathbb{E}_{n \in [N]} f_i(n) - \delta > \epsilon/2 - \delta. \]
By Lemma 3.7(i) we see that
\[ \mathbb{E}_{n \in [N]} f^*_i(n) \leq \frac{1}{N} \sum_{n \in A_i} (1 + O_\delta(\eta)) + \mathbb{E}_{n \in [N]} \lambda. \]
Thus by (6),
\[ (1 + O_\delta(\eta)) |A_i| > (\epsilon/2 - \delta - \lambda) N > (\epsilon/4) N. \]
Similarly, using (5) in place of (6), we get
\[ s(1 + \epsilon)/2 - s\delta \leq \mathbb{E}_{n \in [N]} f_1^*(n) + \cdots + f_s^*(n) \]
\[ \leq \frac{1}{N} \sum_{i=1}^{s} \sum_{n \in A_i} (1 + O_\delta(\eta)) + s \mathbb{E}_{n \in [N]} \lambda \]
and so
\[ (1 + O_\delta(\eta)) \sum_{i=1}^{s} |A_i| > (s(1 + \epsilon)/2 - s\delta - s\lambda)N > (s(1 + \epsilon/4)/2)N. \]

We can assume that \( \eta \) is small enough in terms of \( \epsilon \) and \( \delta \), since otherwise the conclusion is trivial. Hence \( \sum_i |A_i| > (s(1 + \lambda)/2)N \) and \( |A_i| > (\lambda/2)N \) for all \( i \in \{1, \ldots, s\} \). Let \( c'(\lambda, s) \) be the constant in Lemma 3.3. Then the inequality (7) and Lemma 3.3 imply that
\[ f_1^* \cdots f_s^*(n) \geq (\lambda^s c'(\lambda, s) - 2^s \delta s^{-q} K^q) N^{s-1} \]
for all \( n \in ((1 - \lambda^2/4)^{sn/2}, (1 + \lambda^2/4)^{sn/2}) \). The result now follows by choosing \( \delta \) sufficiently small in terms of \( \epsilon, s, K, q \).

In the previous lemma the condition (5) is sharp: If \( \mathbb{E}_{n \in [N]} f_1(n) + \cdots + f_s(n) \leq s/2 \), then the sets \( A_1, \ldots, A_s \) in the proof of Proposition 3.9 can all be subsets of the same non-trivial arithmetic progression, which means that also the sumset \( A_1 + \cdots + A_s \) is the subset of a non-trivial arithmetic progression and so it is not true that \( f_1^* \cdots f_s^*(n) > 0 \) for all \( n \in ((1 - \lambda^2/2) sN, (1 + \lambda^2/2) sN) \).

4. Proof of the main theorem. In this section, we will prove Theorem 1.1 using the transference lemma (Proposition 3.9), assuming some lemmas which we will prove later. We will also prove Theorem 1.3.

4.1. Definitions. Let \( A \subseteq \mathbb{N}^{(k)} \), \( N \in \mathbb{N} \), \( w = \log \log \log N \) and
\[ W := \prod_{p \leq w} p^k. \]

Let \( b \in [W] \) be such that \( b \in \mathbb{Z}_{W}^{(k)} \). Define \( \sigma_W(b) := \{|z \in \mathbb{Z}_W : z^k \equiv b \ (\text{mod } W)\}| \). Define \( f_b, \nu_b : [N] \to \mathbb{R}_{\geq 0} \) by
\[ f_b(n) := \begin{cases} \frac{k}{\sigma_W(b)} t^{k-1} & \text{if } Wn + b = t^k \in A, \\ 0 & \text{otherwise,} \end{cases} \]
\[ \nu_b(n) := \begin{cases} \frac{k}{\sigma_W(b)} t^{k-1} & \text{if } Wn + b = t^k \in \mathbb{N}^{(k)}, \\ 0 & \text{otherwise.} \end{cases} \]

Clearly \( f_b(n) \leq \nu_b(n) \) for all \( n \in [N] \). The purpose of the \( W \)-trick in the definitions of \( f_b \) and \( \nu_b \) is to make pseudorandomness of \( \nu_b \) possible. The
normalization of \( f_b \) and \( \nu_b \) is used to ensure that \( \mathbb{E}_{n \in [M]} \nu_b(n) \sim 1 \) when \( b \in \mathbb{Z}_W^{(k)} \).

Define \( Z(q) := \{a \in \mathbb{Z}_q^{(k)} : (a, q) = 1\} \). Define also \( g : [W] \times \mathbb{N} \to \mathbb{R}_{\geq 0} \) by
\[
g(b, M) := \mathbb{E}_{n \in [M]} f_b(n).
\]
The notation of this subsection will be in force for the rest of the paper.

### 4.2. Key lemmas.

We will apply Proposition \([3.9]\) to the function \( f_b \). The following three lemmas (to be proved later) show that \( f_b \) is \( \eta \)-pseudorandom, \( q \)-restricted and satisfies the mean condition of Proposition \([3.9]\).

**Proposition 4.1 (Mean value lemma).** Let \( \epsilon \in (0, 1/6) \) and let \( N \) be sufficiently large depending on \( \epsilon \). Let \( \delta_A > (1 - (1/2 - 3\epsilon)Z^{-1}_k)^{1/k} \) and \( s \geq 16k\omega(k) + 4k + 4 \). Then, for all \( n \in \mathbb{Z}_W \) with \( n \equiv s \pmod{R_k} \), there exist \( b_1, \ldots, b_s \in Z(W) \) such that \( n \equiv b_1 + \cdots + b_s \pmod{W} \), \( g(b_i, N) > \epsilon/2 \) for all \( i \in \{1, \ldots, s\} \) and
\[
g(b_1, N) + \cdots + g(b_s, N) > s(1 + \epsilon)/2.
\]

We will prove Proposition \([4.1]\) in Section \([6]\). The main ingredient in the proof is a local density version of Waring’s problem. We will deal with this local problem in Section \([5]\).

**Proposition 4.2 (Pseudorandomness).** Let \( \alpha \in \mathbb{T} \). Assume \( \sigma_W(b) \neq 0 \) and \( (b, W) = 1 \). Then
\[
|\widehat{\nu_b}(\alpha) - \frac{1}{|N|}(\alpha)| = o_k(N).
\]

We will prove Proposition \([4.2]\) in Section \([7]\). The proof uses a standard circle method analysis of major and minor arcs.

**Proposition 4.3 (Restriction estimate).** Let \( s > \max(k^2 + k, 4k) \). Assume that \( \sigma_W(b) \neq 0 \) and \( (b, W) = 1 \). Then there exists \( q \in (s - 1, s) \) such that
\[
\|\widehat{f_b}\|_q \ll_k N^{1-1/q}.
\]

We will prove Proposition \([4.3]\) in Section \([8]\). The proof is based on Vinogradov’s mean value theorem and the \( \epsilon \)-removal technique.

### 4.3. Conclusion.

Now we are ready to prove Theorem \([1.1]\) assuming the propositions of the previous subsection.

Proof of Theorem \([1.1]\) assuming Propositions \([4.1]\) \([4.3]\). Let \( n_0 \) be a natural number with \( n_0 \equiv s \pmod{R_k} \). Our goal is to prove that \( n_0 \in sA \) provided that \( n_0 \) is sufficiently large.

Let \( N := \lfloor 2n_0/(sW) \rfloor \). Choose \( \epsilon \in (0, 1/6) \) with \( \delta_A > (1 - (1/2 - 3\epsilon)Z^{-1}_k)^{1/k} \).

By Proposition \([4.1]\) there exist \( b_1, \ldots, b_s \in [W] \) such that \( n_0 \equiv b_1 + \cdots + b_s \pmod{W} \), \( (b_i \pmod{W}) \in Z(W) \), for all \( i \in \{1, \ldots, s\} \), and the mean conditions \((5)\) and \((6)\) of Proposition \([3.9]\) hold for the functions \( f_{b_1}, \ldots, f_{b_s} \). By
Propositions 4.2 and 4.3 also the pseudorandomness and restriction conditions of Proposition 3.9 hold for \( f_{b_1}, \ldots, f_{b_s} \) for some \( q \in (s-1, s) \), \( K > 0 \) and for any \( \eta > 0 \). Now assume \( N \) is sufficiently large depending on \( \epsilon \), and \( \eta \) is sufficiently small depending on \( \epsilon, K, q, s \). Then by Proposition 3.9, for all \( n \in (\frac{1-k^2}{2}sN, \frac{1+k}{2}sN) \), where \( k = \epsilon/32 \). This means that, for all such \( n \),

\[
W_n + b_1 + \cdots + b_s \in sA.
\]

Let \( N \in \mathbb{N} \) be sufficiently large and \( A' = A \cap [N] \). By the Cauchy–Schwarz inequality and \cite[Theorem 5.7]{Nat96},

\[
\left( \sum_{n \in sA'} r_{sA'}(n) \right)^2 \leq |sA'| \sum_{n \in sA'} r_{sA'}(n)^2 \\
\leq |sA'| \sum_{n \in [N]} r_{n(k) \cap [N]}(n)^2 \ll_k |sA'|N^{2s/k-1}
\]

provided that \( s > 2^k \). On the other hand

\[
\sum_{n \in sA'} r_{sA'}(n) \geq |A \cap [N/s]|^s \gg s N^{s/k}.
\]

Hence

\[
|s(A \cap [N])| > c(s)N
\]

for all large \( N \in \mathbb{N} \) and for some small constant \( c(s) > 0 \) that depends only on \( s \).

Let \( A' = A - \min A \). We see that \( 0 \in A' \). Since \( A \) is not a subset of any non-trivial arithmetic progression, \( A' \) is not such a subset either, and therefore there exist \( a, b \in A' \) such that \((a, b) = 1. This implies with Bézout’s identity that \( sA' \) and hence also \( sA = s \min A + sA' \) contains two consecutive integers provided that \( s \) is large enough. Hence there exists \( N' \in \mathbb{N} \) such that, for \( B := sA - N' \), we have \( 0, 1 \in B \) and \( B \) has positive lower density by (13).

We define Shnirel’man density by

\[
\sigma(A) := \inf_{N \in \mathbb{N}} \frac{|A \cap [N]|}{N}.
\]
We can see that $\sigma(B) > 0$. Thus by [Nat96, Theorem 7.7] there exists $s' \in \mathbb{N}$ such that $s'B = \mathbb{N}$. Therefore all sufficiently large natural numbers belong to $(s's)A$. ■

5. Local problem. In this section, we study the local density version of Waring’s problem. This problem is the key new ingredient in solving the density version of Waring’s problem. Recall that $Z(q) = \{ a \in \mathbb{Z}_q^{(k)} : (a,q) = 1 \}$.

For a prime $p$ and $e \geq 1$, we see by [IR90, Chapter 4, §2] that

$$|Z(p^e)| = \frac{\phi(p^e)}{(k, p^e)}. \quad (14)$$

Also, for $n \in \mathbb{N}$, recalling the notation $\tau(n,p)$ from Section 1.1, we have

$$|Z(n)| = \prod_{p|n} \frac{\phi(p^{\tau(n,p)})}{(k, p^{\tau(n,p)})}. \quad (15)$$

We also note by Fermat’s little theorem and the Chinese remainder theorem that if $a \in Z(q)$, then

$$a \equiv 1 \pmod{(R_k,q)}, \quad (16)$$

where $R_k$ is as in (2). This congruence is the reason why we have the congruence condition in Theorem 1.1 because we are restricted to those elements of $A$ which are coprime to $W$.

**Definition 5.1.** Let $q,s \in \mathbb{N}$. We say that $(q,s)$ is a Waring pair if, for any $A \subseteq Z(q)$ with $|A| > \frac{1}{2} |Z(q)|$, we have $sA = \{ a \in \mathbb{Z}_q : a \equiv s \pmod{(R_k,q)} \}$.

Our aim is to prove the following proposition.

**Proposition 5.2.** $(W,s)$ is a Waring pair for any $s \geq 8k\omega(k) + 2k + 2$.

We conjecture that $(W,s)$ is a Waring pair for some $s = O(k)$, but we are satisfied with the number of summands being $o(k^2)$, because the restriction estimate (Proposition 4.3) gives us a lower bound for the number of summands that is of order $k^2$.

One of the main reasons why we are able to solve the local problem is the fact that the Waring pairs have multiplicative-like structure. This behaviour is captured by the following lemma.

**Lemma 5.3.** Let $q, r, s, t \in \mathbb{N}$ and $(q,r) = 1$. If $(q,s)$ and $(r,t)$ are Waring pairs, then $(qr, s + t)$ is a Waring pair.

**Proof.** Let $A \subseteq Z(qr)$ with $|A| > \frac{1}{2} |Z(qr)|$. By the pigeonhole principle there exists a congruence class $a^* \in Z(q)$ such that the set $B := \{ b \in A : b \equiv a^* \pmod{q} \}$ satisfies $|B| > \frac{1}{2} |Z(r)|$. Let $n \in Z_{qr}$ be such that $n \equiv s + t \pmod{(R_k,qr)}$. Since $(q,s)$ is a Waring pair, we have
\[ n \equiv ta^* + a_1 + \cdots + a_s \pmod{q}. \]

for some \( a_1, \ldots, a_s \in A \) (note that \( a \in \mathbb{Z}(q) \) implies \( a \equiv 1 \pmod{(R_k, q)} \)). Since \((r, t)\) is a Waring pair, we also see that

\[ n \equiv b_1 + \cdots + b_t + a_1 + \cdots + a_s \pmod{r} \]

for some \( b_1, \ldots, b_s \in B \). Hence by the Chinese remainder theorem and definition of \( B \),

\[ n \equiv b_1 + \cdots + b_t + a_1 + \cdots + a_s \pmod{qr}. \]

We are going to use this lemma to deal separately with two parts of \( W \), \( \prod_{p \leq w, p \mid k} p^k \) and \( \prod_{p \leq w, p \mid k} p^k \).

5.1. Single moduli. In this subsection, we study the local problem in \( \mathbb{Z}_{p^k} \). For that purpose, we need the following lemma that tells us how the elements in \( Z(p^k) \) are distributed in certain cosets of \( p \cdot \mathbb{Z}_{p^k} \).

**Lemma 5.4.** Let \( p \) be a prime. For all \( a \in Z(p) \), we have

\[ |\{b \in Z_p^{(k)} : b \equiv a \pmod{p}\}| = p^{k-1-\tau(k,p)}. \]

**Proof.** For \( c \in Z(p) \) set \( B(c) := \{b \in Z_p^{(k)} : b \equiv c \pmod{p}\} \). For \( b, c \in Z(p) \) and \( d \in B(c) \), we see that \( bd \in B(bc) \). Hence \( |B(c)| \leq |B(bc)| \). Since \( Z(p) \) is a group, it follows that \( |B(b)| = |B(c)| \) for all \( b, c \in Z(p) \). Furthermore \( |Z(p^k)| = \sum_{b \in Z(p)} |B(b)| \) and so \( |B(b)| = |Z(p^k)|/|Z(p)| \) for all \( b \in Z(p) \) and the claim follows from (14). \( \blacksquare \)

We will also need the following generalization of the Cauchy–Davenport theorem from [COS19, Theorem 1.1].

**Lemma 5.5.** Let \( n \geq 1 \), and \( A_1, \ldots, A_n \) be finite, non-empty subsets of an abelian group \( G \) such that no \( A_i \) is contained in a coset of a proper subgroup of \( G \). Then

\[ |A_1 + \cdots + A_n| \geq \min\left(|G|, \left(1 + \frac{1}{2n}\right) \sum_{i=1}^{n} |A_i|\right). \]

Essentially this means that if \( G \) is finite and \( A \subseteq G \) satisfies the coset condition, then \( A \) is a basis of order \([2|G|/|A|] - 1\).

Now we can prove the local problem for prime power moduli.

**Lemma 5.6.** Let \( p \) be a prime. Then \((p^k, s)\) is a Waring pair for all \( s \geq 8k \).

**Proof.** Let \( A \subseteq Z(p^k) \) with \(|A| > \frac{1}{2}|Z(p^k)|\). If \((p - 1) | k\), then we see by (16) that \( A \subseteq \{a \in Z_{p^k} : a \equiv 1 \pmod{p^{\eta(p,k)}}\} \), where \( \eta(p, k) \) is as in (3). Define \( A' = \{a \in Z_{p^{k-\eta(k,p)}} : (ap^{\eta(k,p)} + 1 \pmod{p^k}) \in A\} \).
Since \(|A'| = |A| > \frac{1}{2}|Z(p^k)| = \frac{1}{2}p^{k-\tau(k,p)} \geq \frac{1}{2}p^{k-\eta(k,p)}\) it follows that \(A'\) is not contained in any coset of a proper subgroup of \(\mathbb{Z}_{p^k-\eta(k,p)}\). Hence by Lemma \(5.5\) we get

\[sA' = \mathbb{Z}_{p^k-\eta(k,p)} \text{ for all } s \geq 4.\]

Similarly if \(p - 1 \nmid k\), then \(|A| > \frac{1}{2}|Z(p^k)| = \frac{1}{2}p^{k-\tau(k,p)} \cdot \frac{p-1}{(k,p-1)} \geq p^{k-\tau(k,p)}\). Thus by Lemma \(5.4\) \(A\) is not contained in any coset of a proper subgroup of \(\mathbb{Z}_{p^k}\). Again by Lemma \(5.5\) we get

\[sA = \mathbb{Z}_{p^k} \text{ for all } s \geq \left\lceil \frac{2|Z_{p^k}|}{|A|} \right\rceil - 1.\]

By (14) and the definition of \(A\) we see that

\[\left\lceil \frac{2|Z_{p^k}|}{|A|} \right\rceil - 1 < \frac{p^k}{2\phi(p^k)/(k,\phi(p^k))} \leq 4k \cdot \frac{p}{p-1} \leq 8k. \]

Using Lemmas \(5.3\) and \(5.6\) we can already see that \((W,s)\) is a Waring pair provided that \(s \geq \omega(W)8k\), but this is not sufficient since we want to have \(s = o(k^2)\). This means that we cannot use Lemma \(5.3\) too many times.

\[5.2. \text{Large moduli.}\] In this subsection, we deal with the local problem for the \(k\)-coprime part of \(W\). First we use Hensel’s lemma to reduce the moduli of the problem to be square-free. Then we use a downset idea from [Mat13, Section 4] to simplify the problem.

We start with the moduli reduction argument.

**LEMMA 5.7.** Let \(e, s \in \mathbb{N}\). Let \(q\) be a square-free natural number with \((q, k) = 1\). If \((q, s)\) is a Waring pair, then \((q^e, s + 2)\) is also a Waring pair.

**Proof.** Let \(A \subseteq \mathbb{Z}(q^e)\) be any set with \(|A| > \frac{1}{2}|Z(q^e)|\) and let \(a \in \mathbb{Z}(q)\). Then by the Chinese remainder theorem and Hensel’s lemma (see e.g. IR90 Proposition 4.2.3)) we find that the equation

\[a + bq \equiv x^k \pmod{q^e}\]

is soluble for all \(b \in \mathbb{Z}_{q^e-1}\). Hence we can partition \(\mathbb{Z}(q^e)\) into sets \(a + q\mathbb{Z}_{q^e-1}\), where \(a\) runs through all elements in \(\mathbb{Z}(q)\). By the pigeonhole principle, for at least one choice of \(b \in \mathbb{Z}(q)\) we have \(|H| > \frac{1}{2}q^{e-1}\), where \(H = (b + q\mathbb{Z}_{q^e-1}) \cap A\). Therefore \(2H = 2b + q\mathbb{Z}_{q^e-1}\).

Again by the pigeonhole principle there exists an interval \(I := (t, (t+1)q]\) for some \(t \in [0, q^{e-1} - 1]\) such that \(|I \cap A| > \frac{1}{2}|Z(q)|\). Since \((q, s)\) is a Waring pair we can now see that

\[2H + s(I \cap A) = \{a \in \mathbb{Z}_{q^e} : a \equiv s + 2 \pmod{(R_k,q)}\}\]
such that \( a' < b', a' \equiv a \pmod{n} \) and \( b' \equiv b \pmod{n} \). Let \( q \) be a square-free natural number. For \( v \in \mathbb{Z}_q \cong \prod_{p | q} \mathbb{Z}_p \) we define

\[
D(v) := \{ b \in \mathbb{Z}_q : \forall p \mid q, 0 \leq b \leq v \pmod{p} \}. \]

We say that \( A \subseteq \mathbb{Z}_q \) is a downset if \( D(v) \subseteq A \) for all \( v \in A \). We also say that \( u \in \mathbb{Z}_q^* \) is an upper bound for \( a \in \mathbb{Z}_q \) if \( a < u \pmod{p} \) for all \( p \mid q \). We say that \( u \in \mathbb{Z}_q^* \) is an upper bound for \( A \subseteq \mathbb{Z}_q \) if \( u \) is an upper bound for all elements in \( A \). For \( A \subseteq \mathbb{Z}_q \) and \( p \mid q \) define the number of residue classes mod \( p \) that occur in \( A \) by

\[
r(A, p) := |\{ a \in [p] : \exists b \in A, a \equiv b \pmod{p} \}|. \]

We define \( u(A) \in \mathbb{Z}_q \) such that

\[
u(A) \equiv r(A, p) \pmod{p} \quad \text{for all } p \mid q. \]

The following lemma reveals how the downsets can be used to analyse the size of sumsets.

**Lemma 5.8.** Let \( q \) be a square-free natural number. Let \( s \in \mathbb{N} \). Let \( A_1, \ldots, A_s \subseteq \mathbb{Z}_q^* \). Then there exist downsets \( A'_1, \ldots, A'_s \subseteq \mathbb{Z}_q \) such that \( |A'_i| = |A_i| \) and \( u(A_i) \) is an upper bound for \( A'_i \) for all \( i \in \{1, \ldots, s\} \), and

\[
|A'_1 + \cdots + A'_s| \leq |A_1 + \cdots + A_s|. \]

**Proof.** Let \( p \mid q \) be a prime and write \( r = q/p \). For \( A \subseteq \mathbb{Z}_q \) and \( a \in \mathbb{Z}_r \) define sets \( A(a, p), A[a, p], A^{(p)} \subseteq \mathbb{Z}_q \) such that

\[
A(a, p) := (\{ a \} \times \mathbb{Z}_p) \cap A, \\
A[a, p] := \begin{cases} 
\{ a \} \times \{ 0, \ldots, |A(a, p)| - 1 \} & \text{if } A(a, p) \neq \emptyset, \\
\emptyset & \text{otherwise,} 
\end{cases} \\
A^{(p)} := \bigcup_{b \in \mathbb{Z}_r} A[b, p]. \]

In other words, \( A^{(p)} \) has the downset property with respect to the coordinate \( p \) and it has the same number of elements as \( A \). Clearly \( A^{(p)}(a, p) = A[a, p] \).

We also define \( \emptyset + A = \emptyset \). We see that

\[
|A_1 + \cdots + A_s| = \sum_{n \in \mathbb{Z}_r} |(A_1 + \cdots + A_s)(n, p)| \geq \sum_{n \in \mathbb{Z}_r} \max_{a_1, \ldots, a_s \in \mathbb{Z}_r} |A_1(a_1, p) + \cdots + A_s(a_s, p)|. \]
Using the Cauchy–Davenport inequality [TV10, Theorem 5.4] we find that
\[
|A_1 + \cdots + A_s| \\
\geq \sum_{n \in \mathbb{Z}_r} \max_{a_1, \ldots, a_s \in \mathbb{Z}_r} \min(p, |A_1(a_1, p)| + \cdots + |A_s(a_s, p)| - (s - 1)) \\
= \sum_{n \in \mathbb{Z}_r} \max_{a_1, \ldots, a_s \in \mathbb{Z}_r} |A_1[a_1, p] + \cdots + A_s[a_s, p]| \\
= \sum_{n \in \mathbb{Z}_r} |(A_1^{(p)} + \cdots + A_s^{(p)})(n, p)| = |A_1^{(p)} + \cdots + A_s^{(p)}|.
\]

Now the sets $A_1^{(p)}, \ldots, A_s^{(p)}$ have a downset type property with respect to the $p$-coordinate. Applying the same process to each of the remaining coordinates $p' | q$ in turn and noticing that the process does not forget the downset-ness of the coordinates already handled, we finally end up with downsets with the desired properties. ■

Using the previous lemma and simple combinatorial calculations, we can prove the following lemma.

**Lemma 5.9.** Let $q$ be a square-free natural number with $(q, k) = 1$. Then $(q, s)$ is a Waring pair for all $s \geq 2k$.

**Proof.** Let $A \subseteq Z(q)$ with $|A| > \frac{1}{2}|Z(q)|$. For $n \in \mathbb{N}$ set $\sigma(n) := |Z(n)|$. By (15) we see that $\sigma$ is a multiplicative function. Let $u \in \mathbb{Z}_q$ be such that
\[
u \equiv \sigma(p) \pmod{p}
\]
for all $p | q$.

By Lemma 5.8 there exists a downset $A' \subseteq Z(q)$ such that $|A| = |A'|$, $u$ is an upper bound for $A'$ and $|sA'| \leq |sA|$ for all $s \geq 1$. Note that $sA'$ is also a downset.

Now let $S \subseteq \mathbb{Z}_q$ be the set of all elements that have upper bound $u$. We see that $|S| = \sigma(q)$. Also, $A', u - A' \subseteq S$. From $2|A'| > |S|$ it follows that
\[
|\{u = a + b : a, b \in A'\}| = |A' \cap (u - A')| = |A' \setminus (S \setminus (u - A'))| \\
\geq |A'| - (|S| - |A'|) > 0.
\]

Hence $u \in 2A'$. Since $2A'$ is a downset, we see that $D(u) \subseteq 2A'$. Because $kD(u) = Z_q$ we have $2kA' = Z_q$. ■

From Lemmas 5.7 and 5.9 we get the following lemma.

**Lemma 5.10.** $(\prod_{p \leq w, p | k} p^k, s)$ is a Waring pair for all $s \geq 2k + 2$.

**5.3. Conclusion.** Combining the results from the previous subsections, we can now solve the local problem.
Proof of Proposition 5.2. Applying Lemma 5.3 inductively with Lemma 5.6 to the primes dividing $k$, we find that
\((\prod_{p \leq w, p | k} p^k, 8k\omega(k))\) is a Waring pair. The result now follows from Lemmas 5.3 and 5.10. 

6. Mean value estimate. In this section, we will prove the mean condition (Proposition 4.1) required in the transference lemma (Proposition 3.9). 

6.1. Mean value over $g(b, N)$. In this subsection, we establish a lower bound for $\mathbb{E}_{b \in \mathbb{Z}(W)} g(b, N)$, where $g$ is as in (11).

Lemma 6.1. Let $\epsilon \in (0, 1)$. Then
\[
\mathbb{E}_{b \in \mathbb{Z}^{(k)}_W} g(b, N) \geq (1 - \epsilon)\delta_A^k
\]
provided that $N$ is large enough depending on $\epsilon$.

Proof. Let $b \in \mathbb{Z}^{(k)}_W$ and write
\[
\delta_b := \frac{|A \cap (W \cdot \lfloor N \rfloor + b)|}{|N^{(k)} \cap (W \cdot \lfloor N \rfloor + b)|}.
\]
Since $|N^{(k)} \cap (W \cdot \lfloor N \rfloor + b)| \sim \sigma_W(b)(WN)^{1/k}/W$, we have
\[
|A \cap (W \cdot \lfloor N \rfloor + b)| \sim \sigma_W(b)\frac{(WN)^{1/k}}{W}\delta_b.
\]
Note also that $\sum_{t \leq x, t \equiv a \pmod{n}} kt^{k-1} \sim x^{k}/n$. Hence
\[
g(b, N) = \frac{1}{N\sigma_W(b)} \sum_{t^k \leq W N + b} \sum_{t^k \equiv b \pmod{W}} kt^{k-1}
\]
\[
= \frac{1}{N\sigma_W(b)} \sum_{z \in \mathbb{Z}} \sum_{z^k \equiv b \pmod{W}} kt^{k-1}
\]
\[
\geq \frac{1}{N\sigma_W(b)} \sum_{z \in \mathbb{Z}} \sum_{t \equiv z \pmod{W}} \left| \frac{|A \cap (W \cdot \lfloor N \rfloor + b)|}{\sigma_W(b)} \right| kt^{k-1}
\]
\[
\geq (1 - o(1))\frac{1}{WN} \left( \frac{W}{\sigma_W(b)} \right)^k |A \cap (W \cdot \lfloor N \rfloor + b)|^k
\]
\[
\geq (1 - o(1))\delta_b^k.
\]
Since, for any $b \in \mathbb{Z}^{(k)}_W$,
\[
\frac{|N^{(k)} \cap [WN + W - 1]|}{|\mathbb{Z}^{(k)}_W|} = |N^{(k)} \cap [WN + W - 1] \cap (W \cdot N + b)| + O(1),
\]
we observe that

\[
\delta_A \leq \frac{|A \cap [WN + W - 1]|}{|N(k) \cap [WN + W - 1]|}
\]

\[
= \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} \frac{|A \cap [WN + W - 1] \cap (W \cdot N + b)|}{|N(k) \cap [WN + W - 1] \cap (W \cdot N + b)|} + O(1)
\]

\[
= (1 + o(1)) \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} \frac{|A \cap [WN + W - 1] \cap (W \cdot N + b)|}{|N(k) \cap [WN + W - 1] \cap (W \cdot N + b)|}
\]

\[
= (1 + o(1)) \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} \delta_b.
\]

Thus by Hölder’s inequality, (18) and (19),

\[
\mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} g(b, N) \geq (1 - o(1)) \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} \delta_b \geq (1 - \epsilon)\delta_A^k
\]

for any \(\epsilon > 0\) provided that \(N\) is large enough depending on \(\epsilon\).

The lower bound in the previous lemma is essentially the best possible because if \(A = \{n^k : n \leq \delta_A(WN)^{1/k}\}\) then \(g(b, \delta_A^k N) \approx 1\) for all \(b \in Z(W)\), and so \(\mathbb{E}_{b \in Z(W)} g(b, N) \approx \delta_A^k\).

Using the previous lemma, we can now prove a similar result for \(Z(W)\). Recall that

\[
Z_k = \lim_{m \to \infty} \frac{|Z_W^{(k)}|}{|Z(W)|} \ \{a \in Z_P^{(k)} : (a, P(m)) = 1\}.
\]

**Lemma 6.2.** Let \(\epsilon > 0\). Let \(Z_k\) be as in (1). Then

\[
\mathbb{E}_{b \in Z(W)} g(b, N) > (1 - \epsilon)(Z_k \delta_A^k - Z_k + 1)
\]

provided that \(N\) is large enough.

**Proof.** Since \(g(b, N) = 1 + o(1)\), we see by Lemma 6.1 that

\[
\mathbb{E}_{b \in Z(W)} g(b, N) = \frac{|Z_W^{(k)}|}{|Z(W)|} \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} g(b, N) - \frac{1}{|Z(W)|} \sum_{b \in \mathbb{Z}_W^{(k)} \backslash \{b : b \cdot W > 1\}} g(b, N)
\]

\[
\geq (1 - o(1)) Z_k \mathbb{E}_{b \in \mathbb{Z}_W^{(k)}} g(b, N) - (1 + o(1)) Z_k + 1
\]

\[
\geq (1 - o(1))(Z_k \delta_A^k - Z_k + 1).\]

Next we present the following lemma about the size of \(Z_k\).

**Lemma 6.3.** Let \(k > 4\). Then

\[
\frac{\zeta(k)}{\zeta(2k)} \leq Z_k \leq \frac{\zeta(k - \log_2(2k))}{\zeta(2k - 2 \log_2(2k))}.
\]

(2) Here \(\log_2 n = \log n / \log 2\).
Proof. By (15) we see that
\[
\frac{|Z^{(k)}_{W}|}{|Z(W)|} = \prod_{p|W} \frac{|Z(p^k)| + 1}{|Z(p)|} = \prod_{p|W} \left(1 + \frac{1}{|Z(p^k)|}\right) = \sum_{d|W^{1/k}} \frac{1}{|Z(d^k)|}.
\]
Thus
\[
Z_k = \lim_{w \to \infty} \frac{|Z^{(k)}_{W}|}{|Z(W)|} = \sum_{n=1}^{\infty} \frac{1}{|Z(n^k)|}.
\]
Let \(n\) be square-free. By (15) we have \(|Z(n^k)| = \prod_{p|n} p^{k-1}(p-1)\). Since \(\omega(n) \leq \log_2 n\) for all \(n > 1\), we have \(k^{\omega(n)} \leq n^{\log_2 k}\). Hence \(n^{k-1-\log_2 k} \leq |Z(n^k)| \leq n^k\). It follows that
\[
\frac{\zeta(k)}{\zeta(2k)} \leq Z_k \leq \frac{\zeta(k-\log_2(2k))}{\zeta(2k - 2\log_2(2k))},
\]
because
\[
\sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{\zeta(s)}{\zeta(2s)} \quad \text{for all } s > 1. \]

The previous lemma implies that \(\lim_{k \to \infty} Z_k = 1\). The table below illustrates the convergence of \(Z_k\).

| Values of \(Z_k\) for small \(k\) |
|---|---|
| \(k\) | \(Z_k\) |
| 2 | 3.279 |
| 3 | 1.493 |
| 4 | 1.570 |
| 5 | 1.071 |
| \(k\) | \(Z_k\) |
| 6 | 1.075 |
| 7 | 1.016 |
| 8 | 1.062 |
| 9 | 1.004 |

6.2. Proof of Proposition 4.1. For the proof, we need the following lemma that is essentially a generalized version of the local problem.

Lemma 6.4. Let \(f : Z(W) \to [0, 1]\) satisfy \(\mathbb{E}_{b \in Z(W)} f(b) > 1/2\). Let \(s \geq 16k\omega(k) + 4k + 4\). Then, for all \(n \in \mathbb{Z}_W\) with \(n \equiv s \pmod{R_k}\), there exist \(b_1, \ldots, b_s \in Z(W)\) such that \(n \equiv b_1 + \cdots + b_s \pmod{W}\), \(f(b_i) > 0\) for all \(i \in \{1, \ldots, s\}\) and
\[
f(b_1) + \cdots + f(b_s) > s/2.
\]

Proof. Write \(M := \mathbb{E}_{b \in Z(W)} f(b)\). Let \(\mu := \max_{b \in Z(W)} f(b)\), \(\lambda := 1 - \mu\) and \(A := \{b \in Z(W) : f(b) > \lambda\}\). Note that \(\mu \geq M > 1/2\) so that \(\lambda < 1/2\) and \(A\) is non-empty.
We see that
\[ M \leq \frac{1}{|Z(W)|} \sum_{b \in Z(W) \setminus A} \lambda + \frac{1}{|Z(W)|} \sum_{b \in A} \mu = \lambda \frac{|Z(W)| - |A|}{|Z(W)|} + \mu \frac{|A|}{|Z(W)|}. \]

Hence
\[ |A| \geq \frac{M - \lambda}{\mu - \lambda} |Z(W)| = \frac{M + \mu - 1}{2\mu - 1} |Z(W)| > M |Z(W)|, \]
since \( M > 1/2 \). Thus
\[ (20) \quad |A| > \frac{1}{2} |Z(W)|. \]

Now it follows from Proposition \[5.2\] that
\[ s'A = \{ a \in Z_W : a \equiv s' \pmod{R_k} \} \]
for all \( s' \geq 8k\omega(k) + 2k + 2 \).

Let \( b \in Z(W) \) be such that \( f(b) = \mu \) and let \( s'' \geq s' \). Then, for each \( n \in Z_W \) with \( n \equiv s' + s'' \pmod{R_k} \), there exist \( b_1, \ldots, b_{s'} \in A \) such that
\[ n - s''b \equiv b_1 + \cdots + b_{s'} \pmod{W} \]
and
\[ s''f(b) + f(b_1) + \cdots + f(b_{s'}) > s''\mu + s'\lambda = (s'' - s')\mu + s' (\mu + \lambda) \geq \frac{s'' - s'}{2} + s' = \frac{s' + s''}{2}. \]

**Proof of Proposition \[4.1\]** Since the condition \( \delta_A > (1 + (3\epsilon - 1/2)Z_k^{-1})^{1/k} \) is equivalent to \( Z_k\delta_A^k - Z_k + 1 > 1/2 + 3\epsilon \), we see by Lemma \[6.2\] that
\[ \mathbb{E}_{b \in Z(W)} g(b, N) > (1 - \epsilon)(1/2 + 3\epsilon) > 1/2 + 2\epsilon \]
provided that \( N \) is large enough depending on \( \epsilon \).

For \( b \in Z(W) \), define
\[ f(b) := \max \left( 0, \frac{1}{1 + \epsilon} (g(b, N) - \epsilon/2) \right). \]

Provided that \( N \) is large enough depending on \( \epsilon \), we have \( f(b) \in [0, 1) \). We also see that
\[ \mathbb{E}_{b \in Z(W)} f(b) \geq \frac{1}{1 + \epsilon} \mathbb{E}_{b \in Z(W)} (g(b, N) - \epsilon/2) > 1/2. \]

Hence by Lemma \[6.4\] for all \( n \in Z_W \) with \( n \equiv s \pmod{R_k} \), there exist \( b_1, \ldots, b_s \in Z(W) \) such that \( n \equiv b_1 + \cdots + b_s \pmod{W} \), \( f(b_i) > 0 \) for all \( i \in \{1, \ldots, s\} \) and
\[ f(b_1) + \cdots + f(b_s) > \frac{s}{2}. \]

By definition of \( f \) we have \( g(b_i, N) > \epsilon/2 \) for all \( i \in \{1, \ldots, s\} \) and
\[ g(b_1, N) + \cdots + g(b_s, N) > \frac{s(1 + \epsilon)}{2} + \frac{s\epsilon}{2}. \]
7. Pseudorandomness condition. In this section, we will establish the pseudorandomness of the function $f_b$ (Proposition 4.2). We use the standard circle method machinery.

Let us first introduce the Hardy and Littlewood decomposition. Let

\[ Q := N^\rho \quad \text{and} \quad T := N^{1-\rho} \]

for $\rho > 0$ to be chosen later. For $a, q \in \mathbb{N}$ and $(a, q) = 1$, write $\mathcal{M}(q, a) := \{\alpha : |\alpha - a/q| \leq 1/T\}$. Let

\[ \mathcal{M} := \bigcup_{a=0}^{q-1} \mathcal{M}(q, a). \]

If $\rho$ is suitably small and $N$ is sufficiently large, then $T > 2Q^2$ and thus the intervals $\mathcal{M}(q, a)$ are the disjoint. Let also $m = \mathbb{T} \setminus \mathcal{M}$. We call $\mathcal{M}$ the major arcs and $m$ the minor arcs.

From (10) we have

\[ \hat{\nu}_b(\alpha) = \sum_n \nu_b(n)e(n\alpha) = e_W(-b\alpha) \sum_{z \in [W]} F(\alpha, z), \]

where

\[ F(\alpha, z) := \sum_{t^k \leq W(N+b)} kt^{k-1}e_W(\alpha t^k). \]

7.1. Minor arcs. In this subsection, we establish Proposition 4.2 in the minor arcs using Weyl’s inequality.

Lemma 7.1. Let $\alpha \in m$. Then

\[ \hat{\nu}_b(\alpha) \ll_{\rho, k} N^{1-\sigma} \]

for some small $\sigma = \sigma(\rho) > 0$.

Proof. Let

\[ f(X, \alpha, z) = \sum_{t^k \leq X} e_W(\alpha t^k). \]

Trivially $|f(X, z, \alpha)| \leq X^{1/k}/W$. Let $\lambda \in (0, 1)$, to be chosen later. Using partial summation we get

\[ F(\alpha, z) = f(WN+b, \alpha, z)k(WN+b)^{1-1/k} \]

\[ (WN+b)^{1/k} \]

\[ - \int_1 f(x^k, \alpha, z)k(k-1)x^{k-2} dx \]
\[ f(WN + b, \alpha, z)k(WN + b)^{1-1/k} \]

\[ = \sum_{u \leq \frac{X^{1/k} - z}{W}} e(\alpha W^{k-1} u^k + g(u)), \]

where \( g(u) \) is a polynomial with degree at most \( k - 1 \).

Let \( q' = q/(q, W^{k-1}) \) and \( a' = W^{k-1} a/(q, W^{k-1}) \). Then \( (q', a') = 1 \) and

\[
|\alpha W^{k-1} - a'/q'| \leq \frac{W^{k-1}}{(q, W^{k-1})} \frac{1}{q'T} \leq \frac{W^{k-1}}{(q, W^{k-1})} \frac{1}{q'^2}.
\]

Now by Weyl’s inequality (see e.g. [Ove14, proof of Proposition 4.14]), for any \( \epsilon > 0 \),

\[
f(X, \alpha, z) \ll_{\epsilon,k} \left( \frac{X^{1/k}}{W} \right)^{1+\epsilon} \left( \frac{W^{k-1}}{(q, W^{k-1})} \frac{1}{q'} + \frac{W}{X^{1/k}} + \frac{W^{k-1}}{(q, W^{k-1})} \frac{W^{k-1}}{X^{1-1/k}} + \frac{q' W^k}{X} \right)^{\sigma},
\]

where \( \sigma = 1/2^{k-1} \). By (8) we have \( W = o(\log N) \). Since \( q > Q \), we also see by (21) that \( q > N^\rho \). Thus, for \( X \in [(WN)^{1-\lambda}, WN + b] \),

\[
f(X, \alpha, z) \ll_{\epsilon,k} X^{(1+2\epsilon)/k} (N^{-\rho} + X^{-1/k} + X^{1-1/k} + X^{-1} N^{1-\rho})^\sigma
\]

\[
\ll X^{(1-\sigma'+2\epsilon)/k} \ll N^{(1-\sigma'+3\epsilon)/k}
\]

for some \( \sigma' = \sigma'(\rho) > 0 \) provided that \( \lambda \) is sufficiently small depending on \( \rho \). Hence

\[
F(\alpha, z) \ll_{\epsilon,k} N^{1-\sigma''}
\]

for some \( \sigma'' = \sigma''(\rho) > 0 \) provided that \( \epsilon \) is small enough depending on \( \sigma' \).

The result now follows from (22) and (23).

By summing the geometric series (see e.g. [Nat96, Lemma 4.7]) we see that

(24) \( \widehat{1}_{[N]}(\alpha) \ll \|\alpha\|^{-1} \ll N^{1-\rho} \)

when \( \alpha \in \mathfrak{m} \). Hence we get the following lemma.

**Lemma 7.2.** Let \( \alpha \in \mathfrak{m} \). Then

\[
|\widehat{\nu}_b(\alpha) - \widehat{1}_{[N]}(\alpha)| \ll \rho N^{1-\epsilon}
\]

for some \( \epsilon = \epsilon(\rho) > 0 \).
7.2. Major arcs. In this subsection, our aim is to prove Proposition 4.2 on the major arcs. The result we will prove is the following.

Lemma 7.3. Let $\alpha \in \mathfrak{M}$. Assume that $(b, W) = 1$. Then

$$|\hat{\nu}_b(\alpha) - \frac{1}{[N]}(\alpha)| \ll_{k, \epsilon} w^{-1/k + \epsilon} N$$

for any $\epsilon > 0$ provided that $\rho$ is sufficiently small depending on $k$.

Let us first introduce the following two auxiliary functions that we will use to tackle pseudorandomness on the major arcs:

\begin{equation}
G_b(\alpha, N) := \sum_{t^k \leq N \atop t^k \equiv b \pmod{W}} k t^{k-1} e_W(\alpha t^k),
\end{equation}

\begin{equation}
V_q(a, b) := \sum_{h \pmod{Wq} \atop h^k \equiv b \pmod{W}} e_{Wq}(ah^k).
\end{equation}

The function $G_b(\alpha, N)$ is called the generating function. Our first goal is to prove an approximation lemma for the generating function.

Let

$$S(N) := \sum_{t^k \leq N \atop t^k \equiv b \pmod{W}} e_{Wq}(at^k).$$

We see that

\begin{equation}
S(N) = \sum_{h \pmod{Wq} \atop h^k \equiv b \pmod{W}} e_{Wq}(ah^k) \sum_{t^k \leq N \atop t \equiv h \pmod{Wq}} 1 = V_q(a, b) \frac{N^{1/k}}{Wq} + O(Wq).
\end{equation}

The following lemma approximates the generating function on the rational numbers.

Lemma 7.4. Let $a, q \in \mathbb{N}$. Then

$$G_b(a/q, N) = \frac{V_q(a, b)}{Wq} N + O(WqN^{1-1/k}).$$

Proof. Using partial summation we see that

$$G_b(a/q, N) = S(N)kN^{1-1/k} - \int_{1}^{N} S(t)(k - 1)t^{-1/k} dt$$

$$= k \frac{V_q(a, b)}{Wq} N - \frac{V_q(a, b)}{Wq} \int_{1}^{N} (k - 1) dt + O(WqN^{1-1/k})$$

$$= \frac{V_q(a, b)}{Wq} N + O(WqN^{1-1/k}).$$
Using the previous lemma we can now prove an approximation lemma for the generating function for all real numbers.

**Lemma 7.5.** Let \( a, q \in \mathbb{N}, \alpha \in \mathbb{R} \) and \( \beta = \alpha - a/q \). Then
\[
G_b(\alpha, N) - \frac{V_q(a, b)}{q} \sum_{t \leq N \atop t \equiv b \,(\text{mod} \, W)} e_W(\beta t) = O(W q N^{1-1/k} + q |\beta| N^{2-1/k}).
\]

**Proof.** We can write
\[
G_b(\alpha, N) = \frac{V_q(a, b)}{q} \sum_{t \leq N \atop t \equiv b \,(\text{mod} \, W)} e_W(\beta t) = \sum_{t \leq N \atop t \equiv b \,(\text{mod} \, W)} u(t)e_W(\beta t),
\]
where
\[
u(n) = \begin{cases} 
kh^{k-1}e_{Wq}(ah^k) - V_q(a,b)/q & \text{if } n = h^k, \\
-V_q(a,b)/q & \text{otherwise}.
\end{cases}
\]
By Lemma 7.4 we see that
\[
U(X) := \sum_{t \leq X \atop t \equiv b \,(\text{mod} \, W)} u(t) = G_b(a/q, X) - \frac{V_q(a, b)}{q} \sum_{t \leq X \atop t \equiv b \,(\text{mod} \, W)} 1 \ll W q X^{1-1/k}.
\]
Hence, by partial summation,
\[
\sum_{t \leq N \atop t \equiv b \,(\text{mod} \, W)} u(t)e_W(\beta t) = e_W(\beta N)U(N) - \int_1^N U(t) \frac{2\pi i\beta}{W} e_W(\beta t) \, dt \ll W q N^{1-1/k} + q |\beta| N^{2-1/k}.
\]

The following lemma tells us that the rational exponential sum \( V_q \) vanishes for small values of \( q > 1 \). This happens because of the \( w \)-smoothness of \( W \). This is also the reason why we use the \( W \)-trick in the definition of \( f_b \).

**Lemma 7.6.** Let \( a, b, q, k \in \mathbb{N} \) be such that \( k \geq 2 \) and \( (a,q) = (b,W) = 1 \). Let \( \epsilon > 0 \). Then
\[
V_q(a, b) = \begin{cases} 
eq W(ab)\sigma_W(b) & \text{if } q = 1, \\
\sigma_W(b)\sigma_{\epsilon,k}(q^{1-1/k+\epsilon}) & \text{if } q > w, \\
0 & \text{otherwise}.
\end{cases}
\]

**Proof.** Follows from [Sal20, Lemma 21] and [Hua40, Theorem].

For later use we record the following consequence of Lemmas 7.5 and 7.6.

**Lemma 7.7.** Let \( s > 2k \). Then
\[
\int_{\mathfrak{M}} |G_b(\alpha, N)|^s \, d\alpha \ll_k \sigma_W(b)^s(N/W)^{s-1}
\]
provided that \( \rho \) is small enough depending on \( k \).
Proof. Let $\alpha \in \mathcal{M}(q,a)$ and $\beta = \alpha - a/q$. By the definition of Hardy–Littlewood decomposition at the beginning of Section 7, we have $|\beta| < 1/N^{1-\rho}$ for $\rho > 0$. By [Nat96, Lemma 4.7] we see that
\[
\sum_{t \equiv b \pmod{W}} e_W(\beta t) \ll \min(N/W, \|\beta\|^{-1}).
\]
Thus by Lemmas 7.5 and 7.6,
\[
G_b(\alpha, N) \ll \epsilon, k \sigma_W(b) q^{\epsilon-1/k} \min(N/W, \|\beta\|^{-1}) + N^{1-1/k+2\rho}
\]
for any $\epsilon > 0$. Hence
\[
G_b(\alpha, N) \ll \epsilon, k \sigma_W(b) q^{\epsilon-1/k} \begin{cases} 
\|\beta\|^{-1} & \text{if } \beta \in [W/N, 1/N^{1-\rho}], \\
N/W & \text{otherwise}, 
\end{cases}
\]
provided that $\rho$ is small enough depending on $k$. Therefore
\[
\int |G_b(\alpha, N)|^s \, d\alpha = \sum_{\substack{1 \leq q \leq Q \mathcal{M}(q,a) \\ 0 \leq a < q \\ (a,q) = 1}} \int |G_b(\alpha, N)|^s \, d\alpha
\]
\[
\ll \epsilon, k \sigma_W(b)^s \sum_{q \leq Q} q^{1+(\epsilon-1/k)s} \left( \int_0^{W/N} (N/W)^s \, d\beta + \int_{W/N}^{1/N^{1-\rho}} \|\beta\|^{-s} \, d\beta \right)
\]
\[
\ll \sigma_W(b)^s (N/W)^{s-1}
\]
provided that $\epsilon$ is small enough depending on $k$. \qed

7.3. Conclusion. Now we are ready to finish the proof of Proposition 4.2 by tackling the major arc case.

Proof of Lemma 7.3. By (23) and (25) we have
\[
G_b(\alpha, WN + b) = \sum_{\substack{z \in [W] \\ z^k \equiv b \pmod{W}}} F(\alpha, z).
\]
Using (22) and Lemma 7.5 we see that
\[
\hat{v}_b(\alpha) = \frac{e_W(-b\alpha)}{\sigma_W(b)} G_b(\alpha, WN + b)
\]
\[
= \frac{e_W(-b\alpha)}{\sigma_W(b)} V_q(a,b) \sum_{\substack{t \leq WN+b \\ t \equiv b \pmod{W}}} e_W(\beta t)
\]
\[
+ O(W^2 q N^{1-1/k} + q W^2 |\beta| N^{2-1/k}).
\]
As $\alpha \in \mathcal{M}(q,a)$, we see by (21) that $q \leq N^\rho$ and $|\beta| \leq N^\rho-1$. Hence the error term in (28) is $O(N^{1-\epsilon'})$ for some $\epsilon' > 0$ provided that $\rho$ is sufficiently small depending on $k$. 

When \( q > 1 \) it follows from Lemma 7.6 that
\[
\hat{\nu}_b(\alpha) \ll_{\epsilon,k} w^{\epsilon - 1/k} N
\]
for any \( \epsilon > 0 \). By (24) we have
\[
\hat{\nu}_b(\alpha) \ll \|\alpha\|^{-1} \ll N^{-\rho}.
\]
Hence it remains to analyse the case \( q = 1 \) when \( a = 0 \) and \( \alpha = \beta \).
Therefore
\[
\hat{\nu}_b(\alpha) = \sum_{n \leq N} e(n\alpha) = e_W(-b\alpha) \sum_{n \leq WN + b, n \equiv b (\bmod W)} e_W(n\alpha),
\]
so by (28) and Lemma 7.6,
\[
\hat{\nu}_b(\alpha) = \hat{\nu}_b(\alpha) + O(N^{1-\epsilon}). \qed
\]
Proposition 4.2 now follows from Lemmas 7.2 and 7.3 by choosing \( \rho \) to be small enough depending on \( k \).

8. **Restriction estimate.** In this section, we will establish the restriction estimate (Proposition 4.3). We do it by using Vinogradov’s mean theorem and the \( \epsilon \)-removal technique. The following lemma is a consequence of Vinogradov’s mean value theorem.

**Lemma 8.1.** Let \( s \geq k(k+1) \) and \( \epsilon > 0 \). Then
\[
\|\hat{f}_b\|_s^2 \ll_{k,\epsilon} N^{s-1+\epsilon}.
\]

**Proof.** Let \( X = (WN + b)^{1/k} \) and \( t = k(k+1)/2 \). Then
\[
\|\hat{f}_b\|_{2t}^{2t} = \int \|\hat{f}_b(\alpha)\|^{2t} d\alpha = \int_{T} \sum_{n_1, \ldots, n_{2t}} f_b(n_1) \cdots f_b(n_t) f_b(n_{t+1}) \cdots f_b(n_{2t})
\]
\[
\times e(\alpha(n_1 + \cdots + n_t - n_{t+1} - \cdots - n_{2t})) d\alpha
\]
\[
\ll_k X^{2t(k-1)} \sum_{z_1, \ldots, z_{2t}} 1 = X^{2t(k-1)} \int_T \sum_{x \leq X} e(\alpha x^k) \big|^{2t} d\alpha.
\]

Let \( J_t(k)(X) \) denote the number of integral solutions of the system
\[
x_1^i + \cdots + x_t^i = x_{s+1}^i + \cdots + x_{2t}^i, \quad 1 \leq i \leq k,
\]
with \( 1 \leq x_1, \ldots, x_{2t} \leq X \).

Now by a triangle inequality application (see [Pie19, Subsection 2.1]) and Vinogradov’s mean value theorem [BDG16, Theorem 1.1] we have
\[
\left\{ \left| \sum_{\alpha \leq X} e(\alpha x^k) \right|^{2t} \right. d\alpha \ll_{t,k} X^{k(k-1)/2} J_{t,k}(X) \\
\ll_{t,k,\epsilon} X^{k(k-1)/2} X^{2t-k(k+1)/2+\epsilon} \ll X^{2t-k+\epsilon}
\]
for all \( \epsilon > 0 \). Thus
\[
\|\hat{f}_b\|_{2t}^{2t} \ll_{t,k,\epsilon} X^{2tk-k+\epsilon} \ll N^{2t-1+\epsilon/k}.
\]
Since \( |\hat{f}_b(\alpha)| \ll N \), for any \( s \geq 2t \) we have \( \|\hat{f}_b\|_s^s \ll_{k,\epsilon} N^{s-1+\epsilon} \). \( \blacksquare \)

Next we introduce the \( \epsilon \)-removal technique. The \( \epsilon \)-removal can be done using Bourgain’s strategy from \( \text{Bou89} \) Section 4], but here we use an alternative strategy that the author learned from Trevor Wooley.

**Lemma 8.2.** Let \( s_0 \geq 1 \) be such that
\[
\|\hat{f}_b\|_{s_0}^{s_0} \ll N^{s_0-1+\epsilon}
\]
for all \( \epsilon > 0 \). Then there exists \( \gamma \in (0,1) \) such that
\[
\|\hat{f}_b\|_s^s \ll_k N^{s-1} \quad \text{for all } s \geq \max(s_0 + \gamma, 4k + \gamma).
\]

**Proof.** Let \( \gamma \in (0,1) \), to be chosen later, and write \( s' = s_0 + \gamma \). Let
\[
f'(\alpha) := \frac{1}{\sigma_W(b)} \sum_{t^k \in [WN+b] \cap A \atop t^k \equiv b \, (\text{mod } W)} k t^{k-1} e_W(\alpha t^k).
\]
Since \( |f'(\alpha)| = |\hat{f}_b(\alpha)| \), it suffices to bound \( \|f'|_{s'}^s \).

Define \( B := \{ \alpha \in \mathbb{T} : |f'(\alpha)| > N^{1-1/s'} \} \) and \( I_t := \int_B |f'(\alpha)|^t d\alpha \), where \( t > 0 \). Since
\[
\int_{\mathbb{T}\setminus B} |f'(\alpha)|^{s'} d\alpha \ll N^{s'-1},
\]
it suffices to show that
\[
I_{s'} \ll N^{s'-1}.
\]

By (29) and the Cauchy–Schwarz inequality,
\[
I_{s'} = \frac{1}{\sigma_W(b)} \sum_{t^k \in [WN+b] \cap A \atop t^k \equiv b \, (\text{mod } W)} \int |f'(\alpha)|^{s'-2} f'(-\alpha) k t^{k-1} e_W(\alpha t^k) d\alpha \\
\leq \frac{1}{\sigma_W(b)} \left( \sum_{t^k \in [WN+b] \cap A \atop t^k \equiv b \, (\text{mod } W)} k t^{k-1} \right)^{1/2} \\
\times \left( \sum_{t^k \in [WN+b] \atop t^k \equiv b \, (\text{mod } W)} \int |f'(\alpha)|^{s'-2} f'(-\alpha) k^{1/2} t^{(k-1)/2} e_W(\alpha t^k) d\alpha \right)^{1/2}.
\]
Let
\[ J := \sum_{t^k \equiv b \pmod{W}} \left| \int_{t^k \in [WN+b]} \left( \sum_{k \in [WN+b]} t^k \equiv b \pmod{W} \right) \right| \]
\[ \int_{\alpha \in B} |f'(\alpha)|^{s'-2} f'(-\alpha)k^{1/2} t^{k-1/2} e_W(t^k \alpha) \, d\alpha \right|^2 \]
\[ = \int_{\alpha \in B} \int_{\beta \in B} |f'(\alpha)|^{s'-2} f'(-\alpha)|f'(\beta)|^{s'-2} \overline{f'(-\beta)} \times \sum_{t^k \equiv b \pmod{W}} k t^{k-1} e_W((\alpha - \beta)t^k) \, d\alpha \, d\beta \]
\[ \ll \int_{\alpha \in B} \int_{\beta \in B} |f'(\alpha)|^{s'-1} |f'(\beta)|^{s'-1} |g(\alpha - \beta)| \, d\alpha \, d\beta, \]
where
\[ g(\alpha) = \sum_{t^k \equiv b \pmod{W}} k t^{k-1} e_W(\alpha t^k). \]
Since
\[ \sum_{t^k \equiv b \pmod{W}} k t^{k-1} \ll \sigma_W(b)N, \]
we now see that
\[ (30) \quad I_{s'} \ll \frac{1}{\sigma_W(b)} (\sigma_W(b)N)^{1/2} J^{1/2}. \]

Let \( m \) and \( M \) be as in Section 7 with \( \rho > 0 \) to be defined later. We see that
\[ (31) \quad J \ll J_m + J_M, \]
where, for \( M \subseteq T \),
\[ J_M = \int_{\alpha \in B} \int_{\beta \in B} \int_{\alpha - \beta \in M} |f'(\alpha)|^{s'-1} |f'(\beta)|^{s'-1} |g(\alpha - \beta)| \, d\alpha \, d\beta. \]

By Lemma 7.1 we see that, whenever \( \alpha \in m \),
\[ g(\alpha) \ll_{\rho,k} \sigma_W(b)N^{1-\delta} \]
for some small \( \delta = \delta(\rho) > 0 \). Since \( s_0 > s' - 1 \), it follows from the definition of \( I_t \) that
\[ I_{s_0} \geq N^{(1-1/s')(s_0-(s'-1))} I_{s'-1}, \]
and so by the assumption \( I_{s_0} \ll N^{s_0-1+\epsilon} \) we have
\[ I_{s'-1} \leq I_{s_0} N^{(1/s'-1)(s_0-(s'-1))} \ll_{\epsilon} N^{s'-2+\epsilon+(s_0+1-s')/s'} \leq N^{s'-2+\epsilon+(1-\gamma)/s'} \]
for any $\epsilon > 0$. Thus
\begin{equation}
J_{m} \ll_{\rho,k} \sigma_{W}(b)I_{s'-1}^{2}N^{1-\delta} \ll_{\epsilon} \sigma_{W}(b)N^{2s'-3+2\epsilon-\delta+2(1-\gamma)/s'} \\
\ll \sigma_{W}(b)N^{2(s'-1)-1}
\end{equation}
provided that $2\epsilon-\delta+2(1-\gamma)/s \leq 0$. This is true if $\epsilon, 1-\gamma$ are small enough depending on $\delta$.

Let us now turn to the major arcs. Take $\zeta \in (2k, s'/2)$ and choose $h$ such that $s'/2\zeta + h(1 - 1/\zeta) = s' - 1$. Then by Hölder’s inequality,
\[
J_{\mathfrak{M}} \leq \left( \int_{\alpha \in B} \int_{\alpha - \beta \in \mathfrak{M}} |f'(\alpha)|^{s'}|g(\alpha - \beta)|^{\zeta} \, d\alpha \, d\beta \right)^{1/(2\zeta)} \\
\times \left( \int_{\alpha \in B} \int_{\alpha - \beta \in \mathfrak{M}} |f'(\beta)|^{s'}|g(\alpha - \beta)|^{\zeta} \, d\alpha \, d\beta \right)^{1/(2\zeta)} \\
\times \left( \int_{\alpha \in B} \int_{\beta \in B} |f'(\alpha)f'(\beta)|^{h} \, d\alpha \, d\beta \right)^{1-1/\zeta}.
\]
Note that $h > s'$ since $s' > 2\zeta$. By Lemma 7.7 and the definition of $h$ we have
\begin{equation}
J_{\mathfrak{M}} \ll_{k} \left( I_{s'}\sigma_{W}(b)^{\zeta}N^{\zeta-1} \right)^{2/(2\zeta)} I_{h}^{2(1-1/\zeta)} \\
\ll \left( I_{s'}\sigma_{W}(b)^{\zeta}N^{\zeta-1} \right)^{1/\zeta} I_{s'}^{2(1-1/\zeta)} N^{2(h-s')(1-1/\zeta)} \\
= \sigma_{W}(b)I_{s'}^{2-1/\zeta}N^{(s'-1)/\zeta-1},
\end{equation}
provided that $\rho$ is small enough depending on $k$. Combining (30)–(33) we get
\[
I_{s'} \ll_{k} \frac{1}{\sigma_{W}(b)}\left( \sigma_{W}(b)N \right)^{1/2}\left( \sigma_{W}(b)N^{2(s'-1)-1} + \sigma_{W}(b)I_{s'}^{2-1/\zeta}N^{(s'-1)/\zeta-1} \right)^{1/2}.
\]
Hence $I_{s'} \ll_{k} N^{s'-1}$. □

**Acknowledgements.** The author wants to thank Oleksiy Klurman for suggesting this interesting problem. The author thanks Trevor Wooley for showing the alternative way of doing the $\epsilon$-removal that we present in Section 8. The author is also grateful to his supervisor Kaisa Matomäki for many useful discussions. The author thanks the referees for careful reading of the paper and for useful comments. During the work the author was supported by Emil Aaltonen foundation.

**References**

[Bou89] J. Bourgain, *On $\Lambda(p)$-subsets of squares*, Israel J. Math. 67 (1989), 291–311.
J. Bourgain, C. Demeter, and L. Guth, *Proof of the main conjecture in Vinogradov’s mean value theorem for degrees higher than three*, Ann. of Math. (2) 184 (2016), 633–682.

T. Cochrane, M. Ostergaard, and C. Spencer, *Cauchy–Davenport theorem for abelian groups and diagonal congruences*, Proc. Amer. Math. Soc. 147 (2019), 3339–3345.

B. Green, *Roth’s theorem in the primes*, Ann. of Math. (2) 161 (2005), 1609–1636.

B. Green and I. Z. Ruzsa, *Sum-free sets in abelian groups*, Israel J. Math. 147 (2005), 157–188.

L.-K. Hua, *On an exponential sum*, J. Chinese Math. Soc. 2 (1940), 301–312.

K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Grad. Texts in Math. 84, Springer, New York, 1990.

H. Li and H. Pan, *A density version of Vinogradov’s three primes theorem*, Forum Math. 22 (2010), 699–714.

K. Matomäki, *Sums of positive density subsets of the primes*, Acta Arith. 159 (2013), 201–225.

K. Matomäki, J. Maynard, and X. Shao, *Vinogradov’s theorem with almost equal summands*, Proc. London Math. Soc. (3) 115 (2017), 323–347.

M. B. Nathanson, *Additive Number Theory*, Grad. Texts in Math. 164, Springer, New York, 1996.

M. Overholt, *A Course in Analytic Number Theory*, Grad. Stud. Math. 160, Amer. Math. Soc., Providence, RI, 2014.

L. B. Pierce, *The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth]*, Séminaire Bourbaki. Vol. 2016/2017, exp. 1134, Astérisque 407 (2019), 479–564.

J. Salmensuu, *On the Waring–Goldbach problem with almost equal summands*, Mathematika 66 (2020), 255–296.

X. Shao, *A density version of the Vinogradov three primes theorem*, Duke Math. J. 163 (2014), 489–512.

T. Tao and V. H. Vu, *Additive Combinatorics*, Cambridge Stud. Adv. Math. 105, Cambridge Univ. Press, Cambridge, 2010.

Juho Salmensuu
Department of Mathematics and Statistics
University of Turku
FI-20014 University of Turku, Finland
E-mail: juelsa@utu.fi