Forum of Mathematics, Sigma (2014), Vol. 2, e27, 26 pages
doi:10.1017/fms.2014.27

AN L-FUNCTION-FREE PROOF OF VINOGRADOV'S THREE PRIMES THEOREM

XUANCHENG SHAO
Mathematical Institute, University of Oxford, Andrew Wiles Building,
Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK;
email: Xuancheng.Shao@maths.ox.ac.uk

Received 20 November 2013; accepted 15 October 2014

Abstract

We give a new proof of Vinogradov's three primes theorem, which asserts that all sufficiently large odd positive integers can be written as the sum of three primes. Existing proofs rely on the theory of L-functions, either explicitly or implicitly. Our proof is sieve theoretical and uses a transference principle, the idea of which was first developed by Green [*Ann. of Math.* (2) **161** (3) (2005), 1609–1636] and used in the proof of Green and Tao's theorem [*Ann. of Math.* (2) **167** (2) (2008), 481–547]. To make our argument work, we also develop an additive combinatorial result concerning popular sums, which may be of independent interest.

2010 Mathematics Subject Classification: 11P32 (primary); 11P70, 11N35 (secondary)

1. Introduction

In this paper, we study additive problems involving primes. The famous Goldbach conjecture asserts that every even positive integer at least four is the sum of two primes. Although the binary Goldbach problem is considered to be beyond the scope of current techniques, its ternary analog was settled by Vinogradov [29] in 1937.

THEOREM (Vinogradov). *There exists a positive integer V such that every odd positive integer $N \geqslant V$ can be written as the sum of three primes.*

The classical approach to Vinogradov's theorem is to use the circle method, which can be found for example in [23, Ch. 8]. The major arcs analysis in the circle method relies on the equidistribution of primes in arithmetic progressions. These arithmetic progressions can have length roughly N and step some large power of $\log N$. In this regime, the equidistribution of primes in arithmetic progressions is given by the Siegel–Walfisz theorem, whose proof uses Siegel's theorem in the theory of Dirichlet L-functions and is ineffective due to the possible existence of Siegel zeros (see [3, Ch. 22]). Heath-Brown [13] (see also [16, Ch. 19]) gave a different proof of Vinogradov's theorem by directly using certain identities involving primes, but his method also requires the Siegel–Walfisz theorem.

The main purpose of this paper is to present an L-function-free proof of Vinogradov's theorem. The new argument might be interesting for at least two reasons. First, such a proof directly produces a bound for V if one keeps track of explicit constants (Vinogradov's method can be made effective as well with more effort; see [2] and [20]). This advantage now looks much less exciting in view of the recent breakthrough by Helfgott [14] which asserts that one can take $V = 7$. A discussion on obtaining a bound for V from our method is contained in Remark 5.5. Second, our method provides another example where the sieve method produces a lower bound. Sieve methods are extremely effective in giving upper bounds with the correct order of magnitude, but they generally do not provide lower bounds (this is related to the parity problem in sieve theory). Notable exceptions to this phenomenon include [6] and, more closely related to our argument, Green and Tao's theorem on finding arbitrarily long arithmetic progressions in primes [12]. A detailed account of sieve theory can be found in [7].

1.1. The transference principle in $\mathbb{Z}/N\mathbb{Z}$. In this subsection, we explain the idea of the transference principle, which is the main ingredient in our proof of Vinogradov's theorem. The transference principle was first developed by Green [9] in his proof of Roth's theorem in the primes, and has since become a powerful tool for studying additive problems in dense subsets of primes such as Green and Tao's theorem [12]. The formulation of the transference principle we give here is more similar to that appearing in [9].

The main idea of the transference principle is to transfer a problem for a sparse subset to a corresponding problem for a dense subset, as far as the sparse subset is pseudorandom in an appropriate sense. In Vinogradov's theorem, if the set of primes (a sparse set) is replaced by a dense subset of the integers with density exceeding a certain threshold, the conclusion is then a standard result in additive combinatorics.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

THEOREM 1.1 (Quantitative Cauchy–Davenport–Chowla). *Let $0 < \delta < 1$ be given. Let N be a sufficiently large positive integer. For $i = 1, 2, 3$, let $A_i \subset \mathbb{Z}/N\mathbb{Z}$ be a subset with $|A_i| = \alpha_i N$. Suppose that $\alpha_1 + \alpha_2 + \alpha_3 \geqslant 1 + \delta$. Then, for any $x \in \mathbb{Z}/N\mathbb{Z}$, there are at least cN^2 ways to write $x = a_1 + a_2 + a_3$ with $a_i \in A_i$, where $c = c(\delta) > 0$ is a constant depending only on δ.*

The statement and proof of the Cauchy–Davenport–Chowla theorem can be found in [28, Ch. 5], and this robust version is contained in [19]. If the sets A_i have the same density, then the hypothesis above is satisfied when $\alpha_i > \frac{1}{3}$, and it is easy to see that this threshold density $\frac{1}{3}$ is optimal without any assumptions on the sets A_i.

Theorem 1.1 can also be stated in terms of the characteristic functions of A_i, which are bounded by the constant function 1. Its sparse version replaces the constant function 1 by an arbitrary majorant, under certain assumptions. For the precise definitions in the pseudorandomness condition and the discrete majorant property, see Definition 3.1 below.

THEOREM 1.2 (Transference principle in $\mathbb{Z}/N\mathbb{Z}$). *Let $0 < \delta < 1$ be given. Then, for sufficiently small $\eta > 0$ and sufficiently large prime N, the following statement holds. For $i = 1, 2, 3$, let $\nu_i, a_i : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ be arbitrary functions. Let α_i be the average of a_i. Suppose that they satisfy the following assumptions.*

(1) *(Majorization condition) $0 \leqslant a_i(n) \leqslant \nu_i(n)$ for all $n \in \mathbb{Z}/N\mathbb{Z}$.*

(2) *(Mean condition) $\alpha_i \geqslant \delta$ and $\alpha_1 + \alpha_2 + \alpha_3 \geqslant 1 + \delta$.*

(3) *(Pseudorandomness condition) The majorant ν_i is η-pseudorandom.*

(4) *(Discrete majorant property) The function a_i satisfies the discrete majorant property for some $2 < q < 3$.*

Then, for any $n \in \mathbb{Z}/N\mathbb{Z}$,

$$\sum_{\substack{n_1,n_2,n_3 \\ n_1+n_2+n_3 \equiv n \pmod N}} a_1(n_1)a_2(n_2)a_3(n_3) \geqslant cN^2,$$

where $c = c(\delta) > 0$ is a constant depending only on δ.

This directly follows by Green's argument (see [19, 27]). It is usually applied as follows in studying additive problems involving dense subsets of primes. Take ν_i to be the (normalized) characteristic function of the primes, and a_i to be the (normalized) characteristic function of the dense subset of the primes.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

The assumptions are all satisfied with these choices, but the pseudorandomness condition for ν_i relies on the Siegel–Walfisz theorem.

REMARK 1.3. The virtue of working in $\mathbb{Z}/N\mathbb{Z}$ comes from its finite abelian group nature. However, a constant factor will be lost in the process of embedding subsets of integers to subsets of $\mathbb{Z}/N\mathbb{Z}$. This is something we cannot afford to lose here. More precisely, the conclusion of Theorem 1.2 counts the number of solutions to $n_1+n_2+n_3 \equiv n \pmod N$, while we are interested in solutions to $n_1+n_2+n_3 = n$ in the integers. For n close to N, we may demand the function a_i to be supported in the interval $[0, 2N/3]$. In doing so, however, we are effectively reducing the average of a_i by a factor of $\frac{2}{3}$, and thus the threshold for the average of a_i becomes $\frac{1}{2}$ rather than $\frac{1}{3}$.

1.2. Transference principle in $\mathbb{Z}$. In our proof of Vinogradov's theorem, we will choose the majorant ν_i differently so that its pseudorandomness can be established elementarily. This can be achieved by using Selberg's majorant. However, the parity phenomenon in sieve theory suggests that the mean value of Selberg's majorant is necessarily more than twice the mean value of the characteristic function of the primes. Thus Theorem 1.2 barely fails to apply to this choice of a_i and ν_i (see Remark 1.3).

The main innovation of the current paper is a new version of Theorem 1.2, which applies even when the average of a_i is slightly less than $\frac{1}{2}$. To make this possible, we will work directly in $\mathbb{Z}$. Let us first state the combinatorial result when a_i is bounded by the constant function 1. For the precise definition of the regularity condition, see Definition 2.3 below.

THEOREM 1.4 (Transference principle in $\mathbb{Z}$, $\nu = 1$ case). *Let $0 < \delta, \kappa < 1$ be given. Let N be a sufficiently large positive integer. Let $N_3 = N$, and let $N_1 = N_2 = \lfloor N/2 \rfloor$. For $i = 1, 2, 3$, let $a_i : [1, N_i] \to [0, 1]$ be an arbitrary function, and let α_i be the average of a_i. Suppose that they satisfy the following assumptions.*

(1) *(Mean condition) $\alpha_i \geqslant \delta$ and $\frac{1}{2}(\min(1, \alpha_1 + \alpha_2) + \alpha_2) + \alpha_3 \geqslant 1 + \delta$.*

(2) *(Regularity condition for a_1) The function a_1 is $(\delta/50, \kappa)$-regular.*

Then

$$\sum_{\substack{n_1,n_2,n_3 \\ n_1+n_2+n_3=N}} a_1(n_1)a_2(n_2)a_3(n_3) \geqslant cN^2,$$

where $c = c(\delta, \kappa) > 0$ is a constant depending only on δ and κ.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

If the functions a_i all have the same average, then the mean condition above is satisfied when $\alpha_i > \frac{2}{5}$, beating the $\frac{1}{2}$ barrier. Theorem 1.4 will be deduced from a robust version of Freiman's $3k-3$ theorem in Section 2. A certain regularity condition for a_i is necessary for the statement to be true. For example, consider the case when each a_i is supported on even integers and N is odd, or the case when each a_i is supported on the first $0.45N_i$ integers in $[1, N_i]$.

As in Theorem 1.2, the majorant ν can be replaced by any pseudorandom functions as long as a_i satisfies the discrete majorant property. For the precise meanings of these conditions, see Definition 3.1 below.

THEOREM 1.5 (Transference principle in $\mathbb{Z}$). *Let $0 < \delta, \kappa < 1$ be given. Then, for sufficiently small $\eta > 0$ and sufficiently large positive integer N, the following statement holds. Let $N_3 = N$, and let $N_1 = N_2 = \lfloor N/2 \rfloor$. For $i = 1, 2, 3$, let $\nu_i, a_i : [1, N_i] \to \mathbb{R}$ be arbitrary functions. Let α_i be the average of a_i. Suppose that they satisfy the following assumptions.*

(1) *(Majorization condition) $0 \leqslant a_i(n) \leqslant \nu_i(n)$ for all $1 \leqslant n \leqslant N_i$.*

(2) *(Mean condition) $\alpha_i \geqslant \delta$ and $\frac{1}{2}(\min(1, \alpha_1 + \alpha_2) + \alpha_2) + \alpha_3 \geqslant 1 + \delta$.*

(3) *(Pseudorandomness condition) The majorant ν_i is η-pseudorandom.*

(4) *(Discrete majorant property) The function a_i satisfies the discrete majorant property for some $2 < q < 3$.*

(5) *(Regularity condition for a_1) The function a_1 is $(\delta/50, \kappa)$-regular.*

Then

$$\sum_{\substack{n_1,n_2,n_3 \\ n_1+n_2+n_3=N}} a_1(n_1)a_2(n_2)a_3(n_3) \geqslant cN^2,$$

where $c = c(\delta, \kappa) > 0$ is a constant depending only on δ and κ.

This will be proved in Section 3. Working directly in $\mathbb{Z}$ requires some modifications to the traditional argument. In particular, one has to deal with problems coming from the fact that the interval $[1, N]$ is not a genuine group.

REMARK 1.6. The dependence of η on δ in Theorem 1.2 and on δ, κ in Theorem 1.5 is exponential. In the application to Roth's theorem in the primes, this causes an extra layer of logarithm in the lower bound for the density threshold. However, this extra layer of logarithm was removed by Helfgott and de Roton [15] (whose result is further improved by Naslund [21, 22]). Such an improvement comes from using a weaker L^2 estimate instead of an L^∞ estimate, but at the cost of decreasing

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

the relevant density in the dense model. Our argument is quite sensitive to this density, and for this reason we are unable to make it work in our setting.

The rest of the article is organized as follows. In Section 2, we treat an additive combinatorial problem arising from Theorem 1.4, which could be of independent interest. In Section 3, we combine this additive combinatorial result with a modification of traditional arguments to prove Theorem 1.5. In Section 4, we review the construction of Selberg's majorant. The proof that it is pseudorandom is quite standard and is given in the appendix. Finally, in Section 5, we deduce Vinogradov's theorem from Theorem 1.5.

2. Generalization of Freiman's $3k-3$ theorem to popular sums

In this section, we prove a combinatorial result related to the $\nu_i = 1$ case of Theorem 1.5. Consider the case when ν_i is the constant function 1 and a_i is the characteristic function of some subset $A_i \subset [1, N_i]$ (recall that $N_1 = N_2 = \lfloor N/2 \rfloor$ and $N_3 = N$). Theorem 1.5 claims that, if the density of A_i is larger than $\frac{2}{5}$, and if A_1 satisfies some regularity condition, then N can be written, in many ways, as $a_1 + a_2 + a_3$ with $a_i \in A_i$. This is certainly false without the regularity condition: for example, take A_i to be the set of consecutive integers starting from 1.

As an important step towards this conclusion, we need to study the problem of obtaining lower bounds on the number of popular sums in the sumset $A_1 + A_2$. More precisely, for $s \in A_1 + A_2$, let $r(s)$ be the number of ways to write $s = a_1 + a_2$ with $a_1 \in A_1$ and $a_2 \in A_2$. We are interested in lower bounds on the cardinality of the set

$$D_K(A_1, A_2) = \{s \in A_1 + A_2 : r(s) \geqslant K\}.$$

Note that, for $K = 1$, $D_1(A_1, A_2)$ is simply the sumset $A_1 + A_2$. However, we are interested in the regime where K is a small positive constant times the cardinality of A_1 or A_2.

In this direction, Green and Ruzsa [**10**] obtained the following generalization of Kneser's theorem in arbitrary finite abelian groups.

LEMMA 2.1 (Green and Ruzsa). *Let G be a finite abelian group. Let $D = D(G)$ be the size of the largest proper subgroup of G. Let $A_1, A_2 \subset G$ be subsets, and let $K > 0$ be a parameter. Suppose that* $\min(|A_1|, |A_2|) \geqslant \sqrt{K|G|}$. *Then*

$$|D_K(A_1, A_2)| \geqslant \min(|G|, |A_1| + |A_2| - D) - 3\sqrt{K|G|}.$$

When G is a cyclic group, this is almost sharp when A_1 and A_2 are arithmetic progressions of the same step. For our purposes, we would like better bounds once

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

these extreme cases are excluded. For $A_1, A_2 \subset \mathbb{Z}$, Freiman [5] has shown that the lower bound for $|A_1 + A_2| = D_1(A_1, A_2)$ can be improved if the diameters of A_1 and A_2 are large compared to $|A_1|$ and $|A_2|$. For $A \subset \mathbb{Z}$, we define the diameter of A to be the smallest d such that A is contained in an arithmetic progression of length d.

THEOREM 2.2 (Freiman). *Let $A_1, A_2 \subset \mathbb{Z}$ be finite sets with diameters d_1, d_2, respectively. Suppose that $d_1 \leqslant d_2$. Then*

$$|A_1 + A_2| \geqslant \min(|A_1| + d_2, 2|A_1| + |A_2| - 3).$$

When $A_1 = A_2 = A$ and $|A| = k$, the lower bound above reads $|A+A| \geqslant 3k-3$ if the diameter of A is large. For this reason, it is traditionally called Freiman's $3k - 3$ theorem.

Our main result in this section is a generalization of Theorem 2.2 to popular sums, which essentially states that the same lower bound above holds for $D_K(A_1, A_2)$ when $K = \gamma N$ for some small $\gamma > 0$, under some regularity assumption on A_1. Before stating the result, we first describe this regularity condition. For $y \geqslant 2$, let $P(y)$ be the product of all primes up to y.

DEFINITION 2.3. Let $0 < \beta, \kappa < 1$ be parameters. A subset $A \subset [1, N]$ is said to be (β, κ)-regular if

$$|\{(u, v) \in A \times A : u \leqslant \beta N, v \geqslant (1 - \beta)N, (v - u, P(\beta^{-1})) = 1\}| \geqslant \kappa N^2.$$

Roughly speaking, this regularity condition on A ensures that the diameter of A is approximately N, even if a small number of elements are removed from A. This definition is compatible with the (β, κ)-regularity of the characteristic function of A (see Definition 3.1 below). We now state our main result in this section.

THEOREM 2.4. *Let $\beta, \kappa > 0$ be parameters with $\beta < \frac{1}{6}$. Let $A_1, A_2 \subset [1, N]$ be arbitrary subsets with $|A_i| \geqslant 4\beta N$ $(i = 1, 2)$. Suppose that A_1 is (β, κ)-regular. Then, for $\gamma < \min(\kappa^2/(16\beta^2), \beta^2/16)$,*

$$|D_{\gamma N}(A_1, A_2)| \geqslant \min(N, |A_1| + |A_2|) + |A_2| - 9\beta N.$$

Our argument is motivated by Lev and Smelianski's proof [18] of Theorem 2.2. We embed the sets A_1 and A_2 in an appropriately chosen cyclic group and then use Lemma 2.1.

Proof. Consider the bipartite graph $\Gamma = (A_1, A_2, E)$, whose vertices are elements of A_1 and A_2, and whose edges are those pairs (a_1, a_2) $(a_1 \in A_1, a_2 \in A_2)$ with

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

$a_1 + a_2 \in D_{\gamma N}(A_1, A_2)$. Since every element $s \in (A_1 + A_2) \setminus D_{\gamma N}(A_1, A_2)$ yields at most γN edges in the complement of Γ, the edge set E contains all but at most $\gamma N \cdot |A_1 + A_2| \leqslant 2\gamma N^2$ pairs.

Let $A_1' \subset A_1$ be the set of vertices in A_1 with degree at least $|A_2| - \sqrt{\gamma}N$. Then

$$|A_1 \setminus A_1'| \leqslant \frac{2\gamma N^2}{\sqrt{\gamma}N} \leqslant 2\sqrt{\gamma}N.$$

By hypothesis, there are at least κN^2 pairs $(u, v) \in A_1 \times A_1$ with $u \leqslant \beta N$ and $v \geqslant (1-\beta)N$ such that $(v-u, P(\beta^{-1})) = 1$. The number of those pairs with either $u \notin A_1'$ or $v \notin A_1'$ is bounded above by $4\beta\sqrt{\gamma}N^2$, which is less than κN^2 by the choice of γ. Hence there exists such a pair with $u, v \in A_1'$. Let $A_1'' = A_1' \cap [u, v]$. Then

$$|A_1''| \geqslant |A_1'| - 2\beta N \geqslant |A_1| - 2(\beta + \sqrt{\gamma})N.$$

Let $d = v - u$ be the difference between the largest and the smallest elements of A_1''. Let B_1, B_2 be the images of A_1'', A_2, respectively, under the projection map $\mathbb{Z} \to \mathbb{Z}/d\mathbb{Z}$. Then $|B_1| = |A_1''| - 1$ and $|B_2| \geqslant |A_2| - 2\beta N$. We claim that

$$|D_{\gamma N}(A_1'', A_2)| \geqslant |D_{3\gamma N}(B_1, B_2)| + (|A_2| - 2\sqrt{\gamma}N). \tag{2.1}$$

In fact, for each popular sum $\bar{s} \in D_{3\gamma N}(B_1, B_2) \subset \mathbb{Z}/d\mathbb{Z}$, there are at most three different ways to lift $\bar{s}$ to an integer $s \in A_1'' + A_2$ (since $\beta < \frac{1}{6}$ and thus $d > 2N/3$). At least one of those liftings lies in $D_{\gamma N}(A_1'', A_2)$. The additional term $|A_2| - 2\sqrt{\gamma}N$ accounts for the fact that, for all but at most $2\sqrt{\gamma}N$ values of $a_2 \in A_2$, both sums $u + a_2$ and $v + a_2$ lie in $D_{\gamma N}(A_1'', A_2)$, but they are the same modulo d.

It is easy to check that $|B_i| \geqslant \sqrt{3\gamma}N$. We may thus apply Lemma 2.1 to the sets B_1, B_2 inside $G = \mathbb{Z}/d\mathbb{Z}$ to conclude that

$$|D_{3\gamma N}(B_1, B_2)| \geqslant \min(d, |B_1| + |B_2| - D) - 6\sqrt{\gamma}N,$$

where $D = D(\mathbb{Z}/d\mathbb{Z})$ is the size of the largest subgroup of $\mathbb{Z}/d\mathbb{Z}$. It follows from $(d, P(\beta^{-1})) = 1$ that $D \leqslant \beta N$. Combining this with the lower bounds for $|B_1|$, $|B_2|$, and d, we get

$$|D_{3\gamma N}(B_1, B_2)| \geqslant \min(N, |A_1| + |A_2|) - (6\beta + 8\sqrt{\gamma})N.$$

Hence, by (2.1),

$$|D_{\gamma N}(A_1, A_2)| \geqslant |D_{\gamma N}(A_1'', A_2)| \geqslant \min(N, |A_1|+|A_2|)+|A_2|-(6\beta+10\sqrt{\gamma})\beta N.$$

This is enough to conclude the proof by the choice of γ. □

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

REMARK 2.5. A central topic in additive combinatorics is the study of structures of sets with small doubling. For $A \subset \mathbb{Z}$, the doubling of A is the quantity $K = |A + A|/|A|$. Freiman's celebrated theorem gives a classification of the sets with small doubling K: they are dense subsets of generalized arithmetic progressions of rank at most K. See [28] for the precise result and its history. Theorems 2.2 and 2.4 roughly state that, if $K < 3$, then A is efficiently covered by an arithmetic progression. This gives a more precise structure than Freiman's theorem when $K < 3$. In the wider region $K < 4$, see [4] for a recent result.

3. The transference principle

In this section, we prove Theorem 1.5. The precise definitions of the pseudorandomness condition, the discrete majorant property, and the regularity condition are given as follows. For a (compactly supported) function $f : \mathbb{Z} \to \mathbb{R}$, its Fourier transform is defined by

$$\hat{f}(\theta) = \sum_{n \in \mathbb{Z}} f(n) e(n\theta),$$

where $e(n\theta) = \exp(2\pi i n\theta)$. The L^q norm of its Fourier transform is defined by

$$\|\hat{f}\|_q = \left(\int_0^1 |\hat{f}(\theta)|^q \, d\theta \right)^{1/q}.$$

For $y \geqslant 2$, let $P(y)$ be the product of all primes up to y.

DEFINITION 3.1. Let $f : [1, N] \to \mathbb{R}$ be an arbitrary function.

(1) The function f is said to be *η-pseudorandom* if $|\hat{f}(r/N) - \delta_{r,0} N| \leqslant \eta N$ for each $r \in \mathbb{Z}/N\mathbb{Z}$, where $\delta_{r,0}$ is the Kronecker delta.

(2) The function f is said to satisfy the discrete majorant property if $\|\hat{f}\|_q \ll_q N^{1-1/q}$, where the implied constant depends only on q.

(3) The function f is said to be (β, κ)-regular if

$$\sum_{(u,v) \in M} f(u) f(v) \geqslant \kappa N^2,$$

where

$$M = \{(u, v) : u \leqslant \beta N, v \geqslant (1 - \beta)N, (v - u, P(\beta^{-1})) = 1\}.$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

Note that, when f is the characteristic function of a subset $A \subset [1, N]$, (β, κ)-regularity of f is equivalent to (β, κ)-regularity of A (recall Definition 2.3).

The proof of Theorem 1.5 is similar to the arguments in [9] and [11], but with some new ingredients. In the treatment of the case $\nu_i = 1$, we use Theorem 2.4 established in the previous section. In the reduction from arbitrary ν_i to the case $\nu_i = 1$, we work directly in $\mathbb{Z}$ rather than in $\mathbb{Z}/N\mathbb{Z}$.

3.1. Proof of the case $\nu_i = 1$ (Theorem 1.4). Let $\xi > 0$ be a small parameter to be chosen later. Let $A_i \subset [1, N_i]$ be the essential support of a_i:

$$A_i = \{1 \leqslant n \leqslant N_i : a_i(n) \geqslant \xi\},$$

Then

$$|A_i| > (\alpha_i - \xi)N_i.$$

Write $\beta = \delta/50$. It follows from the regularity condition for a_1 that

$$\begin{aligned}&|\{(u, v) \in A_1 \times A_1 : u \leqslant \beta N, v \geqslant (1 - \beta)N, (v - u, P(\beta^{-1})) = 1\}| \\ &\quad \geqslant (\kappa - \xi^2\beta^2)N^2 \geqslant \tfrac{1}{2}\kappa N^2\end{aligned}$$

if ξ is chosen small enough. Hence A_1 is $(\beta, \kappa/2)$-regular. By Theorem 2.4, there exists $\gamma = \gamma(\delta, \kappa) > 0$ such that

$$\begin{aligned}|D_{\gamma N_1}(A_1, A_2)| &\geqslant \min(N_1, |A_1| + |A_2|) + |A_2| - \tfrac{1}{2}\delta N_1 \\ &\geqslant (\min(1, \alpha_1 + \alpha_2) + \alpha_2 - \delta)N_1.\end{aligned}$$

Note that $D_{\gamma N_1}(A_1, A_2)$ and A_3 are both subsets of $[1, N]$, and their densities in $[1, N]$ add up to at least $1 + \delta/4$ by the mean condition, provided that $N > 4\delta^{-1}$ is sufficiently large. Hence

$$|D_{\gamma N_1}(A_1, A_2) \cap (N - A_3)| \geqslant \tfrac{1}{4}\delta N.$$

This shows that there are at least $\delta N/4$ ways to write N as the sum of an element in $D_{\gamma N_1}(A_1, A_2)$ and an element in A_3. Each of these $\delta N/4$ representations gives rise to at least γN_1 ways to write N as $a_1 + a_2 + a_3$ $(a_i \in A_i)$. This shows that

$$\sum_{\substack{n_1, n_2, n_3 \\ n_1 + n_2 + n_3 = N}} a_1(n_1)a_2(n_2)a_3(n_3) \geqslant \xi^3 \sum_{\substack{n_i \in A_i \\ n_1 + n_2 + n_3 = N}} 1 \geqslant \tfrac{1}{4}\xi^3\delta\gamma N^2.$$

This completes the proof.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

3.2. Decomposition of a_i into uniform and anti-uniform parts. For notational convenience, in this subsection we will fix some $i \in \{1, 2, 3\}$, and simply write $a = a_i$, $\nu = \nu_i$, and $N = N_i$. The main idea of reducing from general ν to the case $\nu = 1$ is to decompose the function a into a structured part a' and a random part a''. The precise meanings of these properties are summarized in Lemma 3.3 below.

To construct this decomposition, let $0 < \epsilon < 1$ be a small parameter to be chosen later (which depends only on δ and κ). Let

$$T = T_\epsilon = \{\theta \in \mathbb{T} : |\hat{a}(\theta)| \geqslant \epsilon N\}.$$

Since a satisfies the discrete majorant property, the measure of T_ϵ satisfies the bound

$$\operatorname{meas}(T_\epsilon) \ll_\epsilon N^{-1}. \tag{3.1}$$

Define

$$B = B_\epsilon = \{1 \leqslant b \leqslant \epsilon N : \|b\theta\| < \epsilon \text{ for all } \theta \in T\},$$

where $\|x\|$ denotes the distance from x to its closest integer. The definition of B resembles the definition of Bohr sets in finite abelian groups. In that setting, lower bounds for $|B|$ are available in terms of its rank. The following lemma shows that a similar lower bound holds in our situation as well.

LEMMA 3.2. *With the definitions of $T = T_\epsilon$ and $B = B_\epsilon$ as above, we have $|B| \gg_\epsilon N$.*

Proof. For each $\theta \in T_\epsilon$ and $\ell > 0$, let $I(\theta, \ell) = [\theta - \ell/2, \theta + \ell/2]$ be the interval of length ℓ centered at θ. By compactness, there exists $\theta_1, \ldots, \theta_m \in T$ such that

$$T_\epsilon \subset I(\theta_1, \epsilon/24N) \cup \cdots \cup I(\theta_m, \epsilon/24N).$$

By the Vitali covering lemma, there exists a subcollection $\{I(\theta_j, \epsilon/24N) : j \in J\}$ consisting of disjoint intervals and satisfying

$$T_\epsilon \subset \bigcup_{j \in J} I(\theta_j, \epsilon/8N). \tag{3.2}$$

We claim that $|J| = O_\epsilon(1)$. In fact, for any $\theta \in I(\theta_j, \epsilon/8N)$ $(j \in J)$,

$$|\hat{a}(\theta) - \hat{a}(\theta_j)| \leqslant \sum_{n=1}^{N} a(n)|1 - e(n(\theta - \theta_j))| \leqslant \sum_{n=1}^{N} a(n) \cdot \frac{\epsilon n}{2N} \leqslant \frac{\epsilon}{2} \sum_{n=1}^{N} \nu(n) = \frac{1}{2}\epsilon N.$$

Hence

$$\hat{a}(\theta) \geqslant \hat{a}(\theta_j) - \tfrac{1}{2}\epsilon N \geqslant \tfrac{1}{2}\epsilon N.$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

It follows that

$$\bigcup_{j\in J} I(\theta_j, \epsilon/8N) \subset T_{\epsilon/2}.$$

Using (3.1), we get

$$\frac{\epsilon|J|}{24N} = \sum_{j\in J} \mathrm{meas}(I(\theta_j, \epsilon/24N)) \leqslant \mathrm{meas}(T_{\epsilon/2}) \ll_\epsilon \frac{1}{N}.$$

This proves that $|J| = O_\epsilon(1)$.

Now, let

$$B' = \{1 \leqslant b \leqslant \epsilon N : \|b\theta_j\| < \epsilon/2 \text{ for all } j \in J\}.$$

We claim that $B' \subset B$. To see this, take any $b \in B'$ and $\theta \in T_\epsilon$. By (3.2), $\theta \in I(\theta_j, \epsilon/8N)$ for some $j \in J$. Hence

$$\|b\theta\| \leqslant \|b\theta_j\| + b|\theta_{j_k} - \theta| < \epsilon.$$

This shows that $B' \subset B$. A lower bound for $|B'|$ can be obtained by a simple pigeonhole argument. Divide the $|J|$-dimensional cube $[0,1]^{|J|}$ into small cubes of side length $\epsilon/2$. For each $1 \leqslant b \leqslant \epsilon N$, consider the small cube to which the vector $v_b = (\|b\theta_j\|)_{j\in J}$ belongs. By the pigeonhole principle, there exists a small cube containing at least $(2/\epsilon)^{|J|}\epsilon N$ vectors v_b. For b_1, b_2 with v_{b_1}, v_{b_2} in the same small cube, the difference $|b_1 - b_2|$ is an element of B'. Hence $|B| \geqslant |B'| \gg_\epsilon N$. □

The remaining arguments go along the same line as those of Green [9, 11]. Define

$$a'(n) = \mathbb{E}_{b_1,b_2\in B} a(n+b_1-b_2) = \frac{1}{|B|^2} \sum_{b_1,b_2\in B} a(n+b_1-b_2), \quad a''(n) = a(n) - a'(n).$$

LEMMA 3.3. *Suppose that η is chosen small enough depending on ϵ. The functions a' and a'' defined above have the following properties.*

(1) *(a' is set-like)* $0 \leqslant a'(n) \leqslant 1 + O_\epsilon(\eta)$ *for any n. Moreover,* $\mathbb{E}_{1\leqslant n\leqslant N} a'(n) = \alpha + O(\epsilon)$.

(2) *(a'' is uniform)* $\hat{a}''(\theta) = O(\epsilon N)$ *for all θ.*

(3) *(a_1' is regular)* a_1' *is $(\delta/50, \kappa - O(\epsilon))$-regular.*

(4) $\|\hat{a}'\|_q \leqslant \|\hat{a}\|_q$ *and* $\|\hat{a}''\|_q \leqslant \|\hat{a}\|_q$.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

Proof. To prove (1), note that

$$a'(n) \leqslant \mathbb{E}_{b_1,b_2\in B}\nu(n+b_1-b_2) \leqslant \mathbb{E}_{b_1,b_2\in B}\mathbb{E}_{0\leqslant r<N}\hat{\nu}(r/N)e_N(r(n+b_1-b_2))$$
$$= \mathbb{E}_{0\leqslant r<N}\hat{\nu}(r/N)e_N(rn)|\mathbb{E}_{b\in B}e_N(rb)|^2.$$

The term $r=0$ gives $\hat{\nu}(0)=N(1+O(\eta))$. For $r\neq 0$, the summand is bounded in absolute value by $\eta N|\mathbb{E}_{b\in B}e_N(rb)|^2$. Hence

$$a'(n) \leqslant 1+O(\eta)+\eta N\mathbb{E}_{0\leqslant r<N}|\mathbb{E}_{b\in B}e_N(rb)|^2 = 1+O(\eta)+\eta N|B|^{-1}$$

by Parseval's identity. By Lemma 3.2,

$$a'(n) \leqslant 1+O_\epsilon(\eta).$$

If η is chosen sufficiently small, $a'(n)\leqslant 2$ for all n. The fact that $\mathbb{E}_{1\leqslant n\leqslant N}a'(n) = \alpha+O(\epsilon)$ follows since $\mathbb{E}_{n\in\mathbb{Z}}a'(n)=\alpha$ and the support of a' is contained in $[-\epsilon N,(1+\epsilon)N]$.

To prove (2), note that the Fourier transform of a'' can be written as

$$\hat{a}''(\theta) = \hat{a}(\theta)(1-|\mathbb{E}_{b\in B}e(b\theta)|^2).$$

For $\theta\notin T$, $|\hat{a}''(\theta)|\leqslant|\hat{a}(\theta)|\leqslant\epsilon N$. For $\theta\in T$, we have

$$1-|\mathbb{E}_{b\in B}e(b\theta)|^2 \leqslant 2(1-|\mathbb{E}_{b\in B}e(b\theta)|) \leqslant 2\mathbb{E}_{b\in B}|1-e(b\theta)| \ll \epsilon$$

by the definition of B. Hence $|\hat{a}''(\theta)|\ll\epsilon N$ as well.

To prove (3), write $\beta=\delta/50$. Define

$$M = \{(u,v) : 1\leqslant u\leqslant\beta N,(1-\beta)N\leqslant v\leqslant N,(v-u,P(\beta^{-1}))=1\},$$

and

$$M' = \{(u,v) : -\epsilon N\leqslant u\leqslant(\beta+\epsilon)N,(1-\beta-\epsilon)N\leqslant v\leqslant(1+\epsilon)N,$$
$$(v-u,P(\beta^{-1}))=1\}.$$

Note that

$$\sum_{(u,v)\in M'}a'(u)a'(v) = \mathbb{E}_{b_1,b_2,b_3,b_4\in B}\sum_{(u,v)\in M'}a(u+b_1-b_2)a(v+b_3-b_4)$$
$$\geqslant \mathbb{E}_{b_1,b_2,b_3,b_4\in B}\sum_{(u,v)\in M}a(u)a(v) \geqslant \kappa N^2.$$

Hence

$$\sum_{(u,v)\in M}a'(u)a'(v) \geqslant \sum_{(u,v)\in M'}a'(u)a'(v)-2|M'\setminus M| \geqslant (\kappa-O(\epsilon))N^2.$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

To prove (4), note that, for any θ,

$$\hat{a}'(\theta) = \hat{a}(\theta)|\mathbb{E}_{b\in B}e(b\theta)|^2, \quad \hat{a}''(\theta) = \hat{a}(\theta)(1 - |\mathbb{E}_{b\in B}e(b\theta)|^2),$$

and thus $|\hat{a}'(\theta)| \leqslant |\hat{a}(\theta)|$ and $|\hat{a}''(\theta)| \leqslant |\hat{a}(\theta)|$. □

3.3. Reduction to the case $\nu_i = 1$. For each $i \in \{1, 2, 3\}$, we obtained a decomposition $a_i = a_i' + a_i''$ satisfying the conditions summarized in Lemma 3.3. In this section, we will show that the contributions from a_i'' are negligible, and thus we may essentially replace a_i by a_i'. Now that the functions a_i' are essentially bounded above by 1, we are back in the case $\nu_i = 1$ treated in Theorem 1.4.

LEMMA 3.4. *With the functions a_i, a_i' defined as above, we have*

$$\left|\sum_{n,m} a_1(n)a_2(m)a_3(N-n-m) - \sum_{n,m} a_1'(n)a_2'(m)a_3'(N-n-m)\right| \ll \epsilon^{3-q}N^2.$$

Proof. The difference on the left can be expressed as a sum of several terms, each of the form

$$\sum_{\substack{n_1,n_2,n_3\\ n_1+n_2+n_3=N}} f_1(n_1)f_2(n_2)f_3(n_3) = \int_0^1 \hat{f}_1(\theta)\hat{f}_2(\theta)\hat{f}_3(\theta)e(-N\theta)\,d\theta,$$

where $f_i \in \{a_i, a_i', a_i''\}$, and $f_i = a_i''$ for at least one i. Without loss of generality, assume that $f_3 = a_3''$. By Hölder's inequality, this is bounded above by

$$\|\hat{f}_3\|_\infty^{3-q}\|\hat{f}_3\|_q^{q-2}\|\hat{f}_1\|_q\|\hat{f}_2\|_q.$$

By Lemma 3.3, $\|\hat{f}_3\|_\infty \ll \epsilon N$. By the discrete majorant property together with Lemma 3.3, all of $\|\hat{f}_3\|_q$, $\|\hat{f}_1\|_q$, and $\|\hat{f}_2\|_q$ are bounded above by $O_q(N^{1-1/q})$. Combining these, we get the desired bound. □

We now finish the proof of Theorem 1.5. By Lemma 3.3, the functions a_i' are all bounded above uniformly by $1 + O_\epsilon(\eta)$ with averages $\alpha + O(\epsilon)$, and a_1' is $(\delta/50, \kappa/2)$-regular. If ϵ and η are chosen small enough, Theorem 1.4 then implies that

$$\sum_{n,m} a_1'(n)a_2'(m)a_3'(N-n-m) \geqslant cN^2$$

for some $c = c(\delta, \kappa) > 0$. Combining this with Lemma 3.4, we deduce by choosing ϵ small enough that

$$\sum_{n,m} a_1(n)a_2(m)a_3(N-n-m) \geqslant \tfrac{1}{2}cN^2.$$

This completes the proof of Theorem 1.5.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

4. Pseudorandomness of Selberg's majorant

To apply the transference principle, we need a majorant for the primes whose pseudorandomness can be verified in an elementary way. For this purpose, we will use Selberg's upper bound sieve. It is a basic and important tool in sieve theory, and our notation here will follow that in [7, Ch. 7].

Let $W = \prod_{p \leqslant w} p$ be the product of primes up to some large constant w, and let $b \pmod W$ be a reduced residue class. Fix a small positive constant $\delta > 0$. Let N be sufficiently large depending on w and δ. Let $z = N^{1/2-\delta}$, and let $D = z^2$. Let P be the product of all primes $p < z$ and $(p, W) = 1$. Define $\nu = \nu(N, z, W, b) : [N] \to \mathbb{R}_{\geqslant 0}$ by

$$\nu(n) = \frac{\phi(W)}{W} \log z \left(\sum_{d \mid (Wn+b, P)} \rho_d \right)^2.$$

Here, the weights ρ_d are supported on $d < z$, and they satisfy $|\rho_d| \leqslant 1$ and $\rho_1 = 1$. Moreover, the new variables

$$y_d = \mu(d)\phi(d) \sum_{d \mid m} \frac{\rho_m}{m}$$

satisfy $y_d = J^{-1}$ for $d < z$, where

$$J = \sum_{\substack{d \mid P \\ d < z}} \frac{1}{\phi(d)} = \sum_{\substack{d < z \\ (d, W) = 1}} \frac{1}{\phi(d)}.$$

To see that ν is indeed a majorant for the (W-tricked) primes, note that, if $Wn + b$ is prime and $Wn + b \geqslant z$, then

$$\nu(n) = \frac{\phi(W)}{W} \log z \tag{4.1}$$

since $\rho_1 = 1$.

THEOREM 4.1 (Selberg's majorant is pseudorandom). *Let $\nu : [N] \to \mathbb{R}_{\geqslant 0}$ be defined as above. For any $r \in \mathbb{Z}/N\mathbb{Z}$,*

$$\hat{\nu}(r) = (\delta_{r,0} + O_\epsilon(w^{-1+\epsilon}))N,$$

where $\delta_{r,0}$ is the Kronecker delta. In other words, ν is $O_\epsilon(w^{-1+\epsilon})$-pseudorandom.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

The proof of this is quite standard, and is included in the appendix as we are not able to find in the literature exactly what we need. It divides into two cases depending on whether r/N lies in the major arc or minor arc. In the major arc case when $r/N \approx a/q$ for some reduced fraction a/q with q small, we can get an asymptotic formula for $\hat{\nu}(r)$ whose leading term can be analyzed by standard manipulations using mean value estimates for multiplicative functions. The minor arc case follows from a bilinear form estimate.

5. Proof of Vinogradov's theorem

In this section, we use Theorem 1.5 with Selberg's majorant considered in Section 4 to give a proof of Vinogradov's three primes theorem without using the theory of L-functions. In particular, we will not need the Siegel–Walfisz theorem, although we still use the prime number theorem in arithmetic progressions with constant modulus, which can be proved elementarily. Such an elementary proof was first given by Selberg [**26**].

REMARK 5.1. Recently, Koukoulopoulos [**17**] gave a 'pretentious' proof of the Siegel–Walfisz theorem; namely, the proof uses L-functions only when the defining Dirichlet series is convergent. The bound is still ineffective, due to the potential existence of a Siegel zero causing an extremely small value of $L(1, \chi)$. See also [**8**] for an introduction to the pretentious approach in analytic number theory.

Let M be a sufficiently large odd positive integer. We will prove that M can be written as sum of three primes. Take $\delta = 0.01$ in the statement of Theorem 1.5. Let $W = P(w)$ be a parameter to be chosen later. Choose $0 < b_1, b_2, b_3 < W$ with $(b_i, W) = 1$ such that $b_1 + b_2 + b_3 \equiv M \pmod{W}$ (this can always be done by the Chinese remainder theorem). Let $N = (M - b_1 - b_2 - b_3)/W$. Let $N_3 = N$, and let $N_1 = N_2 = \lfloor N/2 \rfloor$. For $i = 1, 2, 3$, define a function $a_i : [1, N_i] \to \mathbb{R}$ by

$$a_i(n) = \begin{cases} \dfrac{\phi(W)}{W} \log z_i & Wn + b_i \text{ is prime and } Wn + b_i \geqslant z_i \\ 0 & \text{otherwise,} \end{cases}$$

where $z_i = N_i^{0.49}$. Construct $\nu_i = \nu(N_i, z_i, W, b_i)$ as in Section 4.

The majorization condition is satisfied by the observation (4.1). The mean condition is satisfied because the average of a_i is at least 0.48 for sufficiently large N by the prime number theorem in arithmetic progressions of modulus W (which is a constant). The pseudorandomness condition is satisfied by Theorem 4.1, if w is chosen large enough.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

Now consider the regularity condition for a_1. Write $\beta = \delta/50$, $y = \beta^{-1}$, $Y = P(y)$, and let

$$M = \{(u, v) : u \leqslant \beta N_1,\ v \geqslant (1-\beta)N_1,\ (v-u, Y) = 1\}.$$

Also write

$$U = \{1 \leqslant u \leqslant \beta N_1 : Wu + b_1 \text{ is prime}\},$$
$$V = \{(1-\beta)N_1 \leqslant v \leqslant N_1 : Wv + b_1 \text{ is prime}\}.$$

$$\begin{aligned}\sum_{(u,v)\in M} a_1(u)a_1(v) &= \left(\frac{\phi(W)}{W}\log z_1\right)^2 \sum_{\substack{u\in U, v\in V\\ (v-u,Y)=1}} 1\\ &\geqslant \left(\frac{\phi(W)}{W}\log z_1\right)^2 \sum_{\substack{s_1,s_2 \pmod Y\\ (s_2-s_1,Y)=1}} |U\cap(Y\mathbb{Z}+s_1)|\\ &\quad\cdot |V\cap(Y\mathbb{Z}+s_2)|\\ &\geqslant \left(\frac{\phi(W)}{W}\log z_1\right)^2 Y\phi(Y)\left(\frac{\beta N_1}{2\log N_1}\cdot\frac{W}{\phi(W)}\cdot\frac{1}{Y}\right)^2 \geqslant \kappa N^2\end{aligned}$$

for some κ depending only on δ. Here, we have used the prime number theorem in arithmetic progressions of modulus WY, which is again a constant.

Finally, the discrete majorant property for a_i follows from the result of Green and Tao [11].

LEMMA 5.2. *For any $q > 2$,*

$$\left(\int_0^1 |\hat{a}_i(\theta)|^q\, d\theta\right)^{1/q} \ll_q N^{1-1/q}.$$

Proof. Consider the linear function $F(n) = Wn + b_i$ and the exponential sum

$$h(\theta) = \sum_{\substack{n\leqslant N_i\\ F(n)\geqslant z_i\\ F(n)\text{ prime}}} e(n\theta).$$

The argument leading to [11, Theorem 1.1] gives

$$\|h\|_q \ll_q \mathfrak{S}_F N_i^{1-1/q}(\log N_i)^{-1},$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

where the singular series $\mathfrak{G}_F$ is defined by

$$\mathfrak{G}_F = \prod_{p \text{ prime}} \gamma(p) \left(1 - \frac{1}{p}\right)^{-1}$$

and

$$\gamma(p) = p^{-1} |\{n \in \mathbb{Z}/p\mathbb{Z} : (p, F(n)) = 1\}|.$$

(See (1.2) and (1.7) in [11]). In the current case, $\gamma(p) = 1$ for $p \leqslant w$ and $\gamma(p) = 1 - 1/p$ for $p > w$. Hence

$$\mathfrak{G}_F = \prod_{p \leqslant w} \frac{p}{p-1} = \frac{W}{\phi(W)}.$$

Finally, note that

$$\hat{a}_i(\theta) = \left(\frac{\phi(W)}{W} \log z_i\right) h(\theta).$$

It follows that

$$\|\hat{a}_i\|_q \leqslant \left(\frac{\phi(W)}{W} \log z_i\right) \|h\|_q \ll_q N_i^{1-1/q}. \qquad \square$$

REMARK 5.3. Lemma 5.2 was also proved in [9], using the Brun sieve and the Siegel–Walfisz theorem. Bourgain [1] showed how to obtain bounds for $\|\hat{f}\|_q$, where f is a function supported on the primes. The proof in [11] differs from these previous arguments, and solely depends on properties of an enveloping sieve (see also [24, 25]); in particular, the theory of L-functions is not used.

Now that all hypotheses in the statement of Theorem 1.5 are verified, we conclude that there exists $n_i \in [1, N_i]$ with $a_i(n_i) > 0$ such that $N = n_1 + n_2 + n_3$. In particular, $Wn_i + b_i$ is prime, and

$$M = WN + b_1 + b_2 + b_3 = (Wn_1 + b_1) + (Wn_2 + b_2) + (Wn_3 + b_3),$$

proving that M is the sum of three primes.

REMARK 5.4. Our method actually produces a lower bound for the number of representations of M as the sum of three primes, which is of the correct order of magnitude, but with a poor constant in the front. In comparison, the traditional circle method is able to produce an asymptotic formula for this number of representations.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

REMARK 5.5. We make a final remark concerning the explicit bound for M that can be produced from our method. Unfortunately, directly following our arguments only gives $M \geqslant \exp(\exp(\exp(C)))$ for a reasonable constant C. This can be seen as follows. For our choice of δ, the transference principle theorem 1.5 requires the parameter η to be exponential in $1/\delta$. Thus, by the pseudorandomness estimate Theorem 4.1, the parameter w should be taken to be exponential in $1/\delta$. Hence W, being the product of primes up to w, becomes double exponential in $1/\delta$. Finally, in the arguments in this section we used lower bounds on the number of primes up to M in congruence classes modulo W. Such lower bounds are only available when M is exponential in W, and thus triple exponential in $1/\delta$.

Acknowledgements

The author would like to express his gratitude to B. Green for sharing with him the idea of proving Vinogradov's theorem using a transference principle, to Kannan Soundararajan for carefully reading early drafts of the paper and providing many useful comments, and to D. Koukoulopoulos for explaining the work [17]. He is also grateful to the anonymous referee for valuable suggestions.

Appendix A. Proof of Theorem 4.1

This appendix is devoted to proving Theorem 4.1. We follow the notation in Section 4. In particular, recall the construction of $\nu = \nu(N, z, W, b) : [N] \to \mathbb{R}_{\geqslant 0}$ from the weights $\{\rho_d\}$. The following basic estimate will be used multiple times.

LEMMA A.1. *For any $z \geqslant 2$ and positive integer m dividing P,*

$$\sum_{\substack{d<z\\(d,m)=1}} \frac{1}{\phi(d)} \ll \frac{\phi(m)}{m} \log z$$

and

$$\sum_{\substack{d<z\\(d,m)=1}} \frac{1}{\phi(d)} = \frac{\phi(m)}{m}(\log z + O_m(1)).$$

Proof. The upper bound is clear:

$$\sum_{\substack{d<z\\(d,m)=1}} \frac{1}{\phi(d)} \leqslant \prod_{\substack{p<z\\p\nmid m}} \left(1 + \frac{1}{\phi(p)}\right) = \frac{\phi(m)}{m} \prod_{p<z} \left(1 + \frac{1}{p-1}\right) \ll \frac{\phi(m)}{m} \log z.$$

For the asymptotic, see [7, Theorem A.8]. □

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

In particular, Lemma A.1 implies that

$$J = \frac{\phi(W)}{W}(\log z + O_W(1)). \tag{A.1}$$

LEMMA A.2. *For any positive integers q and r dividing P, the sum*

$$J(q,r) = \sum_{\substack{d|P \\ (d,q)=1}} \frac{\rho_{rd}}{d}$$

satisfies

$$|J(q,r)| \leqslant J^{-1}\frac{[q,r]}{\phi([q,r])}.$$

Moreover, $J(q,q) = y_q q\mu(q)/\phi(q)$.

Proof. We write

$$J(q,r) = \sum_{d|P}\frac{\rho_{rd}}{d}\sum_{e|(d,q)}\mu(e) = \sum_{e|q}\mu(e)\sum_{\substack{d|P \\ e|d}}\frac{\rho_{rd}}{d}.$$

Note that $\rho_{rd} = 0$ if rd is not squarefree. Hence we can restrict the sum to those e with $(e,r) = 1$:

$$\begin{aligned} J(q,r) &= \sum_{e|q/(q,r)}\mu(e)\sum_{\substack{d|P \\ e|d}}\frac{\rho_{rd}}{d} = r\sum_{e|q/(q,r)}\mu(e)y_{re}\mu(re)\phi(re)^{-1} \\ &= \frac{r\mu(r)}{\phi(r)}\sum_{e|q/(q,r)}\frac{y_{re}}{\phi(e)}. \end{aligned}$$

If $q = r$, then $q/(q,r) = 1$, and thus

$$J(q,q) = \frac{q\mu(q)}{\phi(q)}y_q.$$

In general, since y_{re} is bounded by J^{-1}, it follows that

$$|J(q,r)| \leqslant J^{-1}\frac{r}{\phi(r)}\sum_{e|q/(q,r)}\frac{1}{\phi(e)} = J^{-1}\frac{r}{\phi(r)}\frac{q/(q,r)}{\phi(q/(q,r))} = J^{-1}\frac{[q,r]}{\phi([q,r])}. \qquad \square$$

LEMMA A.3. *For any positive integer q dividing P, the sum*

$$T(q) = \sum_{\substack{d_1,d_2|P \\ q|[d_1,d_2]}}\frac{\rho_{d_1}\rho_{d_2}}{[d_1,d_2]}$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

satisfies

$$|T(q)| \ll_\epsilon J^{-1} q^{-1+\epsilon}.$$

Moreover, $T(1) = J^{-1}$.

Proof. Write $e_1 = (d_1, q)$, $d_1 = e_1 f_1$, $e_2 = (d_2, q)$, and $d_2 = e_2 f_2$. Then

$$T(q) = \sum_{\substack{e_1, e_2 | q \\ [e_1, e_2] = q}} \sum_{\substack{f_1, f_2 | P \\ (f_1, q) = (f_2, q) = 1}} \frac{\rho_{e_1 f_1} \rho_{e_2 f_2}}{q[f_1, f_2]}.$$

For fixed e_1, e_2, use the identities $(f_1, f_2)[f_1, f_2] = f_1 f_2$ and $(f_1, f_2) = \sum_{g | (f_1, f_2)} \phi(g)$ to rewrite the inner sum as

$$\begin{aligned} &\frac{1}{q} \sum_{\substack{f_1, f_2 | P \\ (f_1, q) = (f_2, q) = 1}} \frac{\rho_{e_1 f_1} \rho_{e_2 f_2}}{f_1 f_2} \sum_{g | (f_1, f_2)} \phi(g) \\ &\quad = \frac{1}{q} \sum_{\substack{g | P \\ (g, q) = 1}} \phi(g) \left(\sum_{\substack{f_1 | P \\ (f_1, q) = 1 \\ g | f_1}} \frac{\rho_{e_1 f_1}}{f_1} \right) \left(\sum_{\substack{f_2 | P \\ (f_2, q) = 1 \\ g | f_2}} \frac{\rho_{e_2 f_2}}{f_2} \right). \end{aligned}$$

The two sums in the parentheses above are $g^{-1} J(qg, e_1 g)$ and $g^{-1} J(qg, e_2 g)$. When $q = 1$, apply Lemma A.2 to get

$$T(1) = \sum_{g | P} \phi(g) (g^{-1} J(g, g))^2 = \sum_{g | P} \frac{y_g^2}{\phi(g)} = J^{-2} \sum_{\substack{g | P \\ g < z}} \frac{1}{\phi(g)} = J^{-1}.$$

In general, Lemma A.2 gives the bounds

$$|g^{-1} J(qg, e_1 g)| \leqslant J^{-1} \frac{q}{\phi(qg)}, \quad |g^{-1} J(qg, e_2 g)| \leqslant J^{-1} \frac{q}{\phi(qg)}.$$

Observe that there are $3^{\omega(q)}$ pairs (e_1, e_2) with $[e_1, e_2] = q$. Note also that we can clearly restrict the sum to $g < z$. Hence, by Lemma A.1,

$$\begin{aligned} |T(q)| &\leqslant 3^{\omega(q)} J^{-2} \frac{q}{\phi(q)^2} \sum_{\substack{g < z \\ (g, qW) = 1}} \frac{1}{\phi(g)} \\ &\ll 3^{\omega(q)} J^{-2} \frac{q}{\phi(q)^2} \frac{\phi(qW)}{qW} \log z \ll J^{-1} \frac{3^{\omega(q)}}{\phi(q)}. \end{aligned}$$

The desired bound for $|T(q)|$ follows because $3^{\omega(q)} \ll_\epsilon q^\epsilon$ and $\phi(q) \gg_\epsilon q^{1-\epsilon}$. □

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

We are now ready to prove Theorem 4.1. Let $R = \lfloor N^{1-\delta/2} \rfloor$ and $Q = \lfloor N^{\delta/4} \rfloor$ be parameters. For $q \leqslant Q$ and $(a, q) = 1$, let

$$\mathfrak{M}(q, a) = \left\{ r \in \mathbb{Z}/N\mathbb{Z} : \left| \frac{r}{N} - \frac{a}{q} \right| \leqslant \frac{1}{qR} \right\}.$$

Let

$$\mathfrak{M} = \bigcup_{q=1}^{Q} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathfrak{M}(q, a), \quad \mathfrak{m} = \mathbb{Z}/N\mathbb{Z} \setminus \mathfrak{M}.$$

A.1. Major arc analysis. In this subsection, we prove Theorem 4.1 for those $r \in \mathfrak{M}$. Suppose that $r \in \mathfrak{M}(q, a)$ for some $q \leqslant Q$ and $(a, q) = 1$. Then r/N is very close to a/q. We first prove a result when they are equal. Recall the quantity $T(q)$ defined in Lemma A.3.

PROPOSITION A.4. *With notation as above, for* $1 \leqslant x \leqslant N$,

$$f(x, a/q) = \sum_{n \leqslant x} \nu(n) e_q(an) = \frac{\phi(W)}{W} \log z (\varepsilon x T(q) + E(x, q)),$$

where $\varepsilon = \varepsilon(a/q, W, b)$ *does not depend on* x, *and* $E(x, q) = O(qN^{1-\delta})$. *Moreover,* $\varepsilon = 1$ *if* $q = 1$, $\varepsilon = 0$ *if* $(q, W) > 1$, *and* $|\varepsilon| = 1$ *if* $(q, W) = 1$.

Proof. By the definition of $\nu(n)$, we can write

$$f(x, a/q) = \frac{\phi(W)}{W} \log z \sum_{d_1, d_2 | P} \rho_{d_1} \rho_{d_2} \sum_{\substack{n \leqslant x \\ [d_1, d_2] | Wn + b}} e_q(an).$$

Split the sum into two parts:

$$f(x, a/q) = \frac{\phi(W)}{W} \log z (S_1 + S_2),$$

where

$$S_1 = \sum_{\substack{d_1, d_2 | P \\ q | [d_1, d_2]}} \rho_{d_1} \rho_{d_2} \sum_{\substack{n \leqslant x \\ [d_1, d_2] | Wn + b}} e_q(an),$$

$$S_2 = \sum_{\substack{d_1, d_2 | P \\ q \nmid [d_1, d_2]}} \rho_{d_1} \rho_{d_2} \sum_{\substack{n \leqslant x \\ [d_1, d_2] | Wn + b}} e_q(an).$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

First, consider S_1. For $q \mid [d_1, d_2]$, the inner sum is zero if $(q, W) > 1$. Take $\epsilon = 0$ in the case when $(q, W) > 1$. If $(q, W) = 1$, then the summand in the inner sum is a constant ϵ with $|\epsilon| = 1$. Moreover, $\epsilon = 1$ when $q = 1$. In either case,

$$S_1 = \epsilon \sum_{\substack{d_1, d_2 | P \\ q | [d_1, d_2]}} \rho_{d_1} \rho_{d_2} \left(\frac{x}{[d_1, d_2]} + O(1) \right)$$
$$= \epsilon x T_q + O\left(\left(\sum_{d < z} |\rho_d| \right)^2 \right) = \epsilon x T_q + O(N^{1-\delta})$$

since $|\rho_d| \leqslant 1$.

Now, consider S_2. For $d_1, d_2 \leqslant z$ with $(d_1 d_2, W) = 1$ and $q \nmid [d_1, d_2]$, the inner sum over n is bounded by q. Hence

$$S_2 \leqslant q \left(\sum_{d \leqslant z} |\rho_d| \right)^2 \leqslant q N^{1-\delta}.$$

The proof is completed by combining the estimates for S_1 and S_2. □

We now use partial summation to complete the major arc estimate. Let $r \in \mathcal{M}(q, a)$ for some $q \leqslant Q$ and $(a, q) = 1$. Then $r/N = a/q + \beta$ for some $|\beta| \leqslant 1/qR$. Note that

$$\hat{\nu}(r) = \sum_{n=1}^{N} \nu(n) e_q(an) e(\beta n) = \int_1^N e(\beta x) d \left(\sum_{n \leqslant x} \nu(n) e_q(an) \right).$$

It follows from Proposition A.4 that

$$\hat{\nu}(r) = \frac{\phi(W)}{W} \log z \left(\epsilon T(q) \int_1^N e(\beta x)\, dx + \int_1^N e(\beta x)\, dE(x, q) \right).$$

Consider the second integral above. By partial summation, it is bounded by

$$E(N, q) + \int_1^N E(x, q)(2\pi i \beta) e(\beta x)\, dx \ll q N^{1-\delta} + |\beta| q N^{2-\delta} \leqslant Q N^{1-\delta} + \frac{N^{2-\delta}}{R}.$$

This is $O(N^{1-\delta/2})$ by the choices of Q and R. Hence

$$\hat{\nu}(r) = \frac{\phi(W)}{W} \log z \left(\epsilon T(q) \int_1^N e(\beta x)\, dx + O(N^{1-\delta/2}) \right).$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

If $q > w$, then Theorem 4.1 follows from Lemma A.3 and (A.1). If $1 < q \leqslant w$, then $(q, W) > 1$, and thus $\epsilon = 0$. If $q = 1$ and $\beta > 0$, then β is an integer multiple of $1/N$, and thus the integral above is zero. Finally, if $q = 1$ and $\beta = 0$, then $\epsilon = 1$. Lemma A.3 and (A.1) give

$$\hat{\nu}(0) = \frac{\phi(W)}{W} \log z (J^{-1} + O(N^{-\delta/2}))N = (1 + O_W((\log z)^{-1}))N.$$

This proves Theorem 4.1 for sufficiently large z.

A.2. Minor arc analysis. Now, consider the case when $r \in \mathfrak{m}$. This means that

$$\left| \frac{r}{N} - \frac{a}{q} \right| \leqslant \frac{1}{q^2}$$

for some $Q \leqslant q \leqslant R$ and $(a, q) = 1$.

By the definition of $\nu(n)$, we can write

$$\hat{\nu}(r) = \frac{\phi(W)}{W} \log z \sum_{d_1, d_2 | P} \rho_{d_1} \rho_{d_2} \sum_{\substack{1 \leqslant n \leqslant N \\ [d_1, d_2] | Wn + b}} e_N(rn).$$

Using the bound $|\rho_d| \leqslant 1$, we obtain

$$|\hat{\nu}(r)| \leqslant \frac{\phi(W)}{W} \log z \sum_{\substack{d < z^2 \\ (d, W) = 1}} \left(\sum_{\substack{d_1, d_2 \\ [d_1, d_2] = d}} 1 \right) \left(\sum_{\substack{1 \leqslant n \leqslant N \\ d | Wn + b}} e_N(rn) \right).$$

For any fixed squarefree $d < z^2$, there are at most $3^{\omega(d)} \ll d^{\delta/8}$ pairs (d_1, d_2) with $[d_1, d_2] = d$. Hence

$$|\hat{\nu}(r)| \ll \frac{\phi(W)}{W} (\log z) N^{\delta/8} \sum_{d < z^2} \left| \sum_{\substack{1 \leqslant n \leqslant N \\ d | Wn + b}} e_N(rn) \right|.$$

The following lemma estimates this double sum.

LEMMA A.5. *Suppose that*

$$\left| \alpha - \frac{a}{q} \right| \leqslant \frac{1}{q^2}$$

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

with $(a, q) = 1$. *For any* $1 \leqslant m \leqslant M$, *let* $c_m \pmod m$ *be an arbitrary residue class. Then*

$$\sum_{1\leqslant m\leqslant M}\left|\sum_{\substack{1\leqslant n\leqslant x\\ n\equiv c_m \pmod m}} e(\alpha n)\right| \ll (M + xq^{-1} + q)\log(2qx).$$

Proof. See [16, Lemma 13.7]. □

It follows that

$$|\hat{\nu}(r)| \ll \frac{\phi(W)}{W}(\log z)N^{\delta/8}(z^2 + NQ^{-1} + R)\log N \ll N^{1-\delta/4},$$

completing the proof of Theorem 4.1 in the minor arc case.

References

[1] J. Bourgain, 'On $\Lambda(p)$-subsets of squares', *Israel J. Math.* **67**(3) (1989), 291–311.
[2] J. R. Chen and T. Z. Wang, 'On the Goldbach problem', *Acta Math. Sin.* **32**(5) (1989), 702–718.
[3] H. Davenport, *Multiplicative Number Theory*, 3rd edn, Graduate Texts in Mathematics, 74 (Springer, New York, 2000), revised and with a preface by H. L. Montgomery.
[4] S. Eberhard, B. Green and F. Manners, 'Sets of integers with no large sum-free subset', *Ann. of Math.* (2) **180**(2) (2014), 621–652.
[5] G. A. Freiman, 'Inverse problems in additive number theory VI. On the addition of finite sets III', *Izv. Vyssh. Uchebn. Zaved. Mat.* **3**(28) (1962), 151–157 (in Russian).
[6] J. Friedlander and H. Iwaniec, 'The polynomial $X^2 + Y^4$ captures its primes', *Ann. of Math.* (2) **148**(3) (1998), 945–1040.
[7] J. Friedlander and H. Iwaniec, *Opera de Cribro*, American Mathematical Society Colloquium Publications, 57 (American Mathematical Society, Providence, RI, 2010).
[8] A. Granville, 'Different approaches to the distribution of primes', *Milan J. Math.* **78**(1) (2010), 65–84.
[9] B. Green, 'Roth's theorem in the primes', *Ann. of Math.* (2) **161**(3) (2005), 1609–1636.
[10] B. Green and I. Z. Ruzsa, 'Sum-free sets in abelian groups', *Israel J. Math.* **147** (2005), 157–188.
[11] B. Green and T. Tao, 'Restriction theory of the Selberg sieve, with applications', *J. Théor. Nombres Bordeaux* **18**(1) (2006), 147–182.
[12] B. Green and T. Tao, 'The primes contain arbitrarily long arithmetic progressions', *Ann. of Math.* (2) **167**(2) (2008), 481–547.
[13] D. R. Heath-Brown, 'The ternary Goldbach problem', *Rev. Mat. Iberoam.* **1**(1) (1985), 45–59.
[14] H. A. Helfgott, 'The ternary Goldbach conjecture is true', Preprint (2013), arXiv:1312.7748.
[15] H. A. Helfgott and A. de Roton, 'Improving Roth's theorem in the primes', *Int. Math. Res. Not. IMRN* **2011**(4) (2011), 767–783.
[16] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, 53 (American Mathematical Society, Providence, RI, 2004).

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press

[17] D. Koukoulopoulos, 'Pretentious multiplicative functions and the prime number theorem for arithmetic progressions', *Compositio Math.* **149**(7) (2013), 1129–1149.
[18] V. F. Lev and P. Y. Smeliansky, 'On addition of two distinct sets of integers', *Acta Arith.* **70**(1) (1995), 85–91.
[19] H. Li and H. Pan, 'A density version of Vinogradov's three primes theorem', *Forum Math.* **22**(4) (2010), 699–714.
[20] M.-Ch. Liu and T. Wang, 'On the Vinogradov bound in the three primes Goldbach conjecture', *Acta Arith.* **105**(2) (2002), 133–175.
[21] E. Naslund, 'On improving Roth's theorem in the primes', *Mathematika*, to appear.
[22] E. Naslund, 'A density increment approach to Roth's theorem in the primes', Preprint (2014), arXiv:1409.3595.
[23] M. B. Nathanson, *Additive Number Theory, The Classical Bases*, Graduate Texts in Mathematics, 164 (Springer, New York, 1996).
[24] O. Ramaré, 'On Šnirel'man's constant', *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (4) **22**(4) (1995), 645–706.
[25] O. Ramaré and I. Z. Ruzsa, 'Additive properties of dense subsets of sifted sequences', *J. Théor. Nombres Bordeaux* **13**(2) (2001), 559–581.
[26] A. Selberg, 'An elementary proof of the prime-number theorem for arithmetic progressions', *Canad. J. Math.* **2** (1950), 66–78.
[27] X. Shao, 'A density version of the Vinogradov three primes theorem', *Duke Math. J.* **163**(3) (2014), 489–512.
[28] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, 105 (Cambridge University Press, Cambridge, 2006).
[29] I. M. Vinogradov, 'The representation of an odd number as a sum of three primes', *Dokl. Akad. Nauk. SSSR* **16** (1937), 139–142.

https://doi.org/10.1017/fms.2014.27 Published online by Cambridge University Press