Functiones et Approximatio
XXXV (2006), 61–77

ADDITIVE DECOMPOSABILITY OF MULTIPLICATIVELY DEFINED SETS

Christian Elsholtz

To Professor Eduard Wirsing
on the occasion of his 75th birthday

Abstract: Let $\mathcal{Q}(\mathcal{T})$ denote the set of integers which are composed of prime factors from a given set of primes $\mathcal{T}$ only. Suppose that $\mathcal{A}+\mathcal{B}\subseteq\mathcal{Q}'(\mathcal{T})$, where $\mathcal{Q}(\mathcal{T})$ and $\mathcal{Q}'(\mathcal{T})$ differ at finitely many elements only. Also assume that $\sum_{p\leqslant x, p\in\mathcal{T}}\frac{\log p}{p}=\tau\log x+O(1)$. We prove that $\mathcal{A}(N)\mathcal{B}(N)=O(N(\log N)^{2\tau})$ holds. In the case $\tau\geqslant\frac{1}{2}$ we give an example where both $\mathcal{A}(N)$ and $\mathcal{B}(N)$ are of order of magnitude $\frac{N^{\frac{1}{2}}}{(\log N)^{\frac{1}{4}}}$, which shows that this is close to best possible.

Keywords: inverse Goldbach problem, additive decompositions of sets, sums of two squares.

1. Introduction

In his two-volume monograph [21] on additive number theory, Ostmann, Eduard Wirsing's PhD advisor, studied the structure of sumsets. A sumset is defined by

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

Other monographs on this topic are those by Mann [18], Freĭman [10], Nathanson [20] and a forthcoming one by Tao and Vu [25].

Inverse questions, e.g. whether a given set can be additively decomposed, lead to difficult problems.

Definition 1.1. (See Ostmann [21], vol. 1, p. 1) Let $\mathcal{S}_1$ and $\mathcal{S}_2$ denote sets of positive integers. We say that $\mathcal{S}_1$ and $\mathcal{S}_2$ are *asymptotically equal*, if there exists an integer N_0 such that $\mathcal{S}_1\cap[N_0,\infty]=\mathcal{S}_2\cap[N_0,\infty]$.

Definition 1.2. (See [21], vol. 1, p. 5) Let $\mathcal{S}$ be a set of positive integers. We say that $\mathcal{S}$ is *additively irreducible*[1] if there are no two sets of positive integers $\mathcal{A}, B$, with at least two elements each, such that $\mathcal{A}+B=S$.

2000 Mathematics Subject Classification: Primary 11P32, 11N36; Secondary 11E25

[1] "primitiv" in the German original

Definition 1.3. (See [21], vol. 1, p. 5, and Wirsing [27]) Let $\mathcal{S}$ be a set of positive integers. We say that $\mathcal{S}$ is *asymptotically additively irreducible*[2] (or we say that *no asymptotic additive decomposition exists*) if there are no two sets of positive integers $\mathcal{A}, B$, with at least two elements each, such that $\mathcal{A}+B$ is asymptotically equal to $\mathcal{S}$.

In his first publication Wirsing [27] proved that "almost all" sets of positive integers are asymptotically additively irreducible. (For a precise definition of "almost all" we refer to Wirsing's paper.)

The set of all primes is additively irreducible. The essential reason is that the set of primes starts with $\{2,3,\ldots\}$. The existence of an additive decomposition implied that a pattern $n, n+1$ of two consecutive primes occurred again, which is of course impossible. Ostmann derived a number of generalizations. Moreover he stated the following

Conjecture 1.4. (Ostmann, [21] vol. 1, p. 13) *The set of primes $\mathcal{P}$ is asymptotically additively irreducible.*

This problem has attracted considerable attention, and several authors have proved that in a conceivable asymptotic decomposition of $\mathcal{P}$ both summands $\mathcal{A}$ and $\mathcal{B}$ would need to be infinite, (see Hornfeck [15], Mann [18], Laffer and Mann [16]). It has repeatedly been mentioned in problem collections such as Erdős [6], [7], Oberwolfach 1998 [31], and Wolke's survey [32]. The problem itself has resisted a solution so far. For partial results (in addition to those discussed below) see for example Hofmann and Wolke [14], Puchta [23] or the present author [3], [4], and [5]. In analogy to Goldbach's problem, it seems that a combination of additive and multiplicative properties leads to very difficult problems. The problem of the asymptotic additive decomposition of the primes has also been called the "inverse Goldbach problem", e.g. in Wirsing's Oberwolfach lecture of 1972 [30] and his problem at the 1998 Oberwolfach conference [31].

Also, several authors independently proved the following.

Theorem 1.5. *If $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$, then the following bound on the counting functions holds:*

$$\mathcal{A}(N)\mathcal{B}(N) \ll N.$$

Here $f(N) \ll g(N)$ is the Vinogradov notation, meaning $f(N) = O(g(N))$.

When taking extra care of finitely many elements, this implies:

Corollary 1.6. *If an asymptotic additive decomposition of the set of primes exists, say $\mathcal{P}' = \mathcal{A}+\mathcal{B}$, then the following bound on the counting functions holds:*

$$\mathcal{A}(N)\mathcal{B}(N) \ll N.$$

Indeed, the first author to prove Theorem 1.5 was Eduard Wirsing. He presented his result at the 1972 Oberwolfach meeting [30], but it was never published.

[2] totalprimitiv in the German original

Independently, similar results were proved by Pomerance, Sárközy, and Stewart [22], by Hofmann and Wolke [14], and by Bshouty and Bshouty [1].

The result of Bshouty and Bshouty is stated in a somewhat different way. A slightly simplified version is as follows:

Theorem 1.7. *Let* $\mathcal{A} = \{a_1, \dots, a_n\}, \mathcal{B} = \{b_1, \dots, b_n\}$ *and* $\mathcal{A} + \mathcal{B} \subset \mathcal{P}$, *then* $n^4 \ll (a_n - a_1)(b_n - b_1)$. *Therefore, if* $\mathcal{A}, \mathcal{B} \subset [1, N]$, *then* $n^2 \ll N$.

Note that this theorem only covers the case $|\mathcal{A}| = |\mathcal{B}| = n$. The implied methods would be weaker when applied to the case of different sizes of $|\mathcal{A}|$ and $|\mathcal{B}|$.

As quite a few different ideas go into these proofs of different authors and as Wirsing's original account is unpublished we decided to discuss these methods when proving the new results below.

Theorem 1.8. *Let* $\mathcal{T}$ *denote a set of primes with*

$$\left| \sum_{\substack{p \leqslant x \\ p \in \mathcal{T}}} \frac{\log p}{p} - \tau \log x \right| < C.$$

Here $0 < \tau < 1$ *denotes a real constant and* C *denotes a possitive real constant. Let*

$$\mathcal{Q}(\mathcal{T}) = \{n \in \mathbb{N} : p | n \Rightarrow p \in \mathcal{T}\}.$$

Let $\mathcal{A} + \mathcal{B} \subseteq \mathcal{Q}'(\mathcal{T})$, *where* $\mathcal{Q}(\mathcal{T}) \cap [N_0, \infty] = \mathcal{Q}'(\mathcal{T}) \cap [N_0, \infty]$, *for sufficiently large* N_0. *Then*

$$\mathcal{A}(N)\mathcal{B}(N) \ll_{\tau, C} N(\log N)^{2\tau}.$$

The upper bound is (at least sometimes) close to best possible, as discussed below.

We will give two proofs, one following Wirsing's method to prove Theorem 1.5. This is based on the analytic version of the large sieve method. The second follows the method of Pomerance, Sárközy, and Stewart [22], and Hofmann and Wolke [14], which are very similar. It is based on the arithmetic version of the large sieve method.

We then discuss very simple proofs of Theorem 1.7. The first proof by Bshouty and Bshouty [1] makes use of a divisibility property of the Vandermonde determinant. We then eliminate the Vandermonde argument and replace it by a count of prime factors which resembles a proof of Gallagher's larger sieve, and eventually we show that Gallagher's larger sieve can be directly applied to prove Theorem 1.7.

We then prove the following new bounds in the original problem of Ostmann.

Theorem 1.9. *Suppose that there is an asymptotic additive decomposition of the set of primes,* $\mathcal{A} + \mathcal{B} = \mathcal{P}'$, *then*

$$\sqrt{N}(\log N)^{-3} \ll \mathcal{A}(N) \ll \sqrt{N}(\log N)^2.$$

The same bounds hold for $\mathcal{B}(N)$.

This improves upon the earlier bounds

$$\sqrt{N}(\log N)^{-5} \ll \mathcal{A}(N) \ll \sqrt{N}(\log N)^4$$

in Elsholtz [4]. Let us recall that bounds of this type imply that there is no ternary asymptotic additive decomposition of the primes $\mathcal{P}' = \mathcal{A}+\mathcal{B}+\mathcal{C}$, where $\mathcal{A}, B$ and $\mathcal{C}$ have at least two elements each.

2. Examples

Example 2.1. Assuming the prime k-tuple conjecture, for any admissible set $\mathcal{A} = \{a_1, \dots, a_k\}$ (i.e. that for no prime p the image $\mathcal{A} \bmod p$ occupies all p residue classes) there exists an infinite set $\mathcal{B}$ with counting function $\mathcal{B}(N) \sim c_{a_1,\dots,a_k} \frac{N}{(\log N)^k}$ such that $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$. Unconditionally, for each finite large N and each k there exists a set $\mathcal{A} = \{a_1, \dots, a_k\} \subset [1, N]$ such that there exists a corresponding set $\mathcal{B} \subset [1, N]$ with $|\mathcal{B}| \gg \frac{N}{k(\log N)^k}$ and $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$ (see [22]). Similarly, let $\mathcal{A}+\mathcal{B} \subset \mathcal{Q}(\mathcal{T})$, where $\mathcal{Q}(\mathcal{T})$ is as in Theorem 1.8. Using the method of [22] (see also [8]) one can show that there exists a set $\mathcal{B} \subset [1, N]$ with $\mathcal{B}(N) \gg_k \frac{N}{(\log N)^{(1-\tau)k}}$. Moreover, for each finite N, one can find sets $\mathcal{A}$ and $\mathcal{B}$ with $\mathcal{A}(N) \geqslant \mathcal{B}(N) \geqslant \frac{1+o(1)}{1-\tau} \frac{\log N}{\log\log N}$.

This shows that for small fixed k the upper bound

$$\mathcal{A}(N)\mathcal{B}(N) = O_{\tau,C}(N(\log N)^{2\tau})$$

of Theorem 1.5 is off by a logarithmic factor only. As k increases one would expect much better upper bounds. In the problem $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$, if $\mathcal{A}(N) > (\log N)^r$ where $r \to \infty$ one can prove (using the methods of [5]) that $\mathcal{B}(N) \ll N^{\frac{1}{2}+o_r(1)}$. But if $\mathcal{A}(N)$ grows as fast as $\sqrt{N}$ the best upper bound on $\mathcal{B}(N)$ that one can currently prove is $O(\sqrt{N})$ only. One might expect much stronger upper bounds if $\mathcal{A}(N)$ and $\mathcal{B}(N)$ are of about the same size. If $\mathcal{A}(N) \leqslant \mathcal{B}(N)$ and $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$ one might expect that for example $\mathcal{A}(N) = O(N^\varepsilon)$ for arbitrary positive ε.

On the other hand, Theorem 1.8 is close to best possible if $\mathcal{A}(N) \approx \sqrt{N}$ and $\tau \geqslant \frac{1}{2}$, as shown by the following example.

Example 2.2. Let $\mathcal{A} = \{n^2 : n \in \mathbb{N}\}$ and $\mathcal{B} = \{n^2 : n \in \mathbb{N} \text{ and } (p \mid n \Rightarrow p \equiv 1 \bmod 4)\}$. Note that $\mathcal{A}(N) \sim N^{\frac{1}{2}}$ and $\mathcal{B}(N) \sim c\frac{N^{\frac{1}{2}}}{(\log N)^{\frac{1}{2}}}$, since by Lemma 2.3 or Lemma 2.4 the number of $n \leqslant \sqrt{N}$ composed of prime factors $p \equiv 1 \bmod 4$ only is asymptotically $c'\frac{N^{\frac{1}{2}}}{(\log \sqrt{N})^{\frac{1}{2}}}$. No element $a_i^2 + b_j^2 \in \mathcal{A}+\mathcal{B}$ contains any prime factor $q \equiv 3 \bmod 4$ since for such q one would have:

$$a_i^2 + b_j^2 \equiv 0 \bmod q$$

implies that both $a_i \equiv 0 \bmod q$ and $b_j \equiv 0 \bmod q$, which is not the case by construction. Hence $\mathcal{A}+\mathcal{B} \subset \mathcal{Q}(\mathcal{T})$, where $\mathcal{T} = \{p \in \mathcal{P} : p = 2 \text{ or } p \equiv 1 \bmod 4\}$, and $\tau = \frac{1}{2}$.

In this example we have $\mathcal{A}(N)\mathcal{B}(N) \sim c\frac{N}{(\log N)^{\frac{1}{2}}}$, which is quite close to the general upper bound $O(N \log N)$ given by Theorem 1.8. If $\tau > \frac{1}{2}$ the same example shows that the constructive lower bound differs from the upper bound by a logarithmic factor only (just allow further primes in $\mathcal{T}$). We do not expect that such examples exist if $\tau < \frac{1}{2}$, and for these cases I expect that if $\mathcal{A}(N) \approx \mathcal{B}(N)$, then the size of these sets is considerably smaller than $\sqrt{N}$.

Since often constructions of large sets with certain properties are more difficult in the case of $\mathcal{A} = \mathcal{B}$, let us note that with $\mathcal{B}$ and $\mathcal{T}$ as above we even have $\mathcal{B} + \mathcal{B} \subset \mathcal{Q}(\mathcal{T})$. Moreover, some of the combinatorial constructions only work for finite sets, whereas the sets considered here are infinite.

Variations of the example above are: partition $\mathcal{P}_{4,3} = \{p \in \mathcal{P} : p \equiv 3 \bmod 4\} = \mathcal{T}_1 \dot\cup \mathcal{T}_2$. Let $\sum_{p\in\mathcal{T}_i, p\leqslant x} \frac{1}{p} = \tau_i \log\log x + C_i + o(1)$, for $i = 1, 2$. Here $\tau_1 + \tau_2 = \frac{1}{2}$. Thus an explicit example with $\tau_1 = \tau_2 = \frac{1}{4}$ is $\mathcal{T}_1 = \mathcal{P}_{8,3} = \{p \in \mathcal{P} : p \equiv 3 \bmod 8\}$ and $\mathcal{T}_2 = \mathcal{P}_{8,7} = \{p \in \mathcal{P} : p \equiv 7 \bmod 8\}$. Let $\mathcal{A}$ be the set of all squares composed of primes of $\mathcal{P}_{4,1} \cup \mathcal{T}_1$ and let $\mathcal{B}$ be the set of squares composed of primes of $\mathcal{P}_{4,1} \cup \mathcal{T}_2$. Then $\mathcal{A}(N) \sim c_1 \frac{N}{(\log N)^{\frac{1}{2}-\tau_1}}$ and $\mathcal{B}(N) \sim c_2 \frac{N}{(\log N)^{\frac{1}{2}-\tau_2}}$, so that again $\mathcal{A}(N)\mathcal{B}(N) \sim c_1 c_2 \frac{N}{(\log N)^{\frac{1}{2}}}$. And again by construction: for all primes $q \equiv 3 \bmod 4$:

$$a_i^2 + b_j^2 \not\equiv 0 \bmod q.$$

For the determination of the counting function we made use of the following lemmas.

Lemma 2.3. (Wirsing, [28]) *Let $\mathcal{T}$ denote a set of primes and $0 < \tau \leqslant 1$. If*

$$\sum_{\substack{p\leqslant x \\ p\in\mathcal{T}}} \frac{1}{p} = \tau \log\log x + C + o(1),$$

then

$$|\{n \leqslant N, p \mid n \Rightarrow p \in \mathcal{T}\}| \sim C_{\mathcal{T}} \frac{N}{(\log N)^{1-\tau}}.$$

In view of the prime number theorem for arithmetic progressions, Wirsing's result contains a well known result by Landau as a special case.

Lemma 2.4. (Landau [17]) *Let $a_1, \ldots, a_r$ be r distinct reduced residues modulo m. Then*

$$|\{n \leqslant N, p \mid n \Rightarrow p \equiv a_1, a_2, \ldots, a_r \bmod m\}| \sim C_{m,a_1,a_2,\ldots,a_r} \frac{N}{(\log N)^{1-\frac{r}{\varphi(m)}}}.$$

3. Proof of Theorem 1.8

3.1. A proof using the analytic large sieve, based on Wirsing's approach. We need the following lemma:

Lemma 3.1. *Let $\mathcal{T}$ denote a set of primes, let $0 < \tau < 1$ be a real constant and let C be a positive real constant. If*

$$\left| \sum_{\substack{p \leqslant x \\ p \in \mathcal{T}}} \frac{\log p}{p} - \tau \log x \right| < C,$$

then

$$\sum_{\substack{n \leqslant N \\ p|n \Rightarrow p \in \mathcal{T}}} \mu^2(n) \geqslant C_{\tau,C} \frac{N}{(\log N)^{1-\tau}},$$

for some positive constant $C_{\tau,C}$.

In order to prove this we make use of another result of Wirsing [29] which we adapt from Schwarz and Spilker [24], page 76.

Lemma 3.2. (Wirsing) *Let f be a real non-negative multiplicative arithmetical function satisfying $f(p) \leqslant G$ for all primes p and*

$$\sum_{p \leqslant x} \frac{f(p) \log p}{p} \sim \tau \log x,$$

with some constants $G > 0, \tau > 0$ and

$$\sum_{p} \sum_{k \geqslant 2} \frac{f(p^k)}{p^k} < \infty.$$

If $0 < \tau \leqslant 1$, then, in addition, the condition

$$\sum_{p} \sum_{k \geqslant 2, p^k \leqslant x} f(p^k) = O(\frac{x}{\log x})$$

is assumed to hold. Then

$$\sum_{n \leqslant x} f(n) = (1 + o(1)) \frac{x}{\log x} \frac{e^{-\gamma\tau}}{\Gamma(\tau)} \prod_{p \leqslant x} \left(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^2} + \dots \right).$$

Here γ is the Euler-Mascheroni constant.

Proof of Lemma 3.1. Now with $f(p) = \begin{cases} 1 & \text{if } p \in \mathcal{T} \\ 0 & \text{otherwise,} \end{cases}$ and $f(p^k) = 0$, if $k \geqslant 2$, all hypotheses of the lemma are satisfied and we are left with analysing

the product:

$$\begin{aligned}\prod_{p\leqslant x}\left(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\right) &= \exp\left(\sum_{p\leqslant x,p\in\mathcal{T}}\log(1+\frac{1}{p})\right)\\ &\geqslant \exp\left(\sum_{p\leqslant x,p\in\mathcal{T}}\frac{1}{p}-\frac{1}{2p^2}\right)\\ &= \exp(\tau\log\log x+O_C(1))\\ &\gg_{\tau,C}(\log x)^{\tau}.\end{aligned}$$

Here we used that from $\left|\sum_{p\leqslant x,p\in\mathcal{T}}\frac{\log p}{p}-\tau\log x\right|<C$ it follows by partial summation that $\sum_{p\leqslant x,p\in\mathcal{T}}\frac{1}{p}=\tau\log\log x+O_C(1)$. ∎

For the remainder of this section we adapt Wirsing's proof of Theorem 1.5 to the current problem.

Let us state the analytic large sieve inequality, see equation (20) of Montgomery [19].

Lemma 3.3. *Let* $\mathcal{A}\subset[1,N]$, $\quad e(x)=\exp(2\pi ix)$ *and*

$$S_{\mathcal{A}}:=\sum_{a\in\mathcal{A}}e(ax).$$

Then

$$\sum_{q=1}^{Q}\sum_{\substack{r \bmod q\\(r,q)=1}}\left|S_{\mathcal{A}}\left(\frac{r}{q}\right)\right|^2\leqslant(N+Q^2)|\mathcal{A}|.$$

Let $\mathcal{S}=\mathcal{P}\setminus\mathcal{T}$ and $\mathcal{Q}(\mathcal{S})=\{n\leqslant Q:p\mid n\Rightarrow p\in\mathcal{S}\}$. Note that by Lemma 3.1 $|\mathcal{Q}(\mathcal{S})|\gg_{\tau,C}\frac{Q}{(\log Q)^{\tau}}$ holds, and that for sufficiently large elements $a+b\in\mathcal{A}+\mathcal{B}$ and $q\in\mathcal{Q}(\mathcal{S})$ are coprime.

We shall give an upper bound and a lower bound of

$$\sum_{q\in\mathcal{Q}(\mathcal{S})}\left|\sum_{\substack{r \bmod q\\(r,q)=1}}S_{\mathcal{A}}\left(\frac{r}{q}\right)S_{\mathcal{B}}\left(\frac{r}{q}\right)\right|,$$

combining these will prove the theorem.

Lemma 3.4.

$$\sum_{q\in\mathcal{Q}(\mathcal{S})}\left|\sum_{\substack{r \bmod q\\(r,q)=1}}S_{\mathcal{A}}\left(\frac{r}{q}\right)S_{\mathcal{B}}\left(\frac{r}{q}\right)\right|\leqslant(N+Q^2)|\mathcal{A}|^{\frac{1}{2}}|\mathcal{B}|^{\frac{1}{2}},$$

Proof. Using the Cauchy-Schwarz inequality and the large sieve inequality above we find:

$$\begin{aligned}
&\sum_{q\in\mathcal{Q}(\mathcal{S})}\left|\sum_{\substack{r \bmod q\\(r,q)=1}} S_{\mathcal{A}}\left(\frac{r}{q}\right) S_{\mathcal{B}}\left(\frac{r}{q}\right)\right| \\
&\leqslant \sum_{q\in\mathcal{Q}(\mathcal{S})}\sum_{\substack{r \bmod q\\(r,q)=1}}\left|S_{\mathcal{A}}\left(\frac{r}{q}\right) S_{\mathcal{B}}\left(\frac{r}{q}\right)\right| \\
&\leqslant \left(\sum_{q\in\mathcal{Q}(\mathcal{S})}\sum_{\substack{r \bmod q\\(r,q)=1}}\left|S_{\mathcal{A}}\left(\frac{r}{q}\right)\right|^2\right)^{\frac{1}{2}}\left(\sum_{q\in\mathcal{Q}(\mathcal{S})}\sum_{\substack{r \bmod q\\(r,q)=1}}\left|S_{\mathcal{B}}\left(\frac{r}{q}\right)\right|^2\right)^{\frac{1}{2}} \\
&\leqslant (N+Q^2)|\mathcal{A}|^{\frac{1}{2}}|\mathcal{B}|^{\frac{1}{2}}.
\end{aligned}$$

■

Lemma 3.5. (See (16.6.4) of Hardy and Wright [13])

$$\sum_{\substack{r \bmod q\\(r,q)=1}} e\left(\frac{r}{q}\right) = \mu(q).$$

This is a well known Ramanujan sum.

Lemma 3.6. *If $a+b\in\mathcal{Q}(\mathcal{T})$ and $a+b>Q$, then*

$$\sum_{q\in\mathcal{Q}(\mathcal{S})}\left|\sum_{\substack{r \bmod q\\(r,q)=1}} S_{\mathcal{A}}\left(\frac{r}{q}\right) S_{\mathcal{B}}\left(\frac{r}{q}\right)\right| \geqslant |\mathcal{A}||\mathcal{B}|C'_{\tau,C}\frac{Q}{(\log Q)^{\tau}}.$$

Proof. Consider the identity:

$$\sum_{\substack{r \bmod q\\(r,q)=1}} S_{\mathcal{A}}\left(\frac{r}{q}\right) S_{\mathcal{B}}\left(\frac{r}{q}\right) = \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\sum_{\substack{r \bmod q\\(r,q)=1}} e\left(\frac{(a+b)r}{q}\right).$$

Since $a+b\in\mathcal{Q}(\mathcal{T})$ with $a+b>Q$ and $q\in\mathcal{Q}(\mathcal{S})$ are coprime by construction, the values $(a+b)r$ are just a rearrangement of the coprime residue classes modulo q so that

$$\sum_{\substack{r \bmod q\\(r,q)=1}} e\left(\frac{(a+b)r}{q}\right) = \sum_{\substack{r \bmod q\\(r,q)=1}} e\left(\frac{r}{q}\right) = \mu(q).$$

Therefore, by Lemma 3.1,

$$\sum_{q\in\mathcal{Q}(\mathcal{S})}\left|\sum_{\substack{r \bmod q\\(r,q)=1}} S_{\mathcal{A}}\left(\frac{r}{q}\right) S_{\mathcal{B}}\left(\frac{r}{q}\right)\right| = |\mathcal{A}||\mathcal{B}| \sum_{q\in\mathcal{Q}(\mathcal{S})}\mu^2(q)$$

$$\geqslant |\mathcal{A}||\mathcal{B}| C'_{\tau,C}\frac{Q}{(\log Q)^{\tau}}. \qquad \blacksquare$$

Proof of Theorem 1.8. Let $\mathcal{A}\subset[1,N], \mathcal{A}_0 := \mathcal{A}\cap[0,\sqrt{N}], \mathcal{A}_1 := \mathcal{A}\cap[\sqrt{N},N]$, and similarly for $\mathcal{B}$. Instead of studying the four products $|\mathcal{A}_i||\mathcal{B}_j|$, with $i,j\in\{0,1\}$, it is sufficient to prove the upper bound $|\mathcal{A}||\mathcal{B}| = O_{\tau,C}(N(\log N)^{2\tau})$ without loss of generality for $\mathcal{A}\subset[\sqrt{N},N], \mathcal{B}\subset[1,N]$. To see this observe that $|\mathcal{A}_0||\mathcal{B}_0| = O(N)$ and that one can assume that $|\mathcal{A}_0||\mathcal{B}_1| \leqslant |\mathcal{A}_1||\mathcal{B}_0|$. Also note that $a+b > \sqrt{N} > N_0$ so that one can assume that $a+b\in\mathcal{Q}(\mathcal{T})$.

Combining the upper and lower bounds in Lemma 3.4 and 3.6 gives, with $Q=\sqrt{N}$:

$$|\mathcal{A}||\mathcal{B}| \ll_{\tau,C} N(\log N)^{2\tau}. \qquad \blacksquare$$

Remark. Since we do not sum over all $q\leqslant Q$ but over $\mathcal{Q}(\mathcal{S})$ only, one might hope for a stronger version of the large sieve inequality replacing $N+Q^2$ by something smaller. This would improve Lemma 3.4. There is some discussion on this issue on page 564 of [19].

3.2. The proof using the arithmetic large sieve, following Pomerance, Sárközy and Stewart, and Hofmann and Wolke. The proofs according to Pomerance, Sárközy, and Stewart [22] and Hofmann and Wolke [14] are very similar. We use the arithmetic version of the large sieve in a version due to Montgomery [19]:

Lemma 3.7. *Let $\mathcal{C}$ denote a set of integers which avoids $\omega(p)$ residue classes modulo p. Here $\omega:\mathcal{P}\to\mathbb{N}$ with $0\leqslant\omega(p)\leqslant p-1$. Then the following upper bound on the counting function holds:*

$$C(N) \leqslant \frac{N+Q^2}{L}, \text{ where } L=\sum_{q\leqslant Q}\mu^2(q)\prod_{p|q}\frac{\omega(p)}{p-\omega(p)}.$$

Lemma 3.8. *If $\mathcal{A},\mathcal{B}\subset\mathbb{Z}/p\mathbb{Z}$, then*

$$|\mathcal{A}+\mathcal{B}| \geqslant \min(p, |\mathcal{A}|+|\mathcal{B}|-1).$$

Assume that $\mathcal{A}+\mathcal{B}\subseteq\mathcal{Q}'(\mathcal{T})$, where $\mathcal{Q}'(\mathcal{T})\cap[N_0,\infty] = \mathcal{Q}(\mathcal{T})\cap[N_0,\infty]$. Let $\nu_{\mathcal{A}}(p) = |\mathcal{A} \bmod p|$ and $\nu_{\mathcal{B}}(p) = |\mathcal{B} \bmod p|$ denote the number of residue classes modulo p that do occur in $\mathcal{A}$ and $\mathcal{B}$, and let $\omega_{\mathcal{A}}(p) = p-\nu_{\mathcal{A}}(p)$ and $\omega_{\mathcal{B}}(p) = p-\nu_{\mathcal{B}}(p)$ denote the number of residue classes modulo p that do not occur in $\mathcal{A}$ and $\mathcal{B}$, respectively.

As above, without loss of generality, we consider $\mathcal{A} \subset [\sqrt{N}, N]$ and $\mathcal{B} \subset [1, N]$. Every $a+b \in \mathcal{A}+\mathcal{B}$ is larger than $\sqrt{N}$ and does not contain any prime factor $p \in \mathcal{S} = \mathcal{P} \backslash \mathcal{T}$, so that $\nu_{\mathcal{A}+\mathcal{B}}(p) \leqslant p-1$ holds for these primes.

The Cauchy-Davenport inequality further implies

$$\nu_{\mathcal{A}+\mathcal{B}}(p) \geqslant \nu_{\mathcal{A}}(p) + \nu_{\mathcal{B}}(p) - 1,$$

so that

$$\nu_{\mathcal{A}}(p) + \nu_{\mathcal{B}}(p) \leqslant p.$$

By definition

$$\omega_{\mathcal{A}}(p) + \nu_{\mathcal{A}}(p) = p, \quad \omega_{\mathcal{B}}(p) + \nu_{\mathcal{B}}(p) = p,$$

and therefore

$$\omega_{\mathcal{A}}(p) + \omega_{\mathcal{B}}(p) \geqslant p, \text{ for } p \in \mathcal{S}.$$

Lemma 3.9. *Let* $p \in \mathcal{S} \cap [1, \sqrt{N}]$, *then*

$$\frac{\omega_{\mathcal{A}}(p)}{p-\omega_{\mathcal{A}}(p)} \frac{\omega_{\mathcal{B}}(p)}{p-\omega_{\mathcal{B}}(p)} \geqslant 1.$$

Proof. This follows from

$$\frac{\omega_{\mathcal{A}}(p)}{p-\omega_{\mathcal{A}}(p)} \frac{\omega_{\mathcal{B}}(p)}{p-\omega_{\mathcal{B}}(p)} \geqslant \frac{\omega_{\mathcal{A}}(p)}{\omega_{\mathcal{B}}(p)} \frac{\omega_{\mathcal{B}}(p)}{\omega_{\mathcal{A}}(p)} = 1. \qquad \blacksquare$$

Proof of Theorem 1.8. Using Montgomery's large sieve, the Cauchy-Schwarz inequality, Lemma 3.1, and putting $Q = \sqrt{N}$, it follows that

$$\begin{aligned}
\mathcal{A}(N)\mathcal{B}(N) &\leqslant \frac{N+Q^2}{\sum_{q \leqslant Q} \mu^2(q) \prod_{p|q} \frac{\omega_{\mathcal{A}}(p)}{p-\omega_{\mathcal{A}}(p)}} \frac{N+Q^2}{\sum_{q \leqslant Q} \mu^2(q) \prod_{p|q} \frac{\omega_{\mathcal{B}}(p)}{p-\omega_{\mathcal{B}}(p)}} \\
&\leqslant \frac{4N^2}{\left(\sum_{\substack{q \leqslant Q \\ q \in \mathcal{Q}(\mathcal{S})}} \mu^2(q) \prod_{p|q} \left(\frac{\omega_{\mathcal{A}}(p)}{p-\omega_{\mathcal{A}}(p)} \frac{\omega_{\mathcal{B}}(p)}{p-\omega_{\mathcal{B}}(p)} \right)^{\frac{1}{2}} \right)^2}. \\
&\leqslant \frac{4N^2}{\left(\sum_{\substack{q \leqslant Q \\ q \in \mathcal{Q}(\mathcal{S})}} \mu^2(q) \right)^2} \ll_{\tau, C} \frac{N^2}{\left(\frac{\sqrt{N}}{(\log \sqrt{N})^\tau} \right)^2} \\
&\ll_{\tau, C} N (\log N)^{2\tau}.
\end{aligned}$$

$\blacksquare$

4. The proof of Theorem 1.7

4.1. The proof of Bshouty and Bshouty. This is a very simple proof. Let us first do some preparation.

Lemma 4.1. *Let* $S = S(a_1, \dots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}$, *where the* a_i *are integers. Then* $\prod_{i=1}^{n-1} i!$ *divides* $\det S$.

Proof. Note that $\prod_{i=1}^{n-1} i^{n-i} = 1^{n-1}2^{n-2}\cdots(n-1)^1 = \prod_{i=1}^{n-1} i!$. Also recall that $\det S = \prod_{1\leqslant i<j\leqslant n}(a_j - a_i)$ is the well known Vandermonde determinant; the lemma states a less well known divisibility property of it. (Compare problem 270 of [9]). It can be proved as follows:

$$\det M = \begin{vmatrix} 1 & \binom{a_1}{1} & \binom{a_1}{2} & \cdots & \binom{a_1}{n-1} \\ 1 & \binom{a_2}{1} & \binom{a_2}{2} & \cdots & \binom{a_2}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \binom{a_n}{1} & \binom{a_n}{2} & \cdots & \binom{a_n}{n-1} \end{vmatrix}$$

$$= \frac{1}{1!2!3!\cdots(n-1)!} \begin{vmatrix} 1 & a_1 & a_1(a_1-1) & \cdots & a_1(a_1-1)\cdots(a_1-n+2) \\ 1 & a_2 & a_2(a_2-1) & \cdots & a_2(a_2-1)\cdots(a_2-n+2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n(a_n-1) & \cdots & a_n(a_n-1)\cdots(a_n-n+2) \end{vmatrix}$$

Expanding the products and adding suitable linear combinations of the preceding columns shows that

$$\det M = \frac{1}{1!2!3!\cdots(n-1)!} \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}.$$

Since $\det M$ is an integer it follows that $\det S$ is divisible by $1!2!3!\cdots(n-1)!$. This proves the Lemma. ∎

Now assume $\mathcal{A} = \{a_1, \dots, a_n\} \subset [1, N]$, $\mathcal{B} = \{b_1, \dots, b_n\} \subset [1, N]$ and $\mathcal{A} + B \subset \mathcal{P}$.

$T := \prod_{i=1}^{2n-1} i!$ divides

$$\det S(a_1, \dots, a_n, -b_1, \dots, -b_n) = \prod_{1\leqslant i<j\leqslant n} (a_j - a_i) \prod_{1\leqslant i<j\leqslant n} (b_j - b_i) \prod_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} (a_i + b_j).$$

We will get the required relation between n and N by finding upper and lower bounds on T. Since T consists of prime factors $p < 2n$ only, T also divides

$$S' := \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i) \prod_{1 \leqslant i < j \leqslant n} (b_j - b_i) \prod_{a_i + b_j < 2n} (a_i + b_j),$$

which implies that $T \leqslant S'$. It is important to note that S' is considerably smaller than $\det S(a_1, \ldots, a_n, -b_1, \ldots, -b_n)$. Let us note that all but at most $O\left(\frac{n^2}{(\log n)^2}\right)$ terms of the form $(a_i + b_j)$ are primes $p > 2n$. This can be seen as follows: for fixed a_i, all $a_i + b_1, \ldots, a_i + b_n$ are distinct primes. By Chebychev's inequality, $\pi(n) \ll \frac{n}{\log n}$ so that only $O\left(\frac{n}{\log n}\right)$ of the above $a_i + b_j$ can be primes less than $2n$. Similarly, the size of the index i is bounded by $i = O\left(\frac{n}{\log n}\right)$, giving $O\left(\frac{n^2}{(\log n)^2}\right)$ possibilities with $a_i + b_j < 2n$.

Estimating now T from above and below (using Stirling's formula) shows that

$$(\frac{n}{e})^{2n^2} \leqslant T \leqslant S' \leqslant N^{n^2} (2n)^{\frac{cn^2}{(\log n)^2}}.$$

Taking the n^2-th root, this implies that $n^2 \ll N$.

4.2. A variation, reminiscent of Gallagher's sieve. The proof in this section is inspired by Bshouty and Bshouty's proof but may be more natural. While studying the product $\prod(a_j - a_i)$ seems natural, studying the determinant of $S(a_1, \ldots, a_n, -b_1, \ldots, -b_n)$ appears to be a trick. Here we study how often the prime factors $p < n$ occur in the product $\prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)(b_j - b_i)$. The proof below is selfcontained and does not make direct use of any sieve method. But in a similar spirit it is possible to prove Gallagher's larger sieve.

Lemma 4.2. (Compare [5], page 425) *If $p \in \mathcal{S}$ be an odd prime, such that no $a + b$ is divisible by p, then*

$$\frac{1}{\nu_{\mathcal{A}}(p)} + \frac{1}{\nu_{\mathcal{B}}(p)} \geqslant \frac{4}{p}.$$

With $\nu_{\mathcal{A}}(p) + \nu_{\mathcal{B}}(p) \leqslant p$ (see section 3.2 above) we find that

$$\frac{1}{\nu_{\mathcal{A}}(p)} + \frac{1}{\nu_{\mathcal{B}}(p)} \geqslant \frac{1}{\nu_{\mathcal{A}}(p)} + \frac{1}{p - \nu_{\mathcal{A}}(p)},$$

which takes (for $p \neq 2$) its minimum at $\nu_{\mathcal{A}}(p) = \frac{p \pm 1}{2}$.

Lemma 4.3. *Let $r_1, \ldots, r_k$ be nonnegative integer parameters with $\sum_{i=1}^{k} r_i = C$, for a positive constant C. Then $f(r_1, \ldots, r_k) = \sum_{i=1}^{k} \binom{r_i}{2}$ is minimised if the r_i are as equal as possible, i.e. $r_i \in (\frac{C}{k} - 1, \frac{C}{k} + 1)$.*

Sketch proof. We consider the r_i to be continuous variables and use Lagrange's multiplier method. Here the minimum is attained for $r_i = \frac{C}{k}$. For the discrete

analogue any deviation from the mean value which is larger than necessary increases the value of f: assume $r_1 = \frac{C}{k} + 1 + \delta_1$, where $\delta_1 \geqslant 0$. There must be some $r_2 = \frac{C}{k} - \delta_2$, where $\delta_2 > 0$. Such a pair (r_1, r_2) cannot be part of the minimal solution since it can be reduced to $r_1' = \frac{C}{k} + \delta_1$ and $r_2' = \frac{C}{k} + 1 - \delta_2$, since $\binom{r_1}{2} + \binom{r_2}{2} - \binom{r_1'}{2} - \binom{r_2'}{2} = 2(\delta_1 + \delta_2) > 0$. Similarly, assuming that a minimal solution would contain $r_1 = \frac{C}{k} - 1 - \delta_1, r_2 = \frac{C}{k} + \delta_2$ with $\delta_1 \geqslant 0, \delta_2 > 0$ leads to a contradiction, which proves the Lemma. ∎

The prime factor p occurs in the product $\prod_{1 \leqslant i < j \leqslant n}(a_j - a_i)$, whenever two a_i's are in the same residue class modulo p. Let $p^w \parallel \prod_{1 \leqslant i < j \leqslant n}(a_j - a_i)$, then $w \geqslant \frac{|\mathcal{A}|}{2}(\frac{|\mathcal{A}|}{\nu_{\mathcal{A}}(p)} - 1)$, as can be seen as follows: put $r_i(p) = |\{a \in \mathcal{A} : a \equiv i \bmod p\}|$. Only $\nu_{\mathcal{A}}(p)$ of the $r_i(p)$ are positive. Further $\sum_{i=0}^{p-1} r_i(p) = n$ is constant and $w \geqslant \sum_{i=0}^{p-1} \binom{r_i(p)}{2}$. By the lemma above, the right hand side takes its minimum if the distribution of $\mathcal{A}$ in the $\nu_{\mathcal{A}}(p)$ residue classes is as equal as possible, i.e.: $w \geqslant \frac{|\mathcal{A}|}{2\nu_{\mathcal{A}}(p)}(\frac{|\mathcal{A}|}{\nu_{\mathcal{A}}(p)} - 1)\nu_{\mathcal{A}}(p)$.

This implies:

$$\prod_{p \leqslant n} p^{\frac{|\mathcal{A}|^2}{2\nu_{\mathcal{A}}(p)} - \frac{|\mathcal{A}|}{2} + \frac{|\mathcal{B}|^2}{2\nu_{\mathcal{B}}(p)} - \frac{|\mathcal{B}|}{2}} \leqslant \prod_{1 \leqslant i < j \leqslant n}(a_j - a_i) \prod_{1 \leqslant i < j \leqslant n}(b_j - b_i)$$
$$\leqslant N^{\frac{|\mathcal{A}|^2}{2} - \frac{|\mathcal{A}|}{2} + \frac{|\mathcal{B}|^2}{2} - \frac{|\mathcal{B}|}{2}}.$$

With $|\mathcal{A}| = |\mathcal{B}| = n$ and Lemma 4.2 it follows that

$$\prod_{p \leqslant n} p^{\frac{2n^2}{p} - n} \leqslant N^{n^2 - n}.$$

Writing $\prod_{p \leqslant n} p^{\frac{2n^2}{p}} = \exp\left(2n^2 \sum_{p \leqslant n} \frac{\log p}{p}\right) = \exp\left(2n^2(\log n + O(1))\right)$, and taking the n^2-th root it follows that $n^2 \ll N$.

4.3. A proof based on Gallagher's larger sieve. Here we show that a direct application of Gallagher's larger sieve also proves Theorem 1.7. Gallagher's larger sieve is essentially already incorporated into the ad hoc argument of the last section. For convenience we use a variant of Gallagher's larger sieve, introduced by Croot and the author in [2].

Lemma 4.4. (Variant 1 of [2]) *Let $\mathcal{S} \subset [2, Q]$ denote a set of primes or powers of primes such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu_{\mathcal{A}}(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Then,*

$$|\mathcal{A}| \leqslant \max\left(Q, \frac{23N \exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{q}\right)}{\exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{\nu_{\mathcal{A}}(q)}\right)}\right).$$

Here Λ is the von Mangoldt function.

Applying this to $|\mathcal{A}| = |\mathcal{B}| = n$ with $Q = \sqrt{N}$ gives: if Q is the maximum of any of the upper bounds on $|\mathcal{A}|$ or $|\mathcal{B}|$, then $n \leqslant Q$ and so $|\mathcal{A}||\mathcal{B}| = n^2 \leqslant Q^2 = N$. Otherwise, combining Lemma 4.2 and Lemma 4.4 with $\mathcal{S} = \mathcal{P} \cap [1, Q]$, gives:

$$|\mathcal{A}||\mathcal{B}| \ll \frac{N \exp\left(\sum_{p\in\mathcal{S}} \frac{\Lambda(p)}{p}\right)}{\exp\left(\sum_{p\in\mathcal{S}} \frac{\Lambda(p)}{\nu_{\mathcal{A}}(p)}\right)} \frac{N \exp\left(\sum_{p\in\mathcal{S}} \frac{\Lambda(p)}{p}\right)}{\exp\left(\sum_{p\in\mathcal{S}} \frac{\Lambda(p)}{\nu_{\mathcal{B}}(p)}\right)}$$

$$\ll \frac{N^2 \exp\left(2\sum_{p\leqslant Q} \frac{\log p}{p}\right)}{\exp\left(\sum_{p\leqslant Q} \frac{\log p}{\nu_{\mathcal{A}}(p)} + \frac{\log p}{\nu_{\mathcal{B}}(p)}\right)}$$

$$\ll \frac{N^2}{\exp\left(\sum_{p\leqslant Q} \frac{2\log p}{p}\right)} \ll N.$$

5. Proof of Theorem 1.9

Assume that there exists an asymptotic additive decomposition of the set of primes: $\mathcal{A} + \mathcal{B} = \mathcal{P}'$. Here we prove sharper sieve bounds than were previously proved in [4]. The main idea of the proof in [4] was that an inverse application of Gallagher's larger sieve applied to the smaller sequence $\mathcal{A}$ (say) shows that this sequence occupies many residue classes modulo many primes $p \leqslant y$. This information can successfully be injected into Montgomery's sieve to give an upper bound of $\mathcal{B}(N) \ll \sqrt{N}(\log N)^4$. One of the obstacles to get a still better bound was that the sieve level up to which we used the distribution modulo primes needed to be the same $y = N^{\frac{1}{2m}} = N^{\frac{1}{4}}$ for both parts of the sieve.

Even though we count in finite intervals, the original problem is of course about infinite sets. So, it is possible to do both sieve parts independently. It was already shown in [4], that $\mathcal{A}(N) \gg_\varepsilon N^{\frac{1}{2}-\varepsilon}$. Since this holds for all N, we can show, up to any level Q, by an inverse application of Gallagher's larger sieve that for a positive proportion of the primes $p \leqslant Q$ the number $\nu_{\mathcal{A}}(p)$ of occupied classes modulo p is large. Let us partition the set $\mathcal{S} = [N_0, Q] \cap \mathcal{P}$ into two parts: $\mathcal{S}_1 = \{p \in \mathcal{S} : \nu_{\mathcal{A}}(p) \leqslant \frac{Q}{20\log Q}\}$, and $\mathcal{S}_2 = \{p \in \mathcal{S} : \nu_{\mathcal{A}}(p) > \frac{Q}{20\log Q}\}$. By the prime number theorem $\pi(Q) = \frac{Q}{\log Q} + \frac{Q}{(\log Q)^2} + O\left(\frac{Q}{(\log Q)^3}\right)$, which implies for sufficiently large Q, $|\mathcal{S}| \geqslant \frac{Q}{\log Q}$. Hence at least one of $|\mathcal{S}_1| \geqslant \frac{Q}{2\log Q}$ or $|\mathcal{S}_2| \geqslant \frac{Q}{2\log Q}$ holds. We show that the first condition cannot hold, whence the second condition holds.

We use again variant 1 of Gallagher's larger sieve, see Lemma 4.4. Let $Q = N^{\frac{1}{4}}$. Assume that $|\mathcal{S}_1| \geqslant \frac{Q}{2\log Q}$, which implies $\frac{Q}{3} \leqslant \sum_{q\in\mathcal{S}_1} \Lambda(q) \leqslant Q$.

$$\begin{aligned}\mathcal{A}(N) &\ll \max\left(Q, \frac{N\exp\left(\sum_{q\in\mathcal{S}_1}\frac{\Lambda(q)}{q}\right)}{\exp\left(\sum_{q\in\mathcal{S}_1}\frac{\Lambda(q)}{\nu_{\mathcal{A}}}\right)}\right)\\ &\ll \max\left(Q, \frac{NQ}{\exp\left(\frac{20\log Q}{Q}\sum_{q\in\mathcal{S}_1}\Lambda(q)\right)}\right)\\ &\ll \max\left(Q, \frac{NQ}{\exp\left(\frac{20\log Q}{Q}\frac{Q}{3}\right)}\right) = Q,\end{aligned}$$

which contradicts $\mathcal{A}(N) \gg_\varepsilon N^{\frac{1}{2}-\varepsilon}$.

For the problem of infinite sets $\mathcal{A}, \mathcal{B}$ and therefore $\mathcal{A}(N) \gg_\varepsilon N^{\frac{1}{2}-\varepsilon}$ for all large N, this implies that $|\mathcal{S}_2| \geqslant \frac{Q}{2\log Q}$ holds for all large Q.

Since $a+b > \sqrt{N}$ is a prime, we know that $a+b \not\equiv 0 \bmod p$, for primes $p \leqslant \sqrt{N}$. This implies that $\omega_{\mathcal{B}}(p) \geqslant \nu_{\mathcal{A}}(p)$. We put $Q = N^{\frac{1}{2}}$ and use Montgomery's large sieve in the special case, where the q in the denominator are primes in $\mathcal{S}_2$:

$$\mathcal{B}(N) \leqslant \frac{N+Q^2}{\sum_{p\in\mathcal{S}_2}\frac{\omega_{\mathcal{B}}(p)}{p}} \ll \frac{N}{\frac{Q}{20\log Q}\frac{Q}{2\log Q}\frac{1}{Q}} \ll N^{\frac{1}{2}}(\log N)^2.$$

Since $\mathcal{A}(N)\mathcal{B}(N) \gg \frac{N}{(\log N)}$ this implies that $\mathcal{A}(N) \gg \frac{N^{\frac{1}{2}}}{(\log N)^3}$. By symmetry, the same upper and lower bounds hold on $\mathcal{A}$ and $\mathcal{B}$.

It should be noted that the parameter N in both parts of the sieve application is not necessarily the same. The information about the distribution of the used classes modulo primes up to Q in the first sieve argument possibly makes use of elements $a+b$ which are outside the interval $[1, N]$ that is used for the second sieve application. This trick obviously cannot work for the finite version of the problem.

I would like to thank Eduard Wirsing for discussions on the problem, for giving me access to his unpublished manuscript [30], and for comments on an earlier version of this paper. I am also grateful to Nigel Watt for a discussion on section 3 of this paper, and to Ernie Croot for many discussions on sieve methods.

References

[1] D. Bshouty and N.H. Bshouty, *A note on prime n-tuples*, Rocky Mountain J. Math. **27** (1997), 775–778.

[2] E. Croot and C. Elsholtz, *On variants of the larger sieve*, Acta Math. Hungarica **103** (2004), 243–254.

[3] C. Elsholtz, *A Remark on Hofmann and Wolke's Additive Decompositions of the Set of Prime*, Arch. Math. **76** (2001), 30-33.

[4] C. Elsholtz, *The Inverse Goldbach Problem*, Mathematika **48** (2001), 151-158.

[5] C. Elsholtz, *Some remarks on the additive structure of the set of primes*, Number theory for the millennium, I (Urbana, IL, 2000), 419–427, (Proceedings of the Millennial Conference on Number Theory (Bennett et.al.), AK Peters, 2002).
[6] P. Erdős, *Quelques problèmes de théorie des nombres*, Monographies de L'Enseignement Mathématique, no. 6, pp 81-135, Université Geneva, 1963.
[7] P. Erdős, *Problems and results in number theory*, Number Theory Day, New York, 1976, Springer, Lecture Notes in Math. 626, pp.43–72.
[8] P. Erdős, C.L. Stewart and R. Tijdeman, *Some Diophantine equations with many solutions*, Compositio Math. **66** (1988), 37–56.
[9] D.K. Faddeev and I.S. Sominskii, Problems in Higher Algebra, W.H. Freeman and Company, San Francisco 1965.
[10] G.A. Freĭman, *Foundations of a structural theory of set addition*, translated from the Russian version (1966). Translations of Mathematical Monographs, Vol 37. American Mathematical Society, Providence, 1973.
[11] P.X. Gallagher, *A Larger Sieve*, Acta Arith. **18** (1971), 77-81.
[12] G.H. Hardy and J.K. Littlewood, *Some Problems of 'Partitio Numerorum', III. On The Expression of a Number as a Sum of Primes*, Acta Math. **44** (1923), 1-70.
[13] G.H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, fifth edition, Clarendon Press, Oxford, 1989.
[14] A. Hofmann and D. Wolke, *On Additive Decompositions of the Set of Primes*, Arch. Math. **67** (1996), 379-382.
[15] B. Hornfeck, *Ein Satz über die Primzahlmenge*, Math. Z. **60** (1954), 271–273, see also the Zentralblatt review by Erdős and the correction in Math. Z. 62, (1955), page 502.
[16] W.B. Laffer and H.B. Mann, *Decomposition of sets of group elements*, Pacific J. Math. **14** (1964) 547–558.
[17] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner Verlag, Leipzig, Berlin (1909).
[18] H.B. Mann, *Addition theorems: The addition theorems of group theory and number theory*, Wiley Interscience, New York-London-Sydney, 1965.
[19] H. Montgomery, *The Analytic Principle of the Large Sieve*, Bull. Amer. Math Soc. **84** (1978), 547-567.
[20] M. Nathanson, *Additive Number Theory, Inverse problems and the geometry of sumsets.* Graduate Texts in Mathematics, 165. Springer-Verlag, New York, 1996.
[21] H.-H. Ostmann, *Additive Zahlentheorie*, 2 volumes, Springer-Verlag, Berlin-Heidelberg-New York, 1956, reprint 1968.
[22] C. Pomerance, C.L. Stewart and A. Sárközy, *On Divisors of Sums of Integers, III*, Pacific J. Math. **133** (1988), 363-379.
[23] J.-P. Puchta, On additive decompositions of the set of primes. Arch. Math. **78** (2002), 24–25.
[24] W. Schwarz and J. Spilker, Arithmetical functions, Cambridge University Press, 1994.

[25] T. Tao, V. Vu, *Additive Combinatorics*, Cambridge University Press, in preparation.
[26] R.C. Vaughan, *Some Applications of Montgomery's Sieve*, J. Number Theory **5** (1973), 64-79.
[27] E. Wirsing, *Ein metrischer Satz über Mengen ganzer Zahlen.* Arch. Math. **4** (1953), 392–398.
[28] E. Wirsing, *Über die Zahlen, deren Primteiler einer gegebenen Menge angehören.* Arch. Math. **7**, 263–272, 1956.
[29] W. Wirsing, *Das asymptotische Verhalten von Summen über multiplikative Funktionen, II*, Acta Math. Acad. Sci. Hung. **18**, 414–476 (1967).
[30] E. Wirsing, *Über additive Zerlegungen der Primzahlmenge*, lecture at Oberwolfach, an abstract can be found in Tagungsbericht 28/1972. An unpublished written version (1998) exists with the title *On the additive decomposability of the set of primes, A Memo.*
[31] E. Wirsing, Problem at the problem session of the Oberwolfach conference Elementare und analytische Zahlentheorie 1998. Abstract booklet of the Oberwolfach Conference 10/1998.
[32] D. Wolke, *Das Goldbach'sche Problem*, Math. Semesterber. 41 (1994), 55–67.

Address: Department of Mathematics, Royal Holloway, Egham, Surrey TW20 0EX, UK
E-mail: christian.elsholtz@rhul.ac.uk
Received: 24 December 2005; **revised:** 12 June 2006