A survey on additive and multiplicative decompositions of sumsets and of shifted sets

Christian Elsholtz

April 20, 2009

Abstract

In this paper we survey results on sumsets with multiplicative properties and the question if a shifted copy of a multiplicatively defined set can again be multiplicatively defined. The methods involved are of analytic nature such as the large sieve, and of combinatorial nature such as extremal graph theory.

1 Introduction

Let $\mathcal{A}$ and $\mathcal{B}$ denote sets of integers and let $\mathcal{A}+\mathcal{B}=\{a+b : a \in \mathcal{A}, b \in \mathcal{B}\}$ be the sumset and $\mathcal{AB}=\{ab : a \in \mathcal{A}, b \in \mathcal{B}\}$ be the product set.

We are interested in the additive and multiplicative structure of particular sets of integers $\mathcal{S}$. One way to study its additive structure, is to see how large sets $\mathcal{A}$ and $\mathcal{B}$ exist with $\mathcal{A}+\mathcal{B} \subset \mathcal{S}$.

Perhaps one can even decompose the whole set so that there exist two sets of integers $\mathcal{A}$ and $\mathcal{B}$ with $|\mathcal{A}| \geq 2$ and $|\mathcal{B}| \geq 2$ such that $\mathcal{A}+\mathcal{B}=\mathcal{S}$ holds? A measure theoretic argument due to Wirsing [44] shows that most sets $\mathcal{S}$ cannot be decomposed in that way; (not even in the asymptotic sense of Definition 2.2 below). But, for any particular set $\mathcal{S}$ it may be of interest to show that the set can indeed not be decomposed.

The structure of sumsets and product sets is in the focus of current research activity, see for example Tao and Vu [41].

Here we are particularly interested if multiplicatively defined sets can be written as sumsets or if a shifted copy of a multiplicatively defined set can also be a product set.

This survey consists of two parts: In the first part we review results which show that if $\mathcal{A}+\mathcal{B}$ or $\mathcal{AB}+1$ is a subset of a multiplicatively defined set, then the counting functions on $\mathcal{A}(N)=\sum_{a\in\mathcal{A}, a\leq N} 1$ and $\mathcal{B}(N)$ cannot be too large. In particular we give a proof that the set of primes $\mathcal{P}$ cannot be written in the form $\mathcal{A}\mathcal{B}+c$, where $|\mathcal{A}|,|\mathcal{B}|\geq 2$, even if finitely many exceptions are allowed.

In the second part we use purely combinatorial counting arguments which show that there do exist sets $\mathcal{A}$ and $\mathcal{B}$ of certain finite sizes, such that $\mathcal{A}+\mathcal{B}\subset\mathcal{P}$, for example.

While the first part introduces sieve methods, the second part is based on pigeonhole principle type arguments, or extremal graph theory.

2 Part I

2.1 Multiplicative decompositions of sumsets

In earlier work such as [10], [11], [13], the author studied the question if multiplicatively defined sets can be additively decomposed. We say a set $\mathcal{S}$ can be additively decomposed if there exist sets $\mathcal{A},\mathcal{B}$ with at least two elements each such that $\mathcal{A}+\mathcal{B}=\mathcal{S}$.

Let us study two examples:

Example 2.1. Let $a\mathbb{Z}=\{an : n\in\mathbb{Z}\}$. Then $a\mathbb{Z}+b\mathbb{Z}=\gcd(a,b)\mathbb{Z}$. Note that $\mathbb{Z}\mathbb{Z}=\mathbb{Z}$, so that this sumset also has a trivial multiplicative decomposition:

$$a\mathbb{Z}+b\mathbb{Z}=\left(\gcd(a,b)\mathbb{Z}\right)\mathbb{Z}.$$

If one allows positive integers only, i.e. if one studies $a\mathbb{N}+b\mathbb{N}$, then all sufficiently large multiples of $\gcd(a,b)$ can be represented. So, while $a\mathbb{N}+b\mathbb{N}=\gcd(a,b)\mathbb{N}$ does not quite hold, it holds apart from finitely many exemptions. This suggests the study of asymptotic additive decompositions.

Definition 2.2 (See [37], vol. 1, p. 5)**.** *Let $\mathcal{S}$ be a set of positive integers. We say that $\mathcal{S}$ is* asymptotically additively irreducible *(or we say that* no asymptotic additive decomposition exists*) if there do not exist two sets of positive integers $\mathcal{A}$ and $\mathcal{B}$, with at least two elements each, such that for a sufficiently large N_0:*

$$(\mathcal{A}+\mathcal{B})\cap[N_0,\infty]=\mathcal{S}\cap[N_0,\infty].$$

Another example of a set which has an additive and a multiplicative structure is the following:

Example 2.3. Let $\mathcal{A}=\{x^2 : x\in\mathbb{Z}\}$, $\mathcal{B}=\{2y^2 : y\in\mathbb{Z}\}$. It is well known that n can be written in the from $n=x^2+2y^2$ if and only if n is zero, or n has a prime factorisation of the from

$$n=2^r\prod_{p_i\text{ prime with }p_i\equiv 1,3\bmod 8}p_i^{s_i}\prod_{q_i\text{ prime with }q_i\equiv 5,7\bmod 8}q_i^{2t_i},$$

where the r, s_i, t_i are nonnegative integers. In other words with $\mathcal{C} = \{n \in \mathbb{N} : p \mid n \Rightarrow (p = 2 \text{ or } p \equiv 1, 3 \bmod 8)\}$ and $\mathcal{D} = \{n^2 : n \in \mathbb{Z}\}$ we find that $\mathcal{A} + \mathcal{B} = \mathcal{C}\mathcal{D}$. Note that such a decomposition is not necessarily unique. Let $\mathcal{D}' = \{n^2 : p \mid n \Rightarrow p \equiv 5, 7 \bmod 8\}$, then $\mathcal{A} + \mathcal{B} = \mathcal{C}\mathcal{D}'$ also holds.

This example 2.3 comes from the theory of quadratic forms. Observe that $\mathcal{A}$ and $\mathcal{B}$ occupy modulo primes $\frac{p+1}{2}$ of the residue classes. The sum of two random sets with about half of the residue classes modulo a prime p would usually cover all residue classes modulo p. In this example, modulo many primes this is not the case: namely if $p \equiv 5, 7 \bmod 8$, and $x, y \not\equiv 0 \bmod p$ then the corresponding sets $\mathcal{A}'$ and $\mathcal{B}'$ occupy $\frac{p-1}{2}$ of the residue classes, but $x^2 + 2y^2 \not\equiv 0 \bmod p$. Moreover, the considered sets are very "large". This makes the example interesting.

It may be conjectured that large sumsets that have a multiplicative property must come from related examples based on algebraic polynomials. For example, Croot and myself formulated the following conjecture:

Problem 2.4. Let $\mathcal{A} \subset [1, N]$ with $|\mathcal{A}| > N^{0.4}$. Assume that $|\mathcal{A} \mod p| \leq \frac{2}{3}p$, for every prime $p \leq N$. Must any such $\mathcal{A}$ be contained in the set of values of a quadratic polynomial, apart from $N^{1/3+\varepsilon}$ exceptions?

The values of a quadratic polynomial $f(x) = ax^2 + bx + c$ are in about one half the residue classes modulo primes.

The philosophy behind this conjecture is: if one removes from the interval $[1, N]$ a positive proportion of the residue classes modulo a positive proportion of the primes, and arrives at a set $\mathcal{A}$ with $|\mathcal{A}| > N^\delta$, where $\delta > 0$, then there should be an algebraic reason for this, since for a random sieve process one would expect that $|\mathcal{A}|$ is very much smaller.

One of the long standing open problems about additive decompositions is Ostmann's conjecture:

Conjecture 2.5 (Ostmann, [37] vol. 1, p. 13)**.** *The set of primes $\mathcal{P}$ is asymptotically additively irreducible.*

As partial results in the direction of this conjecture the author proved (see Elsholtz [10] and [13]):

Theorem 2.6. *Suppose that there is an asymptotic additive decomposition of the set of primes, i.e. $\mathcal{P} \cap [N_0, \infty] = \mathcal{P}' \cap [N_0, \infty]$, for some large N_0, and $\mathcal{A} + \mathcal{B} = \mathcal{P}'$, then*

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

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

Also $\mathcal{A}(N)\mathcal{B}(N) \ll N$ is known, a result that was independently proved by several authors. For a survey on this see [13].

As a consequence of a result such as Theorem 2.6, the author proved in Elsholtz [10]:

Corollary 2.7. *The set of primes is not asymptotically additively decomposable into* three *sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ containing at least two elements each.*

For the additive decomposability of multiplicatively defined sets the following holds, (see [13]).

Theorem 2.8. *Let $\mathcal{T}$ denote a set of primes with*

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

Here $0 < \tau < 1$ denotes a 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

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

Corollary 2.9. *Let $\mathcal{T}$ consist of $p = 2$ and all primes $p \equiv 1, 3 \bmod 8$. Then in the above Theorem $\tau = \frac{1}{2}$ is admissible. Let $\mathcal{A}$ and $\mathcal{B}$ be as above. Then*

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

This upper bound is quite close to the following constructive example:

Example 2.10. Let $\mathcal{A} = \{n^2 : n \in \mathbb{N}\}$ and $\mathcal{B} = \{2n^2 : n \in \mathbb{N}$ and $(p \mid n \Rightarrow p \equiv 1, 3 \bmod 8)\}$. No element $a_i + b_j \in \mathcal{A} + \mathcal{B}$ contains any prime factor $q \equiv 5, 7 \bmod 8$ since for such q one would have:

$$m^2 + 2n^2 \equiv 0 \bmod q, \text{ i.e. } \frac{m^2}{n^2} \equiv -2 \bmod q.$$

Since -2 is a quadratic non-residue modulo primes $p \equiv 5, 7 \bmod 8$, this implies that both $m \equiv 0 \bmod q$ and $n \equiv 0 \bmod q$ must hold, which is not the case by construction. Hence $\mathcal{A} + \mathcal{B} \subset \mathcal{Q}(\mathcal{T})$, where $\mathcal{T} = \{p \in \mathcal{P} : p = 2$ or $p \equiv 1, 3 \bmod 8\}$, and $\tau = \frac{1}{2}$.

Note that $\mathcal{A}(N) \sim N^{1/2}$ and $\mathcal{B}(N) \sim c\frac{N^{1/2}}{(\log N)^{1/2}}$, for some positive constant c. Therefore the general upper bound $N \log N$ on $\mathcal{A}(N)\mathcal{B}(N)$ and the actual value $c\frac{N}{(\log N)^{1/2}}$ in this example is only off by a small logarithmic factor.

2.2 Multiplicative decompositions of shifted sets

In [14] the author studied the multiplicative analogues.

Definition 2.11. *Let $\mathcal{S}$ be a set of positive integers. We say that $\mathcal{S}$ is* asymptotically multiplicatively translation-irreducible *if there is no decomposition of the following type: $\mathcal{S}' = \mathcal{A}\mathcal{B} + c$, where $\mathcal{A}, \mathcal{B}$ are sets of positive integers with at least two elements each, $c \neq 0$ an integer, and where $\mathcal{S}'$ is asymptotically equal to $\mathcal{S}$.*

Theorem 2.12. *The set of primes is asymptotically multiplicatively translation-irreducible.*

The following Theorem proves the corresponding result for sets of integers composed of certain prime factors only.

Theorem 2.13. *Let $\mathcal{T} \subset \mathcal{P}$ be a set of primes with the property that*

$$\sum_{p \leq N,\ p \in \mathcal{T}} 1 = \tau \frac{N}{\log N} + O\left(\frac{N}{(\log N)^2}\right),$$

for some constant $0 < \tau < 1$. Let

$$\mathcal{Q}(\mathcal{T}) = \{1\} \cup \{n \in \mathbb{N} : p \mid n \Rightarrow p \in \mathcal{T}\}.$$

Then $\mathcal{Q}(\mathcal{T})$ is asymptotically multiplicatively translation-irreducible.

In this survey we will give a proof of Theorem 2.12. The proof of Theorem 2.13 uses the same methods.

As a motivation for this type of results we observe that famous open problems are closely related. It is not known if there are infinitely Sophie Germain primes, or infinitely Carmichael numbers with a given number of prime factors. A Sophie Germain prime is a prime p where $2p + 1$ is also prime. Let $\mathcal{A} = \{1, 2\}$ and $\mathcal{AB} - 1 \subset \mathcal{P}$. If p is a Sophie Germain prime, then $p + 1 \in \mathcal{B}$. So, the question if an infinite set of shifted primes $\mathcal{P}' + 1$ can be multiplicatively decomposed into $\mathcal{AB}$, where $\mathcal{A} = \{1, 2\}$ is equivalent to the question whether there are infinitely Sophie Germain primes or not.

A Carmichael number n is a composite number such that $a^n \equiv a \bmod n$. Thus a Carmichael number is a pseudo prime to all bases a. By Korselt's criterion a squarefree number n is a Carmichael number if for all prime factors $p \mid n$: $p - 1 \mid n - 1$.

Let $\mathcal{A} = \{6, 10, 12, 40\}$ and $\mathcal{AB} + 1 \subset \mathcal{P}$. If it is possible to find infinitely many b such that $6b + 1, 10b + 1, 12b + 1$ and $40b + 1$ are simultaneously prime, then the product $n = (180b + 7)(300b + 11)(360b + 13)(1200b + 41)$ is a Carmichael number. This follows since $n-1$ is divisible by $120(30b+1)$ so that it is divisible by $180b+6, 300b+10, 360b+12, 1200b+40$. The smallest example of this type is $n = 7 \cdot 11 \cdot 13 \cdot 41 = 41041$. For further details on Carmichael numbers see [21].

2.3 Some background from sieve methods

The two main ingredients of the proof are the large sieve inequality, in a form due to Montgomery, and Gallagher's larger sieve.

Lemma 2.14 (Montgomery [35]). *Let $\mathcal{P}$ denote the set of primes. Let p be a prime. Let $\mathcal{A}$ denote a set of integers which avoids $\omega_{\mathcal{A}}(p)$ residue classes modulo p. Here $\omega_{\mathcal{A}} : \mathcal{P} \to \mathbb{N}$ with $0 \leq \omega_{\mathcal{A}}(p) \leq p-1$. Let $\mathcal{A}(N)$ denote the counting function $\mathcal{A}(N) = \sum_{a \leq N, a \in \mathcal{A}} 1$. Let μ denote the Möbius function. Then the following upper bound on the counting function holds:*

$$\mathcal{A}(N) \leq \frac{N+Q^2}{L}, \text{ where } L = \sum_{q \leq Q} \mu^2(q) \prod_{p \mid q} \frac{\omega_{\mathcal{A}}(p)}{p - \omega_{\mathcal{A}}(p)}.$$

One typically chooses $Q = N^{1/2}$. There are many excellent expositions of a proof of this statement (or variants of it), including those by Montgomery [35], Brüdern [2], Davenport [8], Gallagher [18], Tenenbaum [42].

Vaughan has found a suitable lower bound of L if $\sum_{p \leq y} \dfrac{\omega(p)}{p}$ is known.

Lemma 2.15 (Vaughan [43]). *For sufficiently large Q*

$$L \geq \sum_{m=1}^{\infty} \exp\left(m \log \left(\frac{1}{m} \sum_{p \leq Q^{1/m}} \frac{\omega(p)}{p} \right) \right).$$

The sum $\sum_{m=1}^{\infty}$ is in fact a finite sum only. The parameter m denotes the number of prime factors of q in the definition of L. Hence $1 \leq m \leq \dfrac{\log Q}{\log 2}$. As there are at most $O(\log N)$ summands, a lower bound on L can be found by choosing a suitable value of m, and replacing the sum by this one summand. The loss of a factor of size at most $O(\log N)$ is small in typical large sieve applications.

Lemma 2.16 (Gallagher's larger sieve, [19]). *Let $\mathcal{S}$ denote a set of primes such that $\mathcal{A}$ lies modulo p (for $p \in \mathcal{S}$) in at most $\nu_{\mathcal{A}}(p)$ residue classes. Then the following bound holds, provided the denominator is positive:*

$$\mathcal{A}(N) \leq \frac{-\log N + \sum_{p \in \mathcal{S}} \log p}{-\log N + \sum_{p \in \mathcal{S}} \dfrac{\log p}{\nu_{\mathcal{A}}(p)}}.$$

This sieve has a very elementary proof:

Proof. Let $\mathcal{A} = \{a_1, a_2, \ldots, a_{|\mathcal{A}|}\} \subset [1, N]$, where $a_1 < a_2 < \ldots$. We study upper and lower bounds of $\prod_{1\le i<j\le|\mathcal{A}|}(a_j - a_i)$. The following upper bound is trivial, since $a_j - a_i < a_j \le N$.

$$\prod_{1\le i<j\le|\mathcal{A}|}(a_j - a_i) \le N^{\frac{1}{2}|\mathcal{A}|(|\mathcal{A}|-1)}.$$

To provide a lower bound we observe that the product is divisible by many small primes to a high power. Let $p^{s(p)} \parallel \prod_{1\le i<j\le|\mathcal{A}|}(a_j - a_i)$. A factor of p arises whenever a_i and a_j are in the same residue class modulo p. (Some additional factors of p arise modulo prime powers, leading to a slightly sharper sieve inequality, but for simplicity we ignore this here). Let $t_k = |\{a_i : a_i \equiv k \bmod p\}|$, so that $\sum_k t_k = |\mathcal{A}|$. Then $s(p) \ge \sum_k \frac{1}{2} t_k(t_k - 1)$. The smallest that this latter sum can be is if the a_i are as equidistributed among the $\nu(p)$ residue classes as possible, i.e. if $t_i \approx \frac{|\mathcal{A}|}{\nu(p)}$. This is a consequence of Cauchy's inequality:

$$\nu(p)\sum_k t_k^2 = \sum_{k:t_k>0} 1^2 \sum_{k:t_k>0} t_k^2 \ge \left(\sum_{k:t_k>0} 1\cdot t_k\right)^2 = |\mathcal{A}|^2.$$

This implies that

$$\sum_{k=0}^{p-1} t_k(t_k - 1) \ge |\mathcal{A}|\left(\frac{|\mathcal{A}|}{\nu(p)} - 1\right).$$

Hence $s(p) \ge \frac{|\mathcal{A}|}{2}\left(\frac{|\mathcal{A}|}{\nu(p)} - 1\right)$.

Combining the upper and lower bounds gives:

$$\prod_{p\in\mathcal{S}} p^{\frac{|\mathcal{A}|}{2}(\frac{|\mathcal{A}|}{\nu(p)}-1)} \le N^{\frac{|\mathcal{A}|}{2}(|\mathcal{A}|-1)}.$$

Simplifying and taking logarithms gives

$$|\mathcal{A}|\sum_{p\in\mathcal{S}}\frac{\log p}{\nu(p)} - \sum_{p\in\mathcal{S}}\log p \le (|\mathcal{A}| - 1)\log N.$$

$$|\mathcal{A}|\left(\sum_{p\in\mathcal{S}}\frac{\log p}{\nu(p)} - \log N\right) \le -\log N + \sum_{p\in\mathcal{S}}\log p$$

which proves the larger sieve inequality, provided $\sum_{p\in\mathcal{S}}\frac{\log p}{\nu(p)} > \log N$ holds. □

2.4 Proof of Theorem 2.12

Proposition 2.17. *Let $\mathcal{A}, \mathcal{B}$ be sets of positive integers with at least two elements each. Suppose that $\mathcal{P}' = \mathcal{A}\mathcal{B} + c$, where $\mathcal{P}' \cap [N_0, \infty) = \mathcal{P} \cap [N_0, \infty)$. For sufficiently large N the following holds:*

$$\mathcal{A}(N) \ll N^{\frac{1}{2}+\frac{1}{7}}.$$

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

Proof. Suppose that $\mathcal{P}' = \mathcal{A}\mathcal{B} + c$, where $\mathcal{P}' \cap [N_0, \infty) = \mathcal{P} \cap [N_0, \infty)$. Let $b_1 < b_2$ be the least two elements of $\mathcal{B}$. Without loss of generality we may assume that $N_0 > \max(b_2, c)$; otherwise, we just increase N_0. Let N be a sufficiently large integer. We begin by showing that for any prime $q \in [N_0, N^{1/2}]$ the set $\mathcal{A}_1 = \mathcal{A} \cap [N^{1/2}, N]$ avoids at least two residue classes modulo q. If $a \in \mathcal{A}_1$, then $ab_1 + c = p$ is a prime with $b_1 < N_0 < q < N^{1/2} < p$. Now $b_1 \not\equiv 0 \bmod q$, so $b_1^{-1} \bmod q$ exists and $a \not\equiv -b_1^{-1} c \bmod q$. Similarly $a \not\equiv -b_2^{-1} c \bmod q$. It follows that $a \in \mathcal{A}_1$ avoids the two residue classes $-b_1^{-1}c$ and $-b_2^{-1}c$ modulo q. These are distinct since $0 < b_1 < b_2 < q$ and $0 < c < q$.

Lemma 2.14 applied with $\omega(q) = 2$ gives the upper bound:

$$\begin{aligned}
\mathcal{A}_1(N) &\le \frac{2N}{\sum_{q \le N^{1/2}} \mu^2(q) \prod_{p|q} \frac{2}{p-2}} \\
&\le \frac{2N}{\sum_{q \le N^{1/2}} \mu^2(q) \prod_{p|q} \frac{2}{p}} \\
&= \frac{2N}{\sum_{q \le N^{1/2}} \mu^2(q) \frac{d(q)}{q}} \\
&\ll \frac{N}{(\log N)^2},
\end{aligned}$$

where $d(q)$ denotes the number of divisors of q. So $\mathcal{A}(N) \le \sqrt{N} + \mathcal{A}_1(N) \ll \frac{N}{(\log N)^2}$. In view of $\mathcal{P}(N) \gg \frac{N}{\log N}$ we must have $\mathcal{B}(N) \gg \log N$. This trivially implies that $\mathcal{B}(N) \to \infty$, as $N \to \infty$.

Let $b_1 < b_2 < \ldots < b_k$ be the first $k = 8$ elements of $\mathcal{B}$. We adapt the argument above with the change that $N_0 > b_8$. A sieve with $\omega(q) = 8$, for $q \in [N_0, N^{\frac{1}{2}}]$ implies that $\mathcal{A}(N) \ll \frac{N}{(\log N)^8}$ and therefore $\mathcal{B}(N) \ge c'(\log N)^7$, for some positive constant c'.

To iterate this further we need a lower bound on $\omega(q)$ on average. Since a residue class $b \mod q$ forbids a class $a \bmod q$ for $\mathcal{A}$ we actually count those classes modulo primes that occur in $\mathcal{B}$.

Let $\nu_{\mathcal{B}}(p) = |\mathcal{B} \mod p|$; i.e., $\nu_{\mathcal{B}}(p)$ is the number of residue classes modulo p that contain at least one element of $\mathcal{B}$.

Let

$$y = c'(\log N)^8 \tag{1}$$

and $\mathcal{S} = \{p : N_0 < p \leq y\}$. Assume that $\sum_{p\in\mathcal{S}} \frac{\log p}{\nu_{\mathcal{B}}(p)} > 3 \log N$. This means that $\nu_{\mathcal{B}}(p)$ is small, say of size at most $p^{7/8}$ modulo many small primes $p \leq y$. We therefore apply Gallagher's larger sieve. Using (1) and Chebyshev's bound $\sum_{p\leq y} \log p < 2y$ we obtain a contradiction:

$$c'(\log N)^7 \leq \mathcal{B}(N) \leq \frac{-\log N + \sum_{p\in\mathcal{S}} \log p}{-\log N + \sum_{p\in\mathcal{S}} \frac{\log p}{\nu_{\mathcal{B}}(p)}} < \frac{2y}{-\log N + \sum_{p\in\mathcal{S}} \frac{\log p}{\nu_{\mathcal{B}}(p)}} < c'(\log N)^7.$$

Consequently

$$\sum_{p\in\mathcal{S}} \frac{\log p}{\nu_{\mathcal{B}}(p)} \leq 3 \log N.$$

This means that $\nu_{\mathcal{B}}(p)$ is modulo many small primes $p \leq y$ not as small as originally assumed. Using Montgomery's large sieve, this knowledge will give a good upper bound on $\mathcal{A}$, since any class that occurs in $\mathcal{B}$ forbids one in $\mathcal{S}$. But in order to use Montgomery's sieve we first need to transform the information from a measure on $\sum_{p\in\mathcal{S}} \frac{\log p}{\nu_{\mathcal{B}}(p)}$ to a measure on $\sum_{N_0<p\leq y} \frac{\nu_{\mathcal{B}}(p)}{p}$.

From the Cauchy-Schwarz inequality and by partial summation from Chebyshev's bound on the number of primes $\pi(y) \gg \frac{y}{\log y}$ we find that:

$$\left(\sum_{N_0<p\leq y} \frac{\log p}{\nu_{\mathcal{B}}(p)}\right)\left(\sum_{N_0<p\leq y} \frac{\nu_{\mathcal{B}}(p)}{p}\right) \geq \left(\sum_{N_0\leq p\leq y} \left(\frac{\log p}{p}\right)^{1/2}\right)^2 \gg \frac{y}{\log y}.$$

We combine these two inequalities to obtain

$$\sum_{p\in\mathcal{S}} \frac{\nu_{\mathcal{B}}(p)}{p} \gg \frac{y}{(\log y)(3\log N)} \gg \frac{(\log N)^7}{\log\log N}.$$

It then follows by Montgomery's large sieve and by Vaughan's Lemma 2.15 that $\mathcal{A}_1(N) \leq \frac{2N}{L}$, where

$$\begin{aligned} L &= \sum_{q\leq N^{1/2}} \mu^2(q) \prod_{p|q} \frac{\omega_{\mathcal{A}}(p)}{p - \omega_{\mathcal{A}}(p)} \\ &\geq \max_{m\in\mathbb{N}} \exp\left(m \log\left(\frac{1}{m} \sum_{p\leq N^{1/(2m)}} \frac{\omega_{\mathcal{A}}(p)}{p}\right)\right). \end{aligned}$$

We choose the integer m such that $N^{1/(2m)}$ is close to y. Hence

$$m = \frac{1}{16}\frac{\log N}{\log\log N} + O(1).$$

We may choose

$$\omega_{\mathcal{A}}(p) = \begin{cases} \nu_{\mathcal{B}}(p) & \text{for } p \in \mathcal{S} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore we find that for all $\varepsilon' > 0$:

$$\begin{aligned} \log L &\geq \left(\frac{\log N}{16\log\log N} + O(1)\right) \log\left(\frac{16\log\log N}{\log N}\frac{c''(\log N)^7}{\log\log N}\right) \\ &\geq \left(\frac{1}{2} - \frac{1}{8} - \varepsilon'\right)\log N. \end{aligned}$$

$$\mathcal{A}(N) \leq \mathcal{A}_1(N) + N^{\frac{1}{2}} \leq \frac{2N}{L} + N^{\frac{1}{2}} \ll_{\varepsilon'} N^{\frac{1}{2}+\frac{1}{8}+\varepsilon'} \ll_{\varepsilon} N^{\frac{1}{2}+\frac{1}{7}} = N^{\frac{9}{14}}.$$

The same upper bound holds for $\mathcal{B}(N)$, by symmetry. So, the proposition follows. □

Proof of Theorem 2.12. If $ab + c = p \leq N$, then at least one of $a \ll \sqrt{N}$ or $b \ll \sqrt{N}$ must hold. Since all $\pi(N) + O(1) \gg \frac{N}{\log N}$ primes in $[N_0, N]$ have at least one presentation as $ab + c$, we not only have $\mathcal{A}(N)\mathcal{B}(N) \gg \frac{N}{\log N}$ but also

$$\mathcal{A}(N)\mathcal{B}(\sqrt{N}) + \mathcal{A}(\sqrt{N})\mathcal{B}(N) \gg \frac{N}{\log N}.$$

Since our proposition holds for all sufficiently large N, we can apply it independently several times. Applying it once with $N_1 = \sqrt{N}$, a second time with $N_2 = N$ gives $\mathcal{A}(N) \ll N^{9/14}, \mathcal{B}(N) \ll N^{9/14}$ and $\mathcal{A}(\sqrt{N}) \ll N^{9/28}, \mathcal{B}(\sqrt{N}) \ll N^{9/28}$, whence

$$\frac{N}{\log N} \ll \mathcal{A}(N)\mathcal{B}(\sqrt{N}) + \mathcal{A}(\sqrt{N})\mathcal{B}(N) \ll N^{\frac{9}{14}+\frac{9}{28}} \ll N^{\frac{27}{28}}.$$

This contradiction proves the theorem. □

3 Part II

3.1 Sumsets

If one works in finite sets $[1, N]$ one can show that there exist two sets $\mathcal{A}, \mathcal{B}$ with at least $\frac{\log N}{\log\log N}$ many elements each, with $\mathcal{A} + \mathcal{B} \subset \mathcal{P}$. (If one chooses one of the sets smaller, then the other set can be taken considerably larger!) Below we will study some combinatorial counting methods like the pigeonhole principle to prove this lower bound. For results of this type we make essentially use of the counting function $\mathcal{P}(N) \gg \frac{N}{\log N}$ of the set of primes only. So, the same type of results would hold for other sets with such a counting function. But in order to strengthen the constants in the results slightly we take a stronger result on the prime counting function. Recall that by the prime number theorem $\pi(N) = \mathrm{li(N)} + \mathrm{O}(\frac{\mathrm{N}}{(\log \mathrm{N})^{\mathrm{k}}})$ holds for all k. Here $\mathrm{li(N)} = \int_2^{\mathrm{N}} \frac{\mathrm{dt}}{\log \mathrm{t}}$. In particular (see for example Landau [33], page 47), $\pi(N) = \frac{N}{\log N} + \frac{N}{(\log N)^2} + 2\frac{N}{(\log N)^3} + O(\frac{N}{(\log N)^4})$.

On the other hand, using the large sieve method one can show that two sets $\mathcal{A}, \mathcal{B} \subset [1, N]$ of the same size $|\mathcal{A}| = |\mathcal{B}|$ with $\mathcal{A} + \mathcal{B} \subset \mathcal{P}$ can at most be of size $|\mathcal{A}| = O(N^{1/2})$ (compare [13]). There is a huge gap between the lower bound and upper bound, and I conjecture that the upper bound should be $O_\varepsilon(N^\varepsilon)$, for all $\varepsilon > 0$. Observe that examples 2.3 and 2.10 show, that there are square-like sequences, where the sumset essentially avoids half of all prime factors as divisors. The conjecture says that if the sums are required to be primes, then no such square-like sequences exist.

3.2 Counting methods

The following result is due to Erdős, Stewart and Tijdeman [16] and is a very useful counting tool:

Lemma 3.1. *Let N be a positive integer and let $\mathcal{S} \subset [1, N]$ be a nonempty set. Let k be an integer with $1 \le k \le |\mathcal{S}|$. Then there exists a set $\mathcal{A} \subset \mathcal{S}$ and a set of non-negative integers $\mathcal{B} \subset [0, N-1]$ such that*

$$\mathcal{A} + \mathcal{B} \subset \mathcal{S}, \quad |\mathcal{A}| \ge \frac{\binom{|\mathcal{S}|}{k}}{\binom{N-1}{k-1}}, \quad |\mathcal{B}| = k.$$

Proof. There are $\binom{|\mathcal{S}|}{k}$ subsets of $\mathcal{S}$ containing k elements. To each of these subsets $\{s_1, \ldots, s_k\}$ with $s_1 < \ldots < s_k$ we associate the $(k-1)$-element difference-subset

$$\{s_2 - s_1, \ldots, s_k - s_1\} \subset [1, N-1].$$

By the pigeonhole principle there exists a $k-1$ element set $\{h_1, \ldots, h_{k-1}\}$ which is the difference subset of at least

$$t = \frac{\binom{|\mathcal{S}|}{k}}{\binom{N-1}{k-1}}$$

distinct k-element subsets of $\mathcal{S}$. The least elements of these t sets are denoted by $a_1, a_2, \ldots, a_t$. (Note that all a_i are distinct since otherwise two k-element sets would be the same.) Thus the lemma follows with $\mathcal{A} = \{a_1, \ldots, a_t\}$ and $\mathcal{B} = \{0, h_1, \ldots, h_{k-1}\}$.

□

3.3 A result from extremal graph theory

We shall make use of results from extremal graph theory. The applicability of results from graph theory to number theory has been promoted by Erdős. For an early example see Erdős [15].

More recently it was stated in a paper by Győry, Stewart and Tijdeman [28] that M. Simonovits observed the possibility to apply the Kővari-Sós-Turán-theorem to number theory. That theorem was for example applied in Gyarmati [26] to square-free sumsets and by the author in [11] to triples of primes in arithmetic progression.

Let us state the Kővari-Sós-Turán theorem, [31]. (Compare Theorem IV.10 (page 113) of Bollobás [1].)

Theorem 3.2. *Let $G(m, n)$ denote a bipartite graph with m vertices in the first class and n in the second. Let $z(m, n; s, t)$ denote the maximal number of edges of G such that G does not contain a complete bipartite graph $K_{s,t}$ with s vertices in the first class and t in the second. Then, for all natural numbers m, n, s and t we have*

$$z(m, n; s, t) \leq s^{\frac{1}{t}} n m^{1-\frac{1}{t}} + (t-1)m.$$

Proof. Let us consider a bipartite graph $G(V_1 \cup V_2, E)$ with $|V_1| = m, |V_2| = n$ that does not contain a complete bipartite subgraph $K_{s,t}$, with s elements in the first vertex set and t elements in the second. (For convenience, we say a graph $K_{u,v}$ is oriented if the u vertices are in V_1 and the v vertices are in V_2.) Let us count the number of oriented $K_{1,t}$ subgraphs. On the one side, this number is $\sum_{i=1}^{m} \binom{d_i}{t}$, where d_i denotes the degree of the i-th vertex in V_1. On the other side there are $\binom{n}{t}$ choices of t out of n elements and each of these choices of t elements is counted at most $s-1$ times, since there is no oriented $K_{s,t}$ subgraph. This implies that

$$\sum_{i=1}^{m} \binom{d_i}{t} \leq (s-1)\binom{n}{t}.$$

Since $f(z) = \binom{z}{t}$ is a convex function, and since $\sum_i d_i = |E|$, it follows by

Jensen's inequality (see below) that

$$\sum_{i=1}^{m} \binom{d_i}{t} \geq m \binom{\frac{|E|}{m}}{t}.$$

With $|E| \geq m(t-1)$ (for a graph without $K_{s,t}$ but with a maximal number of edges) this implies that

$$(s-1)n^t \geq m(\frac{|E|}{m} - t + 1)^t$$

so that we have

$$z(m,n;s,t) \leq (s-1)^{\frac{1}{t}} n m^{1-\frac{1}{t}} + (t-1)m.$$

□

Let us recall Jensen's inequality for convex functions:

Lemma 3.3. *If f is a convex function on the interval $[a,b]$, then*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i),$$

where (for all $1 \leq i \leq n$) $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^{n} \lambda_i = 1$ and $x_i \in [a,b]$.

An important special case is with $\lambda_i = \frac{1}{n}$.

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leq \frac{1}{n}\sum_{i=1}^{n} f(x_i),$$

that is, the value of the function at the mean of the x_i is less than or equal to the mean of the values of the function at each x_i.

3.4 The case of primes

3.4.1 The difference counting approach

Pomerance, Sárközy and Stewart [38] proved the following results.

Theorem 3.4 (Pomerance, Sárközy and Stewart)**.** *Let N, k be positive integers with $k < \log N$. There is an effectively computable constant c_1 such that if $N > c_1$, then there exist $\mathcal{A}, \mathcal{B} \subset [1,N]$ with $|\mathcal{B}| = k$*

$$|\mathcal{A}| > \frac{N}{k(\log N)^k} \text{ and } \mathcal{A} + \mathcal{B} \subset \mathcal{P} \cap [1,N].$$

Proof. Recall that $\pi(N) \geq \frac{N}{\log N} + \frac{N}{(\log N)^2}$, for sufficiently large N.

By Theorem 3.1 there exists $\mathcal{A} \subset \mathcal{P}$ and $\mathcal{B}$ with $\mathcal{A} + \mathcal{B} \subset \mathcal{P} \cap [1, N]$ and $|\mathcal{B}| = k$.

$$|\mathcal{A}| \geq \frac{\binom{\pi(N)}{k}}{\binom{N-1}{k-1}} \geq \frac{\frac{(\pi(N)-k)^k}{k!}}{\frac{N^{k-1}}{(k-1)!}} \geq \frac{\left(\frac{N}{\log N}\right)^k}{kN^{k-1}} = \frac{N}{k(\log N)^k}.$$

So far we have $\mathcal{A} \subset [2, N], \mathcal{B} \subset [0, N-2]$. Shifting the set $\mathcal{A}$ down by 1, and shifting $\mathcal{B}$ up by 1, proves the theorem. □

Remark 3.5. An application of the purely combinatorial Theorem 3.1 allows to have that one of the sets $\mathcal{A}$ or $\mathcal{B}$ is itself a subset of the primes (or a shifted copy thereof). In fact, the second set can also have a prime restriction. In order to see this recall that Chudakov [4], van der Corput [5], Estermann [17] and Chowla [3] proved that almost all even integers are the sum of two primes with about the expected number of representations. The exceptional set has a counting function of at most $O_k\left(\frac{N}{(\log N)^k}\right)$, for all k. In particular almost all integers of the form $2p$ are of the form $2p = p_1 + p_2$, about the expected number of times. This means that there are infinitely many triples of primes in arithmetic progression, see Chowla [3]. More precisely, the number of solutions of this equation with primes $p, p_1, p_2 \leq N$ is of order of magnitude $\frac{N^2}{(\log N)^3}$. For a closely related problem we refer to Theorem 3.8 in section 3.5.

3.4.2 The graph theoretic approach

Theorem 3.4 is certainly a good approximation to the prime k-tuple problem. We observe that the lower bound can be refined as follows: Let $y = \frac{1}{2}\log N$. Let $P = \prod_{p \leq y} p$. Define a bipartite graph $G(V_1 \cup V_2, E)$, where $V_1 = \{P, 2P, \ldots, \lfloor \frac{N}{P} \rfloor P\}$ and $V_2 = \{n \leq N : (n, P) = 1\}$. The set of edges is defined by $(v_1, v_2) \in E \Leftrightarrow v_1 + v_2 \in \mathcal{P}$. By this construction $v_1 + v_2$ is not divisible by any prime $p \leq y$. This pre-sieving increases the edge density slightly; for each p by a factor of $\frac{1}{1-\frac{1}{p}}$, and for all $p \leq y$ together by a factor of $\prod_{p \leq y} \frac{1}{1-\frac{1}{p}} \sim \log y$. For any given $v_1 \in V_1$, the number of $v_2 \in V_2$ such that $v_1 + v_2$ is prime is $\frac{N}{\log N} \log y$, so that the the total number of edges is, with a positive constant c_1:

$$|E| \geq (c_1 + o(1))|V_1|\,|V_2|\frac{\log\log N}{\log N}.$$

A sumset $\mathcal{A} + \mathcal{B} \subset \mathcal{B}$ corresponds to a complete bipartite graph $K_{s,t}$ with $|\mathcal{A}| = s$ and $|\mathcal{B}| = t$. Let us assume the graph G does not contain any $K_{s,t}$.

This gives bounds on the parameters s and t. By Theorem 3.2

$$(c_1+o(1))|V_1|\,|V_2|\frac{\log\log N}{\log N} \leq |E| \leq z(m,n,s,t) \leq s^{\frac{1}{t}}|V_1||V_2|^{1-\frac{1}{t}} + t|V_2|.$$

This implies that

$$s \geq |V_2| c_2^t \frac{(\log\log N)^t}{(\log N)^t}.$$

With $|V_2| \sim N \prod_{p\leq y}\left(1-\frac{1}{p}\right) \geq \frac{1}{2}\frac{N}{(\log\log N)}$ it follows that

$$s \geq \frac{c_3^t N(\log\log N)^{t-1}}{(\log N)^t}.$$

Now, let us assume that s is smaller than this bound. By Theorem 3.2 there exists a graph $K_{s,t}$ and hence subsets of the required size.

Corollary 3.6. *There exist sets $\mathcal{A}, \mathcal{B}$ with $\mathcal{A}+\mathcal{B} \subset \mathcal{P}$ and with cardinality $|\mathcal{A}|, |\mathcal{B}| \geq \frac{\log N}{\log\log N - \log\log\log N + O(1)} \geq \frac{\log N}{\log\log N}$, for large N.*

This slightly improves upon the bound $(1-\varepsilon)\frac{\log N}{\log\log N}$ proved by Pomerance et. al. [38].

The lower bound above, $s \geq \frac{c_3^t N(\log\log N)^{t-1}}{(\log N)^t}$, may look somewhat surprising: for the prime k tuple problem one thinks of the upper bound as $C_k\frac{N}{(\log N)^k}$, whereas here the lower bound has additional $\log\log N$ factors. This is however no contradiction: since V_1 and V_2 depend on P, and so on N the k-tuples considered above are not “constant”, but vary, as N increases. This example shows that upper bounds of the general prime k-tuple problem, where $n+b_i$, prime, $b_i \in [0,N], i \in \{1,\dots,k-1\}$, i.e. where the coefficients can vary with N, must also include $(\log\log N)$-factors.

3.5 Chains of primes in arithmetic progressions

If one studies chains of primes in arithmetic progressions $a, a+d, a+2d, a+(k-1)d \leq N$, then an upper bound sieve shows there are at most $O_k(\frac{N^2}{(\log N)^k})$ of these. Until recently, a corresponding lower bound was only known in the case $k=3$, (see the above mentioned work by Chudakov, van der Corput, Estermann and Chowla). For a very precise result see Grosswald [24].

Recently, Green and Tao [22] proved:

Theorem 3.7. *Let*

$$G_k(N) := \left|\{(p_1 < p_2 < \ldots < p_k) : p_k \leq N,\ p_i \in \mathcal{P} \text{ and } p_i = p_1 + (i-1)d\}\right|.$$

Then the following lower bound holds, for some constant $C_k > 0$.

$$G_k(N) \geq (C_k + o(1)) \frac{N^2}{(\log N)^k}.$$

An asymptotic was known before for $k = 3$, and Green and Tao [23] more recently established one for $k = 4$.

In this section we combine the Green-Tao result with the graph theoretic counting above and show the following results.

Theorem 3.8. *Let*

$$G_k(N) := |\{(p_1 < p_2 < \ldots < p_k) : p_k \leq N,\ p_i \in \mathcal{P} \text{ and } p_i = p_1 + (i-1)d\}|.$$

Let $G_k(N) \geq (C_k + o(1)) \frac{N^2}{(\log N)^k}$ and $C'_k = \frac{\ln C_k}{k-2}$.

Let N be sufficiently large. For $t \geq 2, \varepsilon > 0$ and $s \geq (C_k + o(1))^t \frac{N}{(\log N)^{t+1}}$. there exist disjoint sets of primes $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P} \cap [1, N]$ with $|\mathcal{A}| = s, |\mathcal{B}| = t$ such that for all $a_i \in \mathcal{A}, b_j \in \mathcal{B}$ and all $\lambda_r \in \{0, \frac{1}{k}, \ldots, \frac{k-1}{k}, 1\}$: all $\lambda_r a_i + (1 - \lambda_r) b_j$ are also prime.

Remark 3.9. Let us remark that Green and Tao actually proved a stronger theorem. Let $\mathcal{S}$ be a set of primes with positive upper density, i.e. $\limsup \frac{|\mathcal{S} \cap [1,N]|}{\pi(N)} > 0$, then $\mathcal{S}$ contains such progressions. Our corollary works in the density situation as well.

Corollary 3.10. *The same theorem holds for any finite set $L = \{\lambda_1, \ldots, \lambda_l\} \subset [0, 1]$ of rational numbers.*

Proof. Let z denote the last common multiple of the denominators λ_r. Then $L \subset \{0, \frac{1}{z}, \frac{2}{z}, \ldots, \frac{z-1}{z}, 1\}$, and Corollary 1.3.10 follows by an application of Theorem 1.3.8. with $k = z$. □

For some other consequences of the Green-Tao theorem see also Granville [20].

For sets $\mathcal{A}, \mathcal{B}$ of equal size $|\mathcal{A}| = |\mathcal{B}|$ the theorem implies:

Theorem 3.11. *For large N there exist disjoint sets of primes $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P} \cap [1, N]$ as above with $|\mathcal{A}|, |\mathcal{B}| \geq \frac{\log N}{(k-2) \log \log N} - (C'_k + \varepsilon) \frac{\log N}{(\log \log N)^2}$.*

For the proof of Theorem 3.8 we use the Green-Tao theorem combined with the counting argument that we already used in the last section. This shows that this counting method is versatile and can be easily adapted to similar situations.

Proof of Theorem 3.8. We define the bipartite graph $G(V_1 \cup V_2, E)$ as follows: The sets of vertices are $V_1 = V_2 = \mathcal{P} \cap [1, N]$, the set of edges is

$$E_k = \{(v_1, v_2) \in V_1 \times V_2 \mid v_1 \neq v_2 \text{ and } \lambda v_1 + (1-\lambda) v_2 \in \mathcal{P}, \lambda \in \{0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1\}\}.$$

Here v_1 corresponds to p_1 and v_2 to p_k. An edge between v_1 and v_2 corresponds to a $(k+1)$-tuple of primes in progression. Note that $p_1 = p_2 = \ldots = p_k$ is not allowed. A complete bipartite graph $K_{s,t}$ corresponds to disjoint sets of primes $\mathcal{A}$ and $\mathcal{B}$ of sizes s and t such that all $\lambda a_i + (1-\lambda) b_j$ are also prime. By the Green-Tao theorem this graph G contains $G_{k+1}(N) \geq (C_{k+1} + o(1)) \frac{N^2}{(\log N)^{k+1}}$ edges. The Kővari-Sós-Turán theorem (Theorem 3.2) guarantees that a bipartite graph with many edges must have a large $K_{s,t}$ as a subgraph.

Suppose that G does not contain a complete bipartite graph $K_{s,t}$ for which the first class of $K_{s,t}$ lies in the first class V_1 of G and the second class in V_2. We then find by Theorem 3.2 and Theorem 3.7 that for sufficiently large N

$$(C_{k+1} + o(1)) \frac{N^2}{(\log N)^{k+1}} \leq |E| \leq s^{\frac{1}{t}} \pi(N)^{2-\frac{1}{t}} + t\pi(N).$$

For large N we have the estimate $\pi(N) \geq \frac{N}{\log N}$. This implies for $t = O(N^\varepsilon)$ that

$$\begin{aligned} \frac{C_k + o(1)}{(\log N)^{k-2}} &\leq s^{\frac{1}{t}} \frac{(\log N)^{1/t}}{N^{1/t}} + t \frac{\log N}{N} \\ &\leq s^{\frac{1}{t}} \frac{(\log N)^{1/t}}{N^{1/t}} \left(1 + \frac{1}{N^{1/3}}\right) \end{aligned}$$

and therefore

$$s \geq (C_k + o(1))^t \frac{N}{(\log N)^{(k-2)t+1}}.$$

Note that, since we can assume w.l.o.g. that $s \geq t$, the choice $t \leq O(N^\varepsilon)$ is not restrictive.

Hence for s smaller than the above bound the graph G contains a complete bipartite graph $K_{s,t}$ which proves Theorem 3.8.

For Theorem 3.11 we want that both sets are of the same size, i.e. $s = t$. An easy computation shows that

$$t = \left\lfloor \frac{\log N}{(k-2) \log \log N} - \left(\frac{\ln C_k}{k-2} + \varepsilon \right) \frac{\log N}{(\log \log N)^2} \right\rfloor$$

is an admissible value, for sufficiently large N. □

The author is grateful to Oriol Serra and Javier Cilleruelo, the organisers of the Doccourse Additive Combinatorics at CRM in Barcelona, for the kind invitation and the interesting discussions. The author also thanks Laurence Rackham for discussions and congratulates him for the excellent results achieved.

References

[1] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, 184. Springer-Verlag, New York, 1998.

[2] J. Brüdern, *Einführung in die analytische Zahlentheorie*, Springer, Berlin, 1995.

[3] S. Chowla, There exists an infinity of 3–combinations of primes in A. P., *Proc. Lahore Philos. Soc.* **6** (1944), 15–16.

[4] N. G. Chudakov, On the Goldbach problem, *Dokl. Akad. Nauk SSSR* **17** (1937), 335-338.

[5] J. G. van der Corput, Sur l'hypothèse de Goldbach pour presque tous les nombres pairs. *Acta Arith.* **2** (1937), 266-290.

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

[7] E. S. Croot, C. Elsholtz, On thin sets of primes expressible as sumsets, *Acta Math. Hungar.* **106** (2005), 197–226.

[8] H. Davenport, *Multiplicative number theory*, second edition. Revised by Hugh L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag, New York-Berlin, 1980.

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

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

[11] 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).

[12] C. Elsholtz, Triples of primes in arithmetic progressions *Quart. J. Math.* **53** (2002), 393-395.

[13] C. Elsholtz, Additive decomposability of multiplicatively defined sets, *Funct. Approx. Comment. Math.* **35** (2006), 61–77.

[14] C. Elsholtz, Multiplicative decomposability of shifted sets, *Bull. London Math. Soc.* **40** (2008), 97–107.

[15] P. Erdős, On sequences of integers no one of which divides the product of two others, *Izr. Inst. Math. and Mech. Univ. Tomsk* **2** (1938), 74–82.

[16] P. Erdős, C.L. Stewart, R. Tijdeman, Some Diophantine equations with many solutions, *Compositio Math.* **66** (1988), 37–56.

[17] T. Estermann, On Goldbach's problem: Proof that almost all even positive integers are sums of two primes, *Proc. London Math. Soc.* (2) **44** (1938), 307–314.

[18] P. X. Gallagher, The large sieve, *Mathematika* **14** (1967), 14–20.

[19] P. X. Gallagher, A larger sieve, *Acta Arith.* **18** (1971), 77–81.

[20] A. Granville, Prime number patterns. *Amer. Math. Monthly* **115** (2008), no. 4, 279–296.

[21] A. Granville, C. Pomerance, Two contradictory conjectures concerning Carmichael numbers, *Math. Comp.* **71** (2002), 883-908.

[22] B. Green and T. Tao, The primes contain aribtrarily long arithmetic progressions, *Annals of Mathematics* **167** (2008), 481-547.

[23] B. Green and T. Tao, Linear equations in primes, to appear in Annals of Mathematics.

[24] E. Grosswald, Arithmetic progressions that consist only of primes, *J. Number Theory* **14** (1982), 9–31.

[25] E. Grosswald, On the number of quadruples of primes in arithmetic progression below a given bound, *Libertas Math.* **2** (1982), 99–112.

[26] K. Gyarmati, On divisibility properties of integers of the form $ab + 1$, *Period. Math. Hungarica* **43** (2001), 71–79.

[27] K. Győry, A. Sárközy, C. L. Stewart, On the number of prime factors of integers of the form $ab + 1$, *Acta Arith.* **74** (1996), 365–385.

[28] K. Győry, C. L. Stewart, R. Tijdeman, On prime factors of sums of integers. III. *Acta Arith.* **49** (1988), 307–312.

[29] A. Hofmann, D. Wolke, On additive decompositions of the set of primes, *Arch. Math.* **67** (1996), 379-382.

[30] 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.

[31] T. Kővari, V. T. Sós, P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.* **3** (1954), 50–57.

[32] W. B. Laffer, H. B. Mann, Decomposition of sets of group elements, *Pacific J. Math.* **14** (1964) 547–558.

[33] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Leipzig, Berlin, 1909, Chelsea reprint New York, 1953.

[34] H. B. Mann, *Addition theorems: The addition theorems of group theory and number theory*, Wiley Interscience, New York-London-Sydney, 1965.

[35] H. Montgomery, The analytic principle of the large sieve, *Bull. Amer. Math Soc.* **84** (1978), 547-567.

[36] M. Nathanson, *Additive Number Theory, Inverse problems and the geometry of sumsets.* Graduate Texts in Mathematics, 165. Springer-Verlag, New York, 1996.

[37] H.-H. Ostmann, *Additive Zahlentheorie*, 2 volumes, Springer-Verlag, Berlin-Heidelberg-New York, 1956, reprint 1968.

[38] C. Pomerance, C. L. Stewart, A. Sárközy, On Divisors of Sums of Integers III, *Pacific J. Math.* **133** (1988), 363-379.

[39] J.-C. Puchta, On additive decompositions of the set of primes. em Arch. Math. **78** (2002), 24–25.

[40] P. Ribenboim, *The new book of prime number records*, 3rd edition, Springer-Verlag, New York, 1995.

[41] T. Tao, V. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.

[42] G. Tenenbaum, *Introduction to analytic and probabilistic number theory.* Cambridge Studies in Advanced Mathematics, 46. Cambridge University Press, Cambridge, 1995.

[43] R. C. Vaughan, Some Applications of Montgomery's Sieve, J. Number Theory **5** (1973), 64-79.

[44] E. Wirsing, Ein metrischer Satz über Mengen ganzer Zahlen, *Arch. Math.* **4** (1953), 392–398.

[45] E. Wirsing, Über die Zahlen, deren Primteiler einer gegebenen Menge angehören. *Arch. Math.* **7**, 263–272, 1956.

Christian Elsholtz
Department of Mathematics
Royal Holloway
Egham
Surrey TW20 0EX
UK
christian.elsholtz@rhul.ac.uk

2000 Mathematics Subject Classification:
11N25, 11N36, 05D10