Algebra & Number Theory

Volume 15

2021

No. 3

The Erdős–Selfridge problem with square-free moduli

Paul Balister, Béla Bollobás, Robert Morris, Julian Sahasrabudhe and Marius Tiba

msp

msp

ALGEBRA AND NUMBER THEORY 15:3 (2021)
https://doi.org/10.2140/ant.2021.15.609

The Erdős–Selfridge problem with square-free moduli

Paul Balister, Béla Bollobás, Robert Morris, Julian Sahasrabudhe and Marius Tiba

A *covering system* is a finite collection of arithmetic progressions whose union is the set of integers. The study of covering systems with distinct moduli was initiated by Erdős in 1950, and over the following decades numerous problems were posed regarding their properties. One particularly notorious question, due to Erdős, asks whether there exist covering systems whose moduli are distinct and all odd. We show that if in addition one assumes the moduli are square-free, then there must be an even modulus.

1. Introduction

Erdős [1950] initiated the study of *covering systems*, i.e., finite collections of arithmetic progressions (we exclude the trivial arithmetic progression $\mathbb{Z}$) that cover the integers, with distinct moduli. Many well-known questions and conjectures have been posed about such systems (some of which appeared frequently in Erdős's collections of open problems), and in recent years there has been significant progress on several of these. A first crucial step was taken by Filaseta, Ford, Konyagin, Pomerance and Yu [2007], who proved that the sum of the reciprocals of the moduli grows quickly with the minimum modulus, and also confirmed a conjecture of Erdős and Graham [1980] on the density of the uncovered set. A further important breakthrough was made by Hough [2015], who resolved the so-called "minimum modulus problem" of Erdős [1950] by showing that the minimum modulus is bounded. More recently, the current authors [BBMST 2018] developed a general method (based on that of [Hough 2015]) for attacking problems of this type, and used it to study the density of the uncovered set, and to prove a conjecture of Schinzel [1967] by showing that there must exist two moduli, one of which divides the other.

In this paper we will further develop the method of [BBMST 2018] in order to make progress on another old and well-known question: does there exist a covering system whose moduli are distinct and all odd? This question appears to have first been asked by Erdős [1965], who later conjectured that there does exist such a system; see [Erdős 1973]. He then went further, conjecturing that there exist covering systems with square-free moduli, all of whose prime factors are arbitrarily large; see [Erdős 1977]. On the other hand (as recounted, for example, by Filaseta, Ford and Konyagin [2000]), Selfridge believed that there do *not* exist such systems, and (perhaps as a result) the question has become known as the Erdős–Selfridge problem. Apart from its intrinsic appeal, the problem is motivated by a theorem of

Balister and Bollobás were partially supported by NSF grant DMS 1600742, Morris was partially supported by CNPq (Proc. 304237/2016-7) and FAPERJ (Proc. 202.993/2017), and Tiba was supported by a Trinity Hall Research Studentship.
MSC2010: primary 11B25; secondary 11A07, 11N35.
Keywords: covering systems, Erdős–Selfridge problem.

Schinzel [1967], who discovered a connection between the nonexistence of such covering systems and the irreducibility of certain polynomials. More precisely, he showed that if no covering system with distinct, odd moduli exists, then for every polynomial $f(x) \in \mathbb{Z}[X]$ with $f \not\equiv 1$, $f(0) \neq 0$ and $f(1) \neq -1$, there exists an infinite arithmetic progression of values of $n \in \mathbb{Z}$ such that $x^n + f(x)$ is irreducible over the rationals.

The first progress on the Erdős–Selfridge problem was made by Simpson and Zeilberger [1991], who proved that the moduli of a covering system with distinct, odd, square-free numbers use at least 18 primes (this was later improved to 22 primes by Guo and Sun [2005]). A major step forward was taken by Hough and Nielsen [2019], who used a refined (and carefully optimized) version of the method of Hough [2015] to prove that every covering system with distinct moduli contains a modulus that is divisible by either 2 or 3. The general method of [BBMST 2018] (which is also based on that of [Hough 2015]) provides a short proof of the following slight strengthening of this result (see [BBMST 2018, Theorem 1.4]): every covering system with distinct moduli contains either an even modulus, a modulus divisible by 3^2, or (possibly equal) moduli d_1 and d_2 with $3 \mid d_1$ and $5 \mid d_2$. Here we will further develop the method of [BBMST 2018], and use it to solve the Erdős–Selfridge problem in the square-free case.

Theorem 1.1. *In any finite collection of arithmetic progressions with distinct square-free moduli > 1 that covers the integers, at least one of the moduli is even.*

We shall prove Theorem 1.1 in a (slightly more general) geometric setting; a second aim of this paper will be to investigate covering systems in this setting. Let $S_1, \dots, S_n$ be finite sets with at least two elements, and set

$$Q = S_1 \times \cdots \times S_n.$$

If $A = A_1 \times \cdots \times A_n \subseteq Q$ with each A_k either equal to S_k or a singleton element of S_k, then we say that A is a *hyperplane*. We will write $A = [x_1, \dots, x_n]$, where $x_k \in S_k \cup \{*\}$ for each $k \in [n]$, and $*$ indicates that $A_k = S_k$. Let us write $F(A) = \{k : x_k \in S_k\}$ for the set of *fixed coordinates* of A, and say that two hyperplanes A and A' are *parallel* if $F(A) = F(A')$. Let us also say that A is *nontrivial* if $F(A) \neq \varnothing$.

Theorem 1.1 is equivalent to the following theorem in this geometric setting.

Theorem 1.2. *For each $k \in [n]$, let p_k be the k-th prime, and set $S_k = [p_{k+1}]$. Any collection of nontrivial hyperplanes that covers $Q := S_1 \times \cdots \times S_n$ contains two parallel hyperplanes.*

To see the equivalence of Theorems 1.1 and 1.2, observe that, by the Chinese remainder theorem, there is a natural equivalence[1] between finite collections $\mathcal{A}$ of arithmetic progressions with square-free, odd, p_{n+1}-smooth moduli that cover the integers, and finite collections $\mathcal{H}$ of hyperplanes that cover the box $Q = [p_2] \times \cdots \times [p_{n+1}]$. Moreover, if the moduli of $\mathcal{A}$ are distinct then the hyperplanes in $\mathcal{H}$ are nonparallel.

[1] To be precise, the progression $a + d\mathbb{Z}$ with $d = \prod_{i \in I} p_i$ corresponds to the hyperplane $A = [x_1, \dots, x_n]$ where $x_i = a \mod p_i$ if $i \in I$, and $x_i = *$ otherwise.

In order to motivate our second main theorem, let us next state, in this geometric setting, a special case (for square-free moduli) of the breakthrough result of Hough [2015] which resolved the Erdős minimum modulus problem.

Theorem 1.3 [Hough 2015]. *Let $p_1, \dots, p_n$ be the first n primes. There exists a constant C such that if $\mathcal{A}$ is a collection of hyperplanes that cover $Q := [p_1] \times \cdots \times [p_n]$, then either two of the hyperplanes are parallel, or there exists a hyperplane $A \in \mathcal{A}$ with $F(A) \subseteq [C]$.*

To deduce Theorem 1.3 from the main result of [Hough 2015], simply note that if every covering system with distinct, square-free moduli contains an arithmetic progression with modulus at most $M \leqslant p_C$, then the progression with minimum modulus corresponds to a hyperplane A with $F(A) \subseteq [C]$. In the other direction, it follows from Theorem 1.3 that a covering system with distinct, square-free moduli contains a progression corresponding to a hyperplane A with $F(A) \subseteq [C]$, and the modulus d of this progression satisfies $d \leqslant \prod_{i=1}^{C} p_i$.

Using our method, we are able to prove the following strengthening of Theorem 1.3.

Theorem 1.4. *For every sequence of integers $(q_k)_{k\geqslant 1}$ such that $q_k \geqslant 2$ for each $k \in \mathbb{N}$ and*

$$\liminf_{k\to\infty} \frac{q_k}{k} > 3,$$

there exists a constant C such that the following holds. Let $\mathcal{A}$ be a collection of hyperplanes that cover $Q := [q_1] \times \cdots \times [q_n]$ for some $n \in \mathbb{N}$. Then either two of the hyperplanes are parallel, or there exists a hyperplane $A \in \mathcal{A}$ with $F(A) \subseteq [C]$.

Note that in Theorem 1.3 the sequence $(p_k)_{k\geqslant 1}$ grows asymptotically as $k \log k$, whereas in Theorem 1.4 we allow the sequence $(q_k)_{k\geqslant 1}$ to grow only linearly. We will show (see Section 4) that Theorem 1.4 is close to best possible, since there exists an example with $\lim q_k/k = 1$ for which the conclusion of the theorem fails.

The rest of the paper is organized as follows. In Section 2 we outline the sieve that we will use in the proofs, and in Section 3 we state and prove our main technical results, Theorems 3.1 and 3.2. In Section 4 we deduce Theorem 1.4. Finally, we dedicate Section 5 to the proof of our main result, Theorem 1.1.

2. Definition of the sieve

In this section we will outline the proofs of Theorems 1.2 and 1.4. In particular, we will generalize the method developed in [BBMST 2018] to the geometric setting, and while doing so we will introduce several new ideas that will prove to be crucial in the proofs. For the convenience of the reader and for completeness, we will include full proofs of all intermediate results, even though several of them are direct adaptations of the corresponding results in [BBMST 2018].

As in the Introduction, let $S_1, \dots, S_n$ be finite sets with at least two elements, set

$$Q := S_1 \times \cdots \times S_n,$$

and let $\mathcal{A}$ be a collection of nontrivial hyperplanes, no two of which are parallel. Set

$$\mathcal{F} = \mathcal{F}(\mathcal{A}) := \big\{F(A) : A \in \mathcal{A}\big\} \subseteq \mathcal{P}([n]) \setminus \{\varnothing\},$$

and (recalling that $F(A) \neq F(A')$ for distinct $A, A' \in \mathcal{A}$) let us index the hyperplanes in $\mathcal{A}$ by the corresponding set of fixed coordinate indices, so $\mathcal{A} = \{A_F : F \in \mathcal{F}\}$. Our goal is to estimate the density (under some probability measure) of the uncovered set

$$R := Q \setminus \bigcup_{F \in \mathcal{F}} A_F.$$

Rather than considering the entire collection of hyperplanes $\mathcal{A}$ all at once, we expose the hyperplanes dimension by dimension and track how the density of the uncovered set evolves. To be more precise, define, for each $1 \leqslant k \leqslant n$,

$$\mathcal{F}_k := \big\{F \in \mathcal{F} : F \subseteq [k]\big\} \quad \text{and} \quad \mathcal{A}_k := \big\{A_F : F \in \mathcal{F}_k\big\}$$

for the family of sets of fixed coordinate indices and the corresponding hyperplanes that are contained in the initial segment $[k]$. Let

$$R_k := Q \setminus \bigcup_{F \in \mathcal{F}_k} A_F = Q \setminus \bigcup_{A_F \in \mathcal{A}_k} A_F$$

be the set of elements not contained in any of the hyperplanes of $\mathcal{A}_k$, so in particular $R_n = R$. We also write $\mathcal{N}_k := \mathcal{F}_k \setminus \mathcal{F}_{k-1}$ for the family of "new" sets of fixed coordinate indices at the k-th stage, i.e., those sets that contain k and are contained in $[k]$, and define

$$B_k := \bigcup_{F \in \mathcal{N}_k} A_F \tag{1}$$

to be the union of the hyperplanes exposed at step k, so that $R_k = R_{k-1} \setminus B_k$.

It will often be convenient to consider R_k, B_k and A_F with $F \in \mathcal{F}_k$ as subsets of

$$Q_k := S_1 \times \cdots \times S_k$$

by identifying $X \subseteq Q_k$ with $X \times S_{k+1} \times \cdots \times S_n \subseteq Q$. We call a set of this form Q_k-*measurable*.

2.1. ***The probability measures*** $\mathbb{P}_k$**.** The construction of the probability measures is similar to that in [BBMST 2018], and no significant new ideas are needed. The main difference is that instead of starting with the uniform measure as our $\mathbb{P}_0$, we allow for possible optimization of the measure on the first few coordinates. In general we will start with some measure $\mathbb{P}_a$, to be determined, which will be supported on $R_a \subseteq Q_a = S_1 \times \cdots \times S_a$.[2]

Our aim is to construct, for each $a < k \leqslant n$, a measure $\mathbb{P}_k$ on Q_k in such a way that $\mathbb{P}_k(B_k)$ is small, but without changing the measure of B_i for any $i < k$. Fix a sequence of constants $\delta_{a+1}, \ldots, \delta_n \in [0, 1/2]$, and assume that we have already defined a probability measure $\mathbb{P}_{k-1}$ on Q_{k-1}. Recall that $Q_k = Q_{k-1} \times S_k$, and hence the elements of Q_k can be written as pairs (x, y), where $x \in Q_{k-1}$ and $y \in S_k$. We may

[2]We emphasize that here we allow $a = 0$, in which case $\mathbb{P}_0$ is the trivial probability measure on the empty product.

view R_{k-1} as a collection of fibres of the form $F_x = \{(x, y) : y \in S_k\} \subseteq Q_k$, where $\mathbb{P}_{k-1}$ is extended uniformly to a measure on Q_k (so is uniform on each fibre), and view R_k as being obtained from R_{k-1} by removing B_k, i.e., by removing the points that are contained in the new hyperplanes of $\mathcal{A}_k \setminus \mathcal{A}_{k-1}$.

Now, for each $x \in Q_{k-1}$, define

$$\alpha_k(x) = \frac{\mathbb{P}_{k-1}\big(F_x \cap B_k\big)}{\mathbb{P}_{k-1}(x)} = \frac{\big|\{y \in S_k : (x, y) \in B_k\}\big|}{|S_k|}, \tag{2}$$

that is, the proportion of the fibre F_x that is removed at stage k. The probability measure $\mathbb{P}_k$ on Q_k is defined as follows:

$$\mathbb{P}_k(x, y) := \begin{cases} \max\left\{0,\ \dfrac{\alpha_k(x) - \delta_k}{\alpha_k(x)(1-\delta_k)}\right\} \cdot \mathbb{P}_{k-1}(x, y) & \text{if } (x, y) \in B_k; \\ \min\left\{\dfrac{1}{1-\alpha_k(x)},\ \dfrac{1}{1-\delta_k}\right\} \cdot \mathbb{P}_{k-1}(x, y) & \text{if } (x, y) \notin B_k. \end{cases} \tag{3}$$

To motivate the definition above, note that if $\alpha_k(x) \leqslant \delta_k$, then $\mathbb{P}_k(x, y) = 0$ for every element of Q_k that is covered in step k, and that the measure is increased proportionally elsewhere to compensate. On the other hand, for those $x \in Q_{k-1}$ for which $\alpha_k(x) > \delta_k$, we "cap" the distortion by increasing the measure at each point not covered in step k by a factor of $1/(1-\delta_k)$, and decreasing the measure on removed points by a corresponding factor.

The measure $\mathbb{P}_k$ satisfies the following simple properties, cf. [BBMST 2018, Lemmas 2.1 and 2.2].

Lemma 2.1. *For any $k > a$ and any Q_{k-1}-measurable set S we have*

$$\mathbb{P}_k(S) = \mathbb{P}_{k-1}(S). \tag{4}$$

For any set $S \subseteq Q$, we have

$$\mathbb{P}_k(S) \leqslant \frac{1}{1-\delta_k} \cdot \mathbb{P}_{k-1}(S). \tag{5}$$

Moreover, if $S \subseteq B_k$ then

$$\mathbb{P}_k(S) \leqslant \mathbb{P}_{k-1}(S). \tag{6}$$

In particular, it follows from Lemma 2.1 that if

$$\sum_{k=a+1}^{n} \mathbb{P}_k(B_k) < 1 \tag{7}$$

then $\mathcal{A}$ does not cover Q, since B_k is a Q_k-measurable set, so by (4) we have $\mathbb{P}_n(B_k) = \mathbb{P}_k(B_k)$. For each $a \leqslant k \leqslant n$, define

$$\mu_k := 1 - \sum_{i=a+1}^{k} \mathbb{P}_i(B_i),$$

and observe that $\mu_k \leqslant \mathbb{P}_k(R_k)$.

3. Bounding the density of the covered set

In this section we will prove two technical results, Theorems 3.1 and 3.2, which together imply Theorems 1.2 and 1.4. We remark that Theorem 3.1 essentially follows from [BBMST 2018, Theorem 3.1], but Theorem 3.2 introduces a new bound that is motivated geometrically, and that will prove to be crucial in the proof of Theorem 1.2.

Given a collection $\mathcal{A}$ of hyperplanes in $Q = S_1 \times \cdots \times S_n$, a probability distribution $\mathbb{P}_a$ supported on R_a, and constants $\delta_{a+1}, \dots, \delta_n \in [0, 1/2]$, let the probability distributions $\mathbb{P}_k$ and functions $\alpha_k : Q_{k-1} \to [0, 1]$ be defined as in (2) and (3), and set

$$M_k^{(1)} := \mathbb{E}_{k-1}\big[\alpha_k(x)\big] \quad \text{and} \quad M_k^{(2)} := \mathbb{E}_{k-1}\big[\alpha_k(x)^2\big],$$

where we write $\mathbb{E}_{k-1}$ to denote expectation with respect to the measure $\mathbb{P}_{k-1}$.

In order to show that $\mathcal{A}$ does not cover Q, it is sufficient, by (7), to show that $\mu_n > 0$. To do so, we will bound $\mathbb{P}_k(B_k)$ in terms of the moments $M_k^{(1)}$ and $M_k^{(2)}$. As noted above, the following theorem was (essentially) proved in [BBMST 2018].

Theorem 3.1. *Let $\mathcal{A}$ be a collection of nontrivial hyperplanes in $Q = S_1 \times \cdots \times S_n$, no two of which are parallel. If*

$$\sum_{k=a+1}^{n} \min\left\{ M_k^{(1)}, \frac{M_k^{(2)}}{4\delta_k(1-\delta_k)} \right\} < 1, \tag{8}$$

then $\mathcal{A}$ does not cover Q.

In order to show that (8) holds in our applications, we need to bound the moments of $\alpha_k(x)$. To state our bounds on $M_k^{(1)}$ and $M_k^{(2)}$, we will need some additional notation. Define a function $c : \mathcal{P}([a]) \to [0, 1]$ by setting

$$c(I) = \max\big\{\mathbb{P}_a(H) : H \text{ is a hyperplane in } Q_a \text{ with } F(H) = I\big\} \tag{9}$$

for each $I \subseteq [a]$, and define a function $\nu : \mathcal{P}([a+1, n]) \to \mathbb{R}_{>0}$, by setting

$$\nu(J) = \prod_{j \in J} \frac{1}{(1-\delta_j)|S_j|} \tag{10}$$

for each $J \subseteq [a+1, n]$. Note that $c(\varnothing) = \nu(\varnothing) = 1$. For each $k \geqslant a$ and $x \in \mathbb{R}$, set

$$c_k(x) = \sum_{I \subseteq [a]} \sum_{J \subseteq [a+1,k]} c(I)\nu(J)x^{|I|+|J|} = \sum_{I \subseteq [a]} c(I)x^{|I|} \prod_{j=a+1}^{k} \left(1 + \frac{x}{(1-\delta_j)|S_j|}\right). \tag{11}$$

The following technical theorem provides general bounds on $M_k^{(1)}$ and $M_k^{(2)}$.

Theorem 3.2. *Let $\mathcal{A}$ be a collection of nontrivial hyperplanes in $Q = S_1 \times \cdots \times S_n$, no two of which are parallel. Then, for each $a < k \leqslant n$,*

$$M_k^{(1)} \leqslant \frac{c_{k-1}(1)}{|S_k|} \quad \text{and} \quad M_k^{(2)} \leqslant \frac{c_{k-1}(3)}{|S_k|^2}. \tag{12}$$

Moreover, if none of the hyperplanes in $\mathcal{N}_k$ has co-dimension 1, *then*

$$M_k^{(2)} \leqslant \frac{1}{|S_k|^2}\big(c_{k-1}(3) - 2c_{k-1}(1) + 1\big). \tag{13}$$

Before embarking on the (straightforward) proofs of Theorems 3.1 and 3.2, let us briefly discuss the bound (13), which will play an important role in the proof of Theorem 1.1. In order to apply it, we first need to remove from $\mathcal{A}$ each of the codimension 1 hyperplanes, each of which is of the form $S_1 \times \cdots \times S_{i-1} \times \{s\} \times S_{i+1} \times \cdots \times S_n$ for some $i \in [n]$ and $s \in S_i$. Note that in doing so we remove the point s from the possible values of the i-th coordinate, effectively replacing S_i by $S_i' := S_i \setminus \{s\}$. After removing these hyperplanes, the remaining elements of $\mathcal{A}$ will all have at least two fixed coordinates, and can be assumed to be hyperplanes in $Q' = S_1' \times \cdots \times S_n'$, where $S_i' = S_i$ if $\{i\} \notin \mathcal{F}(\mathcal{A})$. Removing the codimension 1 hyperplanes in this way makes a significant difference to our estimate on $M_k^{(2)}$, at the expense of (possibly) reducing each $|S_i|$ by 1. In practice, this turns out to often give better bounds on the removed measure.

Proof of Theorem 3.1. Observe first that

$$\mathbb{P}_n(B_k) = \mathbb{P}_k(B_k) \leqslant \mathbb{P}_{k-1}(B_k) = \mathbb{E}_{k-1}\big[\alpha_k(x)\big],$$

where the first two steps follow by Lemma 2.1 (since B_k is Q_k-measurable), and the third follows by the definition (2) of $\alpha_k(x)$. Moreover, by (2) and (3) (the definitions of α_k and $\mathbb{P}_k$), we have

$$\begin{aligned}
\mathbb{P}_k(B_k) &= \sum_{x \in Q_{k-1}} \max\left\{0,\ \frac{\alpha_k(x) - \delta_k}{\alpha_k(x)(1-\delta_k)}\right\} \cdot \mathbb{P}_{k-1}\big(F_x \cap B_k\big) \\
&= \frac{1}{1-\delta_k} \sum_{x \in Q_{k-1}} \max\big\{0,\ \alpha_k(x) - \delta_k\big\} \cdot \mathbb{P}_{k-1}(x) \\
&\leqslant \frac{1}{1-\delta_k} \sum_{x \in Q_{k-1}} \frac{\alpha_k(x)^2}{4\delta_k} \cdot \mathbb{P}_{k-1}(x) = \frac{\mathbb{E}_{k-1}\big[\alpha_k(x)^2\big]}{4\delta_k(1-\delta_k)},
\end{aligned} \tag{14}$$

where we used the elementary inequality $\max\{a - d, 0\} \leqslant a^2/4d$, which is easily seen to hold for all $a, d > 0$ by rearranging the inequality $(a - 2d)^2 \geqslant 0$.

It follows that the uncovered set R satisfies

$$\mathbb{P}_n(R) \geqslant 1 - \sum_{k=a+1}^{n} \mathbb{P}_n(B_k) \geqslant 1 - \sum_{k=a+1}^{n} \min\left\{M_k^{(1)},\ \frac{M_k^{(2)}}{4\delta_k(1-\delta_k)}\right\} > 0,$$

by (8), and hence $\mathcal{A}$ does not cover Q, as required. □

In the proof of Theorem 3.2 we will use the following notation. Given a hyperplane $A = [x_1, \ldots, x_n]$ and $X \subseteq [n]$, we define $A^X = [y_1, \ldots, y_n]$ to be the hyperplane with $y_i = x_i$ for all $i \in F(A) \cap X$, and $y_i = *$ otherwise. Note that $(A^X)^Y = A^{X \cap Y}$ for every $X, Y \subseteq [n]$.

The first step in the proof of Theorem 3.2 is the following easy bound on the $\mathbb{P}_k$-measure of a Q_k-measurable hyperplane.

Lemma 3.3. *Let $a \leqslant k \leqslant n$, and let A be a Q_k-measurable hyperplane. If $F(A) = I \cup J$, where $I \subseteq [a]$ and $J \subseteq [a+1, k]$, then*

$$\mathbb{P}_k(A) \leqslant c(I)\nu(J). \tag{15}$$

Proof. The proof is by induction on k. Note first that for $k = a$ the conclusion follows immediately from the definition (9) of the function c, since $\nu(\varnothing) = 1$. So let $k \in [a+1, n]$, and assume that the claimed bound holds for $\mathbb{P}_{k-1}$.

Note first that if $k \notin F(A)$ then A is Q_{k-1}-measurable, and so the claimed bound follows immediately by (4) and the induction hypothesis. So assume that $k \in F(A)$, and observe that, by (5), we have

$$\mathbb{P}_k(A) \leqslant \frac{1}{1-\delta_k}\,\mathbb{P}_{k-1}(A) = \frac{1}{(1-\delta_k)|S_k|}\,\mathbb{P}_{k-1}\big(A^{[k-1]}\big),$$

since the probability measure $\mathbb{P}_{k-1}$ is extended uniformly on each fibre. Since $A^{[k-1]}$ is Q_{k-1}-measurable, by the induction hypothesis we have

$$\mathbb{P}_{k-1}\big(A^{[k-1]}\big) \leqslant c(I)\nu(J \setminus \{k\}),$$

and so, recalling the definition (10) of the function ν, the claimed bound follows. □

We will next prove the following bound on the t-th moments $\mathbb{E}_{k-1}\big[\alpha_k(x)^t\big]$.

Lemma 3.4. *For each $a < k \leqslant n$ and $t \in \mathbb{N}$ we have*

$$\mathbb{E}_{k-1}\big[\alpha_k(x)^t\big] \leqslant \frac{1}{|S_k|^t} \sum_{F_1,\dots,F_t \in \mathcal{N}_k} c\big((F_1 \cup \dots \cup F_t) \cap [a]\big) \cdot \nu\big((F_1 \cup \dots \cup F_t) \cap [a+1, k-1]\big).$$

Proof. Observe first that, for each $x \in Q_{k-1}$, we have[3]

$$\alpha_k(x) = \frac{1}{|S_k|} \sum_{y \in S_k} \mathbb{1}\big[(x, y) \in B_k\big] \leqslant \frac{1}{|S_k|} \sum_{y \in S_k} \sum_{F \in \mathcal{N}_k} \mathbb{1}\big[(x, y) \in A_F\big],$$

by the union bound, and the definitions (1) and (2) of B_k and α_k. Note that, given $x \in Q_{k-1}$ and $F \in \mathcal{N}_k$, there exists $y \in S_k$ with $(x, y) \in A_F$ if and only if $x \in A_F^{[k-1]}$, and moreover such a y (if it exists) is unique. It follows that

$$\alpha_k(x) \leqslant \frac{1}{|S_k|} \sum_{F \in \mathcal{N}_k} \mathbb{1}\big[x \in A_F^{[k-1]}\big].$$

[3]Here we write $\mathbb{1}[E]$ for the indicator function of an event E, which takes the value 1 if the event holds, and 0 otherwise.

Note also that if A_1 and A_2 are hyperplanes, then $A_1 \cap A_2$ is either the empty set, or a hyperplane whose set of fixed coordinate indices is $F(A_1) \cup F(A_2)$. Therefore, by Lemma 3.3,

$$\begin{aligned}\mathbb{E}_{k-1}\big[\alpha_k(x)^t\big] &\leqslant \frac{1}{|S_k|^t} \sum_{F_1,\dots,F_t \in \mathcal{N}_k} \mathbb{P}_{k-1}\big(A_{F_1}^{[k-1]} \cap \dots \cap A_{F_t}^{[k-1]}\big) \\ &\leqslant \frac{1}{|S_k|^t} \sum_{F_1,\dots,F_t \in \mathcal{N}_k} c\big((F_1 \cup \dots \cup F_t) \cap [a]\big) \cdot \nu\big((F_1 \cup \dots \cup F_t) \cap [a+1, k-1]\big),\end{aligned}$$

as required. □

The claimed bounds on $M_k^{(1)}$ and $M_k^{(2)}$ will now follow easily.

Proof of Theorem 3.2. By Lemma 3.4, we have

$$\begin{aligned}\mathbb{E}_{k-1}\big[\alpha_i(x)^t\big] &\leqslant \frac{1}{|S_k|^t} \sum_{F_1,\dots,F_t \in \mathcal{N}_k} c\big((F_1 \cup \dots \cup F_t) \cap [a]\big) \cdot \nu\big((F_1 \cup \dots \cup F_t) \cap [a+1, k-1]\big) \\ &\leqslant \frac{1}{|S_k|^t} \sum_{I \subseteq [a]} \sum_{J \subseteq [a+1,k-1]} \sum_{\substack{X_1,\dots,X_t \subseteq [k-1] \\ X_1 \cup \dots \cup X_t = I \cup J}} c(I)\nu(J) \\ &= \frac{1}{|S_k|^t} \sum_{I \subseteq [a]} \sum_{J \subseteq [a+1,k-1]} (2^t - 1)^{|I|+|J|} c(I)\nu(J) \;=\; \frac{c_{k-1}(2^t-1)}{|S_k|^t},\end{aligned}$$

which proves (12). To prove (13), suppose that $\mathcal{F}(\mathcal{N}_k)$ contains no singletons (i.e., $\{k\} \notin \mathcal{F}(\mathcal{N}_k)$), and observe that, by Lemma 3.4, we have

$$\begin{aligned}|S_k|^2\, \mathbb{E}_{k-1}\big[\alpha_k(x)^2\big] &\leqslant \sum_{F_1,F_2 \in \mathcal{N}_k} c\big((F_1 \cup F_2) \cap [a]\big) \cdot \nu\big((F_1 \cup F_2) \cap [a+1, k-1]\big) \\ &\leqslant \sum_{I \subseteq [a]} \sum_{J \subseteq [a+1,k-1]} \sum_{\substack{\varnothing \neq X_1, X_2 \subseteq [k-1] \\ X_1 \cup X_2 = I \cup J}} c(I)\nu(J) \\ &= 1 + \sum_{I \subseteq [a]} \sum_{J \subseteq [a+1,k-1]} \big(3^{|I|+|J|} - 2\big) c(I)\nu(J) \\ &= c_{k-1}(3) - 2c_{k-1}(1) + 1,\end{aligned}$$

as required. □

4. Proof of Theorem 1.4

In order to deduce Theorem 1.4 from Theorems 3.1 and 3.2, it will suffice to show that there is an appropriate choice of C and $\delta_1, \delta_2, \dots, \delta_n$ such that $\mu_n > 0$.

Proof of Theorem 1.4. Let $(q_k)_{k \geqslant 1}$ be a sequence of integers with $\liminf_{k\to\infty} q_k/k > 3$, and let $N \in \mathbb{N}$ and $\varepsilon > 0$ be such that $q_k > (3+\varepsilon)k$ for all $k \geqslant N$. Let $C = C(N, \varepsilon)$ be sufficiently large, let $n \in \mathbb{N}$, and for each $k \in [n]$, let S_k be a set of size q_k. We will show that if $\mathcal{A} = \{A_F : F \in \mathcal{F}\}$ is a finite collection of

hyperplanes in $Q = S_1 \times \cdots \times S_n$, no two of which are parallel, and $F(A) \not\subseteq [C]$ for every $A \in \mathcal{A}$, then $\mathcal{A}$ does not cover Q.

Fix $\delta_1 = \cdots = \delta_n = \varepsilon/6$, and assume (without loss of generality) that ε is sufficiently small. We will start with the trivial probability measure $\mathbb{P}_0$ on the empty product Q_0 (which we also think of as the uniform probability measure on Q), and construct inductively the probability measures $\mathbb{P}_k$ as described in Section 2. By Theorem 3.1 it suffices to show that

$$\sum_{k=1}^{n} \frac{M_k^{(2)}}{4\delta_k(1-\delta_k)} < 1.$$

To prove this, note first that $M_k^{(2)} = 0$ for all $1 \leqslant k \leqslant C$, since $F(A) \not\subseteq [C]$ for every $A \in \mathcal{A}$. So let $C < k \leqslant n$, and observe that, by Theorem 3.2, we have

$$M_k^{(2)} \leqslant \frac{c_{k-1}(3)}{|S_k|^2} = \frac{1}{|S_k|^2} \prod_{j=1}^{k-1} \left(1 + \frac{3}{(1-\delta_j)|S_j|}\right).$$

Now, since $|S_j| = q_j \geqslant (3+\varepsilon)\, j$ for all $j \geqslant N$, and by our choice of δ_j, it follows that[4]

$$\prod_{j=N}^{k-1} \left(1 + \frac{3}{(1-\delta_j)|S_j|}\right) \leqslant \exp\left(\sum_{j=N}^{k-1} \frac{3}{(1-\varepsilon/6)(3+\varepsilon)\, j}\right) \leqslant k^{1-\varepsilon/9}.$$

Moreover,

$$\prod_{j=1}^{N-1} \left(1 + \frac{3}{(1-\delta_j)|S_j|}\right) \leqslant 3^N.$$

Thus, assuming that $C \geqslant N$ (so $|S_k| \geqslant 3k$), we have

$$M_k^{(2)} \leqslant \frac{3^N}{|S_k|^2} \cdot k^{1-\varepsilon/9} \leqslant \frac{3^{N-2}}{k^{1+\varepsilon/9}}$$

for every $C < k \leqslant n$, and hence

$$\sum_{k=1}^{n} \frac{M_k^{(2)}}{4\delta_k(1-\delta_k)} \leqslant \frac{3^N}{\varepsilon} \sum_{k=C}^{n} \frac{1}{k^{1+\varepsilon/9}} < 1$$

if $C = C(N, \varepsilon)$ is sufficiently large, as required. □

We will next show that the condition on the sequence $(q_k)_{k\geqslant 1}$ in Theorem 1.4 is close to best possible. To be precise, we will prove the following proposition.

[4]To see the final inequality, note that $\sum_{j=N}^{k-1} 1/j \leqslant \log k$ and $(1-\varepsilon/6)(3+\varepsilon)(1-\varepsilon/9) \geqslant 3$, since we assumed that ε is sufficiently small.

Proposition 4.1. *There exists a sequence of integers $(q_k)_{k\geqslant 1}$ with $q_k \geqslant 2$ for all $k \in \mathbb{N}$ and*

$$\lim_{k\to\infty} \frac{q_k}{k} = 1$$

such that the following holds. For each $C > 0$, there exists $n \in \mathbb{N}$ and a collection $\mathcal{A}$ of nontrivial hyperplanes that cover $Q := [q_1] \times \cdots \times [q_n]$, no two of which are parallel, and with $F(A) \cap [C] = \varnothing$ for every $A \in \mathcal{A}$.

The first step is the following simple lemma; all hyperplanes are assumed to be nontrivial.

Lemma 4.2. *Let $n \geqslant 3$, and let $q_1, \ldots, q_n \geqslant 2$ be a sequence of integers such that*

$$\prod_{k=1}^{n}\left(1+\frac{1}{q_k}\right) \geqslant n\log n.$$

Then $Q = [q_1] \times \cdots \times [q_n]$ can be covered with hyperplanes, no two of which are parallel.

Proof. The proof is by induction on n, so first let $n = 3$, and note that if $2 \leqslant q_1 \leqslant q_2 \leqslant q_3$ satisfy $\prod_{k=1}^{3}(1+q_k^{-1}) \geqslant 3\log 3$, then $q_1 = q_2 = 2$. Now observe that $[2] \times [2]$ (and hence $[2] \times [2] \times [q_3]$) can be covered by three hyperplanes, no two of which are parallel.

For the induction step, observe first that, by the induction hypothesis, if

$$\prod_{k=1}^{n-1}(1+q_k^{-1}) \geqslant (n-1)\log(n-1)$$

then we can find hyperplanes (with fixed coordinates in $[n-1]$) which cover Q. Thus we may assume that

$$1+\frac{1}{q_n} > \frac{n\log n}{(n-1)\log(n-1)} > 1+\frac{1}{n},$$

and hence (without loss of generality) that $2 \leqslant q_1 \leqslant \ldots \leqslant q_n < n$.

We now cover Q greedily: for each set $\varnothing \neq F \subseteq [n]$ in turn we choose a hyperplane A_F with fixed coordinates F so as to cover as much of the remaining (uncovered) subset of Q as possible. Since Q can be partitioned into exactly $\prod_{k\in F} q_k$ such hyperplanes, there must exist some choice of A_F that covers at least a proportion $\prod_{k\in F} q_k^{-1}$ of the remaining set. Thus, after all the hyperplanes have been chosen, the remaining set has size at most

$$|Q| \prod_{\varnothing\neq F\subseteq[n]}\left(1-\prod_{k\in F}\frac{1}{q_k}\right) \leqslant |Q|\exp\left(-\sum_{\varnothing\neq F\subseteq[n]}\prod_{k\in F}\frac{1}{q_k}\right) = \exp\left(1+\sum_{k=1}^{n}\log q_k - \prod_{k=1}^{n}\left(1+\frac{1}{q_k}\right)\right).$$

Now simply observe that

$$1+\sum_{k=1}^{n}\log q_k - \prod_{k=1}^{n}\left(1+\frac{1}{q_k}\right) < 0,$$

since $1+\sum_{k=1}^{n}\log q_k \leqslant 1+n\log(n-1) < n\log n$, whereas $\prod_{k=1}^{n}(1+q_k^{-1}) \geqslant n\log n$, by assumption. It follows that the number of uncovered points is less than 1, as required. □

We can now easily deduce Proposition 4.1.

Proof of Proposition 4.1. Assume that C is sufficiently large, and set

$$q_k := \left\lfloor \left(1 - \frac{2}{\log k}\right) k \right\rfloor$$

for each $k > C$. Observe that $\lim_{k\to\infty} q_k/k = 1$, and that

$$\begin{aligned}\prod_{k=C+1}^{n} \left(1 + \frac{1}{q_k}\right) &= \exp\left(\sum_{k=C+1}^{n} \frac{1}{q_k} + \frac{O(1)}{q_k^2} \right) = \exp\left(\sum_{k=C+1}^{n} \left(\frac{1}{k} + \frac{2}{k \log k} \right) + O(1) \right) \\ &= \exp\big(\log n + 2 \log\log n + O_C(1)\big) = \Omega\big(n(\log n)^2\big).\end{aligned}$$

Thus, for all sufficiently large n, we have

$$\prod_{k=C+1}^{n} \left(1 + \frac{1}{q_k}\right) \geqslant (n - C) \log(n - C),$$

and hence, by Lemma 4.2, we can cover $[q_{C+1}] \times \cdots \times [q_n]$ with hyperplanes, no two of which are parallel. But this implies that we can cover $[q_1] \times \cdots \times [q_n]$ with hyperplanes whose fixed coordinates do not intersect $[C]$, as required. □

5. The Erdős–Selfridge problem

In this section we will prove Theorem 1.2 (and hence also Theorem 1.1). To do so, we will again apply the sieve introduced in Section 2, but this time we will need to choose the various parameters much more carefully. In particular, we will deal with the primes in three groups: first the set $\{3, 5, 7, 11\}$, then the primes between 13 and 73, and finally the primes larger than 73. We will discuss these in reverse order, so as to motivate the bounds we prove.

Let $\mathcal{B}$ be a collection of nontrivial hyperplanes in $P := [3] \times [5] \times \cdots \times [p_n]$, no two of which are parallel. Our aim is to show that $\mathcal{B}$ does not cover P. To do so, we will in fact apply our sieve to a modified collection, obtained by removing the co-dimension 1 hyperplanes, as described after the statement of Theorem 3.2, for the primes $p \leqslant p_{21} = 73$.[5] After doing so, we obtain a collection $\mathcal{A}$ of hyperplanes in $Q = S_2 \times \cdots \times S_n$, where $|S_k| = p_k - 1$ for each $2 \leqslant k \leqslant 21$, and $|S_k| = p_k$ for each $22 \leqslant k \leqslant n$, such that no two hyperplanes in $\mathcal{A}$ are parallel, and if $F(A) = \{i\}$ for some $A \in \mathcal{A}$ then $i \geqslant 22$. We remark that we will use some results from [BBMST 2018] to deal with the large primes, and our indexing of the sets S_k is chosen to avoid a conflict with the notation used there. It will also be convenient (see Section 5.3, below) to assume (as we may) that $A \not\subseteq B$ for any $A, B \in \mathcal{A}$ with $A \neq B$.

[5] For simplicity, we will assume (without loss of generality) that $\mathcal{B}$ contains a co-dimension 1 hyperplane with fixed set $\{i\}$ for each i.

5.1. ***The primes greater than 73.*** For large primes, it will suffice to apply the results of [BBMST 2018, Section 6]. To state the results we will use, let us first recall some notation. Assume that we have chosen $\delta_6, \dots, \delta_{21}$ and some probability distribution $\mathbb{P}_a = \mathbb{P}_5$ supported on

$$R_5 \subseteq Q_5 := S_2 \times \cdots \times S_5.$$

Now, noting that $p_{21} = 73$, set $\kappa := c_{21}(3)$ and define

$$f_k = f_k(\mathcal{A}) := \frac{\kappa}{\mu_k} \prod_{21<i\leqslant k} \left(1 + \frac{3p_i - 1}{(1-\delta_i)(p_i-1)^2}\right) \tag{16}$$

for each $k \geqslant 21$, where the constants $\{\delta_i : i > 21\}$ will be chosen later, and recall that

$$\mu_k = 1 - \sum_{i=6}^{k} \mathbb{P}_i(B_i)$$

for each $5 \leqslant k \leqslant n$. In order to apply the results stated below, we need to check that condition (20) of [BBMST 2018], which states that

$$M_k^{(2)} \leqslant \frac{\kappa}{(p_k-1)^2} \prod_{21<i<k} \left(1 + \frac{3p_i - 1}{(1-\delta_i)(p_i-1)^2}\right) = \frac{\mu_{k-1} f_{k-1}}{(p_k-1)^2}, \tag{17}$$

is satisfied for every $k > 21$. To see this, note that

$$c_k(3) = c_{k-1}(3)\left(1 + \frac{3}{(1-\delta_k)p_k}\right) \leqslant c_{k-1}(3)\left(1 + \frac{3p_k - 1}{(1-\delta_k)(p_k-1)^2}\right)$$

for every $k > 21$, and observe that therefore, by Theorem 3.2, we have

$$M_k^{(2)} \leqslant \frac{c_{k-1}(3)}{|S_k|^2} \leqslant \frac{c_{21}(3)}{|S_k|^2} \prod_{21<i<k} \left(1 + \frac{3p_i - 1}{(1-\delta_i)(p_i-1)^2}\right),$$

as required, since $|S_k| = p_k > p_k - 1$. Indeed, the bound (17) is the bound given by the arguments of [BBMST 2018] when the primes p_k are allowed to occur to higher powers in the moduli when $k > 21$. The following theorem, which gives an almost optimal termination criterion when k is large, was proved in [BBMST 2018].

Theorem 5.1. *Let $k \geqslant 10$. If $\mu_k > 0$ and $f_k(\mathcal{A}) \leqslant (\log k + \log\log k - 3)^2 k$, then the collection of hyperplanes $\mathcal{A}$ does not cover Q.*

Using Theorem 5.1, one can now compute (see the discussion in [BBMST 2018, Section 6]) weaker sufficient conditions on f_k for the event that the uncovered set is nonempty. In particular, we will use the following result; cf. [BBMST 2018, Corollary 6.3].[6]

Corollary 5.2. *If $f_{21}(\mathcal{A}) \leqslant 138.877$, then the collection of hyperplanes $\mathcal{A}$ does not cover Q.*

[6]In order to prove Corollary 5.2, it suffices to set $\delta_i = (1+a_i)/(1+\sqrt{1+a_i(1+a_i)/b_i f_{i-1}})$ for each $i > 21$, where $a_i = (3p_i-1)/(p_i-1)^2$ and $b_i = 1/4(p_i-1)^2$, and apply [BBMST 2018, Lemma 6.2] and Theorem 5.1 (which are both relatively straightforward consequences of (14) and Theorem 3.2), see [BBMST 2018; 2021].

In order to prove Theorem 1.2, it will therefore suffice to show that we can choose the probability distribution $\mathbb{P}_5$ and constants $\delta_6, \dots, \delta_{21}$ such that $f_{21}(\mathcal{A}) \leqslant 138.877$.

5.2. ***The primes between* 13 *and* 73.** We will next deduce from Corollary 5.2 a sufficient condition[7] on $\mathbb{P}_5$ for the event that the uncovered set is nonempty.

Lemma 5.3. *If $\mathbb{P}_5$ satisfies $c_5(3) - 3c_5(1)/4 \leqslant 9.019$, then there exists a choice of the constants $\delta_6, \dots, \delta_{21}$ such that $f_{21}(\mathcal{A}) < 138.874$.*

Proof. Set $\hat{\mu}_5 := \mu_5 = 1$ and define, for $k = 6, \dots, 21$,

$$\hat{\mu}_k := \hat{\mu}_{k-1} - \frac{c_{k-1}(3) - 2c_{k-1}(1) + 1}{4\delta_k(1-\delta_k)|S_k|^2} \tag{18}$$

for each $k \in \{6, \dots, 21\}$. Recall that, by (14) and Theorem 3.2, we have

$$\mathbb{P}_k(B_k) \leqslant \frac{M_k^{(2)}}{4\delta_k(1-\delta_k)} \leqslant \frac{c_{k-1}(3) - 2c_{k-1}(1) + 1}{4\delta_k(1-\delta_k)|S_k|^2},$$

so $\hat{\mu}_k \leqslant \mu_k$, for each $k \in \{6, \dots, 21\}$, and hence $f_{21}(\mathcal{A}) = c_{21}(3)/\mu_{21} \leqslant c_{21}(3)/\hat{\mu}_{21}$.

Now, observe that

$$c_k(x) = c_{k-1}(x)\left(1 + \frac{x}{(1-\delta_k)(p_k - 1)}\right) \tag{19}$$

for each $x \in \{1, 3\}$ and $k \in \{6, \dots, 21\}$, and therefore (for fixed δ_k) we may write $c_{21}(3) = g\big(c_5(3)\big)$ and $\hat{\mu}_{21} = h\big(c_5(1), c_5(3)\big)$ as functions of $c_5(1)$ and $c_5(3)$. Moreover, the function $g(x)$ is increasing, and it follows from (18) and (19) that the function $h(x, y)$ is increasing in x and decreasing in y. It follows that $c_{21}(3)/\hat{\mu}_{21}$ is increasing in $c_5(3)$ and decreasing in $c_5(1)$, provided $\hat{\mu}_{21} > 0$. Thus, to bound $f_{21}(\mathcal{A})$ from above for all pairs $(c_5(1), c_5(3))$ with $c_5(3) - 3c_5(1)/4 \leqslant 9.019$, it is enough to bound $c_{21}(3)/\hat{\mu}_{21}$ for a finite set of pairs $(c_5(1), c_5(3))$ that "dominate" the region $c_5(3) - 3c_5(1)/4 \leqslant 9.019$, and satisfy $\hat{\mu}_{21} > 0$.

To do this, note that $c_5(1) \geqslant 1$ and $c_5(3) - 1 \geqslant 3(c_5(1) - 1)$, by (11). We may therefore assume that $c_5(1) \leqslant 5$, since otherwise $c_5(3) - 3c_5(1)/4 \geqslant 9c_5(1)/4 - 2 > 9.019$. We therefore only need to cover the part of the region $c_5(3) - 3c_5(1)/4 \leqslant 9.019$ with $1 \leqslant c_5(1) \leqslant 5$. We do so by looping through values of $c_5(1)$ from 1 to 5 in steps of 10^{-4}, i.e., we check the point

$$\big(u(i), v(i)\big) := \left(\frac{i}{10^4},\ 9.019 + \frac{3(i+1)}{4 \cdot 10^4}\right)$$

in the $(c_5(1), c_5(3))$-plane for each $10^4 \leqslant i < 5 \cdot 10^4$. Note that $\big(u(i), v(i)\big)$ dominates the set

$$I(i) := \Big\{\big(c_5(1), c_5(3)\big) : c_5(3) - 3c_5(1)/4 \leqslant 9.019 \text{ and } 10^{-4}i \leqslant c_5(1) \leqslant 10^{-4}(i+1)\Big\},$$

in the sense that if $(x, y) \in I(i)$ then $x \geqslant u(i)$ and $y \leqslant v(i)$, so (by the monotonicity properties proved above) any upper bound on $c_{21}(3)/\hat{\mu}_{21}$ that holds at the point $(u(i), v(i))$ applies to all points of $I(i)$.

[7] A rough diagram of the pairs $(c_5(1), c_5(3))$ that were sufficient to prove the required bound on $f_{21}(\mathcal{A})$ was determined. Based on this, the linear combination $c_5(3) - 3c_5(1)/4$ appears to give the best "figure of merit" among simple linear combinations.

Hence, if an upper bound for $c_{21}(3)/\hat{\mu}_{21}$ holds for each pair $(u(i), v(i))$ in the range above, then it holds whenever $\mathbb{P}_5$ satisfies $c_5(3) - 3c_5(1)/4 \leqslant 9.019$.

Now for each $10^4 \leqslant i < 5 \cdot 10^4$ we choose the constants $\delta_6, \ldots, \delta_{21} \in (0, 1/2]$ so as to minimize the ratio $c_{21}(3)/\hat{\mu}_{21}$ after processing the prime 73. The optimization of the δ_k was made by first taking an initial choice $\delta_k = 1/4$. Then the δ_k were changed by performing coordinate-wise optimization: we minimized $c_{21}(3)/\hat{\mu}_{21}$ with respect to each δ_k in turn from $k = 6$ to 21, and then repeated this process a second time, after which it was seen that $c_{21}(3)/\hat{\mu}_{21}$ had converged adequately. We also checked that $\hat{\mu}_{21} > 0$ in each case.

The maximum value of $c_{21}(3)/\hat{\mu}_{21}$ obtained for any of these points was about 138.873682, and hence $f_{21} \leqslant 138.874$ for any $\mathbb{P}_5$ such that $c_5(3) - 3c_5(1)/4 \leqslant 9.019$, as required. Source code for these calculations can be found at [BBMST 2021]. □

5.3. ***Constructing the measure*** $\mathbb{P}_5$**.** It remains to construct a measure on the uncovered set $R_5 \subseteq S_2 \times \cdots \times S_5$ obtained after removing all hyperplanes corresponding to arithmetic progressions whose moduli involve only the primes 3, 5, 7 and 11. By Corollary 5.2 and Lemma 5.3, it will suffice to prove the following lemma.

Lemma 5.4. *For each $\mathcal{A}$ as above, there exists $\mathbb{P}_5$ such that $c_5(3) - 3c_5(1)/4 \leqslant 9.019$.*

Proof. Let us write $\mathcal{F}$ for the collection of sets $F \subseteq \{2, 3, 4, 5\}$ with $|F| \geqslant 2$. For each set $F \in \mathcal{F}$, we need to choose a hyperplane A_F with fixed set F, and for each such family we need to construct a measure $\mathbb{P}_5$ supported on the uncovered set. Not surprisingly, there are far too many configurations to deal with easily, so we need to make a few reductions.

Recall first that (by assumption) no hyperplane in $\mathcal{A}$ is contained in another. Also, we may assume there is a hyperplane for each F with $|F| \geqslant 2$ as including extra hyperplanes only makes covering easier. Moreover, we only need to study configurations 'up to isomorphism', in the following sense. First, let us write $F' \prec F$ if $\sum_{i \in F'} 2^i < \sum_{i \in F} 2^i$ (i.e., F' precedes F in colexicographic order), and for each $F \subseteq \{2, 3, 4, 5\}$ write $A_F = [a_{2,F}, a_{3,F}, a_{4,F}, a_{5,F}]$ with $a_{i,F} \in S_i \cup \{*\}$ so that $A_F \subseteq \{x_i = a_{i,F}\}$ when $i \in F$ and $a_{i,F} = *$ otherwise. Next, for a given F, order the choices of hyperplane A_F lexicographically, so that $A'_F \prec' A_F$ if $a'_{i,F} < a_{i,F}$ when $i = \min\{j \in F : a'_{j,F} \neq a_{j,F}\}$. Now, suppose there exists a pair (F, i) such that $i \in F$ and $a_{i,F} \geqslant a_{i,F'} + 2$ for every $F' \prec F$ with $i \in F'$. Then we can transpose $a_{i,F}$ and $a_{i,F} - 1$ in S_i to obtain an isomorphic configuration $\mathcal{A}'$ which is lexicographically smaller (with respect to the orders $\prec$ and $\prec'$), since $A'_F \prec' A_F$, but $A'_{F'} = A_{F'}$ for all $F' \prec F$.

Applying these reductions reduces the number of configurations to 6,025,640,717 which, while it represents substantial progress, is still too large to conveniently construct optimized probability distributions for each configuration. However, the main contribution to the large number of configurations comes from the choice of the "last" few hyperplanes, namely A_{45}, A_{245}, A_{345}, and A_{2345}.[8] For example, we might have as many as $2 \times 4 \times 6 \times 10 = 480$ choices for A_{2345}, as we are selecting a single point in Q_5 (although in practice the number of choices is reduced somewhat by the comments above). Ignoring the

[8] We have for brevity denoted, e.g., $A_{\{4,5\}}$ by A_{45}.

choices for A_{45}, A_{245}, A_{345}, and A_{2345} reduces the number of configurations to just 7637, which is far more manageable.

Our strategy is therefore as follows. We first consider a choice of the hyperplanes, A_{23}, A_{24}, A_{25}, A_{34}, A_{35}, A_{234} and A_{235}, without including the last four hyperplanes A_{45}, A_{245}, A_{345} and A_{2345}. In order to optimize the probability distribution on the uncovered region $R := Q_5 \setminus (A_{23} \cup \cdots \cup A_{235})$, we construct a linear programming problem with variables x_r for each $r \in R$ representing the probability of the atom $\{r\}$. For each nonempty set $I \subseteq \{2, 3, 4, 5\}$, and each hyperplane H in Q_5 with $F(H) = I$, we include the constraint

$$\sum_{r \in R \cap H} x_r \leqslant c_I,$$

where the c_I are new variables giving upper bounds on the $c(I)$; cf. (9). (Note that we include the constraints corresponding to sets not in $\mathcal{F}$, since we need to bound $c(I)$ for all subsets $I \subseteq \{2, 3, 4, 5\}$.) We also add the constraints

$$x_r \geqslant 0 \quad \text{and} \quad \sum_{r \in R} x_r = 1$$

to ensure that we have a probability measure supported on R, and then minimize

$$\sum_{I \subseteq \{2,3,4,5\}} \big(3^{|I|} - 3/4\big) c_I, \tag{20}$$

where we define $c_\varnothing = 1$. We define $\mathbb{P}_5$ to be the probability distribution corresponding to this minimum, i.e., we set $\mathbb{P}_5(r) = x_r$ for each $r \in R$. Recalling (11), note that

$$c_5(3) - \frac{3c_5(1)}{4} = \sum_{I \subseteq \{2,3,4,5\}} \big(3^{|I|} - 3/4\big) c(I),$$

and observe that the minimum occurs when $c_I = c(I)$.

We are therefore done, except for the (important) fact that we have not restricted the measure to be zero on the set

$$U := A_{45} \cup A_{245} \cup A_{345} \cup A_{2345}.$$

To do so, we simply remove the measure from the (unknown) set U, and uniformly rescale the measure to again give a probability measure. We claim that this can increase the value of $c_5(3) - 3c_5(1)/4$ to at most

$$\frac{c_5(3) - 3c_5(1)/4 - p/4}{1 - p}, \tag{21}$$

where p is the probability assigned to U. To see this, observe that removing the measure on U does not increase any $c(I)$, and decreases $c(\varnothing)$ by p. Since $c_5(3) - 3c_5(1)/4$ is a positive linear combination of the $c(I)$, with $c(\varnothing)$ occurring with coefficient $1/4$, it follows that $c_5(3) - 3c_5(1)/4$ decreases by at least $p/4$. Renormalizing the measure then increases each $c(I)$ (and hence this linear combination of the $c(I)$) by a factor of $1 - p$.

Configuration	$c_5(3)-3c_5(1)/4$	$p\leqslant$	Bound
11**, 2*1*, *22*, 121*, 1**1, *3*2, 13*3	8.772328	0.043227	9.157362

Table 1. One of the 90 partial configurations on Q_5 using just $A_{23},\dots,A_{235}$ with the bound given by (21) greater than 9.018. The full list is given in [BBMST 2021].

Configuration	$c_5(3)-3c_5(1)/4$
11**, 2*1*, *22*, 121*, 1**1, *3*2, 13*3, **34, 2*31, *232, 1233	9.018070
11**, 2*1*, *22*, 121*, 1**1, *3*2, 13*3, **34, 2*33, *232, 1233	9.018070

Table 2. Full configurations on Q_5 with $c_5(3)-3c_5(1)/4\geqslant 9.018$.

To complete the proof, we bound the probability p of the unspecified set U by

$$p\leqslant c_{45}+c_{245}+c_{345}+c_{2345},$$

where the c_I are the bounds on $c(I)$ given by the linear programming problem. We then check if the bound in (21) is less than 9.018.[9] If so, then we proceed to the next configuration. There are 90 (out of 7637) configurations of $(A_{23},\dots,A_{235})$ where this fails. For these, we loop through all choices of A_{45} and perform the above calculation with just A_{245}, A_{345}, and A_{2345} unspecified. From these 90 configurations we obtain 1083 configurations with A_{45} included, but for only 12 of these does our bound still exceed 9.018. These 12 give rise to 312 configurations including A_{245}, of which 3 still exceed our bound. These 3 give rise to 216 configurations where we are forced to include A_{345}, but only 2 which still exceed our bound. Finally, these 2 give 142 configurations where we are forced to include all the A_F, but only two have $c_5(3)-3c_5(1)/4\geqslant 9.018$, and these are listed in Table 2. We deduce that for all choices of the hyperplanes $\{A_F : F\in\mathcal{F}\}$ in Q_5 we can find a probability measure $\mathbb{P}_5$ on R_5 such that $c_5(3)-3c_5(1)/4<9.018071$.

All calculations were performed in `C` using the Gurobi linear optimization package [Gurobi] to solve the LP minimizations. With the strategy as described above, the calculations to determine the worst case configurations took about 40 seconds on a laptop. Source code and results can be found at [BBMST 2021]. □

As noted above, Theorem 1.2 follows immediately from Corollary 5.2 and Lemmas 5.3 and 5.4. In fact, since the results from [BBMST 2018] (which we used to deal with the large primes) did not require the assumption that the moduli are square-free, we actually proved something slightly stronger: if the moduli are distinct and each prime $p\leqslant 73$ in their prime factorization occurs to a power at most 1, then at least one of the moduli must be even. In other words, the "square-free" condition is only needed on the 73-smooth part of the moduli. Of course, if we could reduce "$p\leqslant 73$" to "$p<3$" then the Erdős–Selfridge problem would be solved, so it would be interesting to see to what extent the bound 73 could be reduced.

[9] This bound was chosen, after some experimentation, to be just below the worst case value of $c_5(3)-3c_5(1)/4$ given in Table 2 for configurations using all the hyperplanes.

Acknowledgement

We are grateful to the referees for their careful reading and helpful suggestions.

References

[BBMST 2018] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe, and M. Tiba, "On the Erdős covering problem: the density of the uncovered set", 2018. arXiv

[BBMST 2021] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe, and M. Tiba, "C Source code and results for "The Erdős–Selfridge problem with square-free moduli"", 2021, available at http://people.maths.ox.ac.uk/balister/Erdos-Selfridge.html.

[Erdős 1950] P. Erdős, "On integers of the form $2^k + p$ and some related problems", *Summa Brasil. Math.* **2** (1950), 113–123. MR

[Erdős 1965] P. Erdős, "Some recent advances and current problems in number theory", pp. 196–244 in *Lectures on Modern Mathematics, III*, edited by T. L. Saaty, Wiley, New York, 1965. MR

[Erdős 1973] P. Erdős, "Résultats et problèmes en théorie des nombres", pp. 1–7 Séminaire de Delange–Pisot–Poitou: Théorie des Nombres **14**, Secrétariat mathématique, Paris, 1973. MR Zbl

[Erdős 1977] P. Erdős, "Problems and results on combinatorial number theory, III", pp. 43–72 in *Number theory day* (New York, 1976), edited by M. Nathanson, Lecture Notes in Math **676**, Springer, Berlin, 1977. MR Zbl

[Erdős and Graham 1980] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathématique **28**, Université de Genève, L'Enseignement Mathématique, 1980. MR

[Filaseta, Ford and Konyagin 2000] M. Filaseta, K. Ford, and S. Konyagin, "On an irreducibility theorem of A. Schinzel associated with coverings of the integers", *Illinois J. Math.* **44**:3 (2000), 633–643. MR Zbl

[Filaseta, Ford, Konyagin, Pomerance and Yu 2007] M. Filaseta, K. Ford, S. Konyagin, C. Pomerance, and G. Yu, "Sieving by large integers and covering systems of congruences", *J. Amer. Math. Soc.* **20**:2 (2007), 495–517. MR Zbl

[Guo and Sun 2005] S. Guo and Z.-W. Sun, "On odd covering systems with distinct moduli", *Adv. in Appl. Math.* **35**:2 (2005), 182–187. MR Zbl

[Gurobi] L. Gurobi Optimization, "Gurobi optimizer software", available at http://www.gurobi.com.

[Hough 2015] B. Hough, "Solution of the minimum modulus problem for covering systems", *Ann. of Math.* (2) **181**:1 (2015), 361–382. MR Zbl

[Hough and Nielsen 2019] R. D. Hough and P. P. Nielsen, "Covering systems with restricted divisibility", *Duke Math. J.* **168**:17 (2019), 3261–3295. MR Zbl

[Schinzel 1967] A. Schinzel, "Reducibility of polynomials and covering systems of congruences", *Acta Arith.* **13** (1967), 91–101. MR Zbl

[Simpson and Zeilberger 1991] R. J. Simpson and D. Zeilberger, "Necessary conditions for distinct covering systems with square-free moduli", *Acta Arith.* **59**:1 (1991), 59–70. MR Zbl

Communicated by Andrew Granville
Received 2019-01-31 Revised 2020-08-14 Accepted 2020-09-18

Paul.Balister@maths.ox.ac.uk *University of Oxford, Oxford, United Kingdom*

bb12@cam.ac.uk *University of Cambridge, Cambridge, United Kingdom*

Current address: *University of Memphis, Memphis, TN, United States*

rob@impa.br *IMPA, Rio de Janeiro, Brazil*

jdrs2@cam.ac.uk *University of Cambridge, Cambridge, United Kingdom*

mt576@cam.ac.uk *University of Cambridge, Cambridge, United Kingdom*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2021 is US $415/year for the electronic version, and $620/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 15 No. 3 2021

1937-0652(2021)15:3;1-V