

ARTICLE

Extremal problems for GCDs

Ben Green¹ and Aled Walker^{2*}

¹Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter (550), Woodstock Road, Oxford, OX2 6GG, UK and ²Trinity College, Trinity Street, Cambridge, CB2 1TQ, UK

*Corresponding author. Email: aw530@cam.ac.uk

(Received 16 December 2020; revised 12 March 2021; first published online 8 April 2021)

Abstract

We prove that if $A \subseteq [X, 2X]$ and $B \subseteq [Y, 2Y]$ are sets of integers such that $\gcd(a, b) \geq D$ for at least $\delta|A||B|$ pairs $(a, b) \in A \times B$ then $|A||B| \ll_{\varepsilon} \delta^{-2-\varepsilon} XY/D^2$. This is a new result even when $\delta = 1$. The proof uses ideas of Koukoulopoulos and Maynard and some additional combinatorial arguments.

2020 MSC Codes: Primary 11A05; Secondary 11B75

1. Introduction and proof strategy

Fix $\varepsilon \in (0, 1)$ throughout the paper; implied constants and thresholds may depend on ε but are otherwise absolute. Let p_0 be a threshold which (at the start of Section 3) will be taken to be sufficiently large. Given a finite set $S \subset \mathbb{N}$, write $\mathcal{P}(S)$ for the set of primes dividing some element of S , and $\mathcal{P}_{\text{small}}(S)$ for the set of primes $p \leq p_0$ dividing some element of S .

Our main result is the following.

Theorem 1.1. *Let $X, Y, D \in [1, \infty)$, and suppose that $D \leq \min(X, Y)$. Let $\delta \in (0, 1]$. Suppose that $A \subset [X, 2X]$ and $B \subset [Y, 2Y]$ are sets of integers with the following property: for at least $\delta|A||B|$ pairs $(a, b) \in A \times B$, $\gcd(a, b) \geq D$. Then we have the bound*

$$|A||B| \leq (1000)^{1+\#\mathcal{P}_{\text{small}}(A \cup B)} \delta^{-2-\varepsilon} \frac{XY}{D^2}.$$

Let us make some remarks on this theorem.

1. This obviously implies the cruder bound $|A||B| \ll \delta^{-2-\varepsilon} XY/D^2$, mentioned in the abstract. The more precise form we have stated seems of little additional interest in its own right, but is critical for the proof. Perhaps the most natural case is when $A = B$ and $X = Y$, when the result says the following: if, for a proportion δ of all pairs $(a, a') \in A \times A$ we have $\gcd(a, a') \geq D$, then $|A| \ll \delta^{-1-\varepsilon} X/D$.

2. We believe that the result is new even when $\delta = 1$, that is to say when $\gcd(a, b) \geq D$ for all $a \in A$ and $b \in B$. In this case, the result is clearly sharp up to a multiplicative constant. Indeed,

The first-named author is supported by a Simons Investigator Award and is grateful to the Simons Foundation for their support. The second-named author is supported by a Postdoctoral Fellowship with the Centre de Recherches Mathématiques and by a Junior Research Fellowship from Trinity College Cambridge.

assuming that D is an integer, we may take $A = \{x \in [X, 2X] : D|x\}$ and $B = \{y \in [Y, 2Y] : D|y\}$. One might wonder whether all the tight examples have approximately such a structure; however, Chow has constructed a different family of tight examples for $\delta = 1$ (see [1, Section 15]) in which there is no single $d \gg D$ that divides a positive proportion of A and B .

3. For $\delta \in (0, 1)$ the result is also sharp for a wide range of parameters, apart from the factor of $\delta^{-\varepsilon}$. To see this, let $D \geq \delta^{-1}$ be given, set $D_0 := \lfloor \delta D \rfloor$, and consider the sets $A = B = \{x \in [X, 2X] : D_0|x\}$. Evidently, $|A||B| \sim \delta^{-2}X^2/D^2$. However, if $x = D_0m, x' = D_0m'$ with $m, m' \in [X/D_0, 2X/D_0]$ and $\gcd(m, m') \geq D/D_0$, then $x, x' \in A$ and $\gcd(x, x') \geq D$. The proportion of pairs of integers with $\gcd k$ is $1/k^2\zeta(2)$, and so the proportion of pairs of integers with $\gcd \geq k$ is $\gg 1/k$. It follows (at least if X/D_0 is big enough compared to D/D_0) that the number of such pairs (x, x') is $\gg \delta|A||B|$.

4. When $A = B, X = Y$ and $\delta = 1$, the result says the following: if $\gcd(a, a') \geq D$ for all $a, a' \in A$, then $|A| \ll X/D$. However (we are rather embarrassed to admit), Zachary Chase pointed out to the authors that this particular result is trivial, because the assumption implies that $|a - a'| \geq D$ whenever $a \neq a'$. This argument does not, however, appear to extend to the other cases.

5. A straightforward dyadic decomposition argument would allow one to establish similar results under the assumption that $A \subset [X]$ and $B \subset [Y]$. We leave the details to the reader.

Notation. Our notation is standard. If p is a prime and $a \in \mathbb{Z}$, we write $v_p(a)$ for the largest k such that $p^k|a$. We extend this to rationals by $v_p(a/b) = v_p(a) - v_p(b)$. Implied constants in the $O(), \ll$ and \gg notations are absolute (though they may depend on ε , which is fixed throughout the paper).

Strategy. Our strategy for proving Theorem 1.1 is essentially to proceed by induction on $\#\mathcal{P}(A \cup B)$, but we will phrase the argument in terms of a hypothetical counterexample with minimal $\#\mathcal{P}(A \cup B)$. The first main business is to show that such a minimal counterexample has a very specific structure.

Proposition 1.2. *Suppose we have a counterexample to Theorem 1.1 with the set $\mathcal{P}(A \cup B)$ minimal in size. Let $\Omega \subset A \times B$, with $|\Omega| = \delta|A||B|$, be the set of pairs for which $\gcd(a, b) \geq D$. Then there is $\Omega' \subset \Omega$, with $|\Omega'| \geq \frac{1}{2}|\Omega|$, and an integer N such that the following is true. For all primes p and for all $(a, b) \in \Omega'$ we have $|v_p(a/N)| + |v_p(b/N)| \leq 1$.*

Though such a statement does not appear explicitly in their work, this proposition should be considered essentially due to Koukoulopoulos and Maynard [1]. We will give a fairly short, self-contained proof. On some level, this is equivalent to the argument of [1], but we phrase things rather differently.

To complete the proof of Theorem 1.1, we prove the following counterpart to Proposition 1.2.

Proposition 1.3. *Suppose that $A \subset [X, 2X], B \subset [Y, 2Y]$, Ω, D, δ, N are as in Proposition 1.2. Then $|A||B| \leq 1000\delta^{-2}XY/D^2$.*

Evidently, this means that A, B do not in fact give a counterexample to Theorem 1.1. Combining Propositions 1.2 and 1.3 shows that no minimal counterexample to Theorem 1.1 exists, so Theorem 1.1 is true.

The proof of Proposition 1.3 uses some combinatorial arguments and is not found in [1].

2. Concentrated measures on \mathbb{Z}^2

In this section, we prove a result about concentration of probability measures on \mathbb{Z}^2 . It is the key technical ingredient in the proof of Proposition 1.2, where it is used to concentrate the pair of valuation functions $(v_p(a), v_p(b))$ around a diagonal pair (k, k) .

Here, as in the rest of the paper, we write $q = 2 + \varepsilon$ and write q' for the conjugate index to q (i.e. $\frac{1}{q} + \frac{1}{q'} = 1$).

Lemma 2.1. *Let $c \leq 1$ and $\lambda \leq \frac{4}{5}$. Suppose that μ is a finitely supported probability measure on \mathbb{Z}^2 . Suppose that there are sequences $x = (x_i)_{i \in \mathbb{Z}}, y = (y_j)_{j \in \mathbb{Z}}$ of non-negative reals such that $\|x\|_{\ell^{q'}(\mathbb{Z})} = \|y\|_{\ell^q(\mathbb{Z})} = 1$, and such that for all $(i, j) \in \mathbb{Z}^2$ we have*

$$\mu(i, j) \leq c\lambda^{|i-j|}x_iy_j. \tag{2.1}$$

Then $c \geq \frac{1}{9}$, and μ is highly concentrated near some point (k, k) :

$$\sum_{|i-k|+|j-k| \geq 2} \mu(i, j) \ll \lambda^{q+\varepsilon}. \tag{2.2}$$

Proof. We first prove the lower bound on c . Using (2.1), $\sum_{\mathbb{Z}^2} \mu(i, j) = 1$, $\sum_{m \in \mathbb{Z}} \lambda^{|m|} \leq 9$ and $q' < 2$, we have

$$\frac{1}{c} \leq \sum_{i,j} \lambda^{|i-j|}x_iy_j \leq \sum_{i,j} \lambda^{|i-j|}(x_iy_j)^{\frac{q'}{2}} \leq 9 \sup_m \sum_i (x_iy_{i+m})^{\frac{q'}{2}} \leq 9,$$

where the last step follows from the Cauchy–Schwarz inequality and the assumption $\|x\|_{\ell^{q'}(\mathbb{Z})} = \|y\|_{\ell^q(\mathbb{Z})} = 1$. The lower bound on c follows.

Turning to (2.2), write $\sup_i x_iy_i = 1 - \gamma$ for some $\gamma \in [0, 1]$, and suppose this supremum attained when $i = k$. Then, $x_k, y_k \geq 1 - \gamma$, so

$$\sum_{i \neq k} x_i^{q'}, \sum_{j \neq k} y_j^{q'} \ll \gamma \text{ and } x_i, y_j \ll \gamma^{1/q'} \text{ when } i, j \neq k. \tag{2.3}$$

For $n = 1, 2, 3, 4, 5, 6$ write $\Sigma_n := \sum_{(i,j) \in S_n} \mu(i, j)$, where S_1, \dots, S_6 are the following sets, which partition \mathbb{Z}^2 :

$$S_1 := \{(i, j) \in \mathbb{Z}^2 : i \neq j \neq k \neq i\}, \quad S_2 := \{(k, j) : |j - k| \geq 2\},$$

$$S_3 := \{(i, k) : |i - k| \geq 2\}, \quad S_4 := \{(k, k \pm 1), (k \pm 1, k)\}, \quad S_5 := \{(i, i) : i \neq k\},$$

and finally $S_6 := \{(k, k)\}$. We bound the Σ_n in turn.

Bound for Σ_1 . By (2.1) and (2.3), we have

$$\Sigma_1 \leq \sum_{\substack{i,j \neq k \\ i \neq j}} \lambda^{|i-j|}x_iy_j \ll \gamma^{\frac{2}{q'}(1-\frac{q'}{2})} \sum_{\substack{i,j \neq k \\ i \neq j}} \lambda^{|i-j|}(x_iy_j)^{\frac{q'}{2}} = \gamma^{\frac{2}{q'}-1} \sum_{m \neq 0} \lambda^{|m|} \sum_{i, i+m \neq k} (x_iy_{i+m})^{\frac{q'}{2}}.$$

By Cauchy–Schwarz and (2.3), for each fixed m we have

$$\sum_{i, i+m \neq k} (x_iy_{i+m})^{\frac{q'}{2}} \leq \left(\sum_{i \neq k} x_i^{q'} \right)^{1/2} \left(\sum_{j \neq k} y_j^{q'} \right)^{1/2} \ll \gamma.$$

Since $\sum_{m \neq 0} \lambda^{|m|} \ll \lambda$, putting these together gives $\Sigma_1 \ll \lambda\gamma^{\frac{2}{q'}}$.

Bounds for Σ_2, Σ_3 . For Σ_2 , we use the trivial bound $x_k \leq 1$ and (2.3) for y_j . This gives (using the assumption that $\lambda \leq \frac{4}{5}$)

$$\Sigma_2 \leq \sum_{|j-k| \geq 2} (5\lambda/4)^{|j-k|} (4/5)^{|j-k|} y_j \ll \lambda^2 \sum_{|j-k| \geq 2} (4/5)^{|j-k|} y_j \ll \gamma^{\frac{1}{q'}} (1 - \frac{q'}{2}) \lambda^2 \sum_{|j-k| \geq 2} (4/5)^{|j-k|} y_j^{\frac{q'}{2}}.$$

By Cauchy–Schwarz and (2.3),

$$\sum_{|j-k| \geq 2} (4/5)^{|j-k|} y_j^{\frac{q'}{2}} \ll \left(\sum_{j \neq k} (4/5)^{2|j-k|} \right)^{1/2} \left(\sum_{j \neq k} y_j^{q'} \right)^{1/2} \ll \gamma^{1/2}.$$

Combining these bounds gives $\Sigma_2 \ll \gamma^{\frac{1}{q'}} \lambda^2$, and an essentially identical argument yields $\Sigma_3 \ll \gamma^{\frac{1}{q'}} \lambda^2$.

Bound for Σ_4 . From (2.1), (2.3) we immediately get $\Sigma_4 \ll \gamma^{\frac{1}{q'}} \lambda$.

Bound for Σ_5 . A trivial modification to the argument used for Σ_1 (allowing $i = j$, which gives just a term with $m = 0$) shows that $\Sigma_5 \ll \gamma^{\frac{2}{q'}}$.

Bound for $\Sigma_5 + \Sigma_6$. By (2.1) and the fact that $\sup_i x_i y_i = 1 - \gamma$,

$$\Sigma_5 + \Sigma_6 \leq \sum_i x_i y_i \leq (1 - \gamma)^{1 - \frac{q'}{2}} \sum_i (x_i y_i)^{\frac{q'}{2}} \leq (1 - \gamma)^{1 - \frac{q'}{2}} \leq 1 - \left(1 - \frac{q'}{2}\right) \gamma,$$

where we used Cauchy–Schwarz yet again.

Putting all this together gives

$$1 = \sum_{n=1}^6 \Sigma_n \leq 1 - \left(1 - \frac{q'}{2}\right) \gamma + O\left(\lambda \gamma^{\frac{1}{q'}}\right).$$

This implies that $\lambda \gg \gamma^{1 - \frac{1}{q'}}$, i.e. $\gamma \ll \lambda^q$. Finally, we see that

$$\sum_{|i-k| + |j-k| \geq 2} \mu(i, j) = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_5 \ll \lambda^{\frac{2q}{q'}} + \lambda^{2 + \frac{q}{q'}} \ll \lambda^{\frac{2q}{q'}},$$

(since $q \leq 3$). The result follows, noting that $\frac{2q}{q'} = q + \varepsilon$. □

3. Properties of a minimal counterexample

We turn now to the proof of Proposition 1.2. We first reduce matters to the following “local” statement at a single prime p .

Proposition 3.1. *Suppose we have a counterexample to Theorem 1.1 with the set $\mathcal{P}(A, B)$ minimal in size. Let $\Omega \subset A \times B$, with $|\Omega| = \delta|A||B|$, be the set of pairs for which $\gcd(a, b) \geq D$. Let $p \in \mathcal{P}(A, B)$ be a prime. Then, $p > p_0(\varepsilon)$, and there is $k_p \in \mathbb{Z}_{\geq 0}$ and $\Omega_p \subset \Omega$ such that for all $(a, b) \in \Omega_p$ we have $|v_p(a) - k_p| + |v_p(b) - k_p| \leq 1$, and such that $|\Omega \setminus \Omega_p| \ll p^{-1 - \varepsilon/3} |\Omega|$.*

Proposition 1.2 follows quickly from this by taking $N = \prod_p p^{k_p}$ and $\Omega' := \bigcap_p \Omega_p$. We have

$$|\Omega'| \geq |\Omega| \left(1 - O \left(\sum_{p > p_0} p^{-1-\varepsilon/3} \right) \right) \geq \frac{1}{2} |\Omega|,$$

if p_0 is big enough (this is the point at which p_0 is constrained).

It remains, then, to establish Proposition 3.1. Fix, for the rest of this section, the prime p . For $i, j \in \mathbb{Z}_{\geq 0}$, we define $A_i := \{a \in A : v_p(a) = i\}$, $B_j := \{b \in B : v_p(b) = j\}$, and write $\alpha_i := \frac{|A_i|}{|A|}$ and $\beta_j := \frac{|B_j|}{|B|}$ for the relative densities of these sets. Write $\mu(i, j) := \frac{|\Omega \cap (A_i \times B_j)|}{|\Omega|}$, thus μ is a finitely supported probability measure on $\mathbb{Z}_{\geq 0}^2$.

For any i, j , consider the sets $\bar{A}_i := p^{-i} \cdot A_i$ and $\bar{B}_j := p^{-j} \cdot B_j$. These are sets of integers, coprime to p , with $\bar{A}_i \subset [\frac{X}{p^i}, \frac{2X}{p^i}]$, $\bar{B}_j \subset [\frac{Y}{p^j}, \frac{2Y}{p^j}]$ and $\text{gcd}(x, y) \geq \frac{D}{p^{\min(i,j)}}$ whenever $x = p^{-i}a$, $y = p^{-j}b$ with $(a, b) \in \Omega$.

By the minimality assumption, these sets cannot be a counterexample to Theorem 1.1, and therefore we have the inequality

$$|\bar{A}_i| |\bar{B}_j| \leq (1000)^{1+\#\mathcal{P}_{\text{small}}(\bar{A}_i, \bar{B}_j)} \left(\frac{\delta \mu(i, j)}{\alpha_i \beta_j} \right)^{-2-\varepsilon} \frac{\frac{X}{p^i} \frac{Y}{p^j}}{\left(\frac{D}{p^{\min(i,j)}} \right)^2}. \tag{3.1}$$

On the other hand,

$$|\bar{A}_i| |\bar{B}_j| = \alpha_i \beta_j |A| |B| \geq (1000)^{1+\#\mathcal{P}_{\text{small}}(A, B)} \alpha_i \beta_j \delta^{-2-\varepsilon} \frac{XY}{D^2}. \tag{3.2}$$

Note also that $\mathcal{P}(\bar{A}_i, \bar{B}_j) \subset \mathcal{P}(A, B) \setminus \{p\}$, and so

$$\#\mathcal{P}_{\text{small}}(\bar{A}_i, \bar{B}_j) \leq \#\mathcal{P}_{\text{small}}(A, B) - 1_{p \leq p_0}. \tag{3.3}$$

Comparing (3.1), (3.2), and (3.3) gives, for all i and j ,

$$\mu(i, j) \leq 10^{-1_{p \leq p_0}} (\alpha_i \beta_j)^{\frac{1+\varepsilon}{2+\varepsilon}} p^{-\frac{|i-j|}{2+\varepsilon}}, \tag{3.4}$$

since $\varepsilon < 1$. This puts us in the situation covered by Lemma 2.1, with (in that lemma)

$$q = 2 + \varepsilon, q' = \frac{2 + \varepsilon}{1 + \varepsilon}, \lambda = p^{-\frac{1}{q}}, c = \left(\frac{1}{10} \right)^{1_{p \leq p_0}}, x_i := \alpha_i^{1/q'}, y_j := \beta_j^{1/q'}.$$

The hypotheses of the lemma are satisfied, since $p^{-\frac{1}{q}} \leq 2^{-\frac{1}{3}} \leq 4/5$. The lemma implies, first of all, that $c > \frac{1}{10}$; this immediately tells us that $p > p_0$. We conclude that there is some k such that

$$\sum_{|i-k|+|j-k| \geq 2} \mu(i, j) \ll \lambda^{q+\varepsilon} \ll p^{-1-\varepsilon/3}. \tag{3.5}$$

This is precisely what is needed in Proposition 3.1, taking

$$\Omega_p = \bigcup_{|i-k|+|j-k| \leq 1} (\Omega \cap (A_i \times B_j)).$$

4. Finishing the argument

In this section, we complete the proof of Theorem 1.1 by establishing Proposition 1.3. That is, our task is as follows. Suppose that $A \subset [X, 2X]$, $B \subset [Y, 2Y]$, that $\Omega \subset A \times B$ has size $\frac{\delta}{2}|A||B|$, and that $\text{gcd}(a, b) \geq D$ whenever $(a, b) \in \Omega$. Suppose that there is some positive integer N such that

$$|v_p(a/N)| + |v_p(b/N)| \leq 1, \tag{4.1}$$

for all primes p and for all $(a, b) \in \Omega$. We are to show that, under these assumptions, we have the bound

$$|A||B| \leq \frac{1000 XY}{\delta^2 D^2}. \tag{4.2}$$

Let us begin the proof. In the course of the argument, it will be convenient to use a little of the language of graph theory. Thus, if $a \in A$, then we write $\deg(a) := \#\{b \in B : (a, b) \in \Omega\}$, and analogously for $b \in B$. Write $A' := \{a \in A : \deg(a) > 0\}$ and $B' := \{b \in B : \deg(b) > 0\}$.

If $a \in A'$ then, by (4.1), $v_p(a/N) \in \{-1, 0, 1\}$ for all primes p . We define the *defect* a_* to be the product of all primes for which $v_p(a/N) \neq 0$. Now we make the crucial observation that if $(a, b) \in \Omega$ then

$$a_* b_* = \frac{ab}{\gcd(a, b)^2}. \tag{4.3}$$

To prove this, we take p -adic valuations. It is easily seen that

$$v_p(ab / \gcd(a, b)^2) = |v_p(a) - v_p(b)| = |v_p(a/N) - v_p(b/N)|,$$

so we need only prove that

$$v_p(a_*) + v_p(b_*) = |v_p(a/N) - v_p(b/N)|, \tag{4.4}$$

whenever $(a, b) \in \Omega$. This follows immediately from (4.1), noting that $v_p(a_*) = 1$ if $v_p(a/N) = \pm 1$ and $v_p(a_*) = 0$ otherwise, and similarly for $v_p(b_*)$.

As a consequence of (4.3) and our assumptions, we see that

$$a_* b_* \leq \frac{4XY}{D^2}, \tag{4.5}$$

whenever $(a, b) \in \Omega$. This would allow us to conclude very quickly, were it not for the fact that the map $a \mapsto a_*$ need not be injective (see Section 5 for some further remarks on this point). Fortunately, we have the following substitute for injectivity.

Lemma 4.1. *Let $T \in \mathbb{R}_{>0}$. The number of $a \in A'$ for which $a_* \leq T$ is at most $2T$. Similarly, the number of $b \in B'$ for which $b_* \leq T$ is at most $2T$.*

Proof. If $a \in A'$, write a_+ for the product of all primes with $v_p(a/N) = 1$, and a_- for the product of all primes with $v_p(a/N) = -1$. Thus,

$$a_* = a_+ a_-. \tag{4.6}$$

Since $v_p(a/N) \in \{-1, 0, 1\}$, we have

$$\frac{a_+}{a_-} = \frac{a}{N}. \tag{4.7}$$

Since $A \subset [X, 2X]$, it follows by multiplying (4.6) and (4.7) that if $a_* \leq T$ then

$$a_+ \leq \left(\frac{aT}{N}\right)^{1/2} \leq \left(\frac{2XT}{N}\right)^{1/2}. \tag{4.8}$$

Similarly, dividing (4.6) by (4.7), we see that if $a_* \leq T$ then

$$a_- \leq \left(\frac{NT}{a}\right)^{1/2} \leq \left(\frac{NT}{X}\right)^{1/2}. \tag{4.9}$$

It follows from (4.8), (4.9) that the number of choices for the pair (a_+, a_-) is at most $2T$. However, if we know a_+, a_- and N then we can recover a uniquely, so the map $a \mapsto (a_+, a_-)$ is injective. The proof for B' is the same. \square

Now we finish the argument. By a standard averaging argument there is a set $\tilde{A} \subset A$ with $|\tilde{A}| \geq \delta|A|/4$ such that $\deg(a) \geq \delta|B|/4$ for all $a \in \tilde{A}$. Clearly, $\tilde{A} \subset A'$, so by Lemma 4.1 there is some $a \in \tilde{A}$ such that

$$a_* \geq \delta|A|/8. \tag{4.10}$$

Set $\tilde{B} := \{b \in B : (a, b) \in \Omega\}$. Thus, $|\tilde{B}| \geq \delta|B|/4$. Clearly $\tilde{B} \subset B'$, so by Lemma 4.1 there is some $b \in \tilde{B}$ such that

$$b_* \geq \delta|B|/8. \tag{4.11}$$

By construction we have $(a, b) \in \Omega$, so we have the upper bound (4.5).

Comparing (4.5), (4.10), (4.11) immediately yields (4.2).

5. Further results and remarks

Suppose that A and B are finite sets of square-free positive integers. In this instance, we may assume that the positive integer N that satisfies (4.1) is also square-free, and thus the map $a \mapsto a_*$ is injective, since knowing N and a_* determines a . This enables us to circumvent Lemma 4.1, and prove the following theorem.

Theorem 5.1. *Let $Q \in [1, \infty)$ and $\delta \in (0, 1]$. Suppose that A, B are finite sets of square-free positive integers with the following property: for at least $\delta|A||B|$ pairs $(a, b) \in A \times B$, $ab/\gcd(a, b)^2 \leq Q$. Then we have the bound*

$$|A||B| \leq (1000)^{1+\#\mathcal{P}_{\text{small}}(A \cup B)} \delta^{-2-\varepsilon} \frac{Q}{4}.$$

Of course this result implies the cruder bound $|A||B| \ll \delta^{-2-\varepsilon} Q$. It also implies Theorem 1.1, upon taking $Q = 4XY/D^2$.

Proof. The proof of Proposition 1.2 holds mutatis mutandis. Analysing the minimal counterexample as before, we conclude that $a_*b_* \leq Q$ (by analogy with (4.5)). Using graph theoretic language as before, there exists a set $\tilde{A} \subset A$ with $|\tilde{A}| \geq \delta|A|/4$ such that $\deg(a) \geq \delta|B|/4$ for all $a \in \tilde{A}$. Since $a \mapsto a_*$ is injective, there is some $a \in \tilde{A}$ for which $a_* \geq \delta|A|/4$. Letting $\tilde{B} := \{b \in B : (a, b) \in \Omega\}$, we have $|\tilde{B}| \geq \delta|B|/4$ and for all $b \in \tilde{B}$ we have $b_* \leq Q/a_*$. Therefore, since $b \mapsto b_*$ is injective, we have

$$\frac{\delta|B|}{4} \leq |\tilde{B}| \leq \max_{b_* : b \in \tilde{B}} b_* \leq \frac{Q}{a_*} \leq \frac{4Q}{\delta|A|}.$$

This rearranges to $|A||B| \leq 16\delta^{-2}Q$, which shows that the minimal counterexample is not in fact a counterexample, thus settling the theorem. \square

One might wonder whether the bound $|A||B| \ll \delta^{-2-\varepsilon} Q$ holds for general finite sets of integers A and B (not just for square-frees). However, there is a counterexample to this assertion, even with $\delta = 1$, given by

$$A = B = \left\{ \left(\prod_{p \leq X} p \right) \frac{m}{n} : mn \leq X, \mu^2(m) = \mu^2(n) = 1, \gcd(m, n) = 1 \right\}.$$

One may establish that for all $(a_1, a_2) \in A \times A$ one has the bound $a_1 a_2 / \gcd(a_1, a_2)^2 \leq X^2$. Yet $|A| \gg X \log X$.

By this, one notes that the use of dyadic ranges in the proof of Lemma 4.1 was critical.

Acknowledgements

We would like to thank Andrew Granville, Dimitris Koukoulopoulos, and James Maynard for interesting conversations related to this work.

Reference

- [1] Koukoulopoulos, D. and Maynard, J. (2020) On the Duffin-Schaeffer conjecture. *Ann. Math.* **192**(1) 251–307.