



COMPOSITIO MATHEMATICA

On the metric theory of approximations by reduced fractions: a quantitative Koukoulopoulos–Maynard theorem

Christoph Aistleitner, Bence Borda and Manuel Hauke

Compositio Math. **159** (2023), 207–231.

[doi:10.1112/S0010437X22007837](https://doi.org/10.1112/S0010437X22007837)



FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865





On the metric theory of approximations by reduced fractions: a quantitative Koukoulopoulos–Maynard theorem

Christoph Aistleitner, Bence Borda and Manuel Hauke

ABSTRACT

Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be given. The Duffin–Schaeffer conjecture, recently resolved by Koukoulopoulos and Maynard, asserts that for almost all reals α there are infinitely many coprime solutions (p, q) to the inequality $|\alpha - p/q| < \psi(q)/q$, provided that the series $\sum_{q=1}^{\infty} \varphi(q)\psi(q)/q$ is divergent. In the present paper, we establish a quantitative version of this result, by showing that for almost all α the number of coprime solutions (p, q) , subject to $q \leq Q$, is of asymptotic order $\sum_{q=1}^Q 2\varphi(q)\psi(q)/q$. The proof relies on the method of GCD graphs as invented by Koukoulopoulos and Maynard, together with a refined overlap estimate from sieve theory, and number-theoretic input on the ‘anatomy of integers’. The key phenomenon is that the system of approximation sets exhibits ‘asymptotic independence on average’ as the total mass of the set system increases.

1. Introduction and statement of results

A foundational result in Diophantine approximation is Dirichlet’s approximation theorem, which asserts that for every real number α there are infinitely many coprime solutions (p, q) to the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^2}. \quad (1)$$

It is well known that this result is optimal up to constant factors for numbers α whose partial quotients in the continued fraction representation are bounded (so-called badly approximable numbers). Metric number theory asks to what extent (1) can be improved for typical reals α , in the sense that the exceptional set has vanishing Lebesgue measure.

One of the fundamental results of metric Diophantine approximation is Khintchine’s theorem [Khi24]. Let $\psi(q)$ be a non-negative sequence, and suppose that $q\psi(q)$ is non-increasing. Then the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{\psi(q)}{q} \quad (2)$$

has infinitely many integer solutions (p, q) for almost all real numbers α , provided that the series $\sum_{q=1}^{\infty} \psi(q)$ diverges. In contrast, inequality (2) has only finitely many solutions for almost all

Received 28 March 2022, accepted in final form 2 November 2022, published online 3 February 2023.

2020 Mathematics Subject Classification 11J83 (primary), 11A05, 11J04, 11K60 (secondary).

Keywords: Diophantine approximation, metric number theory, Duffin–Schaeffer conjecture, Koukoulopoulos–Maynard theorem.

© 2023 The Author(s). The publishing rights in this article are licensed to Foundation Compositio Mathematica under an exclusive licence.

α if this series converges. Very roughly speaking, this says that for typical reals the Dirichlet approximation theorem can be improved by a factor of logarithmic order. By periodicity, it is sufficient to consider $\alpha \in [0, 1]$. It can easily be seen that Khintchine’s theorem addresses the question whether the set system

$$\bigcup_{p=0}^q \left(\frac{p}{q} - \frac{\psi(q)}{q}, \frac{p}{q} + \frac{\psi(q)}{q} \right) \cap [0, 1], \quad q = 1, 2, \dots,$$

contains a given real α for infinitely or only finitely many values of q . If we assume that $\psi(q) \leq 1/2$ (as we will throughout this paper, to avoid degenerate situations), then the measure of such a set is exactly $2\psi(q)$. Thus, the ‘only finitely many’ part of Khintchine’s theorem is a straightforward application of the convergence part of the Borel–Cantelli lemma. The ‘infinitely many’ part of the theorem, however, is much more delicate since the divergence part of the Borel–Cantelli lemma requires some form of stochastic independence. The purpose of the monotonicity condition in the statement of Khintchine’s theorem is to guarantee this stochastic independence property of the set system.

Duffin and Schaeffer [DS41] showed that Khintchine’s theorem generally fails without the monotonicity condition. More precisely, they constructed a function ψ which is supported on a set of very smooth integers (having a large number of small prime factors), such that $\sum_{q=1}^{\infty} \psi(q)$ diverges, but for almost all α there are only finitely many solutions to (2). From a probabilistic perspective, the counterexample of Duffin and Schaeffer exploits the lack of stochastic independence in the set system, by constructing a special configuration where the overlaps between different sets of the system are too large; the crucial point here is that a fraction p/q can have many different representations as a quotient of integers (as long as non-reduced representations are allowed), and thus may appear in many different elements of the set system.

Duffin and Schaeffer suggested that this lack of independence could be overcome by switching to the coprime setting. More precisely, the Duffin–Schaeffer conjecture asserted that for almost all α there are infinitely many *coprime* solutions (p, q) to (2) if and only if the series $\sum_{q=1}^{\infty} \varphi(q)\psi(q)/q$ diverges, where φ denotes the Euler totient function. Let

$$\mathcal{A}_q := \bigcup_{\substack{0 \leq p \leq q, \\ \gcd(p,q)=1}} \left(\frac{p}{q} - \frac{\psi(q)}{q}, \frac{p}{q} + \frac{\psi(q)}{q} \right) \cap [0, 1], \quad q = 1, 2, \dots \tag{3}$$

Then, writing λ for the Lebesgue measure and again assuming that $\psi(q) \leq 1/2$ for all q , we have

$$\lambda(\mathcal{A}_q) = \frac{2\varphi(q)\psi(q)}{q}.$$

Thus, the ‘only finitely many’ part of the Duffin–Schaeffer conjecture is again a direct consequence of the convergence part of the Borel–Cantelli lemma. However, the divergence part of the Duffin–Schaeffer conjecture has resisted resolution for many decades. After important contributions of Gallagher [Gal61], Erdős [Erd70], Vaaler [Vaa78], Pollington and Vaughan [PV90], and Beresnevich and Velani [BV06], the Duffin–Schaeffer conjecture was finally solved in full generality by Koukoulopoulos and Maynard [KM20] in 2020. Their argument relies on an ingenious construction of what they call ‘GCD graphs’. This allows them to implement a step-by-step quality increment strategy until they finally arrive at a situation where they can completely control the divisor structure which is at the heart of the problem. The final, number-theoretic, input is an ‘anatomy of integers’ statement that quantifies the observation that there are only few integers that have many small prime factors.

In the present paper, we prove a quantitative version of the Koukoulopoulos–Maynard theorem. Their result states that there are infinitely many coprime solutions to (2) for almost all α if the sum of measures diverges. We show that for almost all α the number of solutions in fact grows proportionally to the sum of measures.

THEOREM 1. *Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be a function such that $\sum_{q=1}^{\infty} \varphi(q)\psi(q)/q = \infty$. Write $S(Q) = S(Q, \alpha)$ for the number of coprime solutions (p, q) to the inequality*

$$\left| \alpha - \frac{p}{q} \right| < \frac{\psi(q)}{q}, \quad \text{subject to } q \leq Q,$$

and let

$$\Psi(Q) = \sum_{q=1}^Q \frac{2\varphi(q)\psi(q)}{q}. \tag{4}$$

Let $C > 0$ be arbitrary. Then, for almost all α ,

$$S(Q) = \Psi(Q) \left(1 + O\left(\frac{1}{(\log \Psi(Q))^C} \right) \right) \quad \text{as } Q \rightarrow \infty.$$

It is not clear to what extent the error term in the theorem can be improved. It seems to us that any result which contains a power saving, that is, has a multiplicative error of order $(1 + O(\Psi(Q)^{-\varepsilon}))$ for some $\varepsilon > 0$, would require a substantial improvement of the argument in the present paper. By analogy with other results from metric number theory it is reasonable to assume that Theorem 1 actually holds with an error term $(1 + O(\Psi(Q)^{-1/2+\varepsilon}))$ for any $\varepsilon > 0$, and probably even $(1 + O(\Psi(Q)^{-1/2}(\log \Psi(Q))^c))$ for some appropriate c . We note in passing that very precise metric estimates for the asymptotic order of $S(Q)$ are known when an extra monotonicity assumption is imposed upon ψ , in the spirit of Khintchine’s original result; see, for example, Chapter 3 of [Phi71] and Chapter 4 of [Har98]. However, from a technical perspective, the problem is of a very different nature when this extra monotonicity assumption is made. The results for the monotonic case imply as a corollary that Theorem 1 above cannot hold in general with a multiplicative error of order $(1 + O(\Psi(Q)^{-1/2}))$ or less.

The key problem in the metric theory of approximations by reduced fractions is to control the measure of the overlaps $\mathcal{A}_q \cap \mathcal{A}_r$ in some averaged sense. Pairwise independence $\lambda(\mathcal{A}_q \cap \mathcal{A}_r) = \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)$ would allow a direct application of the second Borel–Cantelli lemma, but it turns out that $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ can exceed $\lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)$ by a factor as large as $\log \log(qr)$ for some configurations of q, r, ψ . Such an exceedingly large overlap can happen if there are many small prime factors dividing q but not dividing r , or vice versa, and if simultaneously the greatest common divisor of q and r lies in a certain critical range (which is determined by the values of $\psi(q)$ and $\psi(r)$). The crucial point then is to show that such large extra factors appear only for a small number of pairs q, r . Consider the quotient

$$\frac{\sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{(\sum_{q=1}^Q \lambda(\mathcal{A}_q))^2} = \frac{\sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{\Psi(Q)^2}.$$

Without imposing an absolute lower bound on $\Psi(Q)$, this quotient can be arbitrarily large. The main breakthrough of Koukoulopoulos and Maynard was to prove that

$$\frac{\sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{\Psi(Q)^2} \ll 1, \quad \text{provided that } \Psi(Q) \geq 1.$$

This property is called *quasi-independence on average*, and is sufficient for an application of the second Borel–Cantelli lemma (in the Erdős–Rényi formulation of the lemma); cf. [BV23]. In the

present paper, we show that even more is true: we have

$$\frac{\sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{\Psi(Q)^2} \rightarrow 1 \quad \text{as } \Psi(Q) \rightarrow \infty.$$

Thus, the set system $(\mathcal{A}_q)_{q \geq 1}$ moves towards pairwise independence on average as the total mass of the set system (the sum of measures of the approximation sets) tends towards infinity. Since we consider this fact, which is the key ingredient in our proof of Theorem 1, to be very interesting in its own right, we state it below as a separate theorem.

THEOREM 2. *Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be a function. Let the sets \mathcal{A}_q , $q = 1, 2, \dots$, be defined as in (3), and let $\Psi(Q)$ be defined as in (4). Let $C > 0$ be arbitrary. For any $Q \in \mathbb{N}$ such that $\Psi(Q) \geq 3$, we have*

$$\sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) - \Psi(Q)^2 = O\left(\frac{\Psi(Q)^2}{(\log \Psi(Q))^C}\right)$$

with an implied constant depending only on C .

The rest of this paper is organized as follows. In §2 we show how Theorem 2 implies Theorem 1. The following seven sections are concerned with the proof of Theorem 2. Section 3 contains an estimate of the measure of the overlap $\mathcal{A}_q \cap \mathcal{A}_r$ for given q and r . This estimate exploits information on the divisor structure of q and r in order to bound the difference between $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ and $\lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)$, thus addressing the issue of the ‘stochastic dependence’ between \mathcal{A}_q and \mathcal{A}_r . In §4 we reduce Theorem 2 to two second-moment bounds. Section 5 contains a brief introduction to the ‘GCD graph’ machinery developed by Koukoulopoulos and Maynard [KM20]. In §6 we show how the second-moment bounds follow from the existence of a ‘good’ GCD subgraph. In the final two sections, we establish the existence of such a good GCD subgraph, using a modification of the iteration procedure of [KM20]. Our argument requires a careful balancing of the ‘quality gain’ against the potential ‘density loss’ coming from this iterative procedure, in such a way that information on the ‘anatomy of integers’ can be exploited beyond a certain threshold. This threshold is determined by the order of the error terms coming from sieve theory (which translate into the error terms of the overlap estimate in §3).

For the rest of the paper, $\psi : \mathbb{N} \rightarrow [0, 1/2]$ is an arbitrary function, \mathcal{A}_q , $q \in \mathbb{N}$, is as in (3), and $\Psi(Q)$, $Q \in \mathbb{N}$, is as in (4).

2. Proof of Theorem 1

Let $C > 4$ be fixed, and assume that Theorem 2 holds. Let $\mathbb{1}_A$ denote the indicator function of a set A . Formulated in probabilistic language, Theorem 2 controls the variance of the random variables $\mathbb{1}_{\mathcal{A}_1}, \dots, \mathbb{1}_{\mathcal{A}_Q}$, and we obtain

$$\int_0^1 \left(\sum_{q=1}^Q \mathbb{1}_{\mathcal{A}_q}(\alpha) - \Psi(Q) \right)^2 d\alpha = \sum_{q,r \leq Q} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) - \Psi(Q)^2 = O\left(\frac{\Psi(Q)^2}{(\log \Psi(Q))^C}\right). \quad (5)$$

Define

$$Q_k = \min \{ Q : \Psi(Q) \geq e^{k^{1/\sqrt{C}}} \}, \quad k \geq 1,$$

and let

$$\mathcal{B}_k = \left\{ \alpha \in [0, 1] : \left| \sum_{q=1}^{Q_k} \mathbb{1}_{\mathcal{A}_q}(\alpha) - \Psi(Q_k) \right| \geq \frac{\Psi(Q_k)}{(\log \Psi(Q_k))^{C/4}} \right\}.$$

By Chebyshev’s inequality and (5), we have

$$\lambda(\mathcal{B}_k) \ll (\log \Psi(Q_k))^{-C/2} \ll k^{-\sqrt{C}/2}.$$

Since we assumed that $C > 4$, we have $\sum_{k=1}^\infty \lambda(\mathcal{B}_k) < \infty$, and the Borel–Cantelli lemma implies that almost all α are contained in at most finitely many sets \mathcal{B}_k . Thus, for almost all α ,

$$\left| \sum_{q=1}^{Q_k} \mathbb{1}_{\mathcal{A}_q}(\alpha) - \Psi(Q_k) \right| \leq \frac{\Psi(Q_k)}{(\log \Psi(Q_k))^{C/4}}$$

holds for all $k \geq k_0(\alpha)$. Clearly, for any $Q \geq 3$ there exists a k such that $Q_k \leq Q < Q_{k+1}$, which also implies that

$$\sum_{q=1}^{Q_k} \mathbb{1}_{\mathcal{A}_q}(\alpha) \leq \sum_{q=1}^Q \mathbb{1}_{\mathcal{A}_q}(\alpha) \leq \sum_{q=1}^{Q_{k+1}} \mathbb{1}_{\mathcal{A}_q}(\alpha).$$

Since $\psi \leq 1/2$ by assumption, we have $\Psi(Q_k) \in [e^{k^{1/\sqrt{C}}}, e^{k^{1/\sqrt{C}}} + 1/2]$, and so

$$\Psi(Q_{k+1})/\Psi(Q_k) = 1 + O(k^{-1+1/\sqrt{C}}) = 1 + O\left((\log \Psi(Q_k))^{-\sqrt{C}+1}\right).$$

From the previous three formulas and the triangle inequality, we deduce that for almost all α there exists a $Q_0 = Q_0(\alpha)$ such that for all $Q \geq Q_0$,

$$\left| \sum_{q=1}^Q \mathbb{1}_{\mathcal{A}_q}(\alpha) - \Psi(Q) \right| = O\left(\frac{\Psi(Q)}{(\log \Psi(Q))^{\sqrt{C}-1}}\right).$$

As C can be chosen arbitrarily large, this proves Theorem 1.

3. The overlap estimate

In this section we develop a new estimate for the measure of the overlaps $\mathcal{A}_q \cap \mathcal{A}_r$. For the rest of the paper, let

$$D(q, r) := \frac{\max(r\psi(q), q\psi(r))}{\gcd(q, r)}, \quad q, r \in \mathbb{N}. \tag{6}$$

The standard bound for the measure of $\mathcal{A}_q \cap \mathcal{A}_r$ is due to Pollington and Vaughan [PV90]: for any $q \neq r$,

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) \prod_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p > D(q,r)}} \left(1 + \frac{1}{p}\right), \tag{7}$$

with an absolute implied constant. Clearly, because of the presence of the implied constant this standard bound cannot be sufficient to deduce Theorem 2. Below we will use a more refined argument from sieve theory which allows us to isolate a main term, and prove an upper bound of the form

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)(1 + [\text{error}]),$$

with an error term that becomes small if there are not too many small primes which divide q and r with different multiplicities (see Lemma 5 below for details).

The following lemma is called the fundamental lemma of sieve theory. We state it in the formulation of [Kou19, Theorem 18.11].

LEMMA 3 (Fundamental lemma of sieve theory). *Let $(a_n)_{n \geq 1}$ be non-negative reals, such that $\sum_{n=1}^\infty a_n < \infty$. Let \mathcal{P} be a finite set of primes, and write $P = \prod_{p \in \mathcal{P}} p$. Set $y = \max \mathcal{P}$ and $A_d = \sum_{n \equiv 0 \pmod d} a_n$. Assume that there exist a multiplicative function g such that $0 \leq g(p) < p$ for all $p \in \mathcal{P}$, a real number x , and positive constants κ, C such that*

$$A_d =: x \frac{g(d)}{d} + r_d, \quad d \mid P,$$

and

$$\prod_{p \in (y_1, y_2] \cap \mathcal{P}} \left(1 - \frac{g(p)}{p}\right)^{-1} < \left(\frac{\log y_2}{\log y_1}\right)^\kappa \left(1 + \frac{C}{\log y_1}\right), \quad 3/2 \leq y_1 \leq y_2 \leq y.$$

Then, uniformly in $u \geq 1$, we have

$$\sum_{(n, P)=1} a_n = (1 + O(u^{-u/2}))x \prod_{p \in \mathcal{P}} \left(1 - \frac{g(p)}{p}\right) + O\left(\sum_{d \leq y^u, d \mid P} |r_d|\right).$$

We will also need an estimate for the order of the partial sums of a particular multiplicative function.

LEMMA 4. *Let \mathcal{P} be a set of odd primes, and define*

$$f(n) = \prod_{\substack{p \mid n, \\ p \in \mathcal{P}}} \left(1 + \frac{1}{p-2}\right).$$

Then, for any $x \geq 2$,

$$\sum_{n \leq x} f(n) = x \prod_{p \in \mathcal{P}} \left(1 + \frac{1}{p(p-2)}\right) + O(\log x),$$

where the implied constant is absolute.

Proof. Define $g(n) = \sum_{d \mid n} \mu(d) f(n/d)$, where μ is the Möbius function. Note that f and g are multiplicative functions. For $p \in \mathcal{P}$,

$$g(p) = f(p) - 1 = \frac{1}{p-2} \quad \text{and} \quad g(p^m) = 0, \quad m \geq 2,$$

whereas for $p \notin \mathcal{P}$, we have $g(p^m) = 0$ for all $m \geq 1$. By the definition of g ,

$$\begin{aligned} \sum_{n \leq x} f(n) &= \sum_{n \leq x} \sum_{d \mid n} g(d) \\ &= \sum_{d \leq x} g(d) \left\lfloor \frac{x}{d} \right\rfloor \\ &= x \sum_{d \leq x} \frac{g(d)}{d} + O\left(\sum_{d \leq x} g(d)\right) \\ &= x \sum_{d=1}^\infty \frac{g(d)}{d} + O\left(x \sum_{d > x} \frac{g(d)}{d} + \sum_{d \leq x} g(d)\right). \end{aligned} \tag{8}$$

Here

$$\sum_{d=1}^\infty \frac{g(d)}{d} = \prod_{p \in \mathcal{P}} \left(1 + \frac{1}{p(p-2)}\right),$$

and it remains to estimate the error term in (8).

Note that $p^m g(p^m) \leq p/(p-2) \leq 3$ for all prime powers p^m . Hence, by a general upper bound for the order of partial sums of multiplicative functions (see, for example, [Kou19, Theorem 14.2]), the partial sums of $dg(d)$ satisfy

$$\sum_{d \leq x} dg(d) \ll x \exp\left(\sum_{p \leq x} \frac{pg(p) - 1}{p}\right) \ll x \exp\left(\sum_{p > 2} \frac{2}{p(p-2)}\right) \ll x.$$

In particular, $\sum_{x \leq d \leq 2x} g(d)/d \ll x^{-2} \sum_{d \leq 2x} dg(d) \ll x^{-1}$, and the first error term in (8) is $x \sum_{d > x} g(d)/d \ll 1$. Further, $\sum_{x \leq d \leq 2x} g(d) \leq x^{-1} \sum_{d \leq 2x} dg(d) \ll 1$, and the second error term in (8) is $\sum_{d \leq x} g(d) \ll \log x$, as claimed. All implied constants are absolute. \square

LEMMA 5 (Overlap estimate). *Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be a function and \mathcal{A}_q , $q = 1, 2, \dots$, be defined as in (3). For any positive integers $q \neq r$ and any reals $u \geq 1$ and $T \geq 2$, we have*

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) \left(1 + O\left(u^{-u/2} + \frac{T^u \log(D+2) \log T}{D}\right)\right) \prod_{\substack{p | \frac{qr}{\gcd(q,r)^2}, \\ p > T}} \left(1 + \frac{1}{p-1}\right) \quad (9)$$

with an absolute implied constant, where $D = D(q, r)$ is as in (6). In particular, for any $C \geq 1$,

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) (1 + O((\log(D+2))^{-C})) \prod_{\substack{p | \frac{qr}{\gcd(q,r)^2}, \\ p > A}} \left(1 + \frac{1}{p-1}\right)$$

with an implied constant depending only on C , where

$$A = A_C(q, r) := \exp\left(\frac{\log(D+100) \log \log \log(D+100)}{8C \log \log(D+100)} + 1\right). \quad (10)$$

Proof. We follow the general strategy of Pollington and Vaughan in [PV90, § 3]. If $D < 1/2$, then $\psi(q)/q + \psi(r)/r < 1/\text{lcm}(q, r)$, hence $\mathcal{A}_q \cap \mathcal{A}_r = \emptyset$, and the claim trivially holds. We may thus assume throughout the rest of the proof that $D \geq 1/2$.

We set

$$\delta = \min\left(\frac{\psi(q)}{q}, \frac{\psi(r)}{r}\right) \quad \text{and} \quad \Delta = \max\left(\frac{\psi(q)}{q}, \frac{\psi(r)}{r}\right),$$

and define the piecewise linear function

$$w(y) = \begin{cases} 2\delta & \text{if } 0 \leq y \leq \Delta - \delta, \\ \Delta + \delta - y & \text{if } \Delta - \delta < y \leq \Delta + \delta, \\ 0 & \text{otherwise.} \end{cases}$$

We can express the measure of $\mathcal{A}_q \cap \mathcal{A}_r$ as

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) = \sum_{\substack{0 \leq a \leq q, \\ \gcd(a,q)=1}} \sum_{\substack{0 \leq b \leq r, \\ \gcd(b,r)=1}} w\left(\left|\frac{a}{q} - \frac{b}{r}\right|\right).$$

For any prime p , let $u = u(p, q)$ and $v = v(p, r)$ be defined by $q = \prod_p p^u$ and $r = \prod_p p^v$, and let

$$l = \prod_{p: u=v} p^u, \quad m = \prod_{p: u \neq v} p^{\min(u,v)}, \quad n = \prod_{p: u \neq v} p^{\max(u,v)}.$$

If $q, r \geq 2$, we have

$$\sum_{\substack{0 \leq a \leq q, \\ \gcd(a,q)=1}} \sum_{\substack{0 \leq b \leq r, \\ \gcd(b,r)=1}} w\left(\left|\frac{a}{q} - \frac{b}{r}\right|\right) = \sum_{\substack{1 \leq a \leq q, \\ \gcd(a,q)=1}} \sum_{\substack{1 \leq b \leq r, \\ \gcd(b,r)=1}} w\left(\left\|\frac{a}{q} - \frac{b}{r}\right\|\right),$$

which follows from the assumption that $\psi(q), \psi(r) \leq \frac{1}{2}$.

Following the argument on p. 195 of [PV90] (an application of the Chinese remainder theorem, together with a simple counting argument) thus leads for $q, r \geq 2$ to

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) = \sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n)=1}} 2w\left(\frac{c}{\ln}\right) \varphi(m) l \prod_{p|\gcd(l,c)} \left(1 - \frac{1}{p}\right) \prod_{\substack{p|l, \\ p \nmid c}} \left(1 - \frac{2}{p}\right),$$

and the formula above follows immediately also in the case $q = 1$ or $r = 1$.

Assume first that l is odd. By rewriting the right-hand side of the previous formula we see that $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ equals

$$\begin{aligned} & 2\varphi(m) \frac{\varphi(l)^2}{l} \sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n)=1}} w\left(\frac{c}{\ln}\right) \prod_{p|\gcd(l,c)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{p|l, \\ p \nmid c}} \left(\left(1 - \frac{2}{p}\right) \left(1 - \frac{1}{p}\right)^{-2}\right) \\ & = 2\varphi(m) \frac{\varphi(l)^2}{l} \prod_{p|l} \left(1 - \frac{1}{(p-1)^2}\right) \sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n)=1}} w\left(\frac{c}{\ln}\right) \prod_{p|\gcd(l,c)} \left(1 + \frac{1}{p-2}\right). \end{aligned}$$

We now find an upper bound for this expression. First, we replace the condition $\gcd(c, n) = 1$ by $\gcd(c, n^*) = 1$, where n^* denotes the T -smooth part of n (i.e. $n^* = \prod_{p \leq T, u \neq v} p^{\max(u,v)}$). Next, we fix a large positive integer K , and divide $[\Delta - \delta, \Delta + \delta]$ into K subintervals of equal length. Observe that the piecewise constant function

$$w^*(y) = \frac{2\delta}{K} \left(\left\lfloor \frac{K(\Delta + \delta - y)}{2\delta} \right\rfloor + 1 \right) = \frac{2\delta}{K} \sum_{k=0}^{K-1} \mathbb{1}_{[0, \Delta + \delta - 2k\delta/K]}(y)$$

satisfies $w(y) \leq w^*(y)$ for all $y \geq 0$. Therefore $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ is bounded above by

$$2\varphi(m) \frac{\varphi(l)^2}{l} \prod_{p|l} \left(1 - \frac{1}{(p-1)^2}\right) \sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n^*)=1}} w^*\left(\frac{c}{\ln}\right) \prod_{p|\gcd(l,c)} \left(1 + \frac{1}{p-2}\right). \tag{11}$$

Here

$$\sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n^*)=1}} w^*\left(\frac{c}{\ln}\right) \prod_{p|\gcd(l,c)} \left(1 + \frac{1}{p-2}\right) = \frac{2\delta}{K} \sum_{k=0}^{K-1} \sum_{\substack{1 \leq c \leq \ln(\Delta + \delta - 2k\delta/K), \\ \gcd(c,n^*)=1}} \prod_{p|\gcd(l,c)} \left(1 + \frac{1}{p-2}\right).$$

Now fix $k \in \{0, \dots, K-1\}$, and set

$$a_c = \prod_{p|\gcd(l,c)} \left(1 + \frac{1}{p-2}\right), \quad 1 \leq c \leq \ln(\Delta + \delta - 2k\delta/K),$$

and $a_c = 0$ for $c > \ln(\Delta + \delta - 2k\delta/K)$. Note that for $d \mid n^*$ we have $a_{dc} = a_c$ as long as $dc \leq \ln(\Delta + \delta - 2k\delta/K)$. By Lemma 4, for any $d \mid n^*$ we thus have

$$\begin{aligned} \sum_{c \equiv 0 \pmod d} a_c &= \sum_{1 \leq c \leq \ln(\Delta + \delta - 2k\delta/K)/d} \prod_{p \mid \gcd(l,c)} \left(1 + \frac{1}{p-2}\right) \\ &= \frac{\ln(\Delta + \delta - 2k\delta/K)}{d} \prod_{p \mid l} \left(1 + \frac{1}{p(p-2)}\right) + O(\log(D+2)). \end{aligned}$$

We have

$$\sum_{\gcd(c,n^*)=1} a_c = \sum_{\substack{1 \leq c \leq \ln(\Delta + \delta - 2k\delta/K), \\ \gcd(c,n^*)=1}} \prod_{p \mid \gcd(l,c)} \left(1 + \frac{1}{p-2}\right),$$

and by an application of Lemma 3 (with \mathcal{P} the set of prime divisors of n^* , $\max \mathcal{P} \leq T$ and $|r_d| \ll \log(D+2)$) this is

$$(1 + O(u^{-u/2})) \ln \left(\Delta + \delta - \frac{2k\delta}{K} \right) \frac{\varphi(n^*)}{n^*} \prod_{p \mid l} \left(1 + \frac{1}{p(p-2)}\right) + O(T^u \log(D+2)).$$

Since

$$\prod_{p \mid l} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p \mid l} \left(1 + \frac{1}{p(p-2)}\right) = 1,$$

formula (11) thus yields that $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ is bounded above by

$$2\varphi(m) \frac{\varphi(l)^2}{l} \cdot \frac{2\delta}{K} \sum_{k=0}^{K-1} \left(\left((1 + O(u^{-u/2})) \ln \left(\Delta + \delta - \frac{2k\delta}{K} \right) \frac{\varphi(n^*)}{n^*} + O(T^u \log(D+2)) \right) \right).$$

Letting $K \rightarrow \infty$, and using $D = \Delta l n$ and $\varphi(n^*)/n^* \geq \prod_{p \leq T} (1 - 1/p) \gg 1/\log T$, we obtain

$$\begin{aligned} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) &\leq 2\varphi(m) \frac{\varphi(l)^2}{l} 2\delta \left((1 + O(u^{-u/2})) \ln \Delta \frac{\varphi(n^*)}{n^*} + O(T^u \log(D+2)) \right) \\ &= 4\varphi(m) \varphi(l)^2 n \frac{\varphi(n^*)}{n^*} \delta \Delta \left(1 + O \left(u^{-u/2} + \frac{T^u \log(D+2) \log T}{D} \right) \right) \\ &= \lambda(\mathcal{A}_q) \lambda(\mathcal{A}_r) \frac{\varphi(n^*)/n^*}{\varphi(n)/n} \left(1 + O \left(u^{-u/2} + \frac{T^u \log(D+2) \log T}{D} \right) \right). \end{aligned}$$

Finally, observe that

$$\frac{\varphi(n^*)/n^*}{\varphi(n)/n} = \frac{1}{\prod_{\substack{p \mid n, \\ p > T}} (1 - \frac{1}{p})} = \prod_{\substack{p \mid n, \\ p > T}} \left(1 + \frac{1}{p-1}\right).$$

This establishes (9) for odd l .

Assume next that l is even. Then

$$\prod_{p \mid \gcd(l,c)} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \mid l, \\ p \nmid c}} \left(1 - \frac{2}{p}\right) = \frac{1}{2} \mathbb{1}_{\{2 \mid c\}} \prod_{\substack{p \mid \gcd(l,c), \\ p > 2}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \mid l, \\ p \nmid c, \\ p > 2}} \left(1 - \frac{2}{p}\right),$$

and similarly to before we obtain that $\lambda(\mathcal{A}_q \cap \mathcal{A}_r)$ equals

$$\begin{aligned} & 4\varphi(m) \frac{\varphi(l)^2}{l} \prod_{\substack{p|l, \\ p>2}} \left(1 - \frac{1}{(p-1)^2}\right) \sum_{\substack{1 \leq c \leq \ln, \\ \gcd(c,n)=1, \\ 2|c}} w\left(\frac{c}{\ln}\right) \prod_{\substack{p|\gcd(l,c), \\ p>2}} \left(1 + \frac{1}{p-2}\right) \\ &= 4\varphi(m) \frac{\varphi(l)^2}{l} \prod_{\substack{p|l, \\ p>2}} \left(1 - \frac{1}{(p-1)^2}\right) \sum_{\substack{1 \leq c \leq \ln/2, \\ \gcd(c,n)=1}} w\left(\frac{c}{\ln/2}\right) \prod_{\substack{p|\gcd(l,c), \\ p>2}} \left(1 + \frac{1}{p-2}\right). \end{aligned}$$

The rest of the proof for odd l applies *mutatis mutandis* to even l . This completes the proof of (9).

Given $C \geq 1$, let us choose

$$u = 4C \frac{\log \log(D+100)}{\log \log \log(D+100)} \quad \text{and} \quad T = \exp\left(\frac{\log(D+100) \log \log \log(D+100)}{8C \log \log(D+100)} + 1\right).$$

One can readily check that $u^{-u/2} \leq (\log(D+100))^{-C}$. Using $4/\log \log \log 100 < 10$, we also have $T^u \leq (D+100)^{1/2} (\log(D+100))^{10C}$, hence

$$\frac{T^u \log(D+2) \log T}{D} \ll \frac{(\log(D+100))^{12C}}{D^{1/2}}$$

is negligible compared to $(\log(D+2))^{-C}$. □

4. Second-moment bounds

In this section we show how two second-moment bounds, stated as Propositions 6 and 7 below, together with the overlap estimate in Lemma 5, imply Theorem 2. These propositions should be compared to the second-moment bound of Koukoulopoulos and Maynard [KM20, Proposition 5.4], which, together with the overlap estimate of Pollington and Vaughan in (7), implies the Duffin–Schaeffer conjecture.

Let $D(q, r)$ be as in (6). For the sake of readability, let

$$L_s(q, r) := \sum_{\substack{p|\frac{qr}{\gcd(q,r)^2}, \\ p \geq s}} \frac{1}{p} \tag{12}$$

and

$$F(x) = F_C(x) := \exp\left(\frac{\log(x+100) \log \log \log(x+100)}{8C \log \log(x+100)} + 1\right). \tag{13}$$

PROPOSITION 6. For any $Q \in \mathbb{N}$ and any real $t \geq 1$, the set

$$\mathcal{E}_t = \left\{ (q, r) \in [1, Q]^2 : D(q, r) \leq \frac{\Psi(Q)}{t} \right\}$$

satisfies

$$\sum_{(q,r) \in \mathcal{E}_t} \frac{\varphi(q)\psi(q)}{q} \cdot \frac{\varphi(r)\psi(r)}{r} \ll \frac{\Psi(Q)^2}{t^{1/5}},$$

with an absolute implied constant.

PROPOSITION 7. Let $C \geq 1$ be arbitrary. For any $Q \in \mathbb{N}$ and any real $t \geq 1$, the set

$$\mathcal{E}_t = \left\{ (q, r) \in [1, Q]^2 : D(q, r) \leq t\Psi(Q) \text{ and } L_{F(t)}(q, r) \geq \frac{1}{F(t)^{1/4}} \right\}$$

satisfies

$$\sum_{(q,r) \in \mathcal{E}_t} \frac{\varphi(q)\psi(q)}{q} \cdot \frac{\varphi(r)\psi(r)}{r} \ll \frac{\Psi(Q)^2}{F(t)^{1/2}}$$

with an implied constant depending only on C .

We now present the proof of Theorem 2, assuming Propositions 6 and 7.

Proof of Theorem 2. Let $Q \in \mathbb{N}$ be such that $\Psi(Q) \geq 3$. We may assume that $C > 0$ is greater than any prescribed absolute constant. Let $K > 1$ be a large constant in terms of C , to be chosen.

We partition the index set $[1, Q]^2$ into the sets

$$\mathcal{E}^1 = \{(q, r) \in [1, Q]^2 : q = r\},$$

$$\mathcal{E}^2 = \left\{ (q, r) \in [1, Q]^2 : q \neq r, D(q, r) \leq \frac{\Psi(Q)}{(\log \Psi(Q))^C}, L_{F(\Psi(Q))}(q, r) \leq 1 \right\},$$

$$\mathcal{E}^3 = \left\{ (q, r) \in [1, Q]^2 : q \neq r, D(q, r) \leq \frac{\Psi(Q)}{(\log \Psi(Q))^C}, L_{F(\Psi(Q))}(q, r) > 1 \right\},$$

$$\mathcal{E}^4 = \left\{ (q, r) \in [1, Q]^2 : q \neq r, D(q, r) > \frac{\Psi(Q)}{(\log \Psi(Q))^C}, L_{F(D(q,r))}(q, r) \leq \frac{K}{(\log \Psi(Q))^C} \right\},$$

$$\mathcal{E}^5 = \left\{ (q, r) \in [1, Q]^2 : q \neq r, D(q, r) > \frac{\Psi(Q)}{(\log \Psi(Q))^C}, L_{F(D(q,r))}(q, r) > \frac{K}{(\log \Psi(Q))^C} \right\}.$$

The contribution of \mathcal{E}^1 is clearly negligible:

$$\sum_{(q,r) \in \mathcal{E}^1} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) = \sum_{q=1}^Q \lambda(\mathcal{A}_q) = \Psi(Q). \tag{14}$$

Now we consider \mathcal{E}^2 . For any $(q, r) \in \mathcal{E}^2$, the condition $L_{F(\Psi(Q))}(q, r) \leq 1$, together with Mertens’s theorem, ensures that

$$\begin{aligned} \prod_{p | \frac{qr}{\gcd(q,r)^2}} \left(1 + \frac{1}{p-1} \right) &\leq \exp \left(\sum_{\substack{p | \frac{qr}{\gcd(q,r)^2}, \\ p < F(\Psi(Q))}} \frac{2}{p} + \sum_{\substack{p | \frac{qr}{\gcd(q,r)^2}, \\ p \geq F(\Psi(Q))}} \frac{2}{p} \right) \\ &\ll \exp(2 \log \log F(\Psi(Q))) \\ &\ll (\log \Psi(Q))^2. \end{aligned}$$

In the last step we used the rough estimate $F(x) \leq x$ for all $x \geq 3$. The overlap estimate (Lemma 5) thus shows that for any $(q, r) \in \mathcal{E}^2$,

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)(\log \Psi(Q))^2.$$

Applying Proposition 6 with $t = (\log \Psi(Q))^C$ leads to

$$\sum_{(q,r) \in \mathcal{E}^2} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \frac{\Psi(Q)^2}{(\log \Psi(Q))^{C/5-2}}. \tag{15}$$

Next we consider \mathcal{E}^3 . For any $(q, r) \in \mathcal{E}^3$, let $j(q, r)$ be the maximal integer j such that $L_{F(\exp \exp(j))}(q, r) > 1$; note that, by construction, $j(q, r) \geq \lfloor \log \log \Psi(Q) \rfloor$. Let $(q, r) \in \mathcal{E}^3$ with $j(q, r) = j$. By definition, $L_{F(\exp \exp(j+1))}(q, r) \leq 1$, hence Mertens's theorem implies

$$\begin{aligned} \prod_{\substack{p \mid \frac{qr}{\gcd(q,r)^2} \\ p < F(\exp \exp(j+1))}} \left(1 + \frac{1}{p-1}\right) &\leq \exp\left(\sum_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p < F(\exp \exp(j+1))}} \frac{2}{p} + \sum_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p \geq F(\exp \exp(j+1))}} \frac{2}{p}\right) \\ &\ll \exp(2 \log \log F(\exp \exp(j+1))) \\ &\ll \exp(2j). \end{aligned}$$

Thus, the overlap estimate gives

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) \exp(2j),$$

and applying Proposition 7 with $t = \exp \exp(j)$ leads to

$$\begin{aligned} \sum_{(q,r) \in \mathcal{E}^3} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) &= \sum_{j \geq \lfloor \log \log \Psi(Q) \rfloor} \sum_{\substack{(q,r) \in \mathcal{E}^3, \\ j(q,r)=j}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \\ &\ll \sum_{j \geq \lfloor \log \log \Psi(Q) \rfloor} \exp(2j) \frac{\Psi(Q)^2}{F(\exp \exp(j))^{1/2}} \\ &\ll \frac{\Psi(Q)^2}{(\log \Psi(Q))^C}. \end{aligned} \tag{16}$$

In the last step we used the fact that $F(x)$ increases faster than any power of $\log x$.

Now we consider \mathcal{E}^4 . For any $(q, r) \in \mathcal{E}^4$,

$$\prod_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p > F(D(q,r))}} \left(1 + \frac{1}{p-1}\right) \leq \exp(2L_{F(D(q,r))}(q, r)) = 1 + O\left(\frac{1}{(\log \Psi(Q))^C}\right).$$

The overlap estimate thus gives

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) \left(1 + O\left(\frac{1}{(\log \Psi(Q))^C}\right)\right),$$

hence

$$\sum_{(q,r) \in \mathcal{E}^4} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \Psi(Q)^2 + O\left(\frac{\Psi(Q)^2}{(\log \Psi(Q))^C}\right). \tag{17}$$

Finally, we consider \mathcal{E}^5 . Set $\kappa = \frac{1}{100}(e/C)^C$. For any $(q, r) \in \mathcal{E}^5$, let $i(q, r)$ be the maximal integer i such that

$$L_{F(\kappa \exp \exp(i/(\log \Psi(Q))^C))}(q, r) > \frac{K/2}{(\log \Psi(Q))^C}.$$

Note that

$$L_{F(\Psi(Q)/(\log \Psi(Q))^C)}(q, r) \geq L_{F(D(q,r))}(q, r) > \frac{K}{(\log \Psi(Q))^C},$$

therefore

$$i(q, r) \geq \left\lfloor (\log \Psi(Q))^C \log \log \frac{\Psi(Q)}{\kappa(\log \Psi(Q))^C} \right\rfloor.$$

Here

$$\frac{\Psi(Q)}{\kappa(\log \Psi(Q))^C} \geq \frac{1}{\kappa} \min_{x \geq 3} \frac{x}{(\log x)^C} = 100.$$

Let $(q, r) \in \mathcal{E}^5$ such that $i(q, r) = i$. By definition,

$$L_{F(\kappa \exp \exp((i+1)/(\log \Psi(Q))^C))}(q, r) \leq \frac{K/2}{(\log \Psi(Q))^C},$$

hence Mertens's theorem shows that

$$\begin{aligned} \prod_{p \mid \frac{qr}{\gcd(q,r)^2}} \left(1 + \frac{1}{p-1}\right) &\leq \exp\left(\sum_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p < F(\kappa \exp \exp \frac{(i+1)}{(\log \Psi(Q))^C})}} \frac{2}{p} + \sum_{\substack{p \mid \frac{qr}{\gcd(q,r)^2}, \\ p \geq F(\kappa \exp \exp \frac{(i+1)}{(\log \Psi(Q))^C})}} \frac{2}{p}\right) \\ &\ll \exp\left(2 \log \log F\left(\kappa \exp \exp \frac{(i+1)}{(\log \Psi(Q))^C}\right)\right) \\ &\ll \exp\left(\frac{2i}{(\log \Psi(Q))^C}\right). \end{aligned}$$

The overlap estimate thus gives

$$\lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r) \exp\left(\frac{2i}{(\log \Psi(Q))^C}\right).$$

Another application of Mertens's theorem, this time with the error term $O((\log x)^{-C})$ due to Landau [Lan09, p. 201], leads to

$$\begin{aligned} \sum_{F(\kappa \exp \exp \frac{i}{(\log \Psi(Q))^C}) \leq p \leq F(\kappa \exp \exp \frac{(i+1)}{(\log \Psi(Q))^C})} \frac{1}{p} &= \log \log F\left(\kappa \exp \exp \frac{(i+1)}{(\log \Psi(Q))^C}\right) \\ &\quad - \log \log F\left(\kappa \exp \exp \frac{i}{(\log \Psi(Q))^C}\right) \\ &\quad + O\left(\left(\log F\left(\kappa \exp \exp \frac{i}{(\log \Psi(Q))^C}\right)\right)^{-C}\right) \\ &\ll \frac{1}{(\log \Psi(Q))^C}. \end{aligned}$$

In the last step we used the facts that $h(x) := \log \log F(\kappa \exp \exp(x))$ satisfies $h'(x) \ll 1$, and

$$\log F\left(\kappa \exp \exp \frac{i}{(\log \Psi(Q))^C}\right) \geq \log F\left(\frac{\kappa^{1/2}\Psi(Q)^{1/2}}{(\log \Psi(Q))^{C/2}}\right) \gg \log \Psi(Q).$$

Choosing $K > 1$ large enough in terms of C , it follows that

$$L_{F(\kappa \exp \exp(i/(\log \Psi(Q))^C))}(q, r) \leq L_{F(\kappa \exp \exp((i+1)/(\log \Psi(Q))^C))}(q, r) + \frac{K/2}{(\log \Psi(Q))^C} \leq \frac{K}{(\log \Psi(Q))^C},$$

hence $D(q, r) \leq \kappa \exp \exp(i/(\log \Psi(Q))^C)$. Applying Proposition 7 with $t = \kappa \exp \exp(i/(\log \Psi(Q))^C)$ thus leads to

$$\sum_{\substack{(q,r) \in \mathcal{E}^5, \\ i(q,r)=i}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \ll \exp\left(\frac{2i}{(\log \Psi(Q))^C}\right) \frac{\Psi(Q)^2}{F(\kappa \exp \exp(i/(\log \Psi(Q))^C))^{1/2}},$$

and by summing over all possible values of i ,

$$\begin{aligned}
 \sum_{(q,r) \in \mathcal{E}^5} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) &\ll \sum_{i \geq \lfloor (\log \Psi(Q))^C \log \log(\Psi(Q)/\kappa(\log \Psi(Q))^C) \rfloor} \frac{\exp(2i/(\log \Psi(Q))^C) \Psi(Q)^2}{F(\kappa \exp \exp(i/(\log \Psi(Q))^C))^{1/2}} \\
 &\ll \sum_{m \geq \log \log(\Psi(Q)/\kappa(\log \Psi(Q))^C)} \frac{e^{2m} \Psi(Q)^2 (\log \Psi(Q))^C}{F(\kappa \exp \exp m)^{1/2}} \\
 &\ll \frac{\Psi(Q)^2 (\log \Psi(Q))^{C+2}}{F(\Psi(Q)/(\log \Psi(Q))^C)} \\
 &\ll \frac{\Psi(Q)^2}{(\log \Psi(Q))^C}. \tag{18}
 \end{aligned}$$

Combining formulas (14)–(18) shows that

$$\sum_{q,r=1}^Q \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \leq \Psi(Q)^2 + O\left(\frac{\Psi(Q)^2}{(\log \Psi(Q))^{C/5-2}}\right),$$

as claimed. □

5. GCD graphs: notation and basic properties

The proof of the Duffin–Schaeffer conjecture given by Koukoulopoulos and Maynard in [KM20] is based on a concept called ‘GCD graphs’, which they introduced in that paper. Very roughly speaking, a GCD graph encodes information on the divisor structure of a set of integers. To each GCD graph a ‘quality’ can be assigned, and the key argument in [KM20] is that one can iteratively pass to subgraphs of the original GCD graph in such a way that in each step the quality increases and/or the divisor structure becomes more regular. At the end of this procedure, one has a graph that has either particularly high quality or a very regular divisor structure. High quality directly implies that the density of the edge set, essentially controlling the influence of the bad pairs (q, r) in such sets as $\mathcal{E}^1, \dots, \mathcal{E}^5$ of the previous section, is small, leading to the desired result. If one cannot achieve high quality, then one obtains a GCD subgraph that has perfect control of the divisor structure of the underlying set of integers; in this case, results on the ‘anatomy of integers’ can be used to show that the problematic factor

$$\prod_{p \mid \frac{qr}{\gcd(q,r)^2}} \left(1 + \frac{1}{p}\right)$$

in the overlap estimate can only be large for a very small proportion of pairs (q, r) , again leading to the desired result.

We do not give a fully detailed presentation of the notion of a GCD graph here, and refer the reader to §6 of [KM20] instead. However, for the convenience of the reader, we will recall the basic definitions and some of the basic properties of GCD graphs.

A GCD graph is a septuple $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$, for which the following properties hold.

- (a) μ is a measure on \mathbb{N} for which $\mu(n) < \infty$ for all n . This measure is extended to \mathbb{N}^2 by defining

$$\mu(\mathcal{N}) = \sum_{(n_1, n_2) \in \mathcal{N}} \mu(n_1)\mu(n_2), \quad \mathcal{N} \subseteq \mathbb{N}^2.$$

- (b) The *vertex sets* \mathcal{V} and \mathcal{W} are finite sets of positive integers.
 (c) The *edge set* \mathcal{E} is a subset of $\mathcal{V} \times \mathcal{W}$.
 (d) \mathcal{P} is a set of primes.
 (e) f and g are functions from \mathcal{P} to $\mathbb{Z}_{\geq 0}$ such that for all $p \in \mathcal{P}$:
 (i) $p^{f(p)} \mid v$ for all $v \in \mathcal{V}$ and $p^{g(p)} \mid w$ for all $w \in \mathcal{W}$;
 (ii) if $(v, w) \in \mathcal{E}$, then $p^{\min(f(p), g(p))} \parallel \gcd(v, w)$;
 (iii) if $f(p) \neq g(p)$, then $p^{f(p)} \parallel v$ for all $v \in \mathcal{V}$ and $p^{g(p)} \parallel w$ for all $w \in \mathcal{W}$.

For two GCD graphs $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ and $G' = (\mu', \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ we say that G' is a GCD subgraph of G , and write $G' \leq G$, if

$$\mu' = \mu, \quad \mathcal{V}' \subseteq \mathcal{V}, \quad \mathcal{W}' \subseteq \mathcal{W}, \quad \mathcal{E}' \subseteq \mathcal{E}, \quad \mathcal{P}' \supseteq \mathcal{P},$$

and if f (respectively, g) coincides with f' (respectively, g') on \mathcal{P} .

For given \mathcal{V} and $k \geq 0$ we define $\mathcal{V}_{p^k} = \{v \in \mathcal{V} : p^k \parallel v\}$. We write $\mathcal{E}_{p^k, p^\ell} = \mathcal{E} \cap (\mathcal{V}_{p^k} \times \mathcal{W}_{p^\ell})$. It turns out that for $p \notin \mathcal{P}$, the GCD graph

$$G_{p^k, p^\ell} := (\mu, \mathcal{V}_{p^k}, \mathcal{W}_{p^\ell}, \mathcal{E}_{p^k, p^\ell}, \mathcal{P} \cup \{p\}, f_{p^k}, g_{p^\ell})$$

is a GCD subgraph of G (where f_{p^k} and g_{p^ℓ} are defined in such a way that they respectively coincide with f and g on \mathcal{P} , and $f_{p^k}(p) = k$ and $g_{p^\ell}(p) = \ell$).

For a GCD graph $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ we make the following definitions.

- (i) The *edge density*

$$\delta(G) = \frac{\mu(\mathcal{E})}{\mu(\mathcal{V})\mu(\mathcal{W})},$$

provided that $\mu(\mathcal{V})\mu(\mathcal{W}) \neq 0$. If $\mu(\mathcal{V})\mu(\mathcal{W}) = 0$, we define $\delta(G)$ to be 0.

- (ii) The *neighborhood sets*

$$\Gamma_G(v) = \{w \in \mathcal{W} : (v, w) \in \mathcal{E}\}, \quad v \in \mathcal{V},$$

and

$$\Gamma_G(w) = \{v \in \mathcal{V} : (v, w) \in \mathcal{E}\}, \quad w \in \mathcal{W}.$$

- (iii) The set $\mathcal{R}(G)$ of primes that have not (yet) been accounted for in the GCD graph:

$$\mathcal{R}(G) = \{p \notin \mathcal{P} : \exists (v, w) \in \mathcal{E} \text{ such that } p \mid \gcd(v, w)\}.$$

- (iv) The *quality*

$$q(G) = \delta(G)^{10} \mu(\mathcal{V})\mu(\mathcal{W}) \prod_{p \in \mathcal{P}} \frac{p^{|f(p)-g(p)|}}{(1 - \mathbb{1}_{f(p)=g(p)} \geq 1/p)^2 (1 - p^{-31/30})^{10}}.$$

This notion of quality of a GCD graph is an ad-hoc definition, which turns out to serve the required purpose for the argument of [KM20]. We refer to [KM20] for the heuristic reasoning which led to this particular definition. It is possible that a modified notion of quality would be better suited to the argument in the present paper. However, we prefer to stick to the original definition of quality from [KM20], since this allows us to directly use a large part of the iteration procedure from [KM20] without the need to adapt it to a modified framework.

We also introduce

$$\mathcal{R}^{\blacktriangleright}(G) := \left\{ p \in \mathcal{R}(G) : \forall k \geq 0 \min \left\{ \frac{\mu(\mathcal{V}_{p^k})}{\mu(\mathcal{V})}, \frac{\mu(\mathcal{W}_{p^k})}{\mu(\mathcal{W})} \right\} \leq 1 - \frac{1}{\sqrt{p}} \right\}.$$

This should be compared to the sets $\mathcal{R}^{\sharp}(G)$ and $\mathcal{R}^{\flat}(G)$ used in [KM20], the latter of which is defined analogous to our $\mathcal{R}^{\blacktriangleright}(G)$ but with $1 - 10^{40}/p$ instead of $1 - 1/\sqrt{p}$. Finally, we define

$$\mathcal{P}_{\text{diff}}(G) := \{p \in \mathcal{P} : f(p) \neq g(p)\}.$$

Among the basic properties of GCD graphs are the facts that $G_1 \preceq G_2$ and $G_2 \preceq G_3$ together imply $G_1 \preceq G_3$ (transitivity), and that $G_1 \preceq G_2$ implies $\mathcal{R}(G_1) \subseteq \mathcal{R}(G_2)$. However, in general $G_1 \preceq G_2$ does not imply that $\mathcal{R}^{\blacktriangleright}(G_1) \subseteq \mathcal{R}^{\blacktriangleright}(G_2)$.

6. Good GCD subgraphs

In this section we state two results on the existence of a ‘good’ GCD subgraph of an arbitrary GCD graph with trivial multiplicative data (i.e. $\mathcal{P} = \emptyset$) in the form of Propositions 8 and 9 below; these should be compared to [KM20, Proposition 7.1]. We then show how Proposition 6 (respectively, 7) follows from Proposition 8 (respectively, 9).

PROPOSITION 8. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \emptyset, f_{\emptyset}, g_{\emptyset})$ be a GCD graph with trivial set of primes and edge density $\delta(G) > 0$. Then there exists a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that the following assertions hold.*

- (a) $\mathcal{R}(G') = \emptyset$.
- (b) For all $v \in \mathcal{V}'$, we have $\mu(\Gamma_{G'}(v)) \geq (9\delta(G')/10)\mu(\mathcal{W}')$.
- (c) For all $w \in \mathcal{W}'$, we have $\mu(\Gamma_{G'}(w)) \geq (9\delta(G')/10)\mu(\mathcal{V}')$.
- (d) $q(G') \gg q(G)$ with an absolute implied constant.

PROPOSITION 9. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \emptyset, f_{\emptyset}, g_{\emptyset})$ be a GCD graph with trivial set of primes, and let $C \geq 1$. Assume that*

$$\mathcal{E} \subseteq \left\{ (v, w) \in \mathcal{V} \times \mathcal{W} : L_{F(t)}(v, w) \geq \frac{1}{F(t)^{1/4}} \right\} \quad \text{and} \quad \delta(G) \geq \frac{1}{F(t)^{1/2}}$$

with some $t \geq 1$ sufficiently large in terms of C . Then there exists a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that the following assertions hold.

- (a) $\mathcal{R}(G') = \emptyset$.
- (b) For all $v \in \mathcal{V}'$, we have $\mu(\Gamma_{G'}(v)) \geq (9\delta(G')/10)\mu(\mathcal{W}')$.
- (c) For all $w \in \mathcal{W}'$, we have $\mu(\Gamma_{G'}(w)) \geq (9\delta(G')/10)\mu(\mathcal{V}')$.
- (d) One of the following assertions holds.
 - (i) $q(G') \gg t^3 q(G)$ with an implied constant depending only on C .
 - (ii) $q(G') \gg q(G)$ with an implied constant depending only on C , and for any $(v, w) \in \mathcal{E}'$, if we write $v = v' \prod_{p \in \mathcal{P}'} p^{f'(p)}$ and $w = w' \prod_{p \in \mathcal{P}'} p^{g'(p)}$, then $L_{F(t)}(v', w') \geq 1/2F(t)^{1/4}$.

Proof of Proposition 6. Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be a function, let $Q \in \mathbb{N}$ and let $t \geq 1$. Consider the GCD graph $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \emptyset, f_{\emptyset}, g_{\emptyset})$ with the measure $\mu(v) = \varphi(v)\psi(v)/v$, the vertex sets $\mathcal{V} = \mathcal{W} = [1, Q]^2$, and the edge set

$$\mathcal{E} = \left\{ (v, w) \in [1, Q]^2 : D(v, w) \leq \frac{\Psi(Q)}{t} \right\}.$$

Note that $\mu(\mathcal{V}) = \mu(\mathcal{W}) = \Psi(Q)/2$. In the language of GCD graphs, the claim of Proposition 6 can equivalently be written as $\mu(\mathcal{E}) \ll \Psi(Q)^2/t^{1/5}$, that is, $\delta(G) \ll t^{-1/5}$.

By Proposition 8, there exists a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G having properties (a)–(d) of the proposition. Following the steps in [KM20, Proof of Proposition 6.3 assuming Proposition 7.1], from properties (a)–(c) we deduce $q(G') \ll \Psi(Q)^2/t^2$. Since G has trivial set of primes, by the definition of quality and property (d),

$$\delta(G)^{10} \mu(\mathcal{V}) \mu(\mathcal{W}) = q(G) \ll q(G') \ll \frac{\Psi(Q)^2}{t^2}.$$

Therefore $\delta(G) \ll t^{-1/5}$, as claimed. □

For the proof of Proposition 7 we will need the following fact about the ‘anatomy of integers’; compare this to [KM20, Lemma 7.3], which is a similar result for a fixed value of c on the right-hand side, rather than allowing $c \rightarrow 0$ as, in view of Lemma 5 above, will be necessary for our application.

LEMMA 10. For any real $x, t \geq 1$ and $c > 0$,

$$\left| \left\{ n \leq x : \sum_{\substack{p|n, \\ p \geq t}} \frac{1}{p} \geq c \right\} \right| \ll x e^{-100ct}$$

with an absolute implied constant.

Proof. An application of the Markov inequality gives

$$\begin{aligned} \left| \left\{ n \leq x : \sum_{\substack{p|n, \\ p \geq t}} \frac{1}{p} \geq c \right\} \right| &= \left| \left\{ n \leq x : \exp \left(100t \sum_{\substack{p|n, \\ p \geq t}} \frac{1}{p} \right) \geq \exp(100ct) \right\} \right| \\ &\leq e^{-100ct} \sum_{n \leq x} \prod_{\substack{p|n, \\ p \geq t}} e^{100t/p}. \end{aligned}$$

Now let f be the multiplicative function defined at prime powers as $f(p^m) = e^{100t/p}$ if $p \geq t$, and $f(p^m) = 1$ if $p < t$. Note that $f(p^m) \leq e^{100}$ at all prime powers. Hence, by [Kou19, Theorem 14.2] the partial sums of f satisfy

$$\begin{aligned} \sum_{\substack{n \leq x \\ p|n, \\ p \geq t}} e^{100t/p} &= \sum_{n \leq x} f(n) \ll x \exp \left(\sum_{p \leq x} \frac{f(p) - 1}{p} \right) = x \exp \left(\sum_{t \leq p \leq x} \frac{e^{100t/p} - 1}{p} \right) \\ &= x \exp \left(O \left(\sum_{p \geq t} \frac{t}{p^2} \right) \right) \\ &\ll x, \end{aligned}$$

where the implied constants are absolute. □

Proof of Proposition 7. Let $\psi : \mathbb{N} \rightarrow [0, 1/2]$ be a function, let $Q \in \mathbb{N}$ and let $t \geq 1$. Consider the GCD graph $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \emptyset, f_\emptyset, g_\emptyset)$ with the measure $\mu(v) = \varphi(v)\psi(v)/v$, the vertex sets $\mathcal{V} = \mathcal{W} = [1, Q]^2$, and the edge set

$$\mathcal{E} = \left\{ (v, w) \in [1, Q]^2 : D(v, w) \leq t\Psi(Q) \text{ and } L_{F(t)}(v, w) \geq \frac{1}{F(t)^{1/4}} \right\}.$$

Note that $\mu(\mathcal{V}) = \mu(\mathcal{W}) = \Psi(Q)/2$. In the language of GCD graphs, the claim can equivalently be written as $\mu(\mathcal{E}) \ll \Psi(Q)^2/F(t)^{1/2}$, that is, $\delta(G) \ll F(t)^{-1/2}$. We may assume in the sequel that $\delta(G) \geq F(t)^{-1/2}$ and that t and $F(t)$ are large enough in terms of C , since otherwise the claim trivially holds.

By Proposition 9, there exists a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G having properties (a)–(d) of the proposition. Let $a = \prod_{p \in \mathcal{P}'} p^{f'(p)}$ and $b = \prod_{p \in \mathcal{P}'} p^{g'(p)}$. By the definition of a GCD graph, $a \mid v$ for all $v \in \mathcal{V}'$ and $b \mid w$ for all $w \in \mathcal{W}'$. Since $\mathcal{R}(G') = \emptyset$, we also have $\gcd(v, w) = \gcd(a, b)$ for all $(v, w) \in \mathcal{E}'$. Following the steps in [KM20, Proof of Proposition 6.3 assuming Proposition 7.1], we deduce from properties (a)–(c) of Proposition 9 that

$$q(G') \ll ab\Psi(Q)^2t^2 \sum_{(v,w) \in \mathcal{E}'} \frac{1}{w_0v_{\max}(w)} \leq \Psi(Q)^2t^2, \tag{19}$$

where $w_0 = \max \mathcal{W}'$ and $v_{\max}(w) = \max\{v \in \mathcal{V}' : (v, w) \in \mathcal{E}'\}$.

Assume first that G' satisfies property (d)(i) in Proposition 9, that is, $q(G') \gg t^3q(G)$. Since G has trivial set of primes, by the definition of quality and (19) we obtain

$$\delta(G)^{10}\mu(\mathcal{V})\mu(\mathcal{W}) = q(G) \ll t^{-3}q(G') \ll \frac{\Psi(Q)^2}{t}.$$

Therefore $\delta(G) \ll t^{-1/10} \ll F(t)^{-1/2}$, as claimed.

Assume next that G' satisfies property d)(ii) in Proposition 9, that is, $q(G') \gg q(G)$, and for any $(v, w) \in \mathcal{E}'$, if we write $v = av'$ and $w = bw'$, then $L_{F(t)}(v', w') \geq 1/2F(t)^{1/4}$. Note that here $\gcd(v', w') = 1$. As in the first case, we have

$$\begin{aligned} \delta(G)^{10}\mu(\mathcal{V})\mu(\mathcal{W}) &= q(G) \ll q(G') \ll ab\Psi(Q)^2t^2 \sum_{(v,w) \in \mathcal{E}'} \frac{1}{w_0v_{\max}(w)} \\ &\leq \frac{ab\Psi(Q)^2t^2}{w_0} \sum_{1 \leq w' \leq w_0/b} \frac{1}{v_{\max}(bw')} \sum_{\substack{1 \leq v' \leq v_{\max}(bw')/a, \\ L_{F(t)}(v', w') \geq 1/(2F(t)^{1/4})}} 1. \end{aligned}$$

For the sake of readability, define $R_s(n) = \sum_{p \mid n, p \geq s} 1/p$ for any $n \in \mathbb{N}$ and $s \geq 1$. Then $1/(2F(t)^{1/4}) \leq L_{F(t)}(v', w') = R_{F(t)}(v') + R_{F(t)}(w')$ implies that $R_{F(t)}(v') \geq 1/(4F(t)^{1/4})$ or $R_{F(t)}(w') \geq 1/(4F(t)^{1/4})$. The previous formula thus shows that $\delta(G)^{10} \ll S_1 + S_2$ with

$$\begin{aligned} S_1 &= \frac{abt^2}{w_0} \sum_{1 \leq w' \leq w_0/b} \frac{1}{v_{\max}(bw')} \sum_{\substack{1 \leq v' \leq v_{\max}(bw')/a, \\ R_{F(t)}(v') \geq 1/(4F(t)^{1/4})}} 1, \\ S_2 &= \frac{abt^2}{w_0} \sum_{\substack{1 \leq w' \leq w_0/b, \\ R_{F(t)}(w') \geq 1/(4F(t)^{1/4})}} \frac{1}{v_{\max}(bw')} \sum_{1 \leq v' \leq v_{\max}(bw')/a} 1. \end{aligned}$$

An application of Lemma 10 with $x = v_{\max}(bw')/a$ and $c = 1/(4F(t)^{1/4})$ yields

$$S_1 \ll \frac{bt^2}{w_0} \sum_{1 \leq w' \leq w_0/b} \exp(-25F(t)^{3/4}) = t^2 \exp(-25F(t)^{3/4}) \ll t^{-100}.$$

Another application of Lemma 10 with $x = w_0/b$ and $c = 1/(4F(t)^{1/4})$ similarly yields

$$S_2 = \frac{bt^2}{w_0} \sum_{\substack{1 \leq w' \leq w_0/b, \\ R_{F(t)}(w') \geq 1/(4F(t)^{1/4}}} 1 \ll t^2 \exp(-25F(t)^{3/4}) \ll t^{-100}.$$

Therefore $\delta(G) \ll (S_1 + S_2)^{1/10} \ll t^{-10} \ll F(t)^{-10}$, as claimed. □

7. Four technical lemmas

In this section we state four lemmas on GCD subgraphs, and show that Propositions 8 and 9 follow from these four lemmas. The key technical improvement in comparison with the iteration argument of [KM20] is in Lemma 11 below, which more carefully balances the quality gain versus the potential density loss of the iteration procedure. The ratio of quality gain to density loss which is necessary for the proof of Theorem 2 is determined by the range of admissible parameters u and A in Lemma 5, and what Lemma 11 provides is just enough for a successful completion of the proof. Lemma 12, which should be compared to [KM20, Lemma 8.4], and Lemma 13 follow from results in [KM20] in a more or less straightforward way. Finally, for the convenience of the reader, we cite [KM20, Lemma 8.5] in the form of Lemma 14.

LEMMA 11. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \emptyset, f_\emptyset, g_\emptyset)$ be a GCD graph with trivial set of primes and $\delta(G) > 0$. Let $C \geq 1$, and let $t \geq 1$ be sufficiently large in terms of C . Then there exists a GCD subgraph $G' \preceq G$ such that $\mathcal{R}^{\blacktriangleright}(G') = \emptyset$, and at least one of the following two statements hold.*

- (a) $q(G') \geq t^3 q(G)$.
- (b) $q(G') \gg q(G)$, $\delta(G')/\delta(G) \geq 1/F(t)^{1/4}$, $|\mathcal{P}_{\text{diff}}(G')| \leq \log t$ with an implied constant depending only on C .

LEMMA 12. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph. Assume that*

$$\delta(G) \geq \frac{1}{s^{1/4}}, \quad \mathcal{R}^{\blacktriangleright}(G) = \emptyset, \quad \mathcal{E} \subseteq \left\{ (v, w) \in \mathcal{V} \times \mathcal{W} : L_s(v, w) \geq \frac{1}{s^{1/4}} \right\}$$

with a sufficiently large $s \geq 1$. Then there exists a GCD subgraph $G' = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}', \mathcal{P}, f, g)$ of G such that

$$q(G') \geq \frac{q(G)}{2} \quad \text{and} \quad \mathcal{E}' \subseteq \left\{ (v, w) \in \mathcal{V} \times \mathcal{W} : \sum_{\substack{p | \frac{vw}{\gcd(v,w)^2}, \\ p \geq s, p \notin \mathcal{R}(G)}} \frac{1}{p} \geq \frac{3}{4s^{1/4}} \right\}.$$

LEMMA 13. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph with $\delta(G) > 0$. Then there exists a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that*

$$\mathcal{P}' \subseteq \mathcal{P} \cup \mathcal{R}(G), \quad \mathcal{R}(G') = \emptyset, \quad q(G') \gg q(G)$$

with an absolute implied constant.

LEMMA 14 [KM20, Lemma 8.5]. *Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph with $\delta(G) > 0$. Then there exists a GCD subgraph $G' = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}', \mathcal{P}, f, g)$ of G such that the following assertions hold.*

- (a) $q(G') \geq q(G)$.
- (b) $\delta(G') \geq \delta(G)$.

(c) For all $v \in \mathcal{V}'$ and $w \in \mathcal{W}'$, we have

$$\mu(\Gamma_{G'}(v)) \geq \frac{9\delta(G')}{10}\mu(\mathcal{W}') \quad \text{and} \quad \mu(\Gamma_{G'}(w)) \geq \frac{9\delta(G')}{10}\mu(\mathcal{V}').$$

We now show how Lemmas 11–14 imply Propositions 8 and 9.

Proof of Proposition 8. Apply Lemma 13 to G to obtain a GCD subgraph $G^{(1)} \preceq G$ with $\mathcal{R}(G^{(1)}) = \emptyset$ and $q(G^{(1)}) \gg q(G)$, satisfying properties (a) and (d). Next, apply Lemma 14 to $G^{(1)}$ to obtain a GCD subgraph $G^{(2)} \preceq G^{(1)}$ which additionally satisfies properties (b) and (c). \square

Proof of Proposition 9. We follow [KM20, Proof of Proposition 7.1], although the ordering of the different stages needs to be changed. It suffices to prove the existence of a GCD subgraph which satisfies properties (a) and (d). Indeed, applying Lemma 14 to such a subgraph, we obtain a GCD subgraph that satisfies all required properties (a)–(d).

We start by applying Lemma 11 to G , and obtain a GCD subgraph $G^{(1)} \preceq G$ such that $\mathcal{R}^{\triangleright}(G^{(1)}) = \emptyset$ and $G^{(1)}$ satisfies at least one of the following properties:

- (A) $q(G^{(1)}) \geq t^3q(G)$;
- (B) $q(G^{(1)}) \gg q(G)$, $\delta(G^{(1)})/\delta(G) \geq 1/F(t)^{1/4}$, $|\mathcal{P}_{\text{diff}}(G^{(1)})| \leq \log t$.

We distinguish between two cases depending on whether (A) or (B) is satisfied.

Case (A). Assume that $q(G^{(1)}) \geq t^3q(G)$. We apply Lemma 13 to obtain a GCD subgraph $G^{(2A)} \preceq G^{(1)}$ with $\mathcal{R}(G^{(2A)}) = \emptyset$ and $q(G^{(2A)}) \gg q(G^{(1)})$. Then $G^{(2A)}$ satisfies properties (a) and (d)(i) in Proposition 9. This finishes the proof for case (A).

Case (B). Assume that $q(G^{(1)}) \gg q(G)$, $\delta(G^{(1)})/\delta(G) \geq 1/F(t)^{1/4}$, $|\mathcal{P}_{\text{diff}}(G^{(1)})| \leq \log t$. First, we remove the effect of the large primes in $\mathcal{R}(G^{(1)})$ on $L_{F(t)}(v, w)$. By the assumption $\delta(G) \geq 1/F(t)^{1/2}$, we have $\delta(G^{(1)}) \geq 1/F(t)^{1/4}$. We can thus apply Lemma 12 to $G^{(1)}$ with $s = F(t)$ to obtain a GCD subgraph $G^{(2B)} \preceq G^{(1)}$ with edge set $\mathcal{E}^{(2B)}$ such that

$$q(G^{(2B)}) \geq \frac{q(G^{(1)})}{2} \quad \text{and} \quad \mathcal{E}^{(2B)} \subseteq \left\{ (v, w) \in \mathcal{V} \times \mathcal{W} : \sum_{\substack{p \mid \frac{vw}{\gcd(v,w)^2}, \\ p \geq F(t), p \notin \mathcal{R}(G^{(1)})}} \frac{1}{p} \geq \frac{3}{4F(t)^{1/4}} \right\}.$$

Now we remove the contribution of the primes in $\mathcal{P}_{\text{diff}}(G^{(1)})$. Using $|\mathcal{P}_{\text{diff}}(G^{(1)})| \leq \log t$, we obtain that for any $(v, w) \in \mathcal{E}^{(2B)}$,

$$\sum_{\substack{p \mid \frac{vw}{\gcd(v,w)^2}, \\ p \geq F(t), p \in \mathcal{P}_{\text{diff}}(G^{(1)})}} \frac{1}{p} \leq \frac{\log t}{F(t)} \leq \frac{1}{4F(t)^{1/4}}$$

for large enough t . Hence, for any $(v, w) \in \mathcal{E}^{(2B)}$,

$$\sum_{\substack{p \mid \frac{vw}{\gcd(v,w)^2}, \\ p \geq F(t), p \notin \mathcal{R}(G^{(1)}) \cup \mathcal{P}_{\text{diff}}(G^{(1)})}} \frac{1}{p} \geq \frac{1}{2F(t)^{1/4}}. \tag{20}$$

Finally, we apply Lemma 13 to $G^{(2B)}$ to obtain a GCD subgraph $G^{(3B)} \preceq G^{(2B)}$ such that

$$\mathcal{R}(G^{(3B)}) = \emptyset \quad \text{and} \quad q(G^{(3B)}) \gg q(G^{(2B)}) \gg q(G).$$

Thus, $G^{(3B)}$ satisfies property (a) in Proposition 9. Following the steps in stage 4b of [KM20, Proof of Proposition 7.1], we deduce from (20) that $G^{(3B)}$ satisfies property (d)(ii) as well. This finishes the proof for case (B). \square

8. Proof of Lemmas 12 and 13

Proof of Lemma 12. Define

$$S(v, w) = \sum_{\substack{p | \frac{vw}{\gcd(v,w)^2}, \\ p \geq s, p \in \mathcal{R}(G)}} \frac{1}{p}.$$

Following the steps in [KM20, Proof of Lemma 8.4], from the assumptions $\mathcal{R}^{\blacktriangleright}(G) = \emptyset$ and $\delta(G) \geq 1/s^{1/4}$ we deduce that

$$\sum_{(v,w) \in \mathcal{E}} \mu(v)\mu(w)S(v, w) \leq \sum_{p \geq s} \frac{2\mu(\mathcal{V})\mu(\mathcal{W})}{p^{3/2}} \leq \frac{\mu(\mathcal{E})}{100s^{1/4}}$$

for large enough s . Consider the edge set

$$\mathcal{E}' := \left\{ (v, w) \in \mathcal{E} : S(v, w) \leq \frac{1}{4s^{1/4}} \right\}.$$

An application of the Markov inequality gives

$$\mu(\mathcal{E} \setminus \mathcal{E}') \leq 4s^{1/4} \sum_{(v,w) \in \mathcal{E}} \mu(v)\mu(w)S(v, w) \leq \frac{\mu(\mathcal{E})}{25},$$

that is, $\mu(\mathcal{E}') \geq \frac{24}{25}\mu(\mathcal{E})$. By the definition of quality, the GCD subgraph $G' := (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}', \mathcal{P}, f, g)$ thus satisfies

$$\frac{q(G')}{q(G)} = \left(\frac{\mu(\mathcal{E}')}{\mu(\mathcal{E})} \right)^{10} \geq \frac{1}{2}.$$

Further, for any $(v, w) \in \mathcal{E}'$ we have

$$\sum_{\substack{p | \frac{vw}{\gcd(v,w)^2}, \\ p \geq s, p \notin \mathcal{R}(G)}} \frac{1}{p} = L_s(v, w) - S(v, w) \geq \frac{3}{4s^{1/4}},$$

as claimed. \square

To prove Lemma 13, we will iteratively apply the following two propositions.

PROPOSITION 15. Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph with $\delta(G) > 0$. Then there is a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that

$$\mathcal{P}' \subseteq \mathcal{P} \cup (\mathcal{R}(G) \cap \{p \leq 10^{2000}\}), \quad \mathcal{R}(G') \subseteq \{p > 10^{2000}\}, \quad \frac{q(G')}{q(G)} \geq \frac{1}{10^{10^{3000}}}.$$

Proof. This is a slight modification of [KM20, Proposition 8.3], the only difference being that in our formulation the set \mathcal{P} can be non-empty. The proof given in [KM20] actually covers the formulation stated above, since it only relies on the iterative application of [KM20, Lemma 13.2], which holds for GCD graphs with an arbitrary set of primes. \square

PROPOSITION 16. Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph with $\delta(G) > 0$ such that $\emptyset \neq \mathcal{R}(G) \subseteq \{p > 10^{2000}\}$. Then there is a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that

$$\mathcal{P} \subsetneq \mathcal{P}' \subseteq \mathcal{P} \cup \mathcal{R}(G), \quad \mathcal{R}(G') \subsetneq \mathcal{R}(G), \quad q(G') \geq q(G).$$

Proof. This follows directly from [KM20, Propositions 8.1 and 8.2]. □

Proof of Lemma 13. First, we apply Proposition 15 to obtain a GCD subgraph $G^{(1)} \preceq G$ with

$$\mathcal{R}(G^{(1)}) \subseteq \{p > 10^{2000}\} \quad \text{and} \quad q(G^{(1)}) \gg q(G).$$

If $\mathcal{R}(G^{(1)}) = \emptyset$, we are done. Otherwise, we apply Proposition 16 to obtain a GCD subgraph $H_1 \preceq G^{(1)}$ with $\mathcal{R}(H_1) \subsetneq \mathcal{R}(G^{(1)})$ and $q(H_1) \geq q(G^{(1)})$. By iterating this argument, we obtain a chain of GCD subgraphs $G^{(1)} \succeq H_1 \succeq H_2 \succeq \dots$ with

$$\mathcal{R}(G^{(1)}) \supsetneq \mathcal{R}(H_1) \supsetneq \mathcal{R}(H_2) \supsetneq \dots \quad \text{and} \quad q(G^{(1)}) \leq q(H_1) \leq q(H_2) \leq \dots.$$

Since $\mathcal{R}(G^{(1)})$ is a finite set, we arrive after finitely many steps at a GCD subgraph $G' \preceq G$ with $\mathcal{R}(G') = \emptyset$ and $q(G') \geq q(G^{(1)}) \gg q(G)$. Furthermore, we have $\mathcal{P}' \subseteq \mathcal{P} \cup \mathcal{R}(G)$ since this property is preserved at each step. □

9. Quality increment versus density loss

The goal of this section is to prove Lemma 11. We start with three preliminary results.

LEMMA 17. Let $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ be a GCD graph with $\delta(G) > 0$, let $p \in \mathcal{R}(G)$, and let

$$\alpha_k = \frac{\mu(\mathcal{V}_{p^k})}{\mu(\mathcal{V})} \quad \text{and} \quad \beta_l = \frac{\mu(\mathcal{W}_{p^l})}{\mu(\mathcal{W})}.$$

Then there exists a pair of non-negative integers $(k, l) = (k_p, l_p)$ such that $\alpha_k, \beta_l > 0$, and

$$\frac{\mu(\mathcal{E}_{p^k, p^l})}{\mu(\mathcal{E})} \geq \begin{cases} (\alpha_k \beta_k)^{9/10} & \text{if } k = l, \\ \frac{\alpha_k(1 - \beta_k) + \beta_k(1 - \alpha_k) + \alpha_l(1 - \beta_l) + \beta_l(1 - \alpha_l)}{40|k - l|^2} & \text{if } k \neq l. \end{cases}$$

Proof. This follows from a straightforward modification of the proof of [KM20, Lemma 12.1], replacing the estimate $\frac{1}{1000} \sum_{|j| \geq 1} 2^{-|j|/20} \leq \frac{1}{10}$ by $\sum_{|j| \geq 1} 1/40j^2 \leq \frac{1}{10}$ in one of the steps. □

LEMMA 18. Let $\alpha_k, \beta_k, \alpha_l, \beta_l \in [0, 1]$ with $\alpha_k, \beta_l > 0$ be such that $\alpha_k + \alpha_l \leq 1$ and $\beta_k + \beta_l \leq 1$, and let

$$S = \alpha_k(1 - \beta_k) + \beta_k(1 - \alpha_k) + \alpha_l(1 - \beta_l) + \beta_l(1 - \alpha_l).$$

If $\min\{\alpha_k, \beta_k\} \leq 1 - R$ and $\min\{\alpha_l, \beta_l\} \leq 1 - R$ with some $R \in [0, 1/\sqrt{2}]$, then $S^2/\alpha_k\beta_l \geq R/2$.

Proof. Clearly,

$$S \geq \alpha_k(1 - \beta_k) + \beta_l(1 - \alpha_l) \geq \alpha_k\beta_l + \beta_l\alpha_k = 2\alpha_k\beta_l. \tag{21}$$

Since the conditions of the lemma and S are invariant under switching α_k with β_l and α_l with β_k , respectively, we may assume that $\alpha_k \geq \beta_l$.

Assume first that $\alpha_k \leq 1/2$. Then $\beta_l \leq 1/2$ as well, hence

$$S = \beta_k(1 - 2\alpha_k) + \alpha_k + \alpha_l(1 - 2\beta_l) + \beta_l \geq \alpha_k + \beta_l \geq 2\sqrt{\alpha_k\beta_l}.$$

Therefore $S^2/(\alpha_k\beta_l) \geq 4 > R/2$, as claimed.

Assume next that $\alpha_k > 1/2$. Formula (21) then gives

$$1 - \beta_k \leq 2\alpha_k(1 - \beta_k) \leq 2S \leq \frac{S^2}{\alpha_k\beta_l}.$$

If $\beta_k \leq 1 - R$, then $R \leq 1 - \beta_k \leq S^2/(\alpha_k\beta_l)$, as claimed. If $\beta_k > 1 - R > 1/4$, then by the assumption $\min\{\alpha_k, \beta_k\} \leq 1 - R$ we have $\alpha_k \leq 1 - R$, and we similarly deduce

$$R \leq 1 - \alpha_k \leq 4\beta_k(1 - \alpha_k) \leq 4S \leq 2\frac{S^2}{\alpha_k\beta_l},$$

which finishes the proof of the statement. □

The following lemma is a variant of [KM20, Lemma 12.2].

LEMMA 19. Consider a GCD graph $G = (\mu, \mathcal{V}, \mathcal{W}, \mathcal{E}, \mathcal{P}, f, g)$ with $\delta(G) > 0$ and a prime $p \in \mathcal{R}^{\triangleright}(G)$. Let $(k, l) = (k_p, l_p)$ be a pair of non-negative integers which satisfies the conclusion of Lemma 17. Then there is a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G with $\mathcal{P}' = \mathcal{P} \cup \{p\}$ and $\mathcal{R}(G') \subseteq \mathcal{R}(G) \setminus \{p\}$ such that

$$\frac{\delta(G')}{\delta(G)} \geq \begin{cases} 1 & \text{if } k = l, \\ \frac{1}{20|k - l|^2} & \text{if } k \neq l, \end{cases}$$

and

$$\frac{q(G')}{q(G)} \geq \begin{cases} 1 & \text{if } k = l, \\ \frac{p^{|k-l|-1/2}}{10^{15}|k - l|^{20}} & \text{if } k \neq l. \end{cases}$$

Proof. We claim that $G' = G_{p^k, p^l}$ satisfies all required properties. Note that $\mathcal{P}' = \mathcal{P} \cup \{p\}$ and $\mathcal{R}(G') \subseteq \mathcal{R}(G) \setminus \{p\}$ hold by the definition of G_{p^k, p^l} . If $k = l$, then by Lemma 17 and the definition of quality,

$$\frac{\delta(G')}{\delta(G)} = \frac{\mu(\mathcal{E}_{p^k, p^k})}{\mu(\mathcal{E})} \cdot \frac{1}{\alpha_k\beta_k} \geq 1,$$

and

$$\frac{q(G')}{q(G)} = \left(\frac{\mu(\mathcal{E}_{p^k, p^k})}{\mu(\mathcal{E})}\right)^{10} (\alpha_k\beta_k)^{-9} \frac{1}{(1 - \mathbb{1}_{k \geq 1/p})^2 (1 - 1/p^{31/30})^{10}} \geq 1,$$

as claimed. Let S be as in Lemma 18. If $k \neq l$, then by Lemma 17 together with (21),

$$\frac{\delta(G')}{\delta(G)} = \frac{\mu(\mathcal{E}_{p^k, p^l})}{\mu(\mathcal{E})} \cdot \frac{1}{\alpha_k\beta_l} \geq \frac{S}{40|k - l|^2\alpha_k\beta_l} \geq \frac{1}{20|k - l|^2}.$$

Furthermore,

$$\begin{aligned} \frac{q(G')}{q(G)} &= \left(\frac{\mu(\mathcal{E}_{p^k, p^l})}{\mu(\mathcal{E})}\right)^{10} (\alpha_k\beta_l)^{-9} \frac{p^{|k-l|}}{(1 - 1/p^{31/30})^{10}} \geq \frac{S^{10}}{(40|k - l|^2)^{10}} \cdot \frac{1}{(\alpha_k\beta_l)^9} p^{|k-l|} \\ &\geq \frac{2^8 p^{|k-l|}}{40^{10}|k - l|^{20}} \cdot \frac{S^2}{\alpha_k\beta_l}. \end{aligned}$$

The assumption $p \in \mathcal{R}^{\triangleright}(G)$ ensures that $\min\{\alpha_k, \beta_k\} \leq 1 - 1/\sqrt{p}$ and $\min\{\alpha_l, \beta_l\} \leq 1 - 1/\sqrt{p}$. Hence, we can apply Lemma 18 with $R = 1/\sqrt{p}$, which shows that

$$\frac{q(G')}{q(G)} \geq \frac{2^7 p^{|k-l|-1/2}}{40^{10}|k - l|^{20}} > \frac{p^{|k-l|-1/2}}{10^{15}|k - l|^{20}},$$

as claimed. □

Proof of Lemma 11. We apply Lemma 19 iteratively to G until we obtain a GCD subgraph $G' = (\mu, \mathcal{V}', \mathcal{W}', \mathcal{E}', \mathcal{P}', f', g')$ of G such that $\mathcal{R}^{\blacktriangleright}(G') = \emptyset$. Note that each prime p is used at most once, and \mathcal{P}' is precisely the set of primes to which Lemma 19 was applied. For each $p \in \mathcal{P}'$, let (k_p, l_p) be the pair of non-negative integers with which Lemma 19 is applied.¹ Since the original graph G had an empty set of primes, we have $\mathcal{P}_{\text{diff}}(G') = \{p \in \mathcal{P}' : k_p \neq l_p\}$. By Lemma 19, the resulting graph G' satisfies

$$\frac{\delta(G')}{\delta(G)} \geq \prod_{p \in \mathcal{P}_{\text{diff}}(G')} \frac{1}{20|k_p - l_p|^2} \quad \text{and} \quad \frac{q(G')}{q(G)} \geq \prod_{p \in \mathcal{P}_{\text{diff}}(G')} \frac{p^{|k_p - l_p| - 1/2}}{10^{15}|k_p - l_p|^{20}}.$$

In particular,

$$\frac{q(G')}{q(G)} \gg \prod_{p \in \mathcal{P}_{\text{diff}}(G')} p^{|k_p - l_p|/4} \gg 1. \tag{22}$$

Fix $C \geq 1$, and let $t \geq 1$ be large enough in terms of C . Let $N = |\mathcal{P}_{\text{diff}}(G')|$, and, for the sake of readability, in the sequel let \log_i denote the i -fold iterated logarithm. It will be enough to show that if $q(G') < t^3 q(G)$ (i.e. property (a) does not hold), then $\delta(G')/\delta(G) \geq 1/F(t)^{1/4}$, and $N \leq \log t$ (i.e. property (b) holds). The latter follows easily from (22) and $q(G') < t^3 q(G)$:

$$(N!)^{1/4} \leq \prod_{p \in \mathcal{P}_{\text{diff}}(G')} p^{|k_p - l_p|/4} \ll \frac{q(G')}{q(G)} < t^3.$$

Hence, $N \ll (\log t)/\log_2 t$, and, in particular, $N \leq \log t$ for large enough t , as claimed. It remains to show that $q(G') < t^3 q(G)$ implies $\delta(G')/\delta(G) \geq 1/F(t)^{1/4}$.

Let $Y = \{p \in \mathcal{P}_{\text{diff}}(G') : |k_p - l_p| \geq \log_3 t\}$. Bounding the sum term by term gives

$$\sum_{p \notin Y} \log(20|k_p - l_p|^2) \ll N \log_4 t \ll \frac{\log t \log_4 t}{\log_2 t}. \tag{23}$$

On the other hand, (22) and the assumption $q(G') < t^3 q(G)$ lead to

$$\log t \gg \sum_{p \in \mathcal{P}_{\text{diff}}(G')} |k_p - l_p| \log p \geq \log_3 t \sum_{p \in Y} \log p \gg (\log_3 t)|Y| \log |Y|,$$

hence $|Y| \ll (\log t)/(\log_2 t \log_3 t)$. The previous formula also shows that $\sum_{p \in Y} |k_p - l_p| \ll \log t$. An application of the inequality of arithmetic and geometric means thus yields

$$\begin{aligned} \sum_{p \in Y} \log(20|k_p - l_p|^2) &\leq 2 \sum_{p \in Y} \log(20|k_p - l_p|) \leq 2|Y| \log \frac{\sum_{p \in Y} 20|k_p - l_p|}{|Y|} \\ &\ll |Y| \log \left(\frac{\log t}{|Y|} \right) \\ &\ll \frac{\log t}{\log_2 t}. \end{aligned}$$

The previous formula and (23) thus give

$$-\log \frac{\delta(G')}{\delta(G)} \leq \sum_{p \in \mathcal{P}_{\text{diff}}(G')} \log(20|k_p - l_p|^2) \ll \frac{\log t \log_4 t}{\log_2 t}.$$

¹ We might use primes $p \notin \mathcal{R}^{\blacktriangleright}(G)$ of the original GCD graph G , since $\mathcal{R}^{\blacktriangleright}$ does not necessarily decrease at each step. However, \mathcal{R} decreases by at least one element at each step, hence the algorithm terminates.

Hence, $-\log(\delta(G')/\delta(G)) \leq \frac{1}{4} \log F(t)$ for large enough t , that is, $\delta(G')/\delta(G) \geq 1/F(t)^{1/4}$, and we obtain the desired result. \square

ACKNOWLEDGEMENTS

CA is supported by the Austrian Science Fund (FWF), projects F-5512, I-3466, I-4945, I-5554, P-34763, P-35322 and Y-901. BB is supported by the Austrian Science Fund (FWF), project F-5510. We wish to thank the referee for a very careful reading of our paper and for many helpful comments.

REFERENCES

- BV23 V. Beresnevich and S. Velani, *The divergence Borel–Cantelli lemma revisited*, J. Math. Anal. Appl. **519** (2023), 126750.
- BV06 V. Beresnevich and S. Velani, *A mass transference principle and the Duffin–Schaeffer conjecture for Hausdorff measures*, Ann. of Math. (2) **164** (2006), 971–992.
- DS41 R. J. Duffin and A. C. Schaeffer, *Khinchine’s problem in metric Diophantine approximation*, Duke Math. J. **8** (1941), 243–255.
- Erd70 P. Erdős, *On the distribution of the convergents of almost all real numbers*, J. Number Theory **2** (1970), 425–441.
- Gal61 P. Gallagher, *Approximation by reduced fractions*, J. Math. Soc. Japan **13** (1961), 342–345.
- Har98 G. Harman, *Metric number theory*, London Mathematical Society Monographs. New Series, vol. 18 (Clarendon Press, Oxford, 1998).
- Khi24 A. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*, Math. Ann. **92** (1924), 115–125.
- Kou19 D. Koukoulopoulos, *The distribution of prime numbers*, Graduate Studies in Mathematics, vol. 203 (American Mathematical Society, Providence, RI, 2019).
- KM20 D. Koukoulopoulos and J. Maynard, *On the Duffin–Schaeffer conjecture*, Ann. of Math. (2) **192** (2020), 251–307.
- Lan09 E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. I (German) (Teubner, Leipzig and Berlin, 1909).
- Phi71 W. Philipp, *Mixing sequences of random variables and probabilistic number theory*, Memoirs of the American Mathematical Society, No. 114 (American Mathematical Society, Providence, RI, 1971).
- PV90 A. D. Pollington and R. C. Vaughan, *The k -dimensional Duffin and Schaeffer conjecture*, Mathematika **37** (1990), 190–200.
- Vaa78 J. D. Vaaler, *On the metric theory of Diophantine approximation*, Pacific J. Math. **76** (1978), 527–539.

Christoph Aistleitner aistleitner@math.tugraz.at

Graz University of Technology, Institute of Analysis and Number Theory, Steyrergasse 30/II,
8010 Graz, Austria

Bence Borda borda@math.tugraz.at

Graz University of Technology, Institute of Analysis and Number Theory, Steyrergasse 30/II,
8010 Graz, Austria

Manuel Hauke hauke@math.tugraz.at

Graz University of Technology, Institute of Analysis and Number Theory, Steyrergasse 30/II,
8010 Graz, Austria