LITTLEWOOD AND DUFFIN–SCHAEFFER-TYPE PROBLEMS IN DIOPHANTINE APPROXIMATION

SAM CHOW AND NICLAS TECHNAU

Dedicated to Andy Pollington

Abstract. Gallagher's theorem describes the multiplicative diophantine approximation rate of a typical vector. We establish a fully-inhomogeneous version of Gallagher's theorem, a diophantine fibre refinement, and a sharp and unexpected threshold for Liouville fibres. Along the way, we prove an inhomogeneous version of the Duffin–Schaeffer conjecture for a class of non-monotonic approximation functions.

Table of Contents

2020 *Mathematics Subject Classification.* 11J83, 11J54, 11H06, 52C07, 11J70.
Key words and phrases. Metric diophantine approximation, geometry of numbers, additive combinatorics, continued fractions.

15 Jul 2021 16:34:42 PDT Alg+NumTh
201210-Technau Version 2 - Submitted to Memoirs of the AMS

1. Introduction

This manuscript concerns two fundamental problems in diophantine approximation. We introduce a method to tackle, in a general context, inhomogeneous versions of Littlewood's conjecture which are metric in at least one parameter. Further, we prove an inhomogeneous version of the Duffin–Schaeffer conjecture for a relatively large class of functions.

Let us begin by explaining the link to Littlewood's conjecture, and defer elaborating on the inhomogeneous Duffin–Schaeffer conjecture to Section 1.2. Around 1930, Littlewood raised the question of whether planar badly approximable vectors exist in a multiplicative sense: That is, if for all $(\alpha_1, \alpha_2) \in \mathbb{R}^2$, we have

$$\liminf_{n\to\infty} n\|n\alpha_1\| \cdot \|n\alpha_2\| = 0, \tag{1.1}$$

where $\|\cdot\|$ denotes the distance to the nearest integer. Until the time of writing, finding non-trivial examples of (α_1, α_2) satisfying (1.1), barring rather special cases, evades the best efforts of the mathematical community. For instance, the problem remains open even for $(\alpha_1, \alpha_2) = (\sqrt{2}, \sqrt{3})$. Here non-trivial means that α_1, α_2 lie in the set

$$\mathrm{Bad} = \{\beta \in \mathbb{R} : \exists_{c>0} \quad n\|n\beta\| > c \quad \text{for all}\, n \in \mathbb{N}\}$$

of *badly approximable numbers*. This set has Lebesgue measure zero, by the Borel–Cantelli lemma, but full Hausdorff dimension, by the Jarník–Besicovitch theorem [12, Theorem 3.2].

Remark 1.1. By Dirichlet's approximation theorem, the inequality $n\|n\beta\| < 1$ holds infinitely often for *any* $\beta \in \mathbb{R}$. So badly approximable numbers are precisely those numbers for which Dirichlet's approximation theorem is optimal, up to a constant.

The study of the measure theory and fractal geometry centering around (1.1) has turned out to be a fruitful endeavour. Indeed, it has led to various exciting developments in homogeneous dynamics and in diophantine approximation. We presently expound upon this.

Homogeneous dynamics. Let $k \in \mathbb{N}$, define

$$G = \mathrm{SL}_{k+1}(\mathbb{R}), \qquad \Gamma = \mathrm{SL}_{k+1}(\mathbb{Z}),$$

and let D be the subgroup of diagonal matrices in G. There is a classical correspondence—the *Dani correspondence* [23]—between the diophantine properties of in $(\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$ and the dynamical properties of the orbit of

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \alpha_1 \\ & 1 & & & \alpha_2 \\ & & \ddots & & \vdots \\ & & & 1 & \alpha_k \\ & & & & 1 \end{pmatrix} \in G/\Gamma$$

under the action of D. By virtue of Dani's correspondence, Littlewood's conjecture is closely linked to a conjecture of Margulis [45, Conjecture 1]. Below we state a special case, namely [45, Conjecture 9].

Conjecture 1.2. *Let $k \geqslant 2$, and let G, Γ, D be as above. Let $z \in G/\Gamma$, and assume that $Dz \subseteq G/\Gamma$ has compact closure. Then Dz is closed.*

If true, Conjecture 1.2 would imply Littlewood's conjecture. See [45] for further details.

With this dynamical perspective, Einsiedler, Katok, and Lindenstrauss [26] established, *inter alia*, the striking result that the set of putative counterexamples to Littlewood's conjecture (1.1) has Hausdorff dimension zero. A crucial ingredient was a deep result of Ratner.

From a similar departure point, Shapira [53] established a measure-theoretic, uniform version of an inhomogeneous Littlewood-type problem, solving an old problem of Cassels. To state it, we stress that 'for almost all' (and similarly for 'almost every') in this manuscript always means with respect to the Lebesgue measure on the ambient space, expressing that the complement of the set under consideration has Lebesgue measure zero. Shapira proved that for almost all $(\alpha, \beta) \in \mathbb{R}^2$ the relation

$$\liminf_{n\to\infty} n\|n\alpha - \gamma\| \cdot \|n\beta - \delta\| = 0$$

holds for any $\gamma, \delta \in \mathbb{R}$. Gorodnik and Vishe [33] improved this by a factor of $(\log\log\log\log\log n)^{\lambda}$, for some constant $\lambda > 0$.

While the above results suggest that Littlewood's conjecture might be correct, there is an indication against it: Adiceam, Nesharim, and Lunnon [1] proved that a certain function field analogue of the Littlewood conjecture is false.

In what follows, the word metric is used to mean 'measure-theoretic', as is customary in this area. The goal of metric number theory is to classify behaviour up to exceptional sets of measure zero.

Metric multiplicative diophantine approximation. Alongside the theory of homogeneous dynamics linked to Littlewood's conjecture, there have been significant advances towards the corresponding metric theory. The first systematic result in this direction is a famous theorem of Gallagher [31]:[1] For any non-increasing $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ and almost all $(\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$, we have

$$\|n\alpha_1\| \cdots \|n\alpha_k\| < \psi(n) \tag{1.2}$$

[1]The convergence part was already known, see [14, Remark 1.2]. For this reason, *Gallagher's theorem* sometimes refers to the divergence part alone.

infinitely often if the relevant series of measures diverges, that is, if

$$\sum_{n\geqslant 1} \psi(n)(\log n)^{k-1} = \infty. \tag{1.3}$$

Consequently (1.1) holds true for almost every $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ by a log-squared margin:

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1\| \cdot \|n\alpha_2\| = 0.$$

Continuing this line of research, Pollington and Velani [49] showed the following fibre statement, by exploiting the Fourier decay property of a certain fractal measure: If $\alpha_1 \in \mathrm{Bad}$ then there is a set of numbers $\alpha_2 \in \mathrm{Bad}$, of full Hausdorff dimension, such that

$$n \log n \|n\alpha_1\| \cdot \|n\alpha_2\| < 1$$

holds for infinitely many $n \in \mathbb{N}$.

A further fibre statement concerning (1.1) was established by Beresnevich, Haynes, and Velani in [11, Theorem 2.4]. To state it, we recall that *Liouville numbers* are irrational real numbers α such that for any $w > 0$ the inequality

$$\|n\alpha\| < n^{-w}$$

holds infinitely often. We denote the set of Liouville numbers by $\mathcal{L}$.

Theorem 1.3 ([11, Theorem 2.4])**.** *Let $\alpha_1 \in \mathbb{R} \setminus (\mathbb{Q} \cup \mathcal{L})$ and $\gamma \in \mathbb{R}$. If the Duffin–Schaeffer conjecture is true, then for almost all $\alpha_2 \in \mathbb{R}$ we have*

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1 - \gamma\| \cdot \|n\alpha_2\| = 0.$$

When this theorem was proved, the Duffin–Schaeffer conjecture was still open. The former was then proved, without appealing to the Duffin–Schaeffer conjecture, by the first named author [19] in a stronger form:

Theorem 1.4 ([19])**.** *Let $\alpha_1, \gamma \in \mathbb{R}$ and assume that $\alpha_1 \in \mathbb{R} \setminus (\mathbb{Q} \cup \mathcal{L})$. If $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ is non-increasing and the series*

$$\sum_{n\geqslant 1} \psi(n) \log n \tag{1.4}$$

diverges, then for almost all $\alpha_2 \in \mathbb{R}$ there exist infinitely many $n \in \mathbb{N}$ such that

$$\|n\alpha_1 - \gamma\| \cdot \|n\alpha_2\| < \psi(n).$$

The proof used the structural theory of Bohr sets (see Section 1.2), as well as continued fractions and the geometry of numbers, in a crucial way. This

combinatorial–geometric method was further developed by the authors [21] to prove higher-dimensional results, removing the reliance on continued fractions. Instead, a more versatile framework from the geometry of numbers was brought to bear on the problem. Another interesting facet of the approach is the application of *diophantine transference inequalities* [13, 16, 17, 20, 32, 40] to deal with the inhomogeneous shifts.

By fixing α_1 above, one considers pairs (α_1, α_2) which lie on a vertical line in the plane. With Yang, the first named author showed in [22] that if $\mathcal{L}_0$ is an arbitrary line in the plane then for almost all $(\alpha_1, \alpha_2) \in \mathcal{L}_0$ we have

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1\| \cdot \|n\alpha_2\| = 0.$$

Writing

$$\mathcal{L}_0 = \{(\alpha, \beta) \in \mathbb{R}^2 : \alpha = a\beta + b\},$$

and assuming that that pair (a, b) satisfies the Lebesgue-generic condition

$$\sup\{w \in \mathbb{R} : \exists^\infty (x, y) \in \mathbb{Z}^2 \quad \|xa + yb\| < (|x| + |y|)^{-w}\} < 5, \tag{1.5}$$

it was also shown there that if $\psi : \mathbb{N} \to [0, \infty)$ is non-increasing and

$$\sum_{n\geqslant 1} \psi(n) \log n < \infty$$

then for almost all $(\alpha_1, \alpha_2) \in \mathcal{L}_0$ the inequality

$$\|n\alpha_1\| \cdot \|n\alpha_2\| < \psi(n)$$

has at most finitely many solutions $n \in \mathbb{N}$. The divergence theory was attained via an effective asymptotic equidistribution theorem for unipotent orbits in $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$, whilst the convergence statement involved the correspondence between Bohr sets and generalised arithmetic progressions. All of this sits within the broader context of *metric diophantine approximation on manifolds*, for which there is a vast literature [7, 9, 37, 38, 41, 59].

A common feature of these results is that they are homogeneous in the metric parameter, i.e. they involve $\|n\alpha_2\|$ but not $\|n\alpha_2 - \gamma\|$ with a general parameter γ. Even a weak inhomogeneous version of Gallagher's theorem, akin to Shapira's [53, Theorem 1.2], remains completely open, despite numerous attempts. In light of this, Beresnevich, Haynes, and Velani [11, Problem 2.3] posed the following problem:

Problem 1.5 (A fully-inhomogeneous version of Gallagher's theorem on vertical planar lines, weak form)**.** *Let $\alpha_1, \gamma_1, \gamma_2 \in \mathbb{R}$, and suppose that $\alpha_1 \notin \mathcal{L} \cup \mathbb{Q}$. Prove that*

$$\liminf_{n\to\infty} n\,(\log n)^2 \|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| = 0 \quad \textit{for almost all } \alpha_2 \in \mathbb{R}.$$

They write in the paragraph leading up to [11, Problem 2.3] concerning this problem that it "currently seems well out of reach". Nevertheless, a stronger conjecture was put forth by the first named author [19, Conjecture 1.6]:

Conjecture 1.6 (A fully-inhomogeneous version of Gallagher's theorem on vertical planar lines, strong form)**.** *Let $\alpha_1, \gamma_1, \gamma_2$ be as in Problem 1.5. Suppose $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ is non-increasing and that the series* (1.4) *diverges. Then for almost all $\alpha_2 \in \mathbb{R}$ there exist infinitely many $n \in \mathbb{N}$ such that*

$$\|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| < \psi(n).$$

This manuscript resolves Problem 1.5 *a fortiori* by proving Conjecture 1.6. Additionally, our methods are capable of deducing a higher-dimensional generalisation, as conjectured by the authors in [21, Conjecture 1.7]. Furthermore, we resolve Conjecture 2.1 of Beresnevich, Haynes, and Velani [11]:

Conjecture 1.7 (A fully-inhomogeneous version of Gallagher's theorem in the plane)**.** *Let $\gamma_1, \gamma_2, \in \mathbb{R}$, and let $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ be a non-increasing function such that the series* (1.4) *diverges. Then for almost all $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ the inequality*

$$\|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| < \psi(n)$$

holds infinitely often.

Note that Conjecture 1.6 implies Conjecture 1.7, since $\mathcal{L} \cup \mathbb{Q}$ has Lebesgue measure zero. We now formulate our results in greater detail.

1.1. **Main results.** Recall that the *multiplicative exponent* $\omega^\times(\boldsymbol{\alpha})$ of a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ is the supremum of all $w > 0$ such that

$$\|n\alpha_1\| \cdots \|n\alpha_d\| < n^{-w}$$

infinitely often. The property of $\omega^\times(\boldsymbol{\alpha})$ being finite can be interpreted as a higher-dimensional generalisation of being an irrational, non-Liouville number.

Theorem 1.8 (A fully-inhomogeneous version of Gallagher's theorem on vertical lines, strong form)**.** *Let $k \geqslant 2$. Fix $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{k-1}) \in \mathbb{R}^{k-1}$ and $\gamma_1, \ldots, \gamma_k \in \mathbb{R}$. For $k = 2$, suppose that α_1 is an irrational, non-Liouville number, and for $k \geqslant 3$ suppose that*

$$\omega^\times(\boldsymbol{\alpha}) < \frac{k-1}{k-2}. \tag{1.6}$$

If $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ is non-increasing and satisfies (1.3)*, then for almost all $\alpha_k \in \mathbb{R}$ there are infinitely many $n \in \mathbb{N}$ for which*

$$\|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_k - \gamma_k\| < \psi(n). \tag{1.7}$$

Remark 1.9. The set of $\boldsymbol{\alpha} \in \mathbb{R}^{k-1}$ for which

$$\omega^\times(\boldsymbol{\alpha}) \geqslant \frac{k-1}{k-2}$$

has Lebesgue measure zero and, stronger still, has Hausdorff dimension strictly less than $k-1$. The former follows from the Borel–Cantelli lemma and the latter from the work of Hussain and Simmons [39]. The set of Liouville numbers has Lebesgue measure zero and, stronger still, has Hausdorff dimension 0.

Theorem 1.8 is precisely [21, Conjecture 1.7], and we have the following noteworthy special cases.

Corollary 1.10. *Conjectures 1.6 and 1.7 are true.*

We have the following generalisation of Conjecture 1.7, which includes its complementary convergence theory.

Corollary 1.11 (A fully-inhomogeneous version of Gallagher's theorem)**.** *Let $\gamma_1, \dots, \gamma_k \in \mathbb{R}$, and let $\psi : \mathbb{N} \to (0, \infty)$ be a non-increasing function. Write $\mathcal{W}^\times = \mathcal{W}^\times(\psi, \gamma_1, \dots, \gamma_k)$ for the set of $(\alpha_1, \dots, \alpha_k) \in [0,1]^k$ such that* (1.7) *has infinitely many solutions $n \in \mathbb{N}$. Then*

$$\mu_k(\mathcal{W}^\times) = \begin{cases} 1, & \text{if } \displaystyle\sum_{n=1}^{\infty} \psi(n)(\log n)^{k-1} = \infty \\ 0, & \text{if } \displaystyle\sum_{n=1}^{\infty} \psi(n)(\log n)^{k-1} < \infty, \end{cases}$$

where μ_k denotes k-dimensional Lebesgue measure.

The divergence part of Corollary 1.11 is a consequence of Theorem 1.8 and Remark 1.9. The convergence part requires only classical techniques, and will be proved in Section 3.5. Theorem 1.8 also resolves Problem 1.5 in the following stronger and more general form:

Corollary 1.12. *Let $k \geqslant 2$. Fix a fibre vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{k-1}) \in \mathbb{R}^{k-1}$, and shifts $\gamma_1, \dots, \gamma_k \in \mathbb{R}$. For $k = 2$, suppose that α_1 is an irrational, non-Liouville*

number, and for $k \geqslant 3$ assume (1.6). *Then for almost all $\alpha_k \in \mathbb{R}$ there are infinitely many $n \in \mathbb{N}$ for which*

$$\|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_k - \gamma_k\| < \frac{1}{n(\log n)^k \log\log n}.$$

When $\gamma_k = 0$, the results above follow from [19]. In the planar case, we go beyond the scope of Problem 1.5. Indeed, we also solve it on fibres $(\alpha_1, \mathbb{R})$ where α_1 is a Liouville number:

Theorem 1.13. *Let $\alpha_1 \in \mathcal{L}$, and let $\gamma_1, \gamma_2 \in \mathbb{R}$. Then, for almost all $\alpha_2 \in \mathbb{R}$, we have*

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| = 0.$$

In view of Theorem 1.8, we see that this result holds for any irrational α_1, Liouville or not. To be clear, we obtain the following statement.

Theorem 1.14. *Let $\alpha_1 \in \mathbb{R} \setminus \mathbb{Q}$, and let $\gamma_1, \gamma_2 \in \mathbb{R}$. Then, for almost all $\alpha_2 \in \mathbb{R}$, we have*

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| = 0.$$

However, if $\alpha_1 \in \mathbb{Q}$, then one can easily construct a counterexample by choosing any $\gamma_1 \notin \alpha_1\mathbb{Z}$ and applying Szüsz's theorem [54], see Theorem 1.17. In this sense, and in the sense described in the next two paragraphs, Theorem 1.14 is definitive.

The analysis on Liouville fibres is delicate, owing to the erratic behaviour of the arising sums of reciprocals of fractional parts [11]. In light of our earlier discussion on Problem 1.5 and Conjecture 1.6, one might expect Theorem 1.13 not to be sharp. Surprisingly, the result is sharp, as we now detail.

Let us ask for a strengthening of Theorem 1.13 by considering an approximation function with a faster decay, say

$$\psi_\xi(n) = \frac{1}{n(\log n)^2 \xi(n)}, \tag{1.8}$$

where $\xi : \mathbb{N} \to [1, \infty)$ is an unbounded and non-decreasing function. Then the strengthened statement of Theorem 1.13, with $\psi(n) = \psi_\xi(n)$, is false:

Theorem 1.15. *Let $\xi : \mathbb{N} \to [1, \infty)$ be non-decreasing and unbounded. Then there are continuum many pairs $(\alpha_1, \gamma_1) \in \mathcal{L} \times \mathbb{R}$ such that for any $\gamma_2 \in \mathbb{R}$ and almost all $\alpha_2 \in \mathbb{R}$ the inequality*

$$\|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| < \psi_\xi(n)$$

has at most finitely many solutions $n \in \mathbb{N}$.

Remark 1.16. (1) One could regard this as a result 'in the opposite direction' to Littlewood's conjecture, though there are several differences. A volume heuristic, and the works of Peck [47] and Pollington–Velani [49], suggest that (1.1) can be strengthened by roughly a logarithm—see the discussion in [5]—and meanwhile Badziahin [4] has shown that it cannot be strengthened much further than that. For this reason, we consider that results in the opposite direction to Littlewood's conjecture are worthy of further study.

(2) In the course of our proof, we explicitly construct the pairs (α_1, γ_1). This feature is often not present in results of this flavour.

To conclude this discussion, Theorem 1.14 is sharp, and has no unnecessary restrictions on α_1, γ_1, and γ_2.

By work of Beresnevich and Velani [14, Section 1], as well as Hussain and Simmons [39], 'fractal' Hausdorff measures are known to be insensitive to the multiplicative nature of these types of problems. Fix $k \geqslant 2$. For $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ non-increasing with $\lim_{n\to\infty} \psi(n) = 0$, and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k) \in \mathbb{R}^k$, denote by $\mathcal{W}_k^\times(\psi, \boldsymbol{\gamma})$ the set of $(\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$ satisfying (1.7) for infinitely many n. Further, denote by $\mathcal{W}_k(\psi, \boldsymbol{\gamma})$ the set of $(\alpha_1, \ldots, \alpha_k) \in [0, 1]^k$ for which

$$\max(\|n\alpha_1 - \gamma_1\|, \ldots, \|n\alpha_k - \gamma_k\|) < \psi(n)$$

has infinitely many solutions $n \in \mathbb{N}$. By [39, Corollary 1.4] and [12, Theorem 6.1], for $\boldsymbol{\gamma} \in \mathbb{R}^k$ we have the Hausdorff measure identity

$$H^s(\mathcal{W}_k^\times(\psi, \boldsymbol{\gamma})) = H^{s-(k-1)}(\mathcal{W}_1(\psi, \boldsymbol{\gamma})) \qquad (k - 1 < s < k). \tag{1.9}$$

We interpret from this that multiplicatively approximating k reals using the same denominator is no different to approximating one of the k numbers, except possibly for a set of zero Hausdorff s-measure. This behaviour differs greatly from that of the Lebesgue case $s = k$, where there are extra logarithms for the multiplicative problem. As explained in [14, 39], in the remaining ranges for s the Hausdorff theory trivialises: if $s > k$ then $H^s(\mathcal{W}_k^\times(\psi), \boldsymbol{\gamma}) = 0$, irrespective of ψ, whereas if $s \leqslant k - 1$ then $H^s(\mathcal{W}_k^\times(\psi, \boldsymbol{\gamma})) = \infty$.

1.2. **Key ideas and further results.**

A fully-inhomogeneous fibre refinement of Gallagher's theorem. Owing to the robustness of our method, much of the argument for Theorem 1.8 is transparent already in the planar case $k = 2$. As it is simpler from a technical point of view, we shall outline the proof only in this case, and indicate in passing how to generalise to higher dimensions. To begin, let us isolate the metric parameter $\alpha = \alpha_2$ in (1.7) on the left hand side of the inequality. As $\|\cdot\|$ is 1-periodic, it suffices if we show

$$\|n\alpha - \gamma\| < \Phi(n) := \frac{\psi(n)}{\|n\alpha_1 - \gamma_1\|}, \qquad \gamma := \gamma_2, \tag{1.10}$$

holds for almost every $\alpha \in [0,1]$ infinitely often. If Φ were a non-increasing function, then we could utilise Szüsz's extension of Khintchine's theorem, which grants a sharp description of the inhomogeneous approximation rate of a generic real number:

Theorem 1.17 (Szüsz [54])**.** *If $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$ is non-increasing and $\gamma \in \mathbb{R}$, then the Lebesgue measure of the set of $\alpha \in [0,1]$ for which*

$$\|n\alpha - \gamma\| < \Psi(n) \tag{1.11}$$

holds infinitely often is 1 (resp. 0) if

$$\sum_{n \geqslant 1} \Psi(n) \tag{1.12}$$

diverges (resp. converges).

Since Φ is very much not monotonic, deducing (1.10) infinitely often for almost all α, from the divergence of $\sum_n \Phi(n)$, is far more demanding. In fact, it is known that this naive condition is insufficient, as Duffin and Schaeffer [25] pointed out: There exists $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$ such that (1.12) diverges but for $\gamma = 0$ and almost all $\alpha \in [0,1]$ the inequality (1.11) holds at most finitely often. This was generalised by Ramírez in [50].

To circumvent their counterexamples, Duffin and Schaeffer restricted attention to reduced fractions, and correspondingly imposed the condition that the series

$$\sum_{n \geqslant 1} \Psi(n) \frac{\varphi(n)}{n} \tag{1.13}$$

diverges, where φ is Euler's totient function. The Duffin–Schaeffer conjecture was a major open problem in diophantine approximation since the 1940s. Over the course of the nearly eight decades, various partial results towards the Duffin–Schaeffer conjecture were obtained, by

- Erdős [27] in 1970, Vaaler [58] in 1978
- Pollington–Vaughan [48] in 1990
- Harman [35] in 1990, Haynes–Pollington–Velani [36] in 2012,

and several other authors [2, 3, 10]. Recently, Koukoulopoulos and Maynard (2019) broke through with a complete proof of the Duffin–Schaeffer conjecture:

Theorem 1.18 (Koukoulopoulos–Maynard [43])**.** *If $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$ is such that* (1.13) *diverges then, for almost all $\alpha \in [0,1]$, the inequality*

$$|n\alpha - a| < \Psi(n)$$

holds for infinitely many coprime $a, n \in \mathbb{N}$.

A natural generalisation would be an inhomogeneous version of the Duffin–Schaeffer conjecture:

Conjecture 1.19 (Inhomogeneous Duffin–Schaeffer, see Ramírez [50])**.** *Let $\gamma \in \mathbb{R}$. If $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$ is such that* (1.13) *diverges, then for almost all $\alpha \in [0,1]$ the inequality*

$$|n\alpha - a - \gamma| < \Psi(n)$$

holds for infinitely many coprime $a, n \in \mathbb{N}$.

For us it is enough to establish a similar result for a concrete class of functions of the shape (1.10). We consider sets $\mathcal{A}_n$ that are roughly of the form

$$\{\alpha \in [0,1] : \|n\alpha - \gamma\| < \Phi(n)\}.$$

By standard probabilistic arguments, it suffices to show that the measures of the sets $\mathcal{A}_n$ are not summable, and that the sets are quasi-independent in an averaged sense. The latter property is, as always, the crux of the matter, and involves *overlap estimates* that quantify how large $\mu(\mathcal{A}_n \cap \mathcal{A}_m)$ is compared to $\mu(\mathcal{A}_n)\mu(\mathcal{A}_m)$ on average.

To this end, we may confine our analysis to a reasonably large set $\mathcal{G}$ of 'good' indices n. To simplify matters, we decompose $\mathbb{N}$ into dyadic ranges, wherein $n \asymp N$, and in addition the oscillating factor $\|n\alpha_1 - \gamma_1\|$ has a fixed order of magnitude. Sets of such integers n are essentially *Bohr sets*

$$\mathcal{B} = \mathcal{B}_{\alpha_1}^{\gamma_1}(N; \rho_1) := \{n \in \mathbb{Z} : |n| \leqslant N, \|n\alpha_1 - \gamma_1\| \leqslant \rho_1\}, \tag{1.14}$$

which appear in many areas of mathematics. A novelty of this paper is to show how to handle the overlap estimates via congruences in Bohr sets. Here the structural theory from our previous work [21], constructing the correspondence between Bohr sets and generalised arithmetic progressions—a central pillar of additive combinatorics [55]—in the present context, plays a pivotal role. Previously there was progress made in this direction by Tao–Vu [56], and by the first named author [19]. Further, it will be helpful to group m, n according to the size of the greatest common divisor d of m and n.

We handle the overlap estimates by averaging over indices m, n from different Bohr sets of the shape (1.14). After summing over different dyadic ranges, we can then infer the required quasi-independence on average. In the course of our analysis, we need to count solutions to congruences in generalised arithmetic progressions that are essentially Bohr sets. The range in which d is large requires extra care: For $\gamma \notin \mathcal{L} \cup \mathbb{Q}$, a repulsion stemming from this diophantine assumption enables us to treat this challenging case. For $\gamma \in \mathcal{L} \cup \mathbb{Q}$, an additional argument enables us to crack this devilish final case; the idea is to introduce a counterpart to reduced fractions, which we call 'shift-reduced'.

Definition 1.20. Let $\gamma \in \mathbb{R}$, $\eta \in (0,1)$, and $n \in \mathbb{N}$. Let $c_0/q_0', c_1/q_1', \dots$ be the continued fraction convergents of γ, see Subsection 2.1 for what these are.[2] Denote by c_t/q_t' the continued fraction convergent of γ for which t is maximal satisfying

$$q_t' \leqslant n^\eta. \tag{1.15}$$

The pair $(a, n) \in \mathbb{Z} \times \mathbb{N}$ is (γ, η)-*shift-reduced* if $(q_t' a + c_t, n) = 1$.

To our knowledge, this notion has not hitherto appeared in the diophantine approximation literature.

Remark 1.21. Note that γ can be a rational in the above definition, in which case the sequence $(q_t')_t$ terminates. Indeed, in the case $\gamma = 0$, we recover the traditional notion of reduced fractions: Letting $c_t = 0$ and $q_t' = 1$, the fraction a/n is reduced if and only if the pair (a, n) is $(0, 1/2)$-shift-reduced. Moreover, if $\gamma \in \mathbb{Q}$ and n^η is greater than or equal to the denominator of γ, then $\gamma = c_t/q_t'$. We provide some background on continued fractions in Section 2.1.

[2]In this paper, we will employ the continued fraction expansions of quantities denoted by α and γ. We reserve the more common notation $p_0/q_0, p_1/q_1, \dots$ for the continued fraction convergents of a quantity denoted by α, see Sections 4 and 5.

Our definition of shift-reduced fractions is sensitive to the shift γ. This finesse slightly complicates matters, because we lose measure by not using all fractions. However, by sieve theory we are able to show that shift-reduced fractions are at least as prevalent as reduced fractions, and we are then able to establish the divergence of the relevant series. We suspect that the idea of using shift-reduced fractions will be useful for other arithmetic problems, including perhaps a version of the inhomogeneous Duffin–Schaeffer conjecture, as we will shortly discuss further.

Yet all of this combined yields only that (1.7) holds infinitely often on a set of positive measure. In the absence of a zero–one law for inhomogeneous diophantine approximation, we need to carefully 'localise' the overlap estimates to indeed deduce that (1.7) holds for a set of α_k of full measure; this machinery, though not confined to the realm of metric diophantine approximation, is well-explained in the monograph of Beresnevich, Dickinson, and Velani [8]. This subtlety complicates the analysis non-trivially, as it requires us to keep hold of a factor of $\mu(\mathcal{I})$ throughout. In the regime of m, n in which d is essentially constant it turns out to be surprisingly difficult to do so. To avoid this scenario, we introduce artificial powers of 4 as divisors, which guarantees us that d is not too small. This is an unorthodox manoeuvre, but one that we found useful in practice, and one that may find other uses.

In the course of our proof, we establish the following weakened version of Conjecture 1.19 for a class of non-monotonic approximating functions generalising Φ as defined in (1.10).

Conjecture 1.22 (Weak inhomogeneous Duffin–Schaeffer conjecture)**.** *Let $\gamma \in \mathbb{R}$. If $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$ is such that* (1.13) *diverges, then for almost all $\alpha \in [0, 1]$ the inequality*

$$\|n\alpha - \gamma\| < \Psi(n)$$

has infinitely many solutions $n \in \mathbb{N}$.

The $\gamma = 0$ case of this was coined the *weak Duffin–Schaeffer conjecture* by Ramírez [51, Conjecture 9.1]. The word 'weak' is used for the following reason: Euler's totient function is present in the divergence hypothesis despite there being no coprimality aspect.

Theorem 1.23 (Special case of weak inhomogeneous Duffin–Schaeffer)**.** *Let $k \geqslant 2$. Fix $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{k-1}) \in \mathbb{R}^{k-1}$ and $\gamma_1, \ldots, \gamma_k \in \mathbb{R}$. For $k = 2$, suppose that α_1 is an irrational, non-Liouville number, and for $k \geqslant 3$ assume* (1.6)*.*

Let $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ be a non-increasing function satisfying (1.3)*, and let*

$$\Phi(n) = \frac{\psi(n)}{\|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_{k-1} - \gamma_{k-1}\|} \qquad (n \in \mathbb{N}). \tag{1.16}$$

Then Conjecture 1.22 holds for $\Psi = \Phi$.

We wonder if there is a sharp dichotomy along the lines of Conjecture 1.22.

Question 1.24 (Inhomogeneous Duffin–Schaeffer dichotomy)**.** *Let $\gamma \in \mathbb{R}$ and $\Psi : \mathbb{N} \to \mathbb{R}_{>0}$. Does there exist $\eta \in (0,1)$ with the following property? Denote by $\mathcal{W}(\Psi;\gamma,\eta)$ the set of $\alpha \in [0,1]$ such that*

$$|n\alpha - \gamma - a| < \Psi(n), \qquad (a,n) \text{ is } (\gamma,\eta)\text{-shift-reduced}$$

has infinitely many solutions $(a,n) \in \mathbb{Z} \times \mathbb{N}$. Then

$$\mu(\mathcal{W}(\Psi;\gamma,\eta)) = \begin{cases} 1, & \text{if } \displaystyle\sum_{n \geqslant 1} \frac{\varphi_{\gamma,\eta}(n)}{n}\Psi(n) = \infty \\ 0, & \text{if } \displaystyle\sum_{n \geqslant 1} \frac{\varphi_{\gamma,\eta}(n)}{n}\Psi(n) < \infty, \end{cases}$$

where

$$\varphi_{\gamma,\eta}(n) = \#\{a \in \{1,2,\dots,n\} : (a,n) \text{ is } (\gamma,\eta)\text{-shift-reduced}\}.$$

Remark 1.25. It is not clear at this stage whether η should need to depend on γ or Ψ. Moreover, it could be that the property holds for all $\eta \in (0,\eta_0)$, for some $\eta_0 \in (0,1]$. This question has the appeal of a matching convergence theory, *unlike* Conjecture 1.22.

We are also able to answer Question 1.24 positively for $\Psi = \Phi$, subject to natural assumptions:

Theorem 1.26 (Special case of inhomogeneous Duffin–Schaeffer dichotomy)**.** *Let $k \geqslant 2$. Fix $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_{k-1}) \in \mathbb{R}^{k-1}$ and $\gamma_1,\dots,\gamma_k \in \mathbb{R}$. For $k = 2$, suppose that α_1 is an irrational, non-Liouville number, and for $k \geqslant 3$ assume* (1.6)*. Let $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ be a non-increasing function satisfying* (1.3)*, and let Φ be as in* (1.16)*. Then, with $\gamma = \gamma_k$ and the notation of Question 1.24, there exists $\eta \in (0,1)$ such that*

$$\sum_{n=1}^{\infty} \frac{\varphi_{\gamma,\eta}(n)}{n}\Phi(n) = \infty \tag{1.17}$$

and

$$\mu(\mathcal{W}(\Phi;\gamma,\eta)) = 1. \tag{1.18}$$

In particular, Question 1.24 has a positive answer for $\Psi = \Phi$.

Our method comes close to directly establishing Theorem 1.26. However, it misses a pathological case where an auxiliary function exceeds 1/2 infinitely often, causing the relevant intervals to overlap and their union to have smaller measure. We are able to circumvent this by ad-hoc means, and provide the details in an appendix.

Inhomogeneous, non-monotonic diophantine approximation is also discussed in Harman's book [34, Chapter 3], as well as in recent work of Yu [60, 61].

Our method is robust with respect to the dimension k. If $k \geqslant 3$, the combinatorics for controlling the overlap estimates is relatively similar to the planar case, $k = 2$. The notable differences are that there are more dyadic ranges to sum over and the choice of cutoff parameters needs to be adapted.

An inhomogeneous version of Gallagher's theorem on Liouville fibres. For Theorem 1.13, the overall structure of our proof parallels that of the Duffin–Schaeffer theorem [34, Theorem 2.3], which is a special case of the Duffin–Schaeffer conjecture. This naturally leads us to count solutions to congruences in Bohr sets. In this setting, the latter are rather sparse in the set of positive integers, which presents new difficulties. The partial quotients, see Section 2, grow extremely rapidly infinitely many times, and we use this together with classical continued fraction analysis to deal with the combinatorial aspects of the overlap estimates.

Liouville fibres: sharpness. The prove the sharpness result, Theorem 1.15, we construct pairs $(\alpha,\gamma) \in \mathcal{L} \times \mathbb{R}$ in such a way as to keep $\|n\alpha - \gamma\|$ away from zero. We achieve this via the Ostrowski expansion [11, Section 3], by choosing each Ostrowski coefficient of γ with respect to α to be roughly half times the corresponding partial quotient of α, and by choosing α to have extremely rapidly-growing partial quotients.

1.3. **Open problems.** We have already discussed a few open questions. Here are some that we have yet to cover.

Relaxing the diophantine condition. We expect that the assumption (1.6) in Theorem 1.8 can be somewhat relaxed. This likely requires different methods.

Convergence theory. It would be desirable to advance the convergence theory complementing Theorem 1.8, especially for $k \geqslant 3$. For $k = 2$, progress has been made by Beresnevich, Haynes, and Velani [11].

By partial summation and the Borel–Cantelli lemma, proving the desired convergence statement can be reduced to showing that

$$\sum_{n \leqslant N} \frac{1}{\|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_{k-1} - \gamma_{k-1}\|} \ll N(\log N)^{k-1}.$$

In the case $k = 2$ and $\gamma_1 = 0$, for a generic choice of $\alpha_1 \in \mathbb{R}$ this bound is false, see [11, Example 1.1]. However, the logarithmically averaged sums

$$S_{\boldsymbol{\alpha}}^{\boldsymbol{\gamma}}(N) = \sum_{n \leqslant N} \frac{1}{n\|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_{k-1} - \gamma_{k-1}\|}, \qquad \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{k-1})$$

could be better behaved, and accurately bounding these would lead to a similar outcome. On probabilistic grounds, we expect that

$$S_{\boldsymbol{\alpha}}^{\boldsymbol{\gamma}}(N) \ll_{\boldsymbol{\alpha},\boldsymbol{\gamma}} (\log N)^k.$$

Subject to a generic diophantine condition on $\boldsymbol{\alpha}$, it is less difficult to prove matching lower bounds via dyadic pigeonholing and estimates for the cardinality of a Bohr set, see [21, Lemma 3.1]. So the task is to determine the order of magnitude of $S_{\boldsymbol{\alpha}}^{\boldsymbol{\gamma}}(N)$. Beresnevich, Haynes and Velani [11, Theorem 1.4] showed that if $k = 2$ and $\gamma \in \mathbb{R}$ then $S_{\alpha}^{\gamma}(N) \ll_{\alpha,\gamma} (\log N)^2$ for almost every $\alpha \in \mathbb{R}$. With this in mind, we pose the following problem which, if resolved in a sufficiently positive manner, would entail a coherent convergence theory.

Problem 1.27. *Let $k \geqslant 3$, and let $\mathcal{C}_k$ be the set of $(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$ for which*

$$S_{\boldsymbol{\alpha}}^{\boldsymbol{\gamma}}(N) \ll_{\boldsymbol{\alpha},\boldsymbol{\gamma}} (\log N)^k.$$

Is the set $\mathcal{C}_k$ non-empty? What is its Lebesgue measure?

Empirical evidence mildly supports the assertion that $S_{\boldsymbol{\alpha}}^{\mathbf{0}}(N) \ll_{\boldsymbol{\alpha}} (\log N)^3$ holds for generic values of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$. We randomly generated

$$\alpha_1 \approx 0.957363115715396, \qquad \alpha_2 \approx 0.3049448415027476.$$

With

$$H = 10^6, \qquad c = \frac{S_{\boldsymbol{\alpha}}^{\mathbf{0}}(H)}{(\log H)^3} \approx 1.73475,$$

we plotted $S_{\boldsymbol{\alpha}}^{\mathbf{0}}(N)$ and $c(\log N)^3$ against N for $N = 2, 3, \ldots, H$, see Figure 1. There are 'jumps' when $\|n\alpha_1\| \cdot \|n\alpha_2\|$ is very small, but these do not appear to affect the order of magnitude of $S_{\boldsymbol{\alpha}}^{\mathbf{0}}(N)$. For further discussion, we refer the reader to the article by Lê and Vaaler [44].

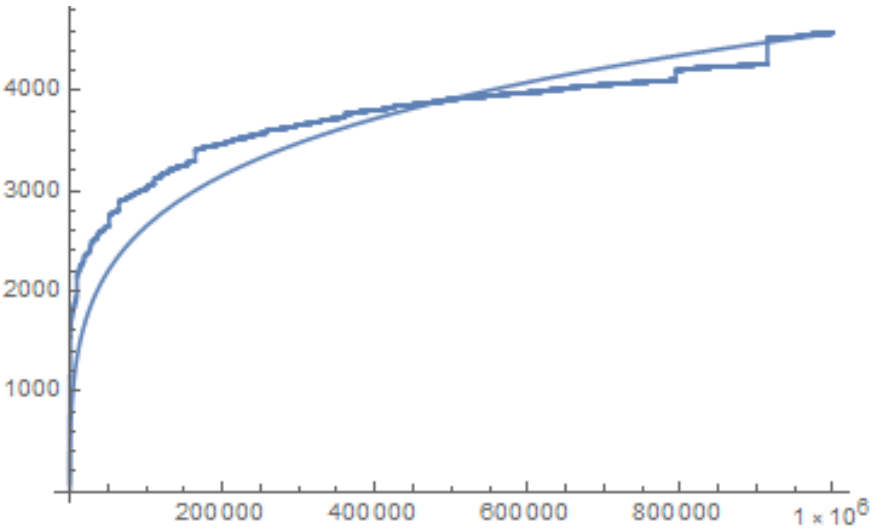

FIGURE 1. $S^{\mathbf{0}}_{\boldsymbol{\alpha}}(N)$ and $c(\log N)^3$ against N

Dual approximation. There are natural dual versions of Gallagher's theorem. Loosely speaking, in the dual framework one studies how close a given vector is to a hyperplane of bounded height, as the height bound increases. In fact, we intend to address the next problem in a future work.

Conjecture 1.28. *Let $k \geqslant 2$, and $(\alpha_1, \dots, \alpha_{k-1}) \in \mathbb{R}^{k-1}$. Then for almost all $\alpha_k \in \mathbb{R}$ there exist infinitely many $(n_1, \dots, n_k) \in \mathbb{Z}^k$ such that*

$$\|n_1\alpha_1 + \cdots + n_k\alpha_k\| < \frac{1}{H(\mathbf{n})(\log H(\mathbf{n}))^k},$$

where $H(\mathbf{n}) = H(n_1, \dots, n_k) = n_1^+ \cdots n_k^+$ and $n^+ = \max(|n|, 2)$.

This is also [21, Conjecture 1.9], which comes with some further discussion.

The dual convergence theory is also of interest, that is, to show that the convergence of the series in (1.3) implies that for almost almost all α_k the inequality

$$\|n_1\alpha_1 + \cdots + n_k\alpha_k\| < \psi(H(\mathbf{n}))$$

holds at most finitely often. As with the usual multiplicative approximation problems described above, the convergence theory in the dual setting is very much open. There is a discussion of the relevant sums, as well as a reference to a possible departure point, in the paragraphs surrounding Conjecture 1.1 of Beresnevich, Haynes, and Velani [11].

1.3.1. *A Hausdorff dimension problem.* In Theorem 1.15, we did not estimate the Hausdorff dimension of the set of pairs $(\alpha_1, \gamma_1) \in \mathbb{R}^2$ such that for any $\gamma_2 \in \mathbb{R}$ and almost all $\alpha_2 \in \mathbb{R}$ the inequality

$$\|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| < \psi_\xi(n)$$

has at most finitely many solutions $n \in \mathbb{N}$. It would be of interest to do so, for ξ increasing slowly to infinity, or even for $\xi(n) = \log\log n$. This problem can be further simplified by fixing $\gamma_2 = 0$.

1.4. **Organisation and notation.**

Organisation of the manuscript. In Section 2, we collect tools and technical lemmata that would otherwise disrupt the course of the main arguments. In Section 3, we prove Theorems 1.8 and 1.23 together, followed by the convergence part of Corollary 1.11. Thereafter, we prove Theorems 1.13 and 1.15 in Sections 4 and 5, respectively. Appendix A describes how the proof of Theorem 1.23 can be adapted to give Theorem 1.26.

Notation. We use the Vinogradov and Bachmann–Landau notations: For functions f and positive-valued functions g, we write $f \ll g$ or $f = O(g)$ if there exists a constant C such that $|f(x)| \leqslant Cg(x)$ for all values of x under consideration. We write $f \asymp g$ or $f = \Theta(g)$ if $f \ll g \ll f$. Throughout this manuscript, the implied constants are allowed to depend on:

- The approximation function ψ
- A fixed vector $\boldsymbol{\alpha} \in \mathbb{R}^{k-1}$, and $\gamma_1, \dots, \gamma_k \in \mathbb{R}$
- A constant $C \geqslant 2$, which in turn only needs to depend on $\boldsymbol{\alpha}, \gamma_1, \dots, \gamma_k$, specifying the ranges $[C^j, C^{j+1}]$ to which we localise various parameters
- A small positive constant ε_0, which only needs to depend on $\boldsymbol{\alpha}, \gamma_1, \dots, \gamma_k$.

These dependencies shall usually not be indicated by a subscript. To be explicit, we do consider the dimension k of the ambient space to be data of the vector $\boldsymbol{\alpha}$ and hence shall not indicate its dependency. If any other dependence occurs, we record this using an appropriate subscript. If $\mathcal{S}$ is a set, we denote the cardinality of $\mathcal{S}$ by $|\mathcal{S}|$ or $\#\mathcal{S}$. The symbol p is reserved for primes. The pronumeral N denotes a positive integer, sufficiently large in terms of $\boldsymbol{\alpha}, \gamma_1, \dots, \gamma_k$, the approximation function ψ, and a bounded interval $\mathcal{I}$ that will arise in the course of some of the proofs. For a vector $\boldsymbol{\alpha} \in \mathbb{R}^{k-1}$, we abbreviate the product of its coordinates to

$$\Pi(\boldsymbol{\alpha}) = \alpha_1 \cdots \alpha_{k-1}. \tag{1.19}$$

When $x \in \mathbb{R}$, we write $\|x\|$ for the distance from x to the nearest integer. Furthermore, μ denotes one-dimensional Lebesgue measure. Given $\mathcal{S} \subseteq \mathbb{R}$, we write $\mathcal{S}_{\leqslant X}$ for $\{x \in \mathcal{S} : x \leqslant X\}$.

Finally, we often have to deal with expressions such as

$$\sum_{n \leqslant N} \frac{1}{(\log n) \log\log n},$$

which are not always well-defined because of finitely many small $n \in \mathbb{N}$. To deal with this, we write $\ln(x)$ for the natural logarithm of a positive real number x, and put $\log(x) = \max(\ln(x), 1)$ to ensure that these logarithms and their iterates are indeed well-defined and positive.

Funding and acknowledgements. SC was supported by EPSRC Fellowship Grant EP/S00226X/2. NT was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 786758), as well as by the Austrian Science Fund (FWF) grant J 4464-N, and is grateful to Rajula Srivastava for her eternal encouragement. We thank Sanju Velani for his enthusiasm towards this topic, Andy Pollington for many enthralling conversations, Victor Beresnevich for further encouragement, and Zeev Rudnick for comments on a draft. Last, but not least, we thank the referee for a thorough reading and for helpful remarks, including a nice proof Lemma 2.14 that we have now included.

2. Preliminaries

In this subsection, we gather together a panoply of tools. We begin with the theory of continued fractions, before continuing to that of Bohr sets. Then we present some standard measure theory and real analysis. The latter will enable us to introduce the artificial divisors to which we alluded earlier, at essentially no cost. After that we discuss some estimates from the geometry of numbers, whose raison d'être is to count solutions to congruences in generalised arithmetic progressions. We then review some basic prime number theory and sieve theory, culminating in the fundamental lemma of sieve theory, which we will later use to count shift-reduced fractions.

2.1. **Continued fractions.** The material here is standard; see for instance [15, Chapter 1, Section 2]. For each irrational number α there exists a unique sequence of integers $a_0, a_1, a_2, \ldots$, the *partial quotients of* α, such that $a_j \geqslant 1$ for $j \geqslant 1$ and

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}.$$

For $j \geqslant 0$, let $[a_0; a_1, \ldots, a_j]$ denote the truncation the above infinite continued fraction at place j, and let $p_j \in \mathbb{Z}$ and $q_j \in \mathbb{N}$ be the coprime integers satisfying

$$[a_0; a_1, \ldots, a_j] = \frac{p_j}{q_j}.$$

This rational number is the called the *j-th convergent to α*. Continued fractions enjoy an impressive portfolio of beautiful properties. One particular feature for which we have ample need is the following recursion: For $j \geqslant 1$, we have

$$q_{j+1} = a_{j+1}q_j + q_{j-1} \quad \text{and} \quad p_{j+1} = a_{j+1}p_j + p_{j-1}. \tag{2.1}$$

The initial values are

$$p_0 = a_0, \qquad q_0 = 1, \qquad p_1 = a_1a_0 + 1, \qquad q_1 = a_1.$$

For each $j \geqslant 0$, the quantity

$$D_j = q_j\alpha - p_j \tag{2.2}$$

satisfies the bound

$$\frac{1}{2} \leqslant |D_j|\, q_{j+1} \leqslant 1. \tag{2.3}$$

Furthermore, it is well-known that the signs $D_j/|D_j|$ are alternating. For an irrational number $\alpha \in \mathbb{R}$, the *diophantine exponent* of α is

$$\omega(\alpha) = \sup\{w > 0 : \|q\alpha\| < q^{-w} \text{ for infinitely many } q \geqslant 1\}.$$

Note that $\omega(\alpha) = \omega^{\times}(\alpha)$. We require the following well-known fact characterising diophantine numbers in terms of the growth of the consecutive continued fraction denominators [11, Lemma 1.1]:

Lemma 2.1. *Let $\alpha \in \mathbb{R}\setminus\mathbb{Q}$, and $(q_k)_k$ be the sequence of its continued fraction denominators. Then*

$$\omega(\alpha) = \limsup_{k\to\infty} \frac{\log q_{k+1}}{\log q_k}.$$

In particular $\alpha \notin \mathcal{L}$ if and only if $\log q_{k+1} \ll \log q_k$.

Continued fractions have been used to prove a highly aesthetic result called the *three gap theorem*. For $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $m \in \mathbb{N}$, this asserts that there are at most three distinct gaps $d_{i+1} - d_i$, where

$$\{d_1, \ldots, d_m\} = \{i\alpha - \lfloor i\alpha \rfloor : 1 \leqslant i \leqslant m\}$$

and

$$0 = d_0 < \cdots < d_{m+1} = 1.$$

Many find this surprising at first. The sizes of the gaps can be computed and described in terms of the continued fraction expansion. We will need only the

size of the largest gap. The following theorem combines parts of Theorem 1 and Corollary 1 of [46], and adopts the typical convention that $q_{-1} = 0$.

Theorem 2.2. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $m \in \mathbb{N}$. Then:*

(a) There is a unique representation

$$m = rq_k + q_{k-1} + s,$$

for some

$$k \geqslant 0, \qquad 1 \leqslant r \leqslant a_{k+1}, \qquad 0 \leqslant s \leqslant q_k - 1.$$

(b) If $s < q_k - 1$ then

$$\max\{d_{i+1} - d_i : 0 \leqslant i \leqslant m\} = \begin{cases} |D_{k+1}| + |D_k|, & \text{if } r = a_{k+1} \\ |D_{k+1}| + (a_{k+1} - r + 1)|D_k|, & \text{if } r < a_{k+1}. \end{cases}$$

If $s = q_k - 1$ then

$$\max\{d_{i+1} - d_i : 0 \leqslant i \leqslant m\} = \begin{cases} |D_k|, & \text{if } r = a_{k+1} \\ |D_{k+1}| + (a_{k+1} - r)|D_k|, & \text{if } r < a_{k+1}. \end{cases}$$

Rational numbers also have continued fraction expansions, however they are finite, taking the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_t}}} \quad .$$

The partial quotients $a_0 \in \mathbb{Z}$ and $a_1, \ldots, a_t \in \mathbb{N}$, as well as the convergents p_j/q_j $(0 \leqslant j \leqslant t)$, are defined in the same way is in the irrational case, except that we impose the additional constraint $a_t > 1$ to be sure that the expansion is unique.

In the next subsection, we fix an irrational number α and describe an expansion, the *Ostrowski expansion*, that allows us to accurately read off for each $n \in \mathbb{N}$ the size of $\|n\alpha\|$, and more generally $\|n\alpha - \gamma\|$ for $\gamma \in \mathbb{R}$.

2.2. **Ostrowski expansions.** Let $n \in \mathbb{N}$, and let $K \geqslant 0$ be such that

$$q_K \leqslant n < q_{K+1}. \tag{2.4}$$

It is known [52, p. 24] that there exists a uniquely determined sequence of non-negative integers $c_k = c_k(n)$ such that

$$n = \sum_{k \geqslant 0} c_{k+1} q_k,$$

satisfying

$$c_{k+1} = 0 \text{ for all } k > K,$$

as well as the following additional constraints:

$$0 \leqslant c_1 < a_1, \qquad 0 \leqslant c_{k+1} \leqslant a_{k+1} \quad (k \in \mathbb{N}), \qquad \text{if } c_{k+1} = a_{k+1} \text{ then } c_k = 0.$$

The structure of the set of integers whose initial Ostrowski digits are prescribed is well-understood; this set is referred to as a *cylinder set*. We require information concerning the size of the gap between consecutive elements, which we retrieve from [11, Lemma 5.1].

Lemma 2.3 (Gaps lemma)**.** *Let $m \geqslant 0$, and let $\mathcal{A}(d_1, \ldots, d_{m+1})$ denote the set of positive integers whose initial Ostrowski digits are $d_1, \ldots, d_{m+1}$. Let $n_1 < n_2 < \ldots$ be the elements of the set $\mathcal{A}(d_1, \ldots, d_{m+1})$, and let $i \in \mathbb{N}$. If $d_{m+1} > 0$ then*

$$n_{i+1} - n_i \geqslant q_{m+1},$$

and if $d_{m+1} = 0$ then $n_{i+1} - n_i \in \{q_{m+1}, q_m\}$. Further, if $d_{m+1} = 0$ and $n_{i+1} - n_i = q_m$ then $c_{m+2}(n_i) = a_{m+2}$ and the gap $n_{i+1} - n_i$ is preceded by a_{m+2} gaps of size q_{m+1}.

With reference to (2.2), we will apply the theory of this subsection to pairs (α, γ), where

$$\gamma = \sum_{k \geqslant 0} b_{k+1} D_k,$$

such that

$$a_k \geqslant 64, \quad \frac{a_k}{4} \leqslant b_k \leqslant \frac{a_k}{2} \quad (k \geqslant 1), \qquad a_0 = 0. \tag{2.5}$$

Before proceeding in earnest, we confirm some technical conditions that will put us into a standard setting.

Lemma 2.4. *If we have* (2.5)*, then*

$$0 < \alpha < \frac{1}{64}, \qquad 0 \leqslant \gamma < 1 - \alpha,$$

and

$$\|n\alpha - \gamma\| \neq 0 \qquad (n \in \mathbb{N}). \tag{2.6}$$

Proof. The first inequality follows from $a_0 = 0$ and $a_1 \geqslant 64$.

We compute that

$$|b_{k+1}D_k| \leqslant \frac{a_{k+1}}{q_{k+1}} \leqslant \frac{1}{q_k}, \quad |b_{k+1}D_k| \geqslant \frac{a_{k+1}/4}{2q_{k+1}} \geqslant \frac{1}{16q_k} \qquad (k \geqslant 0).$$

As $a_{k+1} \geqslant 64$ for all $k \geqslant 0$, we see from these inequalities that $|b_{k+1}D_k|$ is a monotonically decreasing sequence which converges to zero as $k \to \infty$. Using that the signs of the D_k alternate and that $D_0 = \{\alpha\} > 0$, we conclude that $b_{k+1}D_k + b_{k+2}D_{k+1} \geqslant 0$ if k is even and $b_{k+1}D_k + b_{k+2}D_{k+1} \leqslant 0$ if k is odd. Therefore

$$0 \leqslant \sum_{k\geqslant 0}(b_{2k+1}D_{2k} + b_{2k+2}D_{2k+1}) = \gamma$$

supplies the lower bound in the second inequality. For the upper bound, note that

$$\gamma = b_1D_0 + \sum_{k\geqslant 0}(b_{2k+2}D_{2k+1} + b_{2k+3}D_{2k+2}) \leqslant b_1D_0 \leqslant \frac{b_1}{a_1} \leqslant 1/2 < 1 - \alpha.$$

Finally, we turn our attention towards (2.6). Observe that

$$n\alpha - \gamma = \sum_{k\geqslant 0}(c_{k+1}q_k\alpha - b_{k+1}(q_k\alpha - p_k)) \in \Sigma + \mathbb{Z},$$

where

$$\Sigma = \sum_{k\geqslant 0}\delta_{k+1}D_k, \qquad \delta_{k+1} = c_{k+1} - b_{k+1} \quad (k \geqslant 0). \tag{2.7}$$

Using (2.3) and (2.1), as well as (2.5), we compute that

$$|\Sigma| \geqslant |D_0| - \frac{3}{4}\sum_{k\geqslant 1}a_{k+1}|D_k| \geqslant \{\alpha\} - \frac{3}{4}\sum_{k\geqslant 1}\frac{1}{q_k},$$

where $\{\alpha\}$ denotes the fractional part of α, and

$$\sum_{k\geqslant 1}\frac{1}{q_k} \leqslant \frac{1}{q_1}\sum_{r\geqslant 0}64^{-r} = \frac{64}{63q_1}.$$

Since

$$\frac{1}{\{\alpha\}} < a_1 + 1 = q_1 + 1 \leqslant \frac{65}{64}q_1,$$

we have

$$|\Sigma| > \left(\frac{64}{65} - \frac{16}{21}\right)q_1^{-1} > 0.$$

Moreover

$$|\Sigma| \leqslant |D_0| + \sum_{k\geqslant 1}a_{k+1}|D_k| \leqslant \{\alpha\} + \sum_{k\geqslant 1}\frac{1}{q_k} \leqslant \frac{1}{64} + \frac{1}{63} < 1.$$

Verily we have (2.6). □

It turns out that we can quantify the size of $\|n\alpha - \gamma\|$ in terms of the quantity Σ from (2.7), as the next lemma details. The assumption (2.5) simplifies several technicalities. The following combines [11, Lemmata 4.3, 4.4, and 4.5], in this special case.

Lemma 2.5. *If we have* (2.5)*, then*

$$\|n\alpha - \gamma\| = \|\Sigma\| = \min(|\Sigma|\,, 1 - |\Sigma|).$$

Furthermore, let $m = m(n)$ be the smallest $i \geqslant 0$ such that $\delta_{i+1} \neq 0$, and let K be as in (2.4). Then we have the following estimates for $|\Sigma|$ and $1 - |\Sigma|$.

(1)

$$|\Sigma| = (|\delta_{m+1}| - 1)\,|D_m| + u_{m+2}\,|D_{m+1}| + u_{m+3}\,|D_{m+2}| + \Upsilon,$$

where $u_{m+2}, u_{m+3}, \Upsilon$ are non-negative real numbers constrained by

$$u_{m+2} \asymp a_{m+2}, \qquad u_{m+3} \asymp a_{m+3}, \qquad \Upsilon \ll |D_{m+2}|.$$

(2)

$$1 - |\Sigma| = u_1\,|D_0| + u_2\,|D_1| + \tilde{\Upsilon},$$

where $u_1, u_2, \tilde{\Upsilon}$ are non-negative, and constrained by

$$u_1 \asymp a_1, \qquad u_2 \asymp a_2, \qquad \tilde{\Upsilon} \ll |D_1|.$$

2.3. **Bohr sets.** For $\boldsymbol{\alpha}, \boldsymbol{\gamma} \in \mathbb{R}^{k-1}$, we have ample need for bounds on the cardinality of Bohr sets

$$\mathcal{B} = \mathcal{B}^{\gamma}_{\boldsymbol{\alpha}}(N; \boldsymbol{\rho}) := \{n \in \mathbb{Z} : |n| \leqslant N, \|n\alpha_i - \gamma_i\| \leqslant \rho_i \quad (1 \leqslant i \leqslant k-1)\}, \tag{2.8}$$

that are sharp up to multiplication by absolute constants. Further, it turns out to be crucial to have precise control over the ranges in which the Bohr sets have a sufficiently nice structure. If the width parameters δ_i and the length parameters N_i are in a suitable regime, then the Bohr set $\mathcal{B}$ is enveloped—efficiently, as we detail soon—in a *k-dimensional generalised arithmetic progression*

$$\mathcal{P}(b; A_1, \ldots, A_k; N_1, \ldots, N_k) = \{b + A_1 n_1 + \cdots + A_k n_k : |n_i| \leqslant N_i\}, \tag{2.9}$$

where $b, A_1, \ldots, A_k, N_1, \ldots, N_k \in \mathbb{N}$. The thresholds for the admissible regimes depend naturally on the diophantine properties of $\boldsymbol{\alpha}$.

Inside $\mathcal{B}$, we will find a large (asymmetric) generalised arithmetic progression

$$\mathcal{P}^+(b; A_1, \ldots, A_k; N_1, \ldots, N_k) = \{b + A_1 n_1 + \cdots + A_k n_k : 1 \leqslant n_i \leqslant N_i\}. \tag{2.10}$$

This is *proper* if for each $n \in \mathcal{P}^+(b, A_1, \ldots, A_k, N_1, \ldots, N_k)$ there is a unique vector $(n_1, \ldots, n_k) \in \mathbb{N}^k$ for which

$$n = b + A_1 n_1 + \cdots + A_k n_k, \qquad n_i \leqslant N_i \quad (1 \leqslant i \leqslant k).$$

Throughout this subsection we operate under the assumption (1.6), which in the case $k = 2$ simply means that α is irrational and non-Liouville. Set

$$\eta(\boldsymbol{\alpha}) = \frac{1}{\omega^{\times}(\boldsymbol{\alpha})} - \frac{k-2}{k-1} \in (0, 1].$$

In [21, Section 3], we exploited the strict positivity of $\eta(\boldsymbol{\alpha})$ to describe regimes in which the Bohr sets (2.8) contain and are contained in generalised arithmetic progressions of the expected size. We presently outline the key findings from that investigation, as far as they are needed here.

Lemma 2.6 (Inner structure)**.** *Let $\vartheta \geqslant 1$. Then there exists $\tilde{\varepsilon} > 0$ with the following property. If $\varepsilon \in (0, \tilde{\varepsilon}]$ is fixed, and N is large in terms of $\vartheta, \varepsilon, \boldsymbol{\alpha}$, and*

$$N^{-\varepsilon} \leqslant \rho_i \leqslant 1 \quad (1 \leqslant i \leqslant k-1), \tag{2.11}$$

then there exists a proper generalised arithmetic progression

$$\mathcal{P} = \mathcal{P}^{+}(b; A_1, \ldots, A_k; N_1, \ldots, N_k)$$

contained in $\mathcal{B}$, for which

$$|\mathcal{P}| \gg \rho_1 \cdots \rho_{k-1} N, \qquad \min_{i \leqslant k} N_i \geqslant N^{\vartheta\varepsilon}, \qquad N^{\sqrt{\varepsilon}} \leqslant b \leqslant \frac{N}{10},$$

and

$$\gcd(A_1, \ldots, A_k) = 1.$$

Proof. Observe that the statement gets weaker as ε decreases, in the sense that if it holds for $\varepsilon = \tilde{\varepsilon}$ then it holds for any $\varepsilon \in (0, \tilde{\varepsilon}]$. It is almost identical to the statement in [21, Lemma 3.1], but there the variable ϑ was equal to 1. By replacing $\tilde{\varepsilon}$ by $\tilde{\varepsilon}/\vartheta$, thereby shrinking the range of admissible values of ε, we obtain the statement here. □

Next, we quantify the range of the width vector $\boldsymbol{\rho}$ for which the Bohr sets are efficiently contained in generalised arithmetic progressions. The outer structure lemma as stated in [21, Lemma 3.2] is homogeneous and fails to record the feature that $\gcd(A_1, \ldots, A_k) = 1$. Here we require an inhomogeneous version as well as the latter feature, and fortunately we can extract this additional information from the proof of [21, Lemma 3.2], as we now explain.

Lemma 2.7 (Outer structure)**.** *Let $\vartheta \geqslant 1$. Then there exists $\tilde{\varepsilon} > 0$ with the following property. If $\varepsilon \in (0, \tilde{\varepsilon}]$ is fixed and N is sufficiently large in terms of ε, ϑ, and we have* (2.11)*, then there exists a generalised arithmetic progression*

$$\mathcal{P} = \mathcal{P}(b; A_1, \ldots, A_k; N_1, \ldots, N_k),$$

containing $B_{\boldsymbol{\alpha}}^{\gamma}(N;\boldsymbol{\rho})$*, for which*

$$\min_{i\leqslant k} N_i \geqslant N^{\vartheta\varepsilon}, \qquad |\mathcal{P}| \ll N_1\cdots N_k \ll \rho_1\cdots\rho_{k-1}N$$

and

$$\gcd(A_1,\dots,A_k)=1.$$

Proof. Lemma 2.6 implies that the Bohr set $\mathcal{B}_{\boldsymbol{\alpha}}^{\gamma}(N;\boldsymbol{\rho})$ is non-empty, so choose $b\in\mathcal{B}_{\boldsymbol{\alpha}}^{\gamma}(N;\boldsymbol{\rho})$ arbitrarily. For any $n\in\mathcal{B}_{\boldsymbol{\alpha}}^{\gamma}(N;\boldsymbol{\rho})$, the triangle inequality yields $n-b\in\mathcal{B}_{\boldsymbol{\alpha}}^{\mathbf{0}}(N;2\boldsymbol{\rho})$. Now [21, Lemma 3.2] assures us that

$$n-b\in\mathcal{P}(0;A_1,\dots,A_k;N_1,\dots,N_k).$$

The coprimality property $\gcd(A_1,\dots,A_k)=1$ comes out of the proof, but was not recorded in [21, Lemma 3.2] because it was not needed there. Similarly, the inequality $\min_{i\leqslant k}N_i\geqslant N^{\varepsilon}$ arises in the proof. Moreover, it is stated in [21, Lemma 3.2] that $|\mathcal{P}|\ll\rho_1\cdots\rho_{k-1}N$, but its proof contains the more refined inequalities

$$|\mathcal{P}|\ll N_1\cdots N_k\ll\rho_1\cdots\rho_{k-1}N.$$

By replacing $\tilde{\varepsilon}$ by $\tilde{\varepsilon}/\vartheta$, thereby shrinking the range of admissible values of ε, we are able to bootstrap the inequality to $\min_{i\leqslant k}N_i\geqslant N^{\vartheta\varepsilon}$. □

Finally, recall that Beresnevich, Haynes, and Velani [11, Lemma 6.1] determined a convenient condition under which a rank 1 Bohr set has the expected size.

Lemma 2.8. *Let* $\alpha\in\mathbb{R}\setminus\mathbb{Q}$*, let* $(q_k)_k$ *be its sequence of denominators of convergents of continued fractions, and let* $\delta\in(0,\|q_2\alpha\|/2)$*. If there exists* $\ell\in\mathbb{Z}$ *such that*

$$\frac{1}{2\delta}\leqslant q_\ell\leqslant N,$$

then

$$\delta N-1\leqslant\#\mathcal{B}_{\alpha}^{0}(N;\delta)\leqslant 32\delta N.$$

2.4. **Measure theory.** To bound from below the measure of the limit superior sets of interest, we deploy the 'divergence' Borel–Cantelli lemma [34, Lemma 2.3].

Lemma 2.9. *Let* $\mathcal{E}_1,\mathcal{E}_2,\dots$ *be a sequence of Borel subsets of* $[0,1]$ *such that*

$$\sum_{n=1}^{\infty}\mu(\mathcal{E}_n)=\infty,$$

and let $\mathcal{E} = \limsup_{n\to\infty} \mathcal{E}_n$. *Then*

$$\mu(\mathcal{E}) \geqslant \limsup_{N\to\infty} \frac{\left(\sum_{n\leqslant N} \mu(\mathcal{E}_n)\right)^2}{\sum_{n,m\leqslant N} \mu\left(\mathcal{E}_n \cap \mathcal{E}_m\right)}.$$

In many applications in metric diophantine approximation, it suffices to establish that a limit superior set $\mathcal{E}$ has positive measure in order to conclude that it has full measure: Often one of the classical zero–one laws, such as Cassels' or Gallagher's [34, Section 2.2], rules out the possibility that $\mu(\mathcal{E}) \in (0,1)$. However, for our purposes no such zero–one law is available. To establish full measure, we turn instead to another device, which follows from Lebesgue's density theorem. Intuitively, the next lemma—sometimes referred to as Knopp's lemma—states that any set whose local densities are uniformly and positively bounded from below must have full measure. This is a special case of [8, Proposition 1].

Lemma 2.10. *If* $\mathcal{E} \subseteq [0,1]$ *is Borel set and*

$$\mu(\mathcal{E} \cap \mathcal{I}) \gg \mu(\mathcal{I})$$

for any interval $\mathcal{I} \subseteq [0,1]$, *then* $\mu(\mathcal{E}) = 1$.

2.5. **Real analysis.** The next lemma will later enable us to eschew a certain small-GCD regime. In the first instance, it asserts that if $\psi(n)\log n$ is not summable then neither is $\psi(an)\log n$ for any fixed $a \in \mathbb{N}$. In fact, we can allow a to increase very slowly with n.

Lemma 2.11. *Let* $d \in \mathbb{N}$. *Let* $\psi : \mathbb{N} \to \mathbb{R}_{>0}$ *be non-increasing such that*

$$\sum_{n\geqslant 1} \psi(n)(\log n)^d$$

diverges. Then, there exists a strictly increasing sequence $(K_i)_i$ *of positive integers satisfying* $K_i \geqslant \exp(\exp i)$, *for all* $i \geqslant 2$, *such that* $K_1 = 1$, $K_i = o(K_{i+1})$, *and:*

(1) The map f *defined by*

$$f(n) = i \qquad (K_i \leqslant n < K_{i+1})$$

satisfies $f(n) \ll \log\log n$.

(2) If $\hat{\psi}(n) = \psi(4^{f(n)}n)$, for all n, then the series

$$\sum_{n\geqslant 1} \hat{\psi}(n)(\log n)^d$$

diverges.

Proof. We construct $(K_i)_i$ recursively. Write

$$S_\psi(N) = \sum_{n\leqslant N} \psi(n)(\log n)^d, \qquad S_{\hat{\psi}}(N) = \sum_{n\leqslant N} \hat{\psi}(n)(\log n)^d.$$

Set $K_1 = 1$, let $i \geqslant 1$, and suppose that K_i has already been constructed. If $a \in \mathbb{N}$ and $N \geqslant N_a$, where $N_a = N(a, \psi)$ is large, then

$$\begin{aligned} S_\psi(aN) &= S_\psi(a^2) + \sum_{a\leqslant j<N} \sum_{r\leqslant a} \psi\left(aj + r\right)\left(\log\left(aj + r\right)\right)^d \\ &\ll S_\psi(a^2) + \sum_{a\leqslant j<N} \sum_{r\leqslant a} \psi\left(aj\right)\left(\log j\right)^d \ll a \sum_{j\leqslant N} \psi\left(aj\right)\left(\log j\right)^d. \end{aligned}$$

Next, consider $a = 4^i$ and $N \geqslant K_i + N_a$, for some $i \in \mathbb{N}$. Supposing we choose $K_{i+1} = N + 1$, then as $i \geqslant f(j)$ for any $j \leqslant N$, we have

$$\psi(aj) = \psi(4^i j) \leqslant \psi(4^{f(j)} j) = \hat{\psi}(j) \qquad (j \leqslant N),$$

and so

$$S_{\hat{\psi}}(N) \gg a^{-1} S_\psi(aN) = 4^{-i} S_\psi(4^i N).$$

We define K_{i+1} to be one more than the smallest positive integer

$$N \geqslant \exp(\exp(2K_i + N_{4^i}))$$

for which $4^{-i} S_\psi(4^i N) \geqslant i$. By construction, we have

$$S_{\hat{\psi}}(K_{i+1}) \gg i, \qquad K_{i+1} \geqslant \exp(\exp(i + 1)).$$

The latter implies that $f(n) \leqslant \log\log n$ for $n \geqslant K_2$, completing the proof. □

The following lemma helps us with dyadic pigeonholing.

Lemma 2.12. *Let $h : \mathbb{N} \to \mathbb{R}_{>0}$ be non-increasing, and fix $C \geqslant 2$, $\kappa > 0$, as well a positive integer $J_0 \ll 1$. Then, for $N \in \mathbb{N}$ and $J = \lfloor \log N/\log C \rfloor$, we have*

$$\sum_{C^{J_0}\leqslant n\leqslant N} h(n)(\log n)^\kappa \asymp \sum_{j=J_0}^{J} j^\kappa C^j h(C^j). \tag{2.12}$$

Proof. Since h is non-increasing, for $j = 1, 2, \dots, J$ we have

$$\sum_{C^j < n \leqslant C^{j+1}} h(n)(\log n)^\kappa \geqslant C^{j+1}(1-1/2)h(C^{j+1})(j \log C)^\kappa$$
$$\gg C^{j+1}h(C^{j+1})((j+1)\log C)^\kappa$$

and

$$\sum_{C^j \leqslant n \leqslant C^{j+1}} h(n)(\log n)^\kappa \leqslant C^{j+1}h(C^j)((j+1)\log C)^\kappa \ll C^j h(C^j)(j \log C)^\kappa.$$

Therefore

$$\sum_{C^{J_0} \leqslant n \leqslant N} h(n)(\log n)^\kappa \gg h(C^{J_0})(J_0 \log C)^\kappa + \sum_{j=J_0}^{J-1} C^{j+1}h(C^{j+1})((j+1)\log C)^\kappa$$
$$\gg \sum_{j=J_0}^{J} C^j h(C^j)(j \log C)^\kappa$$

and

$$\sum_{C^{J_0} \leqslant n \leqslant N} h(n)(\log n)^\kappa \ll \sum_{j=J_0}^{J} C^j h(C^j)(j \log C)^\kappa.$$

□

2.6. **Geometry of numbers.** The following lattice point counting theorem originates from the work of Davenport [24], see also [6] and [57, p. 244]. Our precise statement follows from [6, Lemmata 2.1 and 2.2].

Theorem 2.13 (Davenport)**.** *Let $d, h \in \mathbb{N}$, and let $\mathcal{S}$ be a compact subset of $\mathbb{R}^d$. Assume that the two following conditions are met:*

(i) Any line intersects $\mathcal{S}$ in a set of points which, if non-empty, comprises at most h intervals.

(ii) The first condition holds, with j in place of d, for the projection of $\mathcal{S}$ onto any j-dimensional subspace.

Let $\lambda_1 \leqslant \cdots \leqslant \lambda_d$ be the successive minima, with respect to the Euclidean unit ball, of a (full-rank) lattice Λ in $\mathbb{R}^d$. Then

$$\left| |\mathcal{S} \cap \Lambda| - \frac{\mathrm{vol}(\mathcal{S})}{\det \Lambda} \right| \ll_{d,h} \sum_{j=0}^{d-1} \frac{V_j(\mathcal{S})}{\lambda_1 \cdots \lambda_j},$$

where $V_j(\mathcal{S})$ is the supremum of the j-dimensional volumes of the projections of $\mathcal{S}$ onto any j-dimensional subspace. We adopt the convention that $V_0(\mathcal{S}) = 1$.

We have not encountered a reference for the following classical result, so we provide a proof.

Lemma 2.14. *Let $k \in \mathbb{N}$, and let $A_1, \dots, A_k, d \in \mathbb{N}$ with*

$$\gcd(A_1, \dots, A_k, d) = 1.$$

Then the congruence

$$A_1 n_1 + \cdots + A_k n_k \equiv 0 \bmod d \tag{2.13}$$

defines a full-rank lattice Λ of determinant d.

Proof. As $\gcd(A_1, \dots, A_k, d) = 1$, there exist integers $m_1, \dots, m_{k+1}$ such that

$$A_1 m_1 + \cdots + A_k m_k + d m_{k+1} = 1,$$

and particular

$$A_1 m_1 + \cdots + A_k m_k \equiv 1 \bmod d.$$

As 1 generates the cyclic group $\mathbb{Z}/d\mathbb{Z}$, it follows that the homomorphism

$$\mathbb{Z}^k \to \mathbb{Z}/d\mathbb{Z}$$
$$(n_1, \dots, n_k) \mapsto A_1 n_1 + \cdots + A_k n_k$$

is surjective. Further, its kernel is $\Lambda \subseteq \mathbb{Z}^k$. By the first isomorphism theorem, the group $\mathbb{Z}^k/\Lambda$ has finite order d, and consequently Λ has full rank. An application of [18, Subsection I.2.2, Lemma 1] completes the proof. □

We need to be able to count elements of a generalised arithmetic progression divisible by a given positive integer d. The previous two facts enable us to accurately do so, provided that d is not too large.

Lemma 2.15.

(i) Let $d \in \mathbb{N}$, and let $\mathcal{P}$ be generalised arithmetic progression given by (2.9), *where* $\gcd(A_1, \dots, A_k) = 1$. *Then*

$$\frac{\#\{n \in \mathcal{P} : d \mid n\}}{N_1 \cdots N_k} \ll d^{-1} + (\min_{i \leqslant k} N_i)^{-1}. \tag{2.14}$$

(ii) Let $\mathcal{P}$ be a proper, asymmetric generalised arithmetic progression given by (2.10)*, where* $\gcd(A_1, \dots, A_k) = 1$*, and let* $d \in \mathbb{N}$*. Then*

$$\frac{\#\{n \in \mathcal{P} : d \mid n\}}{N_1 \cdots N_k} = d^{-1} + O(1/\min_{i \leqslant k} N_i). \tag{2.15}$$

Proof. (i) The quantity $\#\{n \in \mathcal{P} : d \mid n\}$ is bounded above by the number of integer solutions

$$(n_1, \dots, n_k) \in [-N_1, N_1] \times \cdots \times [-N_k, N_k]$$

to

$$b + A_1 n_1 + \cdots + A_k n_k \equiv 0 \bmod d.$$

We may assume that this has a solution

$$(n_1^*, \dots, n_k^*) \in [-N_1, N_1] \times \cdots \times [-N_k, N_k],$$

and then

$$(n_1', \dots, n_k') = (n_1, \dots, n_k) - (n_1^*, \dots, n_k^*)$$

lies in $[-2N_1, 2N_1] \times \cdots \times [-2N_k, 2N_k]$ and satisfies

$$A_1 n_1' + \cdots + A_k n_k' \equiv 0 \bmod d.$$

By Lemma 2.14, this defines a full-rank lattice of determinant d, and this is a sublattice of $\mathbb{Z}^n$ so the successive minima are greater than or equal to 1. By Theorem 2.13, we now have

$$\#\{n \in \mathcal{P} : d \mid n\} \leqslant \frac{2^k N_1 \cdots N_k}{d} + O_k(N_1 \cdots N_k / \min_{i \leqslant k} N_i),$$

giving (2.14).

(ii) As $\mathcal{P}$ is proper, the quantity $\#\{n \in \mathcal{P} : d \mid n\}$ counts integer solutions

$$(n_1, \dots, n_k) \in [1, N_1] \times \cdots \times [1, N_k]$$

to

$$b + A_1 n_1 + \cdots + A_k n_k \equiv 0 \bmod d.$$

Since $\gcd(A_1, \dots, A_k) = 1$, there exist integers $n_1^*, \dots, n_k^*$ such that

$$b + A_1 n_1^* + \cdots + A_k n_k^* = 0.$$

Now

$$(n_1', \dots, n_k') = (n_1, \dots, n_k) - (n_1^*, \dots, n_k^*)$$

lies in $[1 - n_1^*, N_1 - n_1^*] \times \cdots \times [1 - n_1^*, N_k - n_1^*]$ and satisfies

$$A_1 n_1' + \cdots + A_k n_k' \equiv 0 \bmod d.$$

By Lemma 2.14, this defines a full-rank lattice of determinant d, and this is a sublattice of $\mathbb{Z}^n$ so the successive minima are greater than or equal

to 1. By Theorem 2.13, we now have

$$\#\{n \in \mathcal{P} : d \mid n\} = \frac{(N_1 - 1)\cdots(N_k - 1)}{d} + O_k(N_1 \cdots N_k / \min_{i \leqslant k} N_i),$$

giving (2.15).

□

2.7. **Primes and sieves.** We require one of Mertens' three famous, classical estimates [42, Theorem 3.4(c)].

Theorem 2.16 (Mertens' third theorem)**.** *For $x \geqslant 2$, we have*

$$\prod_{p \leqslant x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right),$$

where γ is the Euler–Mascheroni constant.

Sieve theory is a powerful collection of techniques used to study prime numbers. Let $\mathcal{P}$ be set of primes, and let

$$\mathcal{A} = (a_n)_{1 \leqslant n \leqslant x}$$

be a finite sequence of non-negative real numbers. The main object of interest is the *sifting function*

$$S(\mathcal{A}, z) = \sum_{(n, P(z))=1} a_n,$$

where

$$P(z) = \prod_{\substack{p \in \mathcal{P} \\ p < z}} p,$$

for a parameter $z \geqslant 2$. For example, if $\mathcal{J} \subseteq [1, x] \cap \mathbb{Z}$ and $a_n = 1$ for $n \in \mathcal{J}$ and $a_n = 0$ for $n \notin \mathcal{J}$, then $S(\mathcal{A}, z)$ counts elements of $\mathcal{J}$ not divisible by any prime $p \in \mathcal{P}$ for which $p < z$.

In this subsection only, the letter μ denotes the Möbius function, and not a Lebesgue measure. Evaluating the sifting function using the inclusion–exclusion principle yields

$$S(\mathcal{A}, z) = \sum_{d \mid P(z)} \mu(d) A_d(x),$$

where

$$A_d(x) = \sum_{\substack{n \leqslant x \\ n \equiv 0 \bmod d}} a_n \qquad (d \in \mathbb{N}).$$

This gives rise to the main term $XV(z)$, where

- X is typically chosen to approximate $A_1(x)$
- The *density function*, g, is a multiplicative arithmetic function satisfying
$$0 \leqslant g(p) < 1 \qquad (p \in \mathcal{P}),$$
and $g(p)$ is typically chosen to approximate $A_p(x)/X$
- $$V(z) = \prod_{\substack{p \in \mathcal{P} \\ p < z}} (1 - g(p)).$$

However, if z is large then $P(z)$ could have many prime factors, and the error terms may combine to overwhelm the main term. This impasse stood for a long time before Viggo Brun was able to overcome it in many situations. Brun's idea was to approximate μ by a function of smaller support, thereby reducing the number of error terms. One particularly useful outcome is the fundamental lemma, which involves the following additional objects:

- Remainders
$$r_d = A_d(x) - g(d)X \qquad (d \in \mathbb{N})$$
- The *dimension* $\kappa \geqslant 0$, typically chosen to approximate a suitable average of $g(p)p$ over $p \in \mathcal{P}$
- The *level* $D > z$, and the *sifting variable* $s = \log D / \log z$.

The result, for which the lower bound is stated below, is taken from Opera de Cribro [30, Theorem 6.9]. In principle there is a great deal of flexibility, however in practice there is often a natural choice of parameters that reflects the nature of the problem, and any less principled choice tends to produce weaker information.

Theorem 2.17 (Fundamental lemma of sieve theory)**.** *Let* $\kappa \geqslant 0$, $z \geqslant 2$, $D \geqslant z^{9\kappa+1}$, $K > 1$, *and assume that*

$$\prod_{w \leqslant p < z} (1 - g(p))^{-1} \leqslant K \Big(\frac{\log z}{\log w}\Big)^{\kappa} \qquad (2 \leqslant w < z). \tag{2.16}$$

Then

$$S(\mathcal{A}, z) \geqslant XV(z)(1 - e^{9\kappa - s} K^{10}) - \sum_{\substack{d \mid P(z) \\ d < D}} |r_d|.$$

3. A fully-inhomogeneous version of Gallagher's theorem

In this section, we prove Theorem 1.8, along the way establishing Theorem 1.23. At the end we prove the convergence part of Corollary 1.11. The reasoning presented here is sensitive to the diophantine nature of the shift γ_k. We begin by introducing some notation, and reducing the statements of Theorems 1.8 and 1.23 to proving statistical properties of certain sets.

3.1. **Notation and reduction steps.** For ease of exposition, we use the abbreviations

$$\alpha = \alpha_k, \qquad \gamma = \gamma_k$$

throughout the present section. With f as in Lemma 2.11, specialising $d = k-1$ therein, we set

$$\hat{\psi}(n) = \psi(\hat{n}), \quad \text{where} \quad \hat{n} = 4^{f(n)} n.$$

In light of (1.3), we then have

$$\sum_{n=1}^{\infty} \psi(\hat{n}) (\log n)^{k-1} = \infty. \tag{3.1}$$

For $\varepsilon_0 > 0$, we define

$$\hat{\mathcal{G}}_\diamond = \{h \in \mathbb{N} : h^{-4\varepsilon_0} \leqslant \|h\alpha_i - \gamma_i\| \leqslant h^{-2\varepsilon_0} \quad (1 \leqslant i \leqslant k-1)\},$$
$$\mathcal{G}_\diamond = \{n \in \mathbb{N} : \hat{n} \in \hat{\mathcal{G}}_\diamond\},$$

and

$$\mathcal{G} = \left\{ n \in \mathcal{G}_\diamond : \psi(\hat{n}) \geqslant \frac{1}{n(\log n)^{k+1}} \right\}.$$

Remark 3.1. The constant ε_0 is sufficiently small depending on the diophantine nature of the fibre vector $\boldsymbol{\alpha}$ and the final shift γ. Since, throughout this section, we operate under the assumption (1.6) for $k \geqslant 3$ and for $k = 2$ that $\alpha \notin (\mathcal{L} \cup \mathbb{Q})$, the constant ε_0 will always be small enough so that the structural theory of Bohr sets applies. In particular, we will have $\varepsilon_0 \leqslant \tilde{\varepsilon}$ when Lemmata 2.6 and 2.7 are applied with $\vartheta = 20k$. We will also always assume that $\varepsilon_0 < (99k)^{-1}$.

Let us now also introduce a parameter $\eta = \eta(\gamma) \in (0,1)$. We shall in due course be more specific about ε_0 and η, if γ is diophantine, rational, or Liouville. The reader seeking these details instantly may consult (3.37), (3.43), and (3.45).

For $n \in \mathbb{N}$, let

$$\Psi(n) = \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1 - \gamma_1\| \cdots \|\hat{n}\alpha_{k-1} - \gamma_{k-1}\|} 1_{\mathcal{G}}(n),$$

where $1_{\mathcal{G}}$ is the indicator function of $\mathcal{G}$. Fix a non-empty interval $\mathcal{I} \subseteq [0,1]$ and $\gamma \in \mathbb{R}$, and for $n \in \mathbb{N}$ let

$$\mathcal{E}_n^{\mathcal{I},\gamma} := \left\{ \alpha \in [0,1] : \exists a \in \mathbb{Z} \text{ s.t. } \begin{array}{c} a + \gamma \in \hat{n}\mathcal{I}, \\ |\hat{n}\alpha - \gamma - a| < \Psi(n), \\ (a, \hat{n}) \text{ is } (\gamma,\eta)\text{-shift-reduced} \end{array} \right\}. \tag{3.2}$$

We call these *(localised) approximation sets*. Note that if n is large in terms of $\mathcal{I}$ then

$$\mu(\mathcal{E}_n^{\mathcal{I},\gamma}) \ll \mu(\mathcal{I})\Psi(n). \tag{3.3}$$

Observe that

$$\mathcal{G} = \{n \in \mathbb{N} : \mu(\mathcal{E}_n) > 0\}.$$

We will often suppress the dependence on γ and $\mathcal{I}$ in the notation by writing $\mathcal{E}_n$ in place of $\mathcal{E}_n^{\mathcal{I},\gamma}$.

We say that γ is *admissible* if there exists $\eta \in (0,1)$ such that for any interval $\mathcal{I} \subseteq [0,1]$ there are infinitely positive integers X for which the following two properties hold:

(1) We have

$$\sum_{n \leqslant X} \mu(\mathcal{E}_n) \asymp \mu(\mathcal{I}) \sum_{n \leqslant X} \psi(\hat{n})(\log n)^{k-1}. \tag{3.4}$$

(2) The sets $\mathcal{E}_n$ are quasi-independent on average for those X, i.e.

$$\sum_{m,n \leqslant X} \mu(\mathcal{E}_n \cap \mathcal{E}_m) \ll \mu(\mathcal{I}) \left(\sum_{n \leqslant X} \psi(\hat{n})(\log n)^{k-1} \right)^2. \tag{3.5}$$

By (3.1), the right hand side of (3.4) is unbounded as a function of X. It is worth stressing that the implicit constants in (3.4) and (3.5) are only allowed to depend on $\boldsymbol{\alpha}, \gamma_1, \ldots, \gamma_k$, and are uniform in $\mathcal{I}$.

Now we show that Theorems 1.23 and 1.8 can be reduced to showing that every $\gamma \in \mathbb{R}$ is admissible. We assume throughout this section that

$$\Psi(n) \leqslant 1/2 \qquad (n \text{ large}), \tag{3.6}$$

as we may because otherwise these two theorems are trivial.

Remark 3.2. Unless otherwise specified, '(sufficiently) large' means large in terms of $\boldsymbol{\alpha}, \gamma_1, \ldots, \gamma_k, \psi, \mathcal{I}$. Similarly, unless otherwise specified, a positive real number is '(sufficiently) small' if it is small in terms of $\boldsymbol{\alpha}, \gamma_1, \ldots, \gamma_k, \psi, \mathcal{I}$.

Proposition 3.3. *If every $\gamma \in \mathbb{R}$ is admissible, then Theorems 1.23 and 1.8 are true.*

Proof. Let $\gamma \in \mathbb{R}$, and let η be as in the definition of admissibility. We first show that $\mathcal{W}(\Psi; \gamma, \eta) := \limsup_{n\to\infty} \mathcal{A}_n$ has full measure, where

$$\mathcal{A}_n = \left\{ \alpha \in [0,1] : \exists a \in \mathbb{Z} \text{ s.t. } \begin{array}{c} |\hat{n}\alpha - \gamma - a| < \Psi(n), \\ (a, \hat{n}) \text{ is } (\gamma, \eta)\text{-shift-reduced} \end{array} \right\}.$$

Fix a non-empty subinterval $\mathcal{I}'$ of $[0,1]$, and let $\mathcal{I}$ be its dilation by $1/2$ about its centre. Observe using the triangle inequality that if $n \geqslant n_0(\mathcal{I})$ then $\mathcal{E}_n^{\mathcal{I},\gamma} \subseteq \mathcal{I}'$. Let $\mathcal{R}^{\mathcal{I},\gamma} = \limsup\limits_{n\to\infty} \mathcal{E}_n^{\mathcal{I},\gamma}$. Inserting (3.4) and (3.5) into Lemma 2.9, we obtain

$$\mu(\mathcal{W}(\Psi; \gamma, \eta) \cap \mathcal{I}') \geqslant \mu(\mathcal{R}^{\mathcal{I},\gamma}) \gg \mu(\mathcal{I}) \gg \mu(\mathcal{I}'),$$

where the implied constant is independent of $\mathcal{I}'$. Lemma 2.10 now grants us that $\mathcal{W}(\Psi; \gamma, \eta)$ has full measure in $[0,1]$. As $\Psi(n) \leqslant \Phi(\hat{n})$ for all n, we obtain Theorem 1.23. By 1-periodicity of $\|\cdot\|$, we thereby obtain Theorem 1.8. □

For the remainder of this section, we establish that any $\gamma \in \mathbb{R}$ is admissible. In fact, we will verify *a fortiori* that the properties (3.4) and (3.5) of the approximation sets hold for all sufficiently large values of X. Moreover, the first property will hold for all $\eta \in (0,1)$.

3.2. **Divergence of the series.** Let $\gamma \in \mathbb{R}$ and $\eta \in (0,1)$, and let $X \in \mathbb{N}$ be large. Let ε_0 be small, as in Remark 3.1. In this subsection, we establish the property (3.4). We begin by estimating $\mu(\mathcal{E}_n)$.

Lemma 3.4. *Let $n \in \mathcal{G}$ be large. Then*

$$\frac{\varphi(\hat{n})}{\hat{n}} \mu(\mathcal{I})\Psi(n) \ll \mu(\mathcal{E}_n) \ll \mu(\mathcal{I})\Psi(n). \tag{3.7}$$

Proof. The enunciated upper bound follows at once by observing that $\mathcal{E}_n$ is contained in $O(\hat{n}\mu(\mathcal{I}))$ many intervals of length $2\Psi(n)/\hat{n}$. On the other hand, it contains

$$\varphi_0(\hat{n}) := \sum_{\substack{a+\gamma\in\hat{n}\mathcal{I}\\(q_t'a+c_t,\hat{n})=1}} 1$$

many open intervals of length $2\Psi(n)/\hat{n}$ and, by (3.6), these are disjoint. We proceed to show that

$$\varphi_0(\hat{n}) \gg \mu(\mathcal{I})\varphi(\hat{n}), \tag{3.8}$$

which would complete the proof. We will find that the implied constant in (3.8) is absolute.

To infer (3.8), we apply Theorem 2.17 with the following specifications. The set $\mathcal{P}$ of relevant primes is the set of primes that divide $\hat{n}$ but not q_t', so that

$$P(z) = \prod_{p\in\mathcal{P}_{<z}} p,$$

and the sifting sequence is

$$a_m = \begin{cases} 1, & \text{if } \exists a \in (\hat{n}\mathcal{I}-\gamma)\cap\mathbb{Z} \quad q_t'a + c_t = m \\ 0, & \text{otherwise.} \end{cases}$$

We choose

$$x = q_t'(\hat{n}+1), \qquad w = 2, \qquad z = (\log\hat{n})^2.$$

The other relevant data are

$$g(p) = 1/p \quad (p\in\mathcal{P}), \qquad \kappa = 1, \qquad X = \hat{n}\mu(\mathcal{I}) = O(1) + \sum_{m\leqslant x} a_m,$$

$K > 1$ is an absolute constant, and

$$s = 10(1+\log K), \qquad D = z^s.$$

The dimension condition (2.16) follows from Mertens' third theorem (Theorem 2.16). With

$$V(z) = \prod_{p\in\mathcal{P}_{<z}} (1-p^{-1}),$$

we obtain

$$\sum_{(m,P(z))=1} a_m \geqslant (1-e^{9-s}K^{10})XV(z) - \sum_{\substack{d|P(z)\\d<D}} |r_d|,$$

where

$$r_d + d^{-1}X = \sum_{m\equiv 0 \bmod d} a_m.$$

For $d < D$ dividing $P(z)$, we have $(d, q_t') = 1$, so

$$\sum_{m \equiv 0 \bmod d} a_m = d^{-1}X + O(1),$$

ergo $r_d \ll 1$. Therefore

$$\sum_{\substack{d \mid P(z) \\ d < D}} |r_d| \ll D = (\log n)^{O(1)},$$

the upshot being that

$$\sum_{(m, P(z))=1} a_m \geqslant (1 - e^{-1})XV(z) + o(XV(z)) \gg XV(z).$$

Finally, the union bound gives

$$\sum_{(m, P(x)/P(z))>1} a_m \ll X \sum_{\substack{p \mid n \\ p \geqslant z}} p^{-1} \ll \frac{X \log n}{z \log z} = o(XV(z)),$$

and so

$$\varphi_0(\hat{n}) = \sum_{(m, P(x))=1} a_m \gg XV(z) \gg \mu(\mathcal{I})\varphi(\hat{n}).$$

□

Let $C \geqslant 4$ be an integer constant, large in terms of the implied constants in Lemmata 2.6 and 2.7. We assume for a purely technical reason that C is a perfect square, and let N be large in terms of C and the other constants. To estimate $\sum_{n \leqslant X} \frac{\varphi(\hat{n})}{\hat{n}} \mu(\mathcal{I})\Psi(n)$, it will be useful to gather all n on a scale N such that, additionally, the $\|\hat{n}\alpha_i - \gamma_i\|$ are in prescribed C-adic ranges; here and in the sequel, a C-adic range is a subinterval of $(0, \infty)$ whose right endpoint is C times the left endpoint.

Define

$$\begin{aligned} \hat{\mathcal{B}}_{\mathrm{loc}}(N; \boldsymbol{\rho}) &= \left\{ h \in \mathbb{N} : \quad \begin{array}{l} \hat{N} < h \leqslant \widehat{CN}, \\ \rho_j < \|h\alpha_j - \gamma_j\| \leqslant C\rho_j \quad (1 \leqslant j \leqslant k-1) \end{array} \right\} \\ &= \mathcal{B}(\widehat{CN}; C\boldsymbol{\rho}) \setminus \left(\mathcal{B}(\hat{N}; C\boldsymbol{\rho}) \cup \bigcup_{i=1}^{k-1} \mathcal{B}(\widehat{CN}; \boldsymbol{\rho}_i) \right), \end{aligned} \tag{3.9}$$

where $\boldsymbol{\rho}_i = (\rho_{i,1}, \ldots, \rho_{i,k-1})$, and $\rho_{i,j}$ is defined to be ρ_j if $i = j$ and $C\rho_j$ otherwise. For this section $\boldsymbol{\rho}$ denotes a parameter in the hyperrectangle

$$\mathcal{W}(N) := [\hat{N}^{-4.1\varepsilon_0}, \hat{N}^{-1.9\varepsilon_0}]^{k-1}. \tag{3.10}$$

Similarly we will have a large parameter $M \leqslant N$, and $\boldsymbol{\delta}$ will denote a parameter in $\mathcal{W}(M)$. We also write

$$\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}) = \{n \in \mathbb{N} : \hat{n} \in \hat{\mathcal{B}}_{\mathrm{loc}}(N;\boldsymbol{\rho})\}. \tag{3.11}$$

By the construction of f, there is a uniquely determined integer $u = u(N)$ with

$$f(n) \in \{u, u+1\} \qquad (N \leqslant n \leqslant CN). \tag{3.12}$$

Furthermore, since $f(n) \ll \log\log n$, we know that if $n \asymp N$ then $\hat{n}/n = N^{o(1)}$. We have ample use for this estimate, and shall use it without further mention.

We proceed to study these localised Bohr sets, first deriving a cardinality estimate in the range of interest. Heuristically, one might expect that each of the $\Theta(\hat{N})$ many integers in the interval $[\hat{N}, \widehat{CN}]$ lies in $\{\hat{n} : n \in \mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})\}$ with probability roughly $\Pi(\boldsymbol{\rho})/4^{f(N)}$, recalling the notation (1.19). For the ranges occurring implicitly in the set $\mathcal{G}$, this heuristic correctly predicts the order of magnitude of $\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})$:

Lemma 3.5. *We have*

$$\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}) \asymp \Pi(\boldsymbol{\rho})N,$$

uniformly for $\boldsymbol{\rho} \in \mathcal{W}(N)$. Here 'uniformly' means that the implied constants are the same for all $\boldsymbol{\rho} \in \mathcal{W}(N)$.

Proof. In light of (3.12), we have

$$4^{f(N)}, 4^{f(n)} \in \{2^{\nu}, 2^{\nu+2}\} \qquad (n \in \mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})), \tag{3.13}$$

for some positive integer ν, where

$$2^{\nu} \leqslant 4^{O(\log\log N)} = (\log N)^{O(1)}.$$

For the upper bound, each $n \in \mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})$ gives rise to a different element of $\hat{\mathcal{B}}_{\mathrm{loc}}(N;\boldsymbol{\rho}) = \mathcal{B}(\widehat{CN}; C\boldsymbol{\rho})$ that is divisible by 2^{ν}. We can count the latter using Lemmata 2.7 and 2.15 therein. Indeed, the former furnishes a generalised arithmetic progression

$$\mathcal{P} = \mathcal{P}(b; A_1, \ldots, A_k; N_1, \ldots, N_k)$$

containing $\mathcal{B}(\widehat{CN}; C\boldsymbol{\rho})$, where

$$N_1 \cdots N_k \ll \widehat{CN} \cdot \Pi(\boldsymbol{\rho}) \ll \hat{N} \cdot \Pi(\boldsymbol{\rho}), \qquad \min_{i \leqslant k} N_i \geqslant N^{20k\varepsilon_0},$$

and $\gcd(A_1, \ldots, A_k) = 1$. Therefore

$$\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}) \leqslant \#\{h \in \mathcal{P} : 2^{\nu} \mid h\},$$

whereupon (2.14) yields

$$\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}) \ll \frac{N_1\cdots N_k}{2^\nu} \ll \frac{\hat{N}\cdot\Pi(\boldsymbol{\rho})}{2^\nu} \ll \frac{\hat{N}\cdot\Pi(\boldsymbol{\rho})}{4^{f(N)}} = N\cdot\Pi(\boldsymbol{\rho}).$$

For the lower bound, we divide the interval $(N, CN]$ into two subintervals $(N, \sqrt{C}N]$ and $(\sqrt{C}N, CN]$, and observe that f must be constant on at least one of these subintervals. We assume that $4^{f(n)} = 2^\nu$ on $(N, \sqrt{C}N]$; the cases involving $(\sqrt{C}N, CN]$ and/or $2^{\nu+2}$ can be dealt with in the same manner. A lower bound is then given by the number of elements of

$$\begin{aligned}&\hat{\mathcal{B}}_{\mathrm{loc}}(N;\boldsymbol{\rho}) \cap (2^\nu N, \sqrt{C}2^\nu N]\\ &\qquad = \mathcal{B}(\widehat{\sqrt{C}N}; C\boldsymbol{\rho}) \setminus \left(\mathcal{B}(\hat{N}; C\boldsymbol{\rho}) \cup \bigcup_{i=1}^{k-1} \mathcal{B}(\widehat{\sqrt{C}N}; (\rho_{i,1},\ldots,\rho_{i,k-1}))\right)\end{aligned}$$

that are divisible by 2^ν, which we can estimate using Lemmata 2.6, 2.7, and 2.15. The point is that C is large, so the count for $\mathcal{B}(\widehat{\sqrt{C}N}; C\boldsymbol{\rho})$ dominates, as we explain further in the next paragraph.

For the remainder of the proof, our implied constants do not depend on C. The argument that we used for the upper bound yields

$$\#\{h \in \mathcal{B}(\hat{N}; C\boldsymbol{\rho}) : 2^\nu \mid h\} \ll C^{k-1}\Pi(\boldsymbol{\rho})N,$$

as well as

$$\#\{h \in \mathcal{B}(\widehat{\sqrt{C}N}; (\rho_{i,1},\ldots,\rho_{i,k-1})) : 2^\nu \mid h\} \ll \sqrt{C}N \prod_{j\leqslant k-1} \rho_{i,j} = C^{k-3/2}\Pi(\boldsymbol{\rho})N$$

for $i = 1, 2, \ldots, k-1$. On the other hand, Lemma 2.6 furnishes a proper, asymmetric generalised arithmetic progression

$$\mathcal{P}' = \mathcal{P}^+(b'; A_1', \ldots, A_k'; N_1', \ldots, N_k')$$

contained in $\mathcal{B}(\widehat{\sqrt{C}N}; C\boldsymbol{\rho})$, where

$$N_1'\cdots N_k' \gg \widehat{\sqrt{C}N}\cdot C^{k-1}\Pi(\boldsymbol{\rho}) \gg \hat{N}\cdot C^{k-1/2}\Pi(\boldsymbol{\rho}), \qquad \min_{i\leqslant k} N_i' \gg N^{20k\varepsilon_0},$$

and $\gcd(A_1', \ldots, A_k') = 1$. Thus, by (2.15), we have

$$\#\{h \in \mathcal{B}(\widehat{\sqrt{C}N}; C\boldsymbol{\rho}) : 2^\nu \mid h\} \geqslant \#\{h \in \mathcal{P}' : 2^\nu \mid h\} \gg C^{k-1/2}\Pi(\boldsymbol{\rho})N.$$

As C is large, we conclude that

$$\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}) \gg C^{k-1/2}\Pi(\boldsymbol{\rho})N,$$

which completes the proof. □

Before we demonstrate (3.4), we prepare one more lemma. Since (3.4) requires us to determine the order of magnitude of $\sum_{n\leqslant N}\mu(\mathcal{E}_n)$, it is helpful to know that the totient weights on the left hand side of (3.7) 'average well' in suitable C-adic ranges, i.e. that for n in a localised Bohr set $\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})$, the average of $\varphi(\hat{n})/\hat{n}$ has order of magnitude 1.

Remark 3.6. A similar strategy for proving this point was employed by the first named author in [19, Lemma 3.1], and then by both authors in [21, Lemma 4.1]. The additional difficulty here is a mild one: Roughly speaking, we need to average only over elements of the Bohr sets which are divisible by an appropriate power of four.

Lemma 3.7 (Good averaging)**.** *We have*

$$\sum_{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}\frac{\varphi(\hat{n})}{\hat{n}}\gg\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}),$$

uniformly for all $\boldsymbol{\rho}\in\mathcal{W}(N)$.

Proof. By the AM–GM inequality, we have

$$\frac{1}{\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}\sum_{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}\frac{\varphi(\hat{n})}{\hat{n}}\geqslant\left(\prod_{n\in\mathcal{P}}\frac{\varphi(\hat{n})}{\hat{n}}\right)^{\frac{1}{\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}}=\prod_{p\leqslant\widehat{CN}}\left(1-\frac{1}{p}\right)^{\tau_p},$$

where

$$\tau_p=\frac{\#\{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho}):p\mid\hat{n}\}}{\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}.$$

Now

$$-\ln\left(\frac{1}{\#\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}\sum_{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})}\frac{\varphi(\hat{n})}{\hat{n}}\right)\ll\sum_{p\leqslant\widehat{CN}}\frac{\tau_p}{p},$$

so as $\tau_2\leqslant 1$ it suffices to show that

$$\tau_p\ll p^{-\varepsilon_0}\qquad(3\leqslant p\leqslant\widehat{CN}).$$

Suppose $3\leqslant p\leqslant\widehat{CN}$. By Lemma 3.5, we have

$$\tau_p\ll\frac{\#\{h\in\mathcal{B}(\widehat{CN};C\boldsymbol{\rho}):h\equiv 0\bmod 2^{\nu}p\}}{\rho N},$$

where ν is as in (3.13) and $\rho = \Pi(\boldsymbol{\rho})$. We can estimate the numerator via Lemmata 2.7 and 2.15, noting that

$$N_1 \cdots N_k \asymp \widehat{CN} \cdot \rho, \qquad \min_{i \leqslant k} N_i \geqslant N^{20k\varepsilon_0}$$

therein. We obtain

$$\tau_p \ll 2^\nu((2^\nu p)^{-1} + N^{-20k\varepsilon_0}) \ll p^{-\varepsilon_0}.$$

□

We can now estimate the normalised-totient-weighted sums of measures, thereby attaining the main result of this subsection. For $j \in \mathbb{N}$, define

$$\mathcal{D}_j = (C^j, C^{j+1}] \cap \mathcal{G}_\diamond, \qquad \mathcal{G}_j = (C^j, C^{j+1}] \cap \mathcal{G}. \tag{3.14}$$

Lemma 3.8. *The sets $\mathcal{E}_n$ satisfy* (3.4) *for all sufficiently large X.*

Proof. For j large, observe that $\mathcal{D}_j$ is contained in a union of $O(j^{k-1})$ sets

$$\mathcal{B}_{\mathrm{loc}}(C^j; (C^{-t_1}, \ldots, C^{-t_{k-1}}))$$

satisfying the hypotheses of Lemma 3.5, and so

$$T_j := \sum_{n \in \mathcal{D}_j} \frac{1}{\|\hat{n}\alpha_1 - \gamma_1\| \cdots \|\hat{n}\alpha_{k-1} - \gamma_{k-1}\|} \ll j^{k-1} C^j. \tag{3.15}$$

Let J be the largest integer j such that $C^j \leqslant X$. By (3.3), we have

$$\sum_{C^j < n \leqslant C^{j+1}} \mu(\mathcal{E}_n) \ll \sum_{n \in \mathcal{D}_j} \mu(\mathcal{I})\Psi(n) \leqslant \mu(\mathcal{I})\psi(\widehat{C^j})T_j$$

for j large. Applying (3.15) and summing over j yields

$$\sum_{j \leqslant J} \psi(\widehat{C^j})T_j \ll 1 + \sum_{j \leqslant J} j^{k-1} C^j \psi(\widehat{C^j}).$$

By (2.12), we now have

$$\mu(\mathcal{I})^{-1} \sum_{n \leqslant X} \mu(\mathcal{E}_n) \ll 1 + \sum_{j \leqslant J} j^{k-1} C^j \psi(\widehat{C^j}) \ll \sum_{n \leqslant X} \psi(\hat{n})(\log n)^{k-1},$$

recalling that X is large and recalling from (3.1) that the right hand side diverges as $X \to \infty$. We have established the upper bound in (3.4).

For the lower bound, we begin by applying (3.7) to give

$$\sum_{n \leqslant X} \mu(\mathcal{E}_n) \gg \mu(\mathcal{I}) \sum_{C_1 < n \leqslant X} \frac{\varphi(\hat{n})}{\hat{n}} \Psi(n),$$

where C_1 is a large, positive constant. Our strategy is to show that

$$\sum_{\substack{n\in\mathcal{G}_\diamond\\ n\leqslant X}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \gg \sum_{n\leqslant X}\psi(\hat{n})(\log n)^{k-1} \tag{3.16}$$

and

$$\sum_{n\in\mathcal{G}_\diamond\setminus\mathcal{G}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \ll 1. \tag{3.17}$$

By (3.1), the right hand size of (3.16) diverges as $X\to\infty$, and so these two bounds will suffice.

We proceed to bound from below the contribution from $n\in\mathcal{D}_j$, for j large. To this end, consider the localised Bohr sets $\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))$, where

$$\boldsymbol{\rho}(\mathbf{t}) = (C^{-t_1},\dots,C^{-t_{k-1}}).$$

For $n\in\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))$, where $\mathbf{t}=(t_1,\dots,t_{k-1})$ satisfies

$$2.1\varepsilon_0(j+1)\leqslant t_r\leqslant 3.9\varepsilon_0 j \quad (1\leqslant r\leqslant k-1),$$

we deduce that

$$\hat{n}^{-4\varepsilon_0}\leqslant C^{-t_r}\leqslant C^{1-t_r}\leqslant \hat{n}^{-2\varepsilon_0} \qquad (1\leqslant r\leqslant k-1).$$

Hence $n\in\mathcal{D}_j$, and so we see that

$$\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))\subseteq\mathcal{D}_j.$$

Next, observe that

$$\sum_{n\in\mathcal{D}_j} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \gg \psi(\widehat{C^{j+1}})\sum_{\mathbf{t}}\sum_{n\in\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))}\frac{\varphi(\hat{n})}{\hat{n}\rho(\mathbf{t})},$$

where $\mathbf{t}$ runs through all the integer vectors as above and $\rho(\mathbf{t})=\Pi(\boldsymbol{\rho}(\mathbf{t}))$. Lemma 3.5 implies $\#\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))\gg C^j\rho(\mathbf{t})$, and so Lemma 3.7 assures us that

$$\sum_{n\in\mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\rho}(\mathbf{t}))}\frac{\varphi(\hat{n})}{\hat{n}\rho(\mathbf{t})}\gg C^j,$$

uniformly for each of the $\Theta(j^{k-1})$ many choices of $\mathbf{t}$. Whence

$$\sum_{n\in\mathcal{D}_j} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \gg \psi(\widehat{C^{j+1}})C^{j+1}(j+1)^{k-1}. \tag{3.18}$$

Let J_0 be large, positive constant. Summing (3.18) over the range

$$J_0\leqslant j\leqslant\frac{\log X}{\log C}$$

yields

$$\sum_{\substack{n\in\mathcal{G}_\diamond\\ n\leqslant X}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \gg \sum_{J_0\leqslant j\leqslant \frac{\log X}{\log C}} \psi(\widehat{C^{j+1}})C^{j+1}(j+1)^{k-1}.$$

Lemma 2.12 now delivers the inequality (3.16).

Finally, by (3.15) we have

$$\begin{aligned}
&\sum_{n\in\mathcal{G}_\diamond\setminus\mathcal{G}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|}\\
&\leqslant \sum_{j=1}^{\infty}\sum_{n\in\mathcal{D}_j\setminus\mathcal{G}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|}\\
&\ll 1+\sum_{j=1}^{\infty}\frac{j^{k-1}C^j}{C^j(j\log C)^{k+1}} \ll 1+\sum_{j=1}^{\infty} j^{-2} \ll 1,
\end{aligned}$$

which is (3.17). □

Having established (3.4), our final task for this section is to prove (3.5) for some $\varepsilon_0>0$ and $\eta\in(0,1)$ depending on γ.

3.3. **Overlap estimates, localised Bohr sets, and the small-GCD regime.** In the present subsection, we reduce the task of proving (3.5) to demonstrating a uniform estimate in localised Bohr sets, see Lemma 3.9. Thereafter, we recast the desired estimate in terms of counting solutions to diophantine inequalities in localised Bohr sets, see Lemma 3.10. At the end of this subsection, we establish such a counting result in a regime where the arising GCDs are relatively small, see Proposition 3.11.

Let $N\geqslant M$, where M is large. The next lemma asserts that if we have a good bound on

$$R^{\mathcal{I},\gamma}(M,N;\boldsymbol{\rho},\boldsymbol{\delta}) := \sum_{\substack{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})\\ m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})\\ n\neq m}} \mu(\mathcal{E}_n^{\mathcal{I},\gamma}\cap\mathcal{E}_m^{\mathcal{I},\gamma}),$$

sufficiently uniform in $\boldsymbol{\rho}$ and $\boldsymbol{\delta}$, then we can deduce (3.5). As before, we shall drop the dependency on γ and $\mathcal{I}$ for most of the time, and simply write $R(M,N;\boldsymbol{\rho},\boldsymbol{\delta})$ instead.

Lemma 3.9. *If*

$$R(M,N;\boldsymbol{\rho},\boldsymbol{\delta}) \ll \mu(\mathcal{I})MN\psi(\hat{M})\psi(\hat{N}), \tag{3.19}$$

uniformly for M large and $N \geqslant M$, $\boldsymbol{\rho} \in \mathcal{W}(N)$, and $\boldsymbol{\delta} \in \mathcal{W}(M)$, then (3.5) *holds.*

Proof. Let J be the least integer j such that $C^j \geqslant N$, and recall (3.14). Then

$$\sum_{m,n \leqslant X} \mu(\mathcal{E}_n \cap \mathcal{E}_m) \ll 1 + \sum_{j \leqslant i \leqslant J} \sum_{\substack{n \in \mathcal{G}_i \\ m \in \mathcal{G}_j}} \mu(\mathcal{E}_n \cap \mathcal{E}_m).$$

Let $X_0 \in \mathbb{N}$ be a large constant. By Lemma 3.8, the contribution from $j < X_0$ is bounded by a constant times

$$\sum_{n \leqslant X} \mu(\mathcal{E}_n) \ll \mu(\mathcal{I}) \sum_{n \leqslant X} \psi(\hat{n}) (\log n)^{k-1}.$$

Thus, recalling from (3.1) that the completed series diverges, it remains to show that

$$\sum_{X_0 \leqslant j \leqslant i \leqslant J} \sum_{\substack{n \in \mathcal{G}_i \\ m \in \mathcal{G}_j}} \mu(\mathcal{E}_n \cap \mathcal{E}_m) \ll \mu(\mathcal{I}) \left(\sum_{n \leqslant X} \psi(\hat{n}) (\log n)^{k-1} \right)^2. \tag{3.20}$$

Presently, we fix i, j with $X_0 \leqslant j \leqslant i \leqslant J$. Consider the vectors

$$\boldsymbol{\rho}(\mathbf{t}) = (C^{-t_1}, \dots, C^{-t_{k-1}}), \quad \boldsymbol{\delta}(\boldsymbol{\ell}) = (C^{-\ell_1}, \dots, C^{-\ell_{k-1}}), \tag{3.21}$$

wherein ranges for the exponents will be prescribed shortly. Let us account for the contribution of the diagonal first. By (3.1), the summands for which $n = m$ contribute

$$\sum_{X_0 \leqslant i \leqslant J} \sum_{n \in \mathcal{G}_i} \mu(\mathcal{E}_n) \ll \mu(\mathcal{I}) \sum_{n \leqslant X} \psi(\hat{n}) (\log n)^{k-1} \ll \mu(\mathcal{I}) \left(\sum_{n \leqslant X} \psi(\hat{n}) (\log n)^{k-1} \right)^2.$$

Next, we account for the off-diagonal contribution. To this end, we approximately decompose

$$S(i,j) := \sum_{\substack{n \in \mathcal{G}_i \\ m \in \mathcal{G}_j \\ n \neq m}} \mu(\mathcal{E}_n \cap \mathcal{E}_m)$$

into sums of the shape

$$S(i,j;\mathbf{t},\boldsymbol{\ell}) := \sum_{\substack{n \in \mathcal{B}_{\mathrm{loc}}(C^i;\boldsymbol{\rho}(\mathbf{t})) \\ m \in \mathcal{B}_{\mathrm{loc}}(C^j;\boldsymbol{\delta}(\boldsymbol{\ell})) \\ n \neq m}} \mu(\mathcal{E}_n \cap \mathcal{E}_m).$$

Specifically, the sum $S(i,j)$ is bounded above by the sum of $S(i,j;\mathbf{t},\boldsymbol{\ell})$ over the integer vectors $\mathbf{t}$, $\boldsymbol{\ell}$ within the ranges

$$1.9\varepsilon_0 i \leqslant t_1,\dots,t_{k-1} \leqslant 4.1\varepsilon_0 i, \qquad 1.9\varepsilon_0 j \leqslant \ell_1,\dots,\ell_{k-1} \leqslant 4.1\varepsilon_0 j.$$

By the assumed estimate (3.19), we have

$$S(i,j;\mathbf{t},\boldsymbol{\ell}) \ll \mu(\mathcal{I})C^iC^j\hat{\psi}(C^i)\hat{\psi}(C^j),$$

uniformly in $\mathbf{t}$ and $\boldsymbol{\ell}$ as above. Summing over the $O(i^{k-1})$ many choices for $\mathbf{t}$ and the $O(j^{k-1})$ many choices for $\boldsymbol{\ell}$, we see that

$$S(i,j) \ll \mu(\mathcal{I})i^{k-1}C^ij^{k-1}C^j\hat{\psi}(C^i)\hat{\psi}(C^j).$$

Summing over i and j gives

$$\sum_{X_0\leqslant j\leqslant i\leqslant J}\sum_{\substack{n\in\mathcal{G}_i\\ m\in\mathcal{G}_j}}\mu(\mathcal{E}_n\cap\mathcal{E}_m) \ll \mu(\mathcal{I})\Bigg(\sum_{X_0\leqslant j\leqslant J} j^{k-1}C^j\hat{\psi}(C^j)\Bigg)^2.$$

Now Lemma 2.12 delivers (3.20), completing the proof. □

Denote by $N_1,\dots,N_k$ the length parameters, arising from Lemma 2.7, associated to the outer structure of $\hat{\mathcal{B}}_{\mathrm{loc}}(N;\boldsymbol{\rho})$. Specifically, we apply the lemma to $\mathcal{B}(\widehat{CN};C\boldsymbol{\rho})$, with $\varepsilon=\varepsilon_0$ and $\vartheta=20k$. By symmetry, we may assume that

$$N_1 = \min\{N_i : 1\leqslant i\leqslant k\}, \tag{3.22}$$

and we do so in order to simplify notation. Note that

$$N_1 \geqslant N^{20k\varepsilon_0}.$$

Let M be large, let $N\geqslant M$, and let

$$\Delta = \widehat{CM}\frac{\psi(\hat{N})}{\Pi(\boldsymbol{\rho})} + \widehat{CN}\frac{\psi(\hat{M})}{\Pi(\boldsymbol{\delta})}. \tag{3.23}$$

Importantly, if there exists $m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})\cap\mathcal{G}$ then we have a lower bound on Δ of the strength

$$\begin{aligned}\Delta &\geqslant \widehat{CN}\frac{\psi(\hat{M})}{\Pi(\boldsymbol{\delta})} \geqslant \widehat{CN}\frac{\psi(\hat{m})}{\Pi(\boldsymbol{\delta})} \geqslant \frac{\widehat{CN}}{\Pi(\boldsymbol{\delta})m(\log m)^{k+1}}\\ &\geqslant \frac{\widehat{CN}}{CM}\,\frac{1}{\Pi(\boldsymbol{\delta})(\log(CM))^{k+1}} \geqslant \frac{1}{\Pi(\boldsymbol{\delta})(\log(CM))^{k+1}}.\end{aligned}$$

For $\boldsymbol{\delta}\in\mathcal{W}(M)$, recalling (3.10), we find that for all sufficiently large M we have

$$\Delta > M^{(k-1)\varepsilon_0} > 1. \tag{3.24}$$

Let $(q'_\ell)_\ell$ be the sequence of continued fraction denominators of γ. We estimate the contribution to the quantity $R(M,N;\boldsymbol{\rho},\boldsymbol{\delta})$ from the small-GCD regime (including the diagonal), i.e.

$$R_{\gcd\leqslant}(M,N;\boldsymbol{\rho},\boldsymbol{\delta}) := \sum_{\substack{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})\\ m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})\\ \gcd(\hat{n},\hat{m})\leqslant\max\{3\Delta/\|q'_2\gamma\|,N_1\}}} \mu(\mathcal{E}_n\cap\mathcal{E}_m), \tag{3.25}$$

and the contribution from large-GCD regime, i.e.

$$R_{\gcd>}(M,N;\boldsymbol{\rho},\boldsymbol{\delta}) := \sum_{\substack{n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})\\ m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})\\ \gcd(\hat{n},\hat{m})>\max\{3\Delta/\|q'_2\gamma\|,N_1\}\\ n\neq m}} \mu(\mathcal{E}_n\cap\mathcal{E}_m), \tag{3.26}$$

separately. When bounding these quantities, we may assume that $\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})$ intersects $\mathcal{G}$, and so we have (3.24).

Let us write

$$\rho=\Pi(\boldsymbol{\rho}),\qquad \delta=\Pi(\boldsymbol{\delta}).$$

As a first step towards estimating $R_{\gcd\leqslant}(M,N;\boldsymbol{\rho},\boldsymbol{\delta})$, we record a useful relation between the size of $R_{\gcd\leqslant}(M,N;\boldsymbol{\rho},\boldsymbol{\delta})$ and the number of solutions to a diophantine inequality with various constraints.

Lemma 3.10. *Let M be large, let $N\geqslant M$, and let $\boldsymbol{\rho}\in\mathcal{W}(N)$, and $\boldsymbol{\delta}\in\mathcal{W}(M)$. Denote by $D_{\gcd\leqslant}$ the number of quadruples $(n,m,a,b)\in\mathbb{N}^2\times\mathbb{Z}^2$ for which*

$$\frac{a+\gamma}{\hat{n}},\frac{b+\gamma}{\hat{m}}\in\mathcal{I},\quad n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})\cap\mathcal{G},\quad m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})\cap\mathcal{G}, \tag{3.27}$$

as well as

$$\gcd(\hat{n},\hat{m})\leqslant\max(3\Delta/\|q'_2\gamma\|,N_1) \tag{3.28}$$

and

$$|(\hat{n}-\hat{m})\gamma-(\hat{m}a-\hat{n}b)|\leqslant\Delta, \tag{3.29}$$

and finally

$$(a,\hat{n}),(b,\hat{m})\ \text{are}\ (\gamma,\eta)\text{-shift-reduced.} \tag{3.30}$$

Furthermore, let $D_{\gcd>}$ be the number of quadruples $(n,m,a,b)\in\mathbb{N}^2\times\mathbb{Z}^2$ which satisfy (3.27) *,* (3.29)*, and* (3.30)*, but instead of* (3.28) *the reversed inequality* $\gcd(\hat{n},\hat{m})>\max(3\Delta/\|q'_2\gamma\|,N_1)$ *and the constraint $n\neq m$. If*

$$D_{\gcd\leqslant}=O(\mu(\mathcal{I})\rho N\delta M\Delta), \tag{3.31}$$

then $R_{\gcd\leqslant}(M,N;\boldsymbol{\rho},\boldsymbol{\delta})\ll\mu(\mathcal{I})MN\psi(\hat{M})\psi(\hat{N})$. Moreover, if we have

$$D_{\gcd>}=O(\mu(\mathcal{I})\rho N\delta M\Delta), \tag{3.32}$$

then $R_{\gcd>}(M,N;\boldsymbol{\rho},\boldsymbol{\delta})\ll\mu(\mathcal{I})MN\psi(\hat{M})\psi(\hat{N})$.

Proof. We detail the proof only in the case of $D_{\gcd\leqslant}$ since the case $D_{\gcd>}$ can be dealt with in the same way. We begin by observing that each $\mathcal{E}_n$ (resp. $\mathcal{E}_m$) is contained in a union of finitely many intervals

$$\mathcal{I}_{n,a} = \left(\frac{a+\gamma-\Psi(n)}{\hat{n}}, \frac{a+\gamma+\Psi(n)}{\hat{n}}\right)$$

(resp. $\mathcal{I}_{m,b}$), and so

$$\mu(\mathcal{E}_n \cap \mathcal{E}_m) \leqslant \max\{\min(\mu(\mathcal{I}_{n,a}), \mu(\mathcal{I}_{m,b})) : a, b \in \mathbb{Z}\} \cdot \#\{(a,b) : \mathcal{I}_{n,a} \cap \mathcal{I}_{m,b} \neq \emptyset\}.$$

The first factor is bounded above by

$$2\min\Big(\frac{\Psi(n)}{\hat{n}}, \frac{\Psi(m)}{\hat{m}}\Big) \leqslant 2\min\Big(\frac{\psi(\hat{N})}{\rho\hat{N}}, \frac{\psi(\hat{M})}{\delta\hat{M}}\Big).$$

To bound the second factor, we have to count how often the centre $(a+\gamma)/\hat{n}$ of $\mathcal{I}_{n,a}$ is 'sufficiently close' to the centre $(b+\gamma)/\hat{m}$ of an interval $\mathcal{I}_{m,b}$. Here sufficiently close means that the distance of the centres is less than sum of the radii of the intervals, i.e.

$$\Big|\frac{b+\gamma}{\hat{m}} - \frac{a+\gamma}{\hat{n}}\Big| \leqslant \frac{\Psi(n)}{\hat{n}} + \frac{\Psi(m)}{\hat{m}}.$$

Multiplying by $\hat{m}\hat{n}$, we see that

$$|\hat{n}(b+\gamma) - \hat{m}(a+\gamma)| \leqslant \widehat{CM}\Psi(n) + \widehat{CN}\Psi(m) \leqslant \Delta,$$

and the number of integer solutions (n, m, a, b) to the above inequality, subject to our constraints, is at most $D_{\gcd\leqslant}$. Thus, if (3.31) holds then

$$R_{\gcd\leqslant}(M, N; \boldsymbol{\rho}, \boldsymbol{\delta}) \ll \min\Big(\frac{\psi(\hat{N})}{\rho\hat{N}}, \frac{\psi(\hat{M})}{\delta\hat{M}}\Big)\mu(\mathcal{I})\rho N\delta M\Delta.$$

Next, observe that

$$\Delta \leqslant 2\widehat{CM}\widehat{CN}\max\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}}, \frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big).$$

It follows from (3.12) that $\widehat{N} \asymp \widehat{CN}$ and $\widehat{M} \asymp \widehat{CM}$, and so

$$\min\Big(\frac{\psi(\hat{N})}{\rho\hat{N}}, \frac{\psi(\hat{M})}{\delta\hat{M}}\Big) \ll \min\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}}, \frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big).$$

The upshot is that $R_{\gcd\leqslant}(M, N; \boldsymbol{\rho}, \boldsymbol{\delta})$ is at most a constant times

$$\min\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}}, \frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big)\rho N\delta M\mu(\mathcal{I})\widehat{CM}\widehat{CN}\max\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}}, \frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big). \tag{3.33}$$

Applying the simple identity

$$\min\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}},\frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big)\max\Big(\frac{\psi(\hat{N})}{\rho\widehat{CN}},\frac{\psi(\hat{M})}{\delta\widehat{CM}}\Big)=\frac{\psi(\hat{N})}{\rho\widehat{CN}}\frac{\psi(\hat{M})}{\delta\widehat{CM}},$$

we see that (3.33) simplifies to $\mu(\mathcal{I})MN\psi(\hat{M})\psi(\hat{N})$. □

Next, we establish overlap estimates in the regime in which the GCDs are relatively small. Here congruence considerations and the structural theory of Bohr sets will be decisive; a welcome circumstance is that the bound works for any final shift γ, irrespective of its diophantine nature.

Proposition 3.11. *Let M be large, and let $N \geqslant M$. Then*

$$R_{\gcd\leqslant}(M,N;\boldsymbol{\rho},\boldsymbol{\delta}) \ll \mu(\mathcal{I})MN\psi(\hat{M})\psi(\hat{N}),$$

uniformly for $\boldsymbol{\rho}\in\mathcal{W}(N)$ and $\boldsymbol{\delta}\in\mathcal{W}(M)$.

Proof. By Lemma 3.10, it remains to verify (3.31). We distinguish several cases according to the size of the of the greatest common divisor $d := (\hat{n},\hat{m})$. Put

$$\hat{n}=dx \quad\text{and}\quad \hat{m}=dy.$$

As f is non-decreasing, we have

$$d=(\hat{n},\hat{m})=(4^{f(n)}n,4^{f(m)}m)\geqslant 4^{f(M)}\geqslant \mu(\mathcal{I})^{-1},$$

where the last inequality comes from M being large in terms of $\mathcal{I}$. We rewrite the overlap inequality (3.29) as

$$|(x-y)\gamma-(ya-xb)|\leqslant \Delta/d. \tag{3.34}$$

Recall that $(q'_\ell)_\ell$ is the sequence of continued fraction denominators of γ.

Case 1: $\mu(\mathcal{I})^{-1}\leqslant d\leqslant 3\Delta/\|q'_2\gamma\|$

By Lemma 3.5, there are $O(\rho N\delta M)$ possible choices of $n\in\mathcal{B}_{\mathrm{loc}}(N;\boldsymbol{\rho})$ and $m\in\mathcal{B}_{\mathrm{loc}}(M;\boldsymbol{\delta})$. Now suppose we are given n and m, which then uniquely determine d,x,y.

Case 1a: $\Delta<\hat{n}/4$

Since x and y are coprime, the inequality (3.34) restricts the integer a to one of at most $2\Delta/d+1$ residue classes modulo x. Furthermore, the constraint $a+\gamma\in\hat{n}\mathcal{I}$ places the integer a in an interval of length $dx\mu(\mathcal{I})$. Within

this interval, and given $r \in \mathbb{Z}/x\mathbb{Z}$, there can be at most $d\mu(\mathcal{I}) + 1$ integers $a \equiv r \bmod x$. Thus, given n and m, there are at most

$$(d\mu(\mathcal{I}) + 1)(2\Delta/d + 1) \ll \mu(\mathcal{I})\Delta$$

possibilities for a. Finally, the inequality (3.34) constrains the integer b to an interval of length at most $2\Delta/\hat{n} < 1/2$, so b (if it exists at all) is determined by the other variables.

Case 1b: $\Delta \geqslant \hat{n}/4$

We may suppose that (3.29) has a solution $(a, b) = (a_0, b_0)$ with $a_0 + \gamma \in n\mathcal{I}$. By the triangle inequality, any solution to (3.29) for which $a + \gamma \in n\mathcal{I}$ satisfies

$$|\hat{m}a' - \hat{n}b'| \leqslant 2\Delta, \qquad a' \in \hat{n}(\mathcal{I} - \mathcal{I}), \tag{3.35}$$

where

$$a' = a - a_0, \qquad b' = b - b_0, \qquad \mathcal{I} - \mathcal{I} = \{r_1 - r_2 : r_1, r_2 \in \mathcal{I}\}.$$

Denote by $\mathcal{R}$ the closure of the set of $(a', b') \in \mathbb{R}^2$ satisfying (3.35). We apply Theorem 2.13 to the region $\mathcal{R}$ and the lattice $\mathbb{Z}^2$. Observe using the triangle inequality that $\mathcal{R}$ is contained in the rectangle

$$\left\{(a', b') \in \mathbb{R}^2 : a' \in \hat{n}(\mathcal{I} - \mathcal{I}), \quad |b'| \leqslant \frac{2\Delta}{\hat{n}} + 2\mu(\mathcal{I})\hat{m}\right\},$$

and in particular the projection of $\mathcal{R}$ onto any line has length at most

$$2\mu(\mathcal{I})\hat{n} + \frac{4\Delta}{\hat{n}} + 4\mu(\mathcal{I})\hat{m} \ll \hat{n}\mu(\mathcal{I}),$$

where for the final inequality we have used that $\psi(\hat{N}), \psi(\hat{M}) \leqslant \psi(1) \ll 1$, as well as that $\boldsymbol{\rho} \in \mathcal{W}(N)$ and $\boldsymbol{\delta} \in \mathcal{W}(M)$. The area of $\mathcal{R}$ is $O(\Delta\mu(\mathcal{I}))$, and we glean that the number of solutions is

$$O(\Delta\mu(\mathcal{I}) + \hat{n}\mu(\mathcal{I}) + 1) = O(\Delta\mu(\mathcal{I})).$$

We conclude that there are $O(\mu(\mathcal{I})\rho N\delta M\Delta)$ solutions in total coming from Case 1, uniformly in $\boldsymbol{\rho} \in \mathcal{W}(N)$ and $\boldsymbol{\delta} \in \mathcal{W}(M)$.

Case 2: $3\Delta/\|q_2'\gamma\| \leqslant d \leqslant N_1$

By Lemma 3.5, there are $O(\delta M)$ possibilities for $m \in \mathcal{B}_{\mathrm{loc}}(M; \boldsymbol{\delta})$. Next, we choose $d \in [\Delta, N_1]$ dividing m, in one of $M^{o(1)}$ ways. By Lemma 2.7, we have

$$dx = \hat{n} = A_1 n_1 + \cdots + A_k n_k + s,$$

where $|n_i| \leqslant N_i$ $(1 \leqslant i \leqslant k)$ and

$$N_1 \cdots N_k \asymp \rho \hat{N}.$$

As $d \leqslant N_1$, wherein we recall (3.22), Lemma 2.15 assures us that the congruence

$$A_1 n_1 + \cdots + A_k n_k + s \equiv 0 \bmod d$$

has $O(\rho \hat{N}/d)$ solutions $n_1, \ldots, n_k$, and these variables determine x.

Now suppose that we are given $d, x, y \in \mathbb{N}$ with $d \geqslant \Delta$ and $\gcd(x, y) = 1$. Then, as $d \geqslant \Delta$, by the same argument as in Case 1a there are at most

$$(d\mu(\mathcal{I}) + 1)(2\Delta/d + 1) \ll d\mu(\mathcal{I})$$

many choices for the pair $(a, b) \in \mathbb{Z}^2$ subject to (3.34) as well as $a + \gamma \in \hat{n}\mathcal{I}$. In view of (3.24), the total is again $O(\mu(\mathcal{I})\rho N \delta M \Delta)$.

We have established (3.31), completing the proof. □

3.4. **Large GCDs.** Let M be large, $N \geqslant M$, $\boldsymbol{\rho} \in \mathcal{W}(N)$, and $\boldsymbol{\delta} \in \mathcal{W}(M)$. On account of the preceding subsections, our final sub-task for this section is to establish (3.32). We will choose ε_0 and η according to how well γ can be approximated by rationals.

3.4.1. *Diophantine final shift.* Throughout the present subsubsection, the final shift $\gamma \in \mathbb{R}$ is diophantine, i.e. there exists $\lambda \geqslant 2$ such that

$$\|n\gamma\| \gg n^{-\lambda/2} \qquad (n \in \mathbb{N}). \tag{3.36}$$

This is equivalent to γ being irrational and non-Liouville. We fix such a value of λ throughout the current subsubsection.

The approximation sets are as in (3.2), however now we are more specific about ε_0 and η. Let $\tilde{\varepsilon}$ be the minimum of its values when Lemmata 2.6 and 2.7 are applied with $\vartheta = 20k$, and fix a small positive real number ε_0 such that

$$1000k^2\varepsilon_0 < \min\{\lambda^{-2}, \tilde{\varepsilon}\}. \tag{3.37}$$

In this diophantine case, the value of η is of no importance, and we arbitrarily choose $\eta = 1/2$.

Recall that $D_{\gcd>}$ denotes the number of quadruples $(m, n, a, b) \in \mathbb{N}^4$ satisfying

$$\frac{a+\gamma}{\hat{n}}, \frac{b+\gamma}{\hat{m}} \in \mathcal{I}, \quad n \in \mathcal{B}_{\mathrm{loc}}(N; \boldsymbol{\rho}) \cap \mathcal{G}, \quad m \in \mathcal{B}_{\mathrm{loc}}(M; \boldsymbol{\delta}) \cap \mathcal{G} \setminus \{n\},$$

as well as $d := \gcd(\hat{n}, \hat{m}) > \max\{3\Delta/\|q_2'\gamma\|, N_1\}$, (3.29), and (3.30). Put $\hat{n} = dx$, $\hat{m} = dy$, and note that (3.29) entails (3.34) and hence

$$\|(x-y)\gamma\| < \Delta/d \quad \text{and} \quad |x-y| < \widehat{CN}/d. \tag{3.38}$$

Combining the diophantine assumption (3.36) with (3.38) yields

$$\Delta/d \gg |x-y|^{-\lambda} \gg (\hat{N}/d)^{-\lambda},$$

and hence

$$d \leqslant \Delta^{1/(1+\lambda)} N^{\lambda/(1+\lambda)+o(1)}. \tag{3.39}$$

We enlarge the Bohr set implicit in (3.38) to

$$\mathcal{B}' = \Big\{ u \leqslant \frac{\widehat{CN}}{d} : \|u\gamma\| \leqslant \Big(\frac{\hat{N}}{d}\Big)^{-1/\lambda} + \frac{\Delta}{d} \Big\},$$

observing that $u := |x-y| \in \mathcal{B}'$.

We claim that

$$\#\mathcal{B}' \leqslant N^{o(1)} \Big(\Big(\frac{N}{d}\Big)^{1-\frac{1}{\lambda}} + \frac{\Delta N}{d^2} \Big). \tag{3.40}$$

This is clearly true if $d \gg \hat{N}$, so let us now assume that $d \leqslant c\hat{N}$ for some small constant $c > 0$. By (3.36), we have

$$\omega(\gamma) \leqslant \lambda/2,$$

so by Lemma 2.1 there exists $\ell \in \mathbb{N}$ such that

$$(\hat{N}/d)^{1/\lambda} \leqslant q_\ell' \leqslant \hat{N}/d.$$

Thus, we may apply Lemma 2.8 to the enlarged Bohr set $\mathcal{B}'$, reaping (3.40).

We begin by choosing $m \in \mathcal{B}_{\mathrm{loc}}(M; \boldsymbol{\delta})$, and by Lemma 3.5 there are at most $O(\delta M)$ choices. Next, we choose $d \mid m$ with $d > \max\{3\Delta/\|q_2'\gamma\|, N_1\}$, and by the standard divisor function bound there are $M^{o(1)}$ possible choices of d.

Given an element u from the Bohr set $\mathcal{B}'$ as well as a choice of y, the value of x is then determined in at most two ways. Furthermore, we claim that for each choice of x, y the number of possible choices of a, b is $O(d\mu(\mathcal{I}))$. To see this, let v be the integer closest to $(x-y)\gamma$, and note from (3.34) that a, b must satisfy $ya - xb = v$. All integer solutions to this linear diophantine equation have the form

$$a = a_0 + tx, \qquad b = b_0 + ty,$$

for a specific solution a_0, b_0 and a parameter $t \in \mathbb{Z}$. Therefore the number of such a, b constrained by $a + \gamma \in \hat{n}\mathcal{I}$ is

$$O\left(\frac{\hat{n}\mu(\mathcal{I})}{x} + 1\right) = O(d\mu(\mathcal{I})).$$

Hence, by (3.40), the number of choices for u, a, b is at most

$$N^{o(1)}\mu(\mathcal{I})(d^{1/\lambda}N^{1-1/\lambda} + \Delta N/d)$$

which, by (3.39) and the inequality $d > N_1$, is at most

$$N^{o(1)}\mu(\mathcal{I})(\Delta^{1/(\lambda+\lambda^2)}N^{1/(1+\lambda)+1-1/\lambda} + \Delta N/N_1). \tag{3.41}$$

The contribution to $D_{\mathrm{gcd}>}$ coming from the second term in (3.41) is at most

$$\delta M N^{o(1)}\mu(\mathcal{I})\Delta N/N_1 \leqslant \mu(\mathcal{I})\delta M N^{1+o(1)-20k\varepsilon_0}\Delta \leqslant \mu(\mathcal{I})\rho\delta M N\Delta.$$

Finally, the contribution to $D_{\mathrm{gcd}>}$ corresponding to the first term on the right hand side of (3.41) is at most

$$\delta M\mu(\mathcal{I})\Big(\frac{\Delta}{N}\Big)^{\frac{1}{\lambda+\lambda^2}}N^{1+o(1)}.$$

This quantity being at most $\mu(\mathcal{I})\rho\delta M N\Delta$ is equivalent to

$$\rho\Delta^{1-1/(\lambda+\lambda^2)}N^{1/(\lambda+\lambda^2)} \geqslant N^{o(1)}. \tag{3.42}$$

To confirm (3.42), we first recall from (3.24) that $\Delta > 1$. Next, the fact that $\boldsymbol{\rho} \in \mathcal{W}(N)$, together with the bound (3.37) on ε_0, give

$$\rho \geqslant N^{-5k\varepsilon_0} \geqslant N^{-1/(9\lambda^2)}.$$

These considerations bestow (3.42) and thence (3.32).

We now prove Theorems 1.8 and 1.23 in the remaining situation $\gamma \in \mathbb{Q} \cup \mathcal{L}$, recalling that it suffices to establish (3.32). The underpinning mechanisms are easier to grasp when γ is rational, so we begin with this case.

3.4.2. *Rational final shift.* Throughout this subsubsection we assume that γ is a rational number, given in lowest terms by $\gamma = c_0/d_0$. We continue to use the notation introduced in the previous subsubsection, and alert the reader to any deviation that we make. Let $\tilde{\varepsilon}$ be the minimum of its values when Lemmata 2.6 and 2.7 are applied with $\vartheta = 20k$, and fix a small positive real number $\varepsilon_0 < \min\{\tilde{\varepsilon}, (99k)^{-1}\}$.

In this case η is again of little importance, and so we once more choose $\eta = 1/2$ arbitrarily. Observe that if n is large then $c_t = c_0$ and $q_t' = d_0$, whereupon

$\mathcal{E}_n = \mathcal{E}_n^{\mathcal{I},\gamma}$ is the set of $\alpha \in [0,1]$ for which there is an integer a satisfying

$$|\hat{n}\alpha - \gamma - a| < \Psi(n), \quad \frac{a+\gamma}{\hat{n}} \in \mathcal{I}, \quad \text{and} \quad (d_0 a + c_0, \hat{n}) = 1. \tag{3.43}$$

The overlap inequality (3.29) is equivalent to

$$|c_0(\hat{n} - \hat{m}) - d_0(\hat{m}a - \hat{n}b)| \leqslant d_0\Delta.$$

Dividing the above inequality by d gives

$$|c_0(x-y) - d_0(ya - xb)| \leqslant \frac{\Delta d_0}{d}.$$

Assume for a contradiction that $c_0(x-y) = d_0(ya - xb)$. Then

$$x(c_0 + d_0 b) = y(c_0 + d_0 a).$$

As $n \neq m$, we have $\max\{x,y\} \geqslant 2$. Let us assume for simplicity that $x \geqslant 2$; a similar argument handles the case $y \geqslant 2$. Let p be a prime divisor of x. By Euclid's lemma p divides y or $c_0 + d_0 a$. The former is excluded by the coprimality of x and y, and the latter by the shift-reduction. This contradiction means that

$$1 \leqslant |c_0(x-y) - d_0(ya - xb)| \leqslant \frac{\Delta d_0}{d},$$

so $d \leqslant d_0\Delta$.

On the other hand, we have $d > \Delta$ by assumption. So let us fix a non-zero integer $h \in [-d_0, d_0]$ and count integer quadruples (n,m,a,b) with $n \neq m$ satisfying

$$c_0(x-y) - d_0(ya - xb) = h \tag{3.44}$$

as well as the constraints (3.27). There are $O(\rho N \delta M)$ many viable choices for the two integers $n \neq m$. Then by (3.44) there are $O(\mu(\mathcal{I})d) = O(\Delta\mu(\mathcal{I}))$ many choices for the numerators a, b. Summing this bound over the $O(1)$ many choices for h verifies (3.32).

It remains to consider the case in which γ is a Liouville number. Informally, this is a careful interpolation between the diophantine and rational cases.

3.4.3. *Liouville final shift.* Throughout the present subsubsection, we fix a Liouville number γ. In a nutshell, we need to find an way to balance between the regimes in which γ behaves like a rational number, and those in which γ behaves like a diophantine number. This is incarnated in the definition of the approximation sets, and shift-reduced fractions play a crucial role.

Fix a non-empty interval $\mathcal{I}$ in $[0,1]$, and a large positive constant C. Further, let

$$\eta = 9(k-1)\varepsilon_0. \tag{3.45}$$

Here, as before, the positive constant ε_0 is at most the lower value of $\tilde{\varepsilon}$ when Lemmata 2.6 and 2.7 are applied with $\vartheta = 20k$, and moreover $\varepsilon < (99k)^{-1}$. Recall that the approximation set $\mathcal{E}_n = \mathcal{E}_n^{\mathcal{I},\gamma}$ is the set of $\alpha \in [0,1]$ for which there exists $a \in \mathbb{Z}$ satisfying

$$|\hat{n}\alpha - \gamma - a| < \Psi(n) \tag{3.46}$$

and

$$\frac{a+\gamma}{\hat{n}} \in \mathcal{I}, \text{ and the pair } (a,\hat{n}) \text{ is } (\gamma,\eta)-\text{shift-reduced}. \tag{3.47}$$

As in the previous subsubsections, we write $d = \gcd(\hat{n},\hat{m})$, $dx = \hat{n}$, $dy = \hat{m}$.

Recall that q_t' is the greatest continued fraction denominator of γ not exceeding n^η, and that c_t is the corresponding numerator. We separate our argument according to the size of the subsequent denominator q_{t+1}'.

<u>Case 1: $q_{t+1}' \geqslant 10\widehat{CN}/d$</u>

This inequality and the continued fraction approximation (2.3) yield

$$\left|\gamma - \frac{c_t}{q_t'}\right| \leqslant \frac{d}{10q_t'\widehat{CN}}.$$

Moreover, the inequality (3.29) can be written in the form

$$|\hat{n}(b+\gamma) - \hat{m}(a+\gamma)| \leqslant \Delta,$$

or equivalently

$$|x(q_t'b + q_t'\gamma) - y(q_t'a + q_t'\gamma)| \leqslant \frac{q_t'\Delta}{d}.$$

Note that

$$x|q_t'\gamma - c_t| \leqslant x\frac{d}{10\widehat{CN}} \leqslant \frac{1}{10}, \qquad y|q_t'\gamma - c_t| \leqslant \frac{\widehat{CM}}{d}\,\frac{d}{10\widehat{CN}} \leqslant \frac{1}{10},$$

wherefore

$$|x(q_t'b + c_t) - y(q_t'a + c_t)| \leqslant \frac{q_t'\Delta}{d} + \frac{1}{5}.$$

Recall from (3.30) that the pairs $(a,\hat{n})$ and $(b,\hat{m})$ are (γ,η)-shift-reduced, from which we now deduce that $x(q_t'b+c_t)-y(q_t'a+c_t)$ is a non-zero integer. Indeed, suppose for a contradiction that $x(q_t'b+c_t) = y(q_t'a+c_t)$. As $n \neq m$, we have $\max\{x,y\} \geqslant 2$. Let us assume for simplicity that $x \geqslant 2$; a similar argument

handles the case $y \geqslant 2$. Let p be a prime divisor of x. Then p divides y or $q_t'a + c_t$. The former is excluded by the coprimality of x and y, and the latter by the shift-reduction. This contradiction means that

$$1 \leqslant |x(q_t'b + c_t) - y(q_t'a + c_t)| \leqslant \frac{q_t'\Delta}{d} + \frac{1}{5},$$

and consequently $d \leqslant 2q_t'\Delta$. The upshot is that we are in the rather special scenario that $\Delta \leqslant d \leqslant 2q_t'\Delta$.

Setting $h = x(q_t'b+c_t)-y(q_t'a+c_t)$, we see that there are $O(q_t')$ many realisable values for the integer h. Now we are in the position to derive an acceptable bound on the contribution from this case to $D_{\mathrm{gcd}>}$. There are $O(\delta M)$ many options for m and at most $N^{o(1)}$ many ways to choose $d > N_1$ dividing $\hat{m}$. It follows from Lemmata 2.7 and 2.15 that there are

$$O\Big(\rho\widehat{CN}\Big(\frac{1}{d} + \frac{1}{N_1}\Big)\Big) \leqslant \frac{\rho N^{1+o(1)}}{N_1}$$

many choices for $n \in \mathcal{B}_{\mathrm{loc}}(N; \boldsymbol{\rho})$ divisible by d. For each choice of $h \ll q_t'$, the parameters a, b are determined up to $O(\mu(\mathcal{I})d) = O(\mu(\mathcal{I})q_t'\Delta)$ many possibilities. By summing over all choices of h, we obtain

$$D_{\mathrm{gcd}>} = O\left(\delta M \frac{\rho N^{1+o(1)}}{N_1}(q_t')^2\mu(\mathcal{I})\Delta\right).$$

Since $(q_t')^2 N^{o(1)} \leqslant N^{2\eta+o(1)} < N^{20(k-1)\varepsilon_0} \leqslant N_1$, this bound is acceptable.

Case 2: $q_{t+1}' < 10\widehat{CN}/d$

By definition, we have $q_{t+1}' > N^{\eta}$, and therefore

$$d < 10\widehat{CN}N^{-\eta}. \tag{3.48}$$

Now, akin to the proof from Subsubsection 3.4.1, we work with an enlarged Bohr set, namely

$$\mathcal{B} := \{u \leqslant 10\widehat{CN}/d : \|u\gamma\| \leqslant L\},$$

where $L = L(d) = \max(\Delta/d, N^{-\eta})$.

Note that $q_{t+1}' \in [1/(2L), 10\widehat{CN}/d]$, where the lower bound comes from maximality in the definition of shift-reduction. The existence of a continued fraction denominator in this range is a key technical ingredient. As M is large and $d > 3\Delta/\|q_2'\gamma\|$, we have $L < \|q_2'\gamma\|/2$. Therefore Lemma 2.8 is applicable, and hence

$$\#\mathcal{B} \ll \frac{N^{1+o(1)}}{d}L(d).$$

First choose $d > \max(3\Delta/\|q_2'\gamma\|, N_1)$. There are $O(\hat{M}/d)$ many conceivable options for m divisible by d. Then $u = |x-y|$ lies in $\mathcal{B}$, and therefore admits at most $\#\mathcal{B}$ possibilities, and then x is determined up at most two choices. The number of viable choices of a, b is then $O(d\mu(\mathcal{I}))$. So, for a fixed choice of d, the number of valid possibilities for (m, n, a, b) is at most

$$\frac{M}{d} N^{o(1)} \frac{N}{d} L(d) d\mu(\mathcal{I}) = N^{o(1)} \mu(\mathcal{I}) MN \frac{L(d)}{d}.$$

Upon summing over d, we infer that $D_{\gcd>}$ is at most

$$N^{o(1)} \mu(\mathcal{I}) MN \sum_{\max(\Delta, N_1) < d \leqslant \widehat{CN}} \frac{L(d)}{d}.$$

To conclude, it suffices to show that for any d in the range of interest we have

$$N^{o(1)} L(d) \leqslant \Delta \delta \rho. \tag{3.49}$$

If $L(d) = N^{-\eta}$ then

$$\rho\delta \geqslant N^{-4.2(k-1)\varepsilon_0} N^{-4.2(k-1)\varepsilon_0} = N^{-8.4(k-1)\varepsilon_0} > N^{o(1)-\eta} = N^{o(1)} L(d).$$

Since $\Delta > 1$, from (3.24), we conclude that $N^{o(1)} L(d) \leqslant \delta\rho\Delta$. If on the other hand $L(d) = \Delta/d$, then

$$d > N_1 \geqslant N^{20k\varepsilon_0}$$

and therefore

$$L(d) N^{o(1)}/\Delta = d^{-1} N^{o(1)} \leqslant N^{o(1)-20k\varepsilon_0} \leqslant \delta\rho.$$

We have (3.49) in both cases, completing the proofs of Theorems 1.8 and 1.23.

3.5. **A convergence statement.** In this subsection, we prove the convergence side of Corollary 1.11. Assume that

$$\sum_{n=1}^{\infty} \psi(n) (\log n)^{k-1} < \infty. \tag{3.50}$$

Replacing $\psi(n)$ by $\max(\psi(n), n^{-2})$, we may suppose that

$$\psi(n) \geqslant n^{-2} \qquad (n \in \mathbb{N}).$$

We wish to show that $\mu_k(\mathcal{W}^\times) = 0$. By 1-periodicity, we may assume that

$$-1 < \gamma_1, \ldots, \gamma_k \leqslant 0.$$

We abbreviate $\tilde{\boldsymbol{\alpha}} = (\alpha_1, \ldots, \alpha_k)$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{k-1})$ for the remainder of this section. Observe that $\mathcal{W}^\times = \limsup_{n\to\infty} \mathcal{A}_n$, where

$$\mathcal{A}_n = \{\tilde{\boldsymbol{\alpha}} \in [0,1]^k : \|n\alpha_1 - \gamma_1\| \cdots \|n\alpha_k - \gamma_k\| < \psi(n)\} \qquad (n \in \mathbb{N}).$$

For $\mathbf{a} = (a_1, \dots, a_k) \in \mathbb{Z}^k$, let

$$\mathcal{A}_{n,\mathbf{a}} = \left\{ \tilde{\boldsymbol{\alpha}} \in [0,1]^k : |n\alpha_k - \gamma_k - a_k| < \min(1, Q(\boldsymbol{\alpha})) \right\},$$

where

$$Q(\boldsymbol{\alpha}) = Q(\boldsymbol{\alpha}; a_1, \dots, a_{k-1}) = \frac{\psi(n)}{\prod_{i \leqslant k-1} |n\alpha_i - \gamma_i - a_i|}.$$

Lemma 3.12. *For $n \in \mathbb{N}$, we have*

$$\mathcal{A}_n = \bigcup_{0 \leqslant a_1, \dots, a_k \leqslant n+1} \mathcal{A}_{n,\mathbf{a}}.$$

Proof. First suppose $\tilde{\boldsymbol{\alpha}} \in \mathcal{A}_{n,\mathbf{a}}$, for some $\mathbf{a} \in \{0, 1, \dots, n+1\}^k$. Then

$$\prod_{i \leqslant k} \|n\alpha_i - \gamma_i\| \leqslant \prod_{i \leqslant k} |n\alpha_i - \gamma_i - a_i| < \psi(n),$$

and so $\tilde{\boldsymbol{\alpha}} \in \mathcal{A}_n$. Therefore $\bigcup_{0 \leqslant a_1, \dots, a_k \leqslant n+1} \mathcal{A}_{n,\mathbf{a}} \subseteq \mathcal{A}_n$.

Next, suppose $\tilde{\boldsymbol{\alpha}} \in \mathcal{A}_n$, and for $i = 1, 2, \dots, k$ let a_i be an integer for which $\|n\alpha_i - \gamma_i\| = |n\alpha_i - \gamma_i - a_i|$. Now $|n\alpha_k - \gamma_k - a_k| < 1$ and $|n\alpha_k - \gamma_k - a_k| < Q(\boldsymbol{\alpha})$. Moreover, the triangle inequality yields

$$-1/2 \leqslant -\gamma_i - 1/2 \leqslant a_i \leqslant n - \gamma_i + 1/2 \leqslant n + 3/2 \qquad (1 \leqslant i \leqslant k),$$

and therefore $\mathcal{A}_n \subseteq \bigcup_{0 \leqslant a_1, \dots, a_k \leqslant n+1} \mathcal{A}_{n,\mathbf{a}}$. □

By the union bound, if $n \in \mathbb{N}$ then

$$\mu_k(\mathcal{A}_n) \leqslant \sum_{a_1, \dots, a_k = 0}^{n+1} \mu_k(\mathcal{A}_{n,\mathbf{a}}).$$

Further, if $n \in \mathbb{N}$ and $0 \leqslant a_1, \dots, a_k \leqslant n+1$ then

$$\mu_k(\mathcal{A}_{n,\mathbf{a}}) \ll n^{-1} \int_{[0,1]^{k-1}} \min(1, Q(\boldsymbol{\alpha}))\, \mathrm{d}\boldsymbol{\alpha} \leqslant n^{-1}(I_1 + I_2),$$

where

$$I_1 = \mu_{k-1}(\mathcal{R}), \qquad \mathcal{R} = \{\boldsymbol{\alpha} \in [0,1]^{k-1} : \min_{1 \leqslant i \leqslant k-1} |n\alpha_i - \gamma_i - a_i| \leqslant n^{-k}\},$$

and

$$I_2 = \int_{[0,1]^{k-1} \setminus \mathcal{R}} Q(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha}.$$

Here μ_{k-1} denotes $(k-1)$-dimensional Lebesgue measure.

We have $I_1 \ll n^{-1-k}$ by inspection. To estimate I_2, we cover $[0,1]^{k-1} \setminus \mathcal{R}$ by $O((\log n)^{k-1})$ dyadically-restricted regions

$$\{\boldsymbol{\alpha} \in [0,1]^{k-1} : \delta_i < |n\alpha_i - \gamma_i - a_i| \leqslant 2\delta_i \ (1 \leqslant i \leqslant k-1)\}.$$

The integral of Q over such a region is $O(\psi(n)/n^{k-1})$, and so

$$I_2 \ll \frac{\psi(n)(\log n)^{k-1}}{n^{k-1}}.$$

Recalling that $\psi(n) \geqslant n^{-2}$, we therefore have

$$\mu_k(\mathcal{A}_{n,\mathbf{a}}) \ll \frac{I_1 + I_2}{n} \ll \frac{\psi(n)(\log n)^{k-1}}{n^k} \qquad (0 \leqslant a_1, \dots, a_k \leqslant n+1),$$

and finally

$$\mu_k(\mathcal{A}_n) \ll \psi(n)(\log n)^{k-1}.$$

In view of (3.50), we now have

$$\sum_{n=1}^{\infty} \mu_k(\mathcal{A}_n) < \infty,$$

and the first Borel–Cantelli lemma completes the proof that $\mu_k(\mathcal{W}^\times) = 0$.

4. Liouville fibres

In this section, we establish Theorem 1.13.

4.1. **A special case.** For expository purposes, we begin with the 'partially homogeneous' case $\gamma_2 = 0$. Here it is simplest to invoke the resolution of the Duffin–Schaeffer conjecture by Koukoulopoulos and Maynard, though this is by no means an essential ingredient. By Theorem 1.18, it suffices to show that

$$\sum_{n=1}^{\infty} \frac{\varphi(n)}{n} \cdot \frac{1}{n(\log n)^2 \|n\alpha - \gamma\|} = \infty,$$

where $\alpha = \alpha_1$ and $\gamma = \gamma_1$.

By Lemma 2.1, there are infinitely many positive integers such that $q_t > q_{t-1}^9$, where the q_ℓ are continued fraction denominators of α. Let $\mathcal{T}$ be a sparse, infinite set of integers $t \geqslant 9$ for which $q_t > q_{t-1}^9$.

Given $t \in \mathcal{T}$, we begin by using Theorem 2.2, with

$$m = q_t = a_t q_{t-1} + q_{t-2} + 0,$$

to find a small positive integer b_t such that $\|b_t\alpha - \gamma\|$ is small. Recall that this concerns the gaps $d_{i+1} - d_i$, where

$$\{d_1, \dots, d_m\} = \{i\alpha - \lfloor i\alpha \rfloor : 0 \leqslant i \leqslant m\}, \qquad 0 = d_0 < \cdots < d_{m+1} = 1.$$

By (2.3), Theorem 2.2 tells us that

$$\max\{d_{i+1} - d_i : 0 \leqslant i \leqslant m\} \leqslant |D_t| + |D_{t-1}| \leqslant 2/q_t,$$

where we have employed the standard notation (2.2). For some $i \in \{0, 1, \dots, m\}$, the fractional part of γ lies in $[d_i, d_{i+1}]$, and for some $b_t \in \{1, 2, \dots, m\}$ the fractional part of $b_t\alpha$ is either d_i or d_{i+1}. The upshot is that we have a positive integer $b_t \leqslant q_t$ such that $\|b_t\alpha - \gamma\| \leqslant 2/q_t$.

For $t \in \mathcal{T}$, denote by $\mathcal{D}_t$ the set of integers of the form

$$n = b_t + q_{t-1}x + q_t y,$$

where $x, y \in \mathbb{Z}$ satisfy

$$q_t^{1/4} \leqslant y \leqslant q_t^{1/3}, \qquad q_t^{2/3} \leqslant x \leqslant q_t^{3/4}.$$

If n is as above then

$$n \asymp q_t y, \qquad \log n \asymp \log q_t, \qquad \|n\alpha - \gamma\| \asymp \frac{x}{q_t}.$$

Indeed, for the final estimate, observe using (2.3) that

$$\|b_t\alpha - \gamma\| \leqslant \frac{2}{q_t}, \qquad \|xq_{t-1}\alpha\| \asymp \frac{x}{q_t}, \qquad \|yq_t\alpha\| \ll \frac{y}{q_t},$$

and apply the triangle inequality. The sets $\mathcal{D}_t$ ($t \in \mathcal{T}$) are disjoint, since $\mathcal{T}$ is sparse. Moreover, for $n \in \mathcal{T}$ the representation above is unique, since $q_{t-1}q_t^{3/4} < q_t$. Fix $t \in \mathcal{T}$, and let

$$S_t = \sum_{n \in \mathcal{D}_t} \frac{\varphi(n)}{n} v_n,$$

where $v_n^{-1} = n(\log n)^2\|n\alpha - \gamma\|$. As $\mathcal{T}$ is infinite, it suffices to prove that $S_t \gg 1$. Note that in this section our implicit constants do not depend on t.

As a first step, we compute that

$$\sum_{n \in \mathcal{D}_t} v_n \gg (\log q_t)^{-2} \sum_{\substack{q_t^{1/4} \leqslant y \leqslant q_t^{1/3} \\ q_t^{2/3} \leqslant x \leqslant q_t^{3/4}}} (xy)^{-1} \gg 1.$$

Let $C_t = \left(\sum_{n\in\mathcal{D}_t} v_n\right)^{-1}$, and for $n \in \mathcal{D}_t$ let $w_n = C_t v_n \ll v_n$, so that we have $\sum_{n\in\mathcal{D}_t} w_n = 1$. The weighted AM–GM inequality gives

$$S_t \gg \sum_{n\in\mathcal{D}_t} w_n \varphi(n)/n \geqslant \prod_p (1-1/p)^{\tau_p},$$

where

$$\tau_p = \sum_{\substack{n\in\mathcal{D}_t \\ n\equiv 0 \bmod p}} w_n.$$

Now

$$-\ln S_t \leqslant O(1) - \sum_p \tau_p \ln(1-1/p) \ll 1 + \sum_p \tau_p/p,$$

so it remains to show that

$$\sum_p \tau_p/p < \infty. \tag{4.1}$$

Let p be prime, and let X, Y be parameters in the ranges

$$q_t^{1/4} \leqslant Y \leqslant q_t^{1/3}/2, \qquad q_t^{2/3} \leqslant X \leqslant q_t^{3/4}/2.$$

Denote by $N_p(X,Y)$ the number of integer solutions $(x,y) \in [X,2X]\times[Y,2Y]$ to

$$b + q_{t-1}x + q_t y \equiv 0 \bmod p.$$

Let us assume that this congruence has a solution in $[X,2X]\times[Y,2Y]$, forcing $p \leqslant q_t^2$. Then, by Lemma 2.15, we have

$$N_p(X,Y) \ll XY/p + X + Y \ll XYp^{-1/8}. \tag{4.2}$$

If

$$X \leqslant x \leqslant 2X, \qquad Y \leqslant y \leqslant 2Y, \qquad n = b + q_{t-1}x + q_t y$$

then $w_n^{-1} \asymp XY(\log q_t)^2$. Using (4.2), and summing over X, Y that are powers of 2 or the endpoints of their allowed ranges, gives

$$\tau_p \ll p^{-1/8},$$

and in particular (4.1).

4.2. **Diophantine second shift.** In this subsection, we prove Theorem 1.13 in the case that γ_2 is diophantine. Let $\lambda \geqslant 2$ satisfy

$$\|n\gamma_2\| \gg n^{-\lambda/2} \qquad (n \in \mathbb{N}). \tag{4.3}$$

Then $\lambda \geqslant 2\omega(\gamma_2)$, where

$$\omega(\gamma_2) = \sup\{w > 0 : \exists^\infty q \in \mathbb{N}\ \|q\gamma_2\| < q^{-w}\}.$$

Fix a constant $c_1 > 0$, and for $n \in \mathbb{N}$ let

$$\Psi(n) = \frac{c_1}{n(\log n)^2 \|n\alpha_1 - \gamma_1\|} \in (0, +\infty],$$

recalling our notational conventions from Section 1.4. By 1-periodicity of $\|\cdot\|$, our task is to show that for almost all $\alpha_2 \in [0,1]$ the inequality

$$\|n\alpha_2 - \gamma_2\| < \Psi(n)$$

has infinitely many solutions $n \in \mathbb{N}$. Let $\mathcal{I}$ be a non-empty subinterval of $[0,1]$. Let θ_3, θ_4 be constants satisfying

$$\theta_3 = 1 - \lambda^{-3} < \theta_4 < 1.$$

Let T_0 be large in terms of $\mathcal{I}$, and let $\mathcal{T}$ be a sparse, infinite set of integers $t \geqslant T_0$ for which $q_t^{1-\theta_4} > q_{t-1}$.

For $t \in \mathcal{T}$, let $\mathcal{D}_t$ be the set of integers of the form

$$n = b_t + q_{t-1}x_1 + q_t y_1 \tag{4.4}$$

with

$$q_t^{1/4} \leqslant y_1 \leqslant q_t^{1/3}, \qquad q_t^{\theta_3} \leqslant x_1 \leqslant q_t^{\theta_4},$$

where b_t, q_{t-1}, and q_t are as in the previous subsection. Let $t \in \mathcal{T}$, and let $n \in \mathcal{D}_t$. The representation above is unique, since $q_t > q_t^{\theta_4} q_{t-1}$. Moreover, we have

$$\|n\alpha_1 - \gamma_1\| \geqslant \|x_1 q_{t-1}\alpha_1\| - \frac{2}{q_t} - y_1\|q_t\alpha_1\| \geqslant \frac{x_1}{2q_t} - \frac{2}{q_t} - \frac{y_1}{q_{t+1}} \geqslant \frac{x_1}{3q_t},$$

so

$$n \asymp q_t y_1, \qquad \log n \asymp \log q_t, \qquad \|n\alpha_1 - \gamma_1\| \asymp x_1/q_t,$$

and in particular

$$\Psi(n) \asymp \frac{q_t}{x_1}(q_t y_1)^{-1}(\log q_t)^{-2} \asymp \frac{1}{x_1 y_1 (\log q_t)^2}.$$

For $n \in \mathbb{N}$ and $a \in \mathbb{Z}$, define

$$\mathcal{A}_{n,a} = \{\beta \in \mathcal{I} : |n\beta - \gamma_2 - a| < \Psi(n)\}, \qquad \mathcal{A}_n = \bigcup_{a \in \mathbb{Z}} \mathcal{A}_{n,a},$$

and observe that if $n\mu(\mathcal{I}) \geqslant 1$ then

$$\mu(\mathcal{A}_n) \asymp \mu(\mathcal{I})\Psi(n).$$

Let $F(1), F(2), \ldots$ be a sequence of powers of 4, non-increasing, satisfying $F(t) \leqslant \log\log t$ for all t, and such that

$$\sum_{t \in \mathcal{T}} F(t)^{-1} = \infty.$$

For $t \in \mathcal{T}$, let

$$\mathcal{G}_t = \{n \in \mathcal{D}_t : n \equiv 0 \bmod F(t)\}.$$

We will show that if $t, s \in \mathcal{T}$ then

$$\sum_{n \in \mathcal{G}_t} \mu(\mathcal{A}_n) \gg \frac{\mu(\mathcal{I})}{F(t)} \tag{4.5}$$

and

$$\sum_{n \in \mathcal{G}_t} \sum_{\substack{m \in \mathcal{G}_s \\ m < n}} \mu(\mathcal{A}_n \cap \mathcal{A}_m) \ll \frac{\mu(\mathcal{I})}{F(t)F(s)}. \tag{4.6}$$

Then, with

$$\mathcal{X} := \bigcup_{t \in \mathcal{T}} \mathcal{G}_t, \qquad \mathcal{X}_N := \bigcup_{\substack{t \in \mathcal{T} \\ t \leqslant N}} \mathcal{G}_t,$$

we would have

$$\sum_{n \in \mathcal{X}} \mu(\mathcal{A}_n) = \infty$$

and

$$\begin{aligned} \mu(\mathcal{I}) \sum_{n,m \in \mathcal{X}_N} \mu(\mathcal{A}_n \cap \mathcal{A}_m) &\ll \left(\sum_{n \in \mathcal{X}_N} \mu(\mathcal{A}_n) \right)^2 + \sum_{n \in \mathcal{X}_N} \mu(\mathcal{A}_n) \\ &\ll \left(\sum_{n \in \mathcal{X}_N} \mu(\mathcal{A}_n) \right)^2. \end{aligned}$$

At that stage Lemmata 2.9 and 2.10 would give

$$\mu(\limsup_{n \in \mathcal{X}} \{\beta \in [0,1] : \|n\beta - \gamma_2\| < \Psi(n)\}) = 1,$$

which would finish the proof. The upshot is that it remains to establish (4.5) and (4.6).

Let $t \in \mathcal{T}$, and let X and Y be parameters in the ranges

$$q_t^{\theta_3} \leqslant X \leqslant q_t^{\theta_4}/2, \qquad q_t^{1/4} \leqslant Y \leqslant q_t^{1/3}/2. \tag{4.7}$$

Then

$$\begin{aligned} \mu(\mathcal{I})^{-1} \sum_{\substack{n = b_t + q_{t-1}x_1 + q_t y_1 \\ n \equiv 0 \bmod F(t) \\ X < x_1 \leqslant 2X \\ Y < y_1 \leqslant 2Y}} \mu(\mathcal{A}_n) &\gg (\log q_t)^{-2} (XY)^{-1} \sum_{\substack{X < x_1 \leqslant 2X \\ Y < y_1 \leqslant 2Y \\ b_t + q_{t-1}x_1 + q_t y_1 \equiv 0 \bmod F(t)}} 1 \\ &\gg (\log q_t)^{-2} F(t)^{-1}, \end{aligned}$$

by Lemma 2.15. Summing over X and Y that are powers of 2 in the above ranges, we obtain (4.5).

It remains to prove (4.6). Writing

$$n = b_t + q_{t-1}x_1 + q_t y_1, \qquad m = b_s + q_{s-1}x_2 + q_s y_2, \tag{4.8}$$

where

$$q_t^{1/4} \leqslant y_1 \leqslant q_t^{1/3}, \quad q_t^{\theta_3} \leqslant x_1 \leqslant q_t^{\theta_4}, \quad q_s^{1/4} \leqslant y_2 \leqslant q_s^{1/3}, \quad q_s^{\theta_3} \leqslant x_2 \leqslant q_s^{\theta_4}, \tag{4.9}$$

we have

$$\begin{aligned}\mu(\mathcal{A}_{n,a} \cap \mathcal{A}_{m,b}) &\ll \min\left\{\frac{\Psi(n)}{n}, \frac{\Psi(m)}{m}\right\} \\ &\asymp \min\left\{\frac{q_t}{x_1(q_t y_1 \log q_t)^2}, \frac{q_s}{x_2(q_s y_2 \log q_s)^2}\right\}\end{aligned}$$

for $a, b \in \mathbb{Z}$.

Let X_1, Y_1, X_2, Y_2 be parameters in the ranges

$$\begin{aligned} q_t^{1/4} &\leqslant Y_1 \leqslant q_t^{1/3}/2, & q_t^{\theta_3} &\leqslant X_1 \leqslant q_t^{\theta_4}/2, \\ q_s^{1/4} &\leqslant Y_2 \leqslant q_s^{1/3}/2, & q_s^{\theta_3} &\leqslant X_2 \leqslant q_s^{\theta_4}/2. \end{aligned}$$

Suppose $\mathcal{A}_{n,a} \cap \mathcal{A}_{m,b}$ is non-empty, and that we have (4.8) and

$$X_i \leqslant x_i \leqslant 2X_i, \quad Y_i \leqslant y_i \leqslant 2Y_i \qquad (i = 1, 2). \tag{4.10}$$

Then

$$\operatorname{dist}(a + \gamma_2, n\mathcal{I}) < 1/2, \qquad \operatorname{dist}(b + \gamma_2, m\mathcal{I}) < 1/2, \tag{4.11}$$

and there exists $\beta \in \mathcal{A}_{n,a} \cap \mathcal{A}_{m,b}$, so that

$$|n\beta - \gamma_2 - a| < \Psi(n), \qquad |m\beta - \gamma_2 - b| < \Psi(m).$$

The triangle inequality gives

$$|(n-m)\gamma_2 - (ma - nb)| < m\Psi(n) + n\Psi(m) < \Delta, \tag{4.12}$$

where

$$\Delta \asymp \frac{q_s Y_2}{X_1 Y_1 (\log q_t)^2} + \frac{q_t Y_1}{X_2 Y_2 (\log q_s)^2}.$$

Note that

$$\Delta^2 \gg \frac{q_t q_s}{X_1 X_2 (\log q_t)^2 (\log q_s)^2} \geqslant q_t^{1-\theta_4-o(1)} q_s^{1-\theta_4-o(1)}. \tag{4.13}$$

Given $t, s \in \mathcal{T}$, denote by $N(t,s)$ the number of solutions

$$(n, m, a, b) \in \mathcal{G}_t \times \mathcal{G}_s \times \mathbb{Z}^2$$

to the diophantine system given by (4.11), (4.12) and $n > m$ for which we have (4.8) for some x_1, x_2, y_1, y_2 in the ranges (4.10). Our goal is to show that

$$N(t,s) \ll X_1 X_2 Y_1 Y_2 \Delta \mu(\mathcal{I}) F(t)^{-1} F(s)^{-1}. \tag{4.14}$$

Note that the number of solutions with $n = m$ is at most X_1Y_1, which is negligible. Assuming for the time being that we can achieve (4.14), the contribution to $\sum_{n\in\mathcal{D}_t}\sum_{m\in\mathcal{D}_s}\mu(\mathcal{A}_n\cap\mathcal{A}_m)$ from x_1, x_2, y_1, y_2 in these dyadic ranges is

$$O\left(\frac{X_1X_2Y_1Y_2\Delta\mu(\mathcal{I})}{F(t)F(s)}\min\left\{\frac{q_t}{X_1(q_tY_1\log q_t)^2}, \frac{q_s}{X_2(q_sY_2\log q_s)^2}\right\}\right),$$

which is

$$O\Big(\frac{\mu(\mathcal{I})}{F(t)F(s)(\log q_t)^2(\log q_s)^2}\Big).$$

Summing over X_1, X_2, Y_1, Y_2 that are powers of 2 or endpoints of the prescribed ranges, we would thereby deduce (4.6). The upshot is that it remains to prove (4.14).

We partition our solutions according to the value of $d = \gcd(m, n)$, and write $n = dx$, $m = dy$, so that $\gcd(x, y) = 1$ and $x > y$. Then (4.12) becomes

$$|(x - y)\gamma_2 - (ya - xb)| < \Delta/d. \tag{4.15}$$

Note that $d \geqslant \min\{F(t), F(s)\} > \mu(\mathcal{I})^{-1}$. Let $(q'_\ell)_\ell$ be the sequence of continued fraction denominators of γ_2.

<u>Case 1: $\mu(\mathcal{I})^{-1} < d < 3\Delta/\|q'_2\gamma_2\|$</u>

Choose m, n with $\gcd(m, n) < 3\Delta/\|q'_2\gamma_2\|$ in $O(F(t)^{-1}F(s)^{-1}X_1X_2Y_1Y_2)$ ways, via Lemma 2.15. Then there are $O(\Delta/d)$ possibilities for $h = ya - xb$, by (4.15). Given h, the value of a is determined modulo x, and lies in an interval of length $O(dx\mu(\mathcal{I}))$, so there are $O(d\mu(\mathcal{I}) + 1)$ possibilities for a, whereupon b is determined. As $d > \mu(\mathcal{I})^{-1}$, the contribution to $N(t, s)$ from this case is $O(X_1X_2Y_1Y_2\Delta\mu(\mathcal{I})F(t)^{-1}F(s)^{-1})$.

<u>Case 2: $3\Delta/\|q'_2\gamma_2\| \leqslant d < Y_1$</u>

Choose m in $O(X_2Y_2)$ ways, and then choose $d \mid m$ with $d \in [3\Delta/\|q'_2\gamma_2\|, Y_1)$ in $Y_2^{o(1)}$ ways. Then, by Lemma 2.15, choose $n \in \mathcal{D}_t$ such that

$$n \equiv 0 \bmod d$$

in $O(X_1Y_1/d)$ ways. Finally there are $d\mu(\mathcal{I})$ possibilities for a and b so, using (4.13), the contribution to $N(t, s)$ from this case is bounded above by

$$X_2Y_2^{1+o(1)}X_1Y_1\mu(\mathcal{I}) \ll X_1Y_1X_2Y_2\Delta\mu(\mathcal{I})F(t)^{-1}F(s)^{-1}.$$

<u>Case 3: $d \geqslant \max\{Y_1, 3\Delta/\|q'_2\gamma_2\|\}$</u>

Choose m in $O(X_2Y_2)$ ways and $d \geqslant \max\{Y_1, 3\Delta/\|q_2'\gamma_2\|\}$ dividing m in $q_t^{o(1)}$ ways. Put

$$N = q_tY_1.$$

Set $x - y = u \in \mathbb{N}$ and $ya - xb = v \in \mathbb{Z}$, so that

$$u \leqslant 3N/d, \qquad |u\gamma_2 - v| < \Delta/d. \tag{4.16}$$

From our choice of λ, we have

$$\Delta/d \gg (N/d)^{-\lambda},$$

so $d \ll (\Delta N^\lambda)^{1/(1+\lambda)}$. We relax the inequalities above, giving

$$u \leqslant CN/d, \qquad \|u\gamma_2\| \leqslant (N/d)^{-1/\lambda} + \Delta/d, \tag{4.17}$$

where C is a large, positive constant.

We claim that (4.17) has $O((N/d)((N/d)^{-1/\lambda} + \Delta/d))$ solutions $u \in \mathbb{N}$. This is clearly true if $d \gg N$, so let us now assume that $d \leqslant cN$ for some small constant $c > 0$. By (4.3), we have $\omega(\gamma_2) \leqslant \lambda/2$, so by Lemma 2.1 there exists $\ell \in \mathbb{N}$ such that

$$(N/d)^{1/\lambda} \leqslant q_\ell' \leqslant CN/d.$$

Thus, we may apply Lemma 2.8 to the Bohr set defined by (4.17), establishing the claim.

Consequently, there are $O((N/d)((N/d)^{-1/\lambda} + \Delta/d))$ pairs $(u, v) \in \mathbb{N} \times \mathbb{Z}$ satisfying (4.16). Hence, given m and d as above, the number of possibilities for u, v, a, b is at most a constant times

$$\begin{aligned}(N/d)((N/d)^{-1/\lambda} + \Delta/d)d &= N^{1-(1/\lambda)}d^{1/\lambda} + N\Delta/d \\ &\ll \Delta^{1/(\lambda+\lambda^2)}N^{1-(1/\lambda)+1/(1+\lambda)} + N\Delta/Y_1.\end{aligned}$$

Recalling that

$$N = q_tY_1, \qquad X_1 \geqslant q_t^{\theta_3} = q_t^{1-\lambda^{-3}}, \qquad Y_1 \geqslant q_t^{1/4},$$

we find that the contribution to $N(t, s)$ from this case is $O(N_1(t, s) + N_2(t, s))$, where

$$\begin{aligned}N_1(t, s) &\ll X_2Y_2q_t^{o(1)}(\Delta/N)^{1/(\lambda+\lambda^2)}N \ll X_2Y_2\Delta q_t^{1-1/(\lambda+\lambda^2)+o(1)}Y_1 \\ &\ll X_1Y_1X_2Y_2\Delta\mu(\mathcal{I})F(t)^{-1}F(s)^{-1}\end{aligned}$$

and

$$\begin{aligned}N_2(t, s) &\ll X_2Y_2q_t^{o(1)}N\Delta/Y_1 = X_2Y_2q_t^{1+o(1)}\Delta \\ &\ll X_1Y_1X_2Y_2\Delta\mu(\mathcal{I})F(t)^{-1}F(s)^{-1}.\end{aligned}$$

We have considered all possible size ranges for d, confirming (4.14). We conclude that Theorem 1.13 holds in the case that γ_2 is diophantine.

4.3. **Liouville second shift.** Next, we prove Theorem 1.13 in the case that $\gamma_2 \in \mathcal{L}$. This is a more sophisticated variant of the proof given in the previous subsection that works whenever $\gamma_2 \in \mathbb{R} \setminus \mathbb{Q}$. As we discuss in the next subsection, a simpler version of it works when $\gamma_2 \in \mathbb{Q}$.

Fix a constant $c_1 > 0$, and for $n \in \mathbb{N}$ let

$$\Psi(n) = \frac{c_1}{n(\log n)^2 \|n\alpha_1 - \gamma_1\|} \in (0, +\infty].$$

By 1-periodicity of $\|\cdot\|$, it suffices to show that for almost all $\alpha_2 \in [0,1]$ the inequality

$$\|n\alpha_2 - \gamma_2\| < \Psi(n)$$

has infinitely many solutions $n \in \mathbb{N}$. Indeed, the latter would imply that

$$\liminf_{n\to\infty} n(\log n)^2 \|n\alpha_1 - \gamma_1\| \cdot \|n\alpha_2 - \gamma_2\| \leqslant c_1,$$

and $c_1 > 0$ is arbitrary.

Let $\mathcal{I}$ be a non-empty subinterval of $[0,1]$, let T_0 be large in terms of $\mathcal{I}$, and let $\mathcal{T}$ be a sparse, infinite set of integers $t \geqslant T_0$ for which $q_t > q_{t-1}^9$, where again $q_1, q_2, \ldots$ are the denominators of the continued fraction convergents to the Liouville number α_1.

Fix $\lambda \geqslant 2$. For $t \in \mathcal{T}$, we define $\mathcal{D}_t$ as in the previous subsection. The sets $\mathcal{D}_t$ ($t \in \mathcal{T}$) are disjoint, because $\mathcal{T}$ is sparse. For $t \in \mathcal{T}$ and $n \in \mathcal{D}_t$, the representation (4.4) is unique, and moreover

$$n \asymp q_t y_1, \qquad \log n \asymp \log q_t, \qquad \|n\alpha_1 - \gamma_1\| \asymp \frac{x_1}{q_t}, \qquad \Psi(n) \asymp \frac{1}{x_1 y_1 (\log q_t)^2}.$$

We also define $\mathcal{G}_t$ via $\mathcal{D}_t$ as in the previous subsection, for $t \in \mathcal{T}$, using the arithmetic function F.

For $t \in \mathcal{T}$, let $c_{t'}/q'_{t'}$ be the continued fraction convergent to γ_2 for which $q'_{t'} < q_t$ is maximal. For $n \in \mathcal{D}_t$ and $a \in \mathbb{Z}$, we again define

$$\mathcal{A}_{n,a} = \{\beta \in \mathcal{I} : |n\beta - \gamma_2 - a| < \Psi(n)\},$$

but now we let $\mathcal{A}_n$ be the union of $\mathcal{A}_{n,a}$ over integers a for which

$$(q'_{t'} a + c_{t'}, n) = 1.$$

As in the previous subsection, it remains to establish (4.5) and (4.6).

Lemma 4.1. *Let $n \in \mathcal{D}_t$. Then*

$$\frac{\varphi(n)}{n}\mu(\mathcal{I})\Psi(n) \ll \mu(\mathcal{A}_n) \ll \mu(\mathcal{I})\Psi(n).$$

Proof. For the upper bound, observe that there are $O(n\mu(\mathcal{I}))$ integers a for which $\mathcal{A}_{n,a}$ is non-empty, and for each of these $\mu(\mathcal{A}_{n,a}) \ll \Psi(n)/n$. For the lower bound, it suffices to show that

$$\#\{a \in n\mathcal{I} : (q'_{t'}a + c_{t'}, n) = 1\} \gg \frac{\varphi(n)}{n}\mu(\mathcal{I}).$$

This follows routinely from the fundamental lemma of sieve theory, in the same way as (3.8). □

Let $t \in \mathcal{T}$, and let X and Y be parameters in the ranges (4.7). Then

$$\mu(\mathcal{I})^{-1} \sum_{\substack{n=b_t+q_{t-1}x_1+q_ty_1 \\ n\equiv 0 \bmod F(t) \\ X<x_1\leqslant 2X \\ Y<y_1\leqslant 2X}} \mu(\mathcal{A}_n)$$

$$\gg (\log q_t)^{-2}(XY)^{-1} \sum_{\substack{X<x_1\leqslant 2X \\ Y<y_1\leqslant 2Y \\ b_t+q_{t-1}x_1+q_ty_1\equiv 0 \bmod F(t)}} \frac{\varphi(b_t+q_{t-1}x_1+q_ty_1)}{b_t+q_{t-1}x_1+q_ty_1}.$$

We claim that

$$S := \sum_{\substack{X<x_1\leqslant 2X \\ Y<y_1\leqslant 2Y \\ b_t+q_{t-1}x_1+q_ty_1\equiv 0 \bmod F(t)}} \frac{\varphi(b_t+q_{t-1}x_1+q_ty_1)}{b_t+q_{t-1}x_1+q_ty_1} \gg XY/F(t).$$

To show this, we write

$$\mathcal{U} = \{n = b_t + q_{t-1}x_1 + q_ty : X < x_1 \leqslant 2X,\ Y < y_1 \leqslant 2Y,\ n \equiv 0 \bmod F(t)\},$$

and apply the AM–GM inequality to give

$$S \geqslant |\mathcal{U}| \left(\prod_{n\in\mathcal{U}} \frac{\varphi(n)}{n}\right)^{1/|\mathcal{U}|} = |\mathcal{U}| \prod_p \prod_{\substack{n\in\mathcal{U} \\ n\equiv 0 \bmod p}} (1-1/p)^{1/|\mathcal{U}|}.$$

By Lemma 2.15, we have

$$|\mathcal{U}| \gg XY/F(t),$$

so for the claim it suffices to show that

$$S' := \prod_p \prod_{\substack{n\in\mathcal{U}\\ n\equiv 0 \bmod p}} (1-1/p)^{1/|\mathcal{U}|} \gg 1.$$

Next, observe that

$$-\ln(S') = -\sum_p \tau_p \ln(1-1/p) \ll \sum_p \tau_p/p,$$

where

$$\tau_p = |\mathcal{U}|^{-1}\#\{n\in\mathcal{U} : n\equiv 0 \bmod p\},$$

so for the claim it remains to prove that

$$\sum_{p\geqslant 3} \tau_p/p \ll 1. \tag{4.18}$$

Let $p \geqslant 3$, and note that $p \nmid F(t)$ because $F(t)$ is a power of 4. As $\tau_p = 0$ for $p > q_t^2$, let us also assume that $p \leqslant q_t^2$. By Lemma 2.15, we have

$$\tau_p \ll F(t)\left(\frac{1}{pF(t)} + \frac{1}{X} + \frac{1}{Y}\right) \ll \frac{1}{p} + Y^{-1/2} \ll p^{-1/16},$$

giving (4.18).

Thus, we have the claim, and so

$$\mu(\mathcal{I})^{-1} \sum_{\substack{n=b_t+q_{t-1}x_1+q_ty_1\\ n\equiv 0 \bmod F(t)\\ X<x_1\leqslant 2X\\ Y<y_1\leqslant 2X}} \mu(\mathcal{A}_n) \gg (\log q_t)^{-2}F(t)^{-1}.$$

Summing over X and Y that are powers of 2 in the ranges (4.7), we obtain (4.5).

It remains to prove (4.6). To this end, it again suffices to prove (4.14), where X_1, Y_1, X_2, Y_2 are parameters in the ranges

$$q_t^{1/4} \leqslant Y_1 \leqslant q_t^{1/3}/2, \qquad q_t^{\theta_3} \leqslant X_1 \leqslant q_t^{\theta_4}/2,$$
$$q_s^{1/4} \leqslant Y_2 \leqslant q_s^{1/3}/2, \qquad q_s^{\theta_3} \leqslant X_2 \leqslant q_s^{\theta_4}/2,$$

but now in the count $N(t,s)$ we impose the additional restrictions

$$(q'_{t'}a + c_{t'}, n) = 1, \qquad (q'_{s'}b + c_{s'}, m) = 1. \tag{4.19}$$

Cases 1 and 2 from the previous subsection are unaffected, so our task is to count solutions for which

$$d = (m,n) \geqslant \max\{Y_1, 3\Delta/\|q'_2\gamma_2\|\} \qquad \text{(Case 3)}.$$

As $\mathcal{T}$ is sparse, this is only possible if

$$s = t.$$

Put $N = q_t Y_1$. Let us again write $n = dx$ and $m = dy$, so that $x > y$ and $(x, y) = 1$. Let C be a large, positive constant.

Case 3a: γ_2 has a continued fraction denominator in $[(N/d)^{1/\lambda}, CN/d]$, or $d \geqslant N$

In the case the proof from the previous subsection carries through, for in this case we may apply Lemma 2.8 therein.

Case 3b: γ_2 has no continued fraction denominator in $[(N/d)^{1/\lambda}, CN/d]$, and $d < N$

In this case, as

$$q'_{t'+1} \geqslant q_t \geqslant q_t Y_1/d = N/d > (N/d)^{1/\lambda},$$

we must have

$$q'_{t'+1} > CN/d,$$

where $q'_{t'+1}$ is the continued fraction denominator of γ_2 subsequent to $q'_{t'}$. Therefore

$$|q'_{t'}\gamma_2 - c_{t'}| < (q'_{t'+1})^{-1} < d/(CN).$$

As

$$|(n-m)q'_{t'}\gamma_2 - q'_{t'}(ma - nb)| < q'_{t'}\Delta,$$

the triangle inequality now confers

$$|(n-m)c_{t'} - q'_{t'}(ma - nb)| < q'_{t'}\Delta + d/2.$$

We thus have

$$1 \leqslant |x(q'_{t'}b + c_{t'}) - y(q'_{t'}a + c_{t'})| < q'_{t'}\Delta/d + 1/2,$$

owing to the coprimality restrictions (4.19) that we have thrust upon the problem. Whence

$$d < 2q'_{t'}\Delta, \qquad c_{t'}(x-y) \equiv O(q'_{t'}\Delta/d) \bmod q'_{t'}. \tag{4.20}$$

Recall that in this case we have $s = t$. We begin our count by choosing d so that

$$\max\{Y_1, 3\Delta/\|q'_2\gamma_2\|\} \leqslant d < 2q'_{t'}\Delta.$$

Next, we choose x in $O(q_t Y_1/d)$ ways. Then

$$y \ll q_t Y_2/d$$

lies in one of $O(q'_{t'}\Delta/d)$ residue classes modulo $q'_{t'}$, according to (4.20), so there are

$$O\left(\frac{q'_{t'}\Delta}{d}\left(\frac{q_tY_2}{dq'_{t'}}+1\right)\right)$$

possibilities for y. After that, the variable a is then determined modulo x from (4.15), and so there are at most $d\mu(\mathcal{I}) \leqslant d$ possibilities for a and then finally b is uniquely determined. Recalling that $X_1, X_2 \geqslant q_t^{\theta_3} = q_t^{1-\lambda^{-3}}$ and $Y_1, Y_2 \geqslant q_t^{1/4}$, our total count from this case is at most a constant times

$$\sum_{\max\{Y_1, 3\Delta/\|q'_2\gamma_2\|\}\leqslant d<2q'_{t'}\Delta} \frac{q_tY_1q'_{t'}\Delta}{d}\left(\frac{q_tY_2}{dq'_{t'}}+1\right) \leqslant N_1(t,s)+N_2(t,s),$$

where

$$N_1(t,s) = \sum_{d>Y_1}\frac{q_t^2Y_1Y_2\Delta}{d^2} \ll q_t^2\Delta Y_2 \ll \frac{X_1Y_1X_2Y_2\Delta\mu(\mathcal{I})}{F(t)F(s)}$$

and

$$\begin{aligned} N_2(t,s) &= \sum_{d<2q'_{t'}\Delta}\frac{q_tY_1q'_{t'}\Delta}{d} \ll q_tY_1q'_{t'}(\log q'_{t'})\Delta \leqslant q_t^{2+o(1)}Y_1\Delta \\ &\ll \frac{X_1Y_1X_2Y_2\Delta\mu(\mathcal{I})}{F(t)F(s)}. \end{aligned}$$

We have considered all cases, confirming (4.14). We conclude that Theorem 1.13 holds in the case that $\gamma_2 \in \mathcal{L}$.

4.4. **Rational second shift.** Finally, we prove Theorem 1.13 in the case that $\gamma_2 \in \mathbb{Q}$. Let $\gamma_2 = c_0/d_0$, where $c_0 \in \mathbb{Z}$ and $d_0 \in \mathbb{N}$ are fixed and coprime. We follow the previous subsection, but this time we replace $q'_{t'}$ and $c_{t'}$ by d_0 and c_0, respectively, for all $t \in \mathcal{T}$, and the proof carries through.

We have covered all possibilities for γ_2, completing the proof of Theorem 1.13.

5. Obstructions on Liouville fibres

Our proof of Theorem 1.15, via an Ostrowski expansion construction, rests upon the following technical lemma. In the following lemma and its proof, the implied constants are allowed to depend on α.

Lemma 5.1. *Suppose α, γ satisfy* (2.5). *Let $m(n)$ denote the least $i \geqslant 0$ such that $\delta_{i+1}(n) \neq 0$. Define*

$$\mathcal{W}_{u,d} = \{n \in \mathbb{N} : m(n) = u,\ |\delta_{u+1}(n)| = d\},$$

whenever $1 \leqslant d \leqslant a_{u+1} - b_{u+1}$, *as well as*

$$S_{u,d} = \sum_{n \in \mathcal{W}_{u,d}} \frac{1}{n(\log n)^2 \|n\alpha - \gamma\|}.$$

Then $\min \mathcal{W}_{u,d} \gg q_u$, *uniformly in* d. *Moreover, we have*

$$S_{u,d} \ll \begin{cases} \frac{1}{d \log q_{u+1}}, & \text{if } d > b_{u+1} \\ \frac{1}{\log q_u}, & \text{if } d = b_{u+1} \\ \frac{q_{u+1}}{(b_{u+1}-d) q_u (\log((b_{u+1}-d)q_u))^2 d} + \frac{1}{d \log q_{u+1}}, & \text{if } d < b_{u+1}. \end{cases}$$

Proof. For $n \in \mathcal{W}_{u,d}$, Lemma 2.5 implies that

$$\begin{aligned} \|n\alpha - \gamma\| &\gg \min\left((d-1)\left|D_u\right| + a_{u+2}\left|D_{u+1}\right|, a_1\left|D_0\right| + a_2|D_1|\right) \\ &= (d-1)\left|D_u\right| + a_{u+2}\left|D_{u+1}\right|. \end{aligned}$$

By (2.3), we have

$$(d-1)\left|D_u\right| \gg \frac{d-1}{q_{u+1}}$$

and

$$a_{u+2}\left|D_{u+1}\right| \gg \frac{a_{u+2}}{q_{u+2}} = \frac{a_{u+2}}{a_{u+2}q_{u+1} + q_u} \gg \frac{1}{q_{u+1}}.$$

Therefore

$$\|n\alpha - \gamma\| \gg \frac{d}{q_{u+1}} \qquad (n \in \mathcal{W}_{u,d}). \tag{5.1}$$

Using the notation of Lemma 2.3, observe that

$$\mathcal{W}_{u,d} = \mathcal{A}\left(b_1, \ldots, b_u, b_{u+1} + d\right) \cup \mathcal{A}\left(b_1, \ldots, b_u, b_{u+1} - d\right),$$

where $\mathcal{A}\left(b_1, \ldots, b_u, b_{u+1} + d\right)$ is understood to be empty if $b_{u+1} + d > a_{u+1}$, and $\mathcal{A}\left(b_1, \ldots, b_u, b_{u+1} - d\right)$ is empty if $b_{u+1} - d < 0$.

Observe that

$$S_{u,d} = S^+_{u,d} + S^-_{u,d},$$

where

$$S^{\pm}_{u,d} = \sum_{n \in \mathcal{A}(b_1, \ldots, b_u, b_{u+1} \pm d)} \frac{1}{n(\log n)^2 \|n\alpha - \gamma\|}.$$

<u>Case 1: $n \in \mathcal{A}(b_1, \ldots, b_u, b_{u+1} + d)$</u>

Then

$$n \geqslant \sum_{0 \leqslant k < u} b_{k+1} q_k + \left(b_{u+1} + d\right) q_u + \sum_{k > u} c_{k+1} q_k \geqslant b_{u+1} q_u \gg q_{u+1}.$$

From Lemma 2.3, applied with $m = u$ and $d_{u+1} = b_{u+1} + d > 0$, any two distinct elements of $\mathcal{A}(b_1, \ldots, b_u, b_{u+1} + d)$ differ by at least q_{u+1}. So the rth smallest element n_r of $\mathcal{A}(b_1, \ldots, b_u, b_{u+1} + d)$ satisfies

$$n_r \geqslant \min \mathcal{A}(b_1, \ldots, b_u, b_{u+1} + d) + (r-1)q_{u+1} \gg rq_{u+1},$$

uniformly in d. Combining this with (5.1), we deduce that

$$\begin{aligned} S^+_{u,d} &\ll \sum_{r \geqslant 1} \frac{1}{rq_{u+1}(\log(rq_{u+1}))^2 \frac{d}{q_{u+1}}} \\ &= \frac{1}{d}\left(\sum_{r < q_{u+1}} \frac{1}{r(\log(rq_{u+1}))^2} + \sum_{r \geqslant q_{u+1}} \frac{1}{r(\log(rq_{u+1}))^2} \right) \\ &\leqslant \frac{1}{d}\left(\sum_{r < q_{u+1}} \frac{1}{r(\log q_{u+1})^2} + \sum_{r \geqslant q_{u+1}} \frac{1}{r(\log r)^2} \right). \end{aligned}$$

The first sum is $O(1/\log q_{u+1})$, and the second sum is bounded by a constant times

$$\sum_{j \geqslant \log q_{u+1}} e^j \frac{1}{e^j j^2} \ll \int_{\log q_{u+1}}^{\infty} \frac{\mathrm{d}x}{x^2} \ll \frac{1}{\log q_{u+1}}.$$

Therefore

$$S^+_{u,d} \ll \frac{1}{d \log q_{u+1}}.$$

Case 2: $n \in \mathcal{A}(b_1, \ldots, b_u, b_{u+1} - d)$

For the aforementioned set to be non-empty, following our earlier convention, we must have $1 \leqslant d \leqslant b_{u+1}$. We distinguish two sub-cases.

Case 2a: $1 \leqslant d < b_{u+1}$

Note that

$$n \geqslant \sum_{0 \leqslant k < u} b_{k+1} q_k + (b_{u+1} - d)\, q_u + \sum_{k > u} c_{k+1} q_k \geqslant (b_{u+1} - d)\, q_u.$$

Let $n_0 = \min \mathcal{W}_{u,d}$, and for $r \geqslant 1$ denote by n_r the rth smallest element of $\mathcal{W}_{u,d} \setminus \{n_0\}$. It follows from Lemma 2.3 that

$$n_r \geqslant rq_{u+1} + (b_{u+1} - d)\, q_u \qquad (r \geqslant 0).$$

Together with (5.1), this gives

$$\begin{aligned} S^-_{u,d} &\ll \sum_{r\geqslant 0} \frac{1}{(rq_{u+1}+(b_{u+1}-d)\,q_u)(\log(rq_{u+1}+(b_{u+1}-d)\,q_u))^2\frac{d}{q_{u+1}}} \\ &\leqslant \frac{q_{u+1}}{(b_{u+1}-d)\,q_u(\log((b_{u+1}-d)\,q_u))^2 d} + \sum_{r\geqslant 1}\frac{1}{rq_{u+1}(\log(rq_{u+1}))^2\frac{d}{q_{u+1}}}. \end{aligned}$$

As in Case 1, we have

$$\sum_{r\geqslant 1}\frac{1}{rq_{u+1}(\log(rq_{u+1}))^2\frac{d}{q_{u+1}}} \ll \frac{1}{d\log q_{u+1}},$$

and so

$$S^-_{u,d} \ll \frac{q_{u+1}}{(b_{u+1}-d)\,q_u(\log((b_{u+1}-d)\,q_u))^2 d} + \frac{1}{d\log q_{u+1}}.$$

<u>Case 2b: $d = b_{u+1}$</u>

Note that

$$n \geqslant \sum_{0\leqslant k<u} b_{k+1}q_k \gg a_u q_{u-1} \gg q_u.$$

Let $n_0 = \min \mathcal{W}_{u,d}$, and for $r \geqslant 1$ denote by n_r the rth smallest element of $\mathcal{W}_{u,d} \setminus \{n_0\}$. By (5.1), we have

$$S^-_{u,d} \ll T_1 + T_2,$$

where

$$T_j = \sum_{r\geqslant 0}\frac{1}{n_{j+2r}(\log n_{j+2r})^2\,\frac{b_{u+1}}{q_{u+1}}} \qquad (j=1,2).$$

Let $j \in \{1,2\}$. We infer from Lemma 2.3 that if $r \geqslant 0$ then $n_{j+2r} \gg rq_{u+1}+q_u$. Whence

$$\begin{aligned} T_j &\ll \sum_{r\geqslant 0}\frac{1}{(rq_{u+1}+q_u)(\log(rq_{u+1}+q_u))^2\,\frac{b_{u+1}}{q_{u+1}}} \\ &\ll \frac{1}{q_u(\log q_u)^2\frac{b_{u+1}}{q_{u+1}}} + \sum_{r\geqslant 1}\frac{1}{r(\log(rq_{u+1}))^2 b_{u+1}} \\ &\ll \frac{1}{(\log q_u)^2} + \frac{1}{\log q_{u+1}} \ll \frac{1}{\log q_u}. \end{aligned}$$

□

We are now in the position to prove the main result of this section.

Proof of Theorem 1.15. Let $A = (a_n)_{n=1}^{\infty}$ be a sequence in $64\mathbb{N}$, sufficiently rapidly-increasing that

- The sequence defined by
$$q_0 = 1, \qquad q_1 = a_1, \qquad q_{u+1} = a_{u+1}q_u + q_{u-1} \quad (u \geqslant 1)$$
satisfies
$$\sum_{u \geqslant 0} \left(\frac{1}{\log q_u} + \frac{1}{\xi(q_u)} \right) < \infty \tag{5.2}$$
- $a_{u+1} \geqslant q_u!$ $(u \geqslant 0)$.

Let $\mathcal{V}$ be the collection of all such sequences $A = (a_i)_i$. For $A \in \mathcal{V}$, define $\alpha(A) := [0; a_1, a_2, \ldots]$. For $\sigma = (\sigma_i)_i \in \{0,1\}^{\mathbb{N}}$ and $A = (a_i)_i \in \mathcal{V}$, we define a sequence $(b_i(A,\sigma))_i$ by
$$b_i(A,\sigma) = \frac{a_i}{2^{1+\sigma_i}} \qquad (i \in \mathbb{N}),$$
as well as a real number
$$\gamma(A,\sigma) := \sum_{k \geqslant 0} b_{k+1}(A,\sigma) D_k(A),$$
where $D_k(A) = q_k\alpha(A) - p_k$ and p_k/q_k is the kth convergent to $\alpha(A)$. Note that we have (2.5), so by Lemma 2.4 we have $\gamma(A,\sigma) \in [0, 1-\alpha(A))$ and
$$\|n\alpha(A) - \gamma(A,\sigma)\| \neq 0 \qquad (n \in \mathbb{N}).$$
There are continuum many $\alpha(A)$, and they are Liouville by Lemma 2.1. For the rest of the proof, we fix a pair $(\alpha(A), \gamma(A,\sigma))$, where $A \in \mathcal{V}$ and $\sigma \in \{0,1\}^{\mathbb{N}}$, and abbreviate $\alpha = \alpha(A)$, $\gamma = \gamma(A,\sigma)$.

Fix $\gamma_2 \in \mathbb{R}$, and consider $\alpha_1 = \alpha$ and $\gamma_1 = \gamma$. By the Borel–Cantelli lemma, it remains to prove that
$$\sum_{n \geqslant 1} \frac{\psi_\xi(n)}{\|n\alpha - \gamma\|} < \infty, \tag{5.3}$$
where ψ_ξ is as in (1.8). To this end, observe that
$$\sum_{n \geqslant 1} \frac{\psi_\xi(n)}{\|n\alpha - \gamma\|} \leqslant \sum_{u \geqslant 0} \sum_{d \neq b_{u+1}} \max_{n \in \mathcal{W}_{u,d}} \frac{1}{\xi(n)} S_{u,d} + \sum_{u \geqslant 0} \max_{n \in \mathcal{W}_{u,b_{u+1}}} \frac{1}{\xi(n)} S_{u,b_{u+1}},$$
where here and henceforth $d \neq b_{u+1}$ means that
$$d \in \{1, \ldots, a_{u+1} - b_{u+1}\} \setminus \{b_{u+1}\}.$$

Since ξ non-decreasing and unbounded, we infer from Lemma 5.1 that

$$\max_{n\in\mathcal{W}_{u,d}} \frac{1}{\xi(n)} = \frac{1}{\xi(\min \mathcal{W}_{u,d})} \ll \frac{1}{\xi(q_{u-1})},$$

where here $q_{-1} = 1$. By Lemma 5.1, we now have

$$\sum_{n\geqslant 1} \frac{\psi_\xi(n)}{\|n\alpha-\gamma\|} \ll T_1 + T_2 + T_3,$$

where

$$T_1 = \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})}\,\frac{1}{\log q_u},$$

$$T_2 = \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})} \sum_{d\neq b_{u+1}} \frac{1}{d\log q_{u+1}},$$

$$T_3 = \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})} \sum_{d<b_{u+1}} \frac{q_{u+1}}{(b_{u+1}-d)\,q_u(\log((b_{u+1}-d)\,q_u))^2 d}.$$

We see from (5.2) that $T_1 < \infty$. The convergence of T_2 also follows straightforwardly from (5.2), since

$$T_2 \ll \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})}\,\frac{\log a_{u+1}}{\log q_{u+1}} \leqslant \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})} < \infty.$$

Our final task is to establish the convergence of the series defining T_3. We begin with the observation that

$$\begin{aligned} T_3 &\leqslant \sum_{u\geqslant 0} \frac{q_{u+1}}{\xi(q_{u-1})q_u} \sum_{d<b_{u+1}} \frac{1}{d(b_{u+1}-d)(\log(b_{u+1}-d))^2} \\ &\ll \sum_{u\geqslant 0} \frac{b_{u+1}}{\xi(q_{u-1})} \sum_{d<b_{u+1}} \frac{1}{d(b_{u+1}-d)(\log(b_{u+1}-d))^2}. \end{aligned}$$

The inner sum is at most $X_u + Y_u$, where

$$X_u = \sum_{d\leqslant b_{u+1}/2} \frac{1}{d(b_{u+1}-d)(\log(b_{u+1}-d))^2}$$

and

$$Y_u = \sum_{t\leqslant b_{u+1}/2} \frac{1}{t(b_{u+1}-t)(\log t)^2} \geqslant X_u.$$

Finally, we have

$$T_3 \ll \sum_{u\geqslant 0} \frac{b_{u+1}}{\xi(q_{u-1})} Y_u \ll \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})} \sum_{t\geqslant 1} \frac{1}{t(\log t)^2} \ll \sum_{u\geqslant 0} \frac{1}{\xi(q_{u-1})},$$

which converges. □

Appendix A. Pathology

In this appendix, we establish Theorem 1.26. If we have (3.6) then the proof of Theorem 1.23 prevails. We also need to consider the pathological situation in which $\Psi(n) > 1/2$ for infinitely many $n \in \mathbb{N}$. We will slightly alter this dichotomy.

Let c^* be a small, positive constant. Our implicit constants will not depend on c^* unless otherwise stated. The first idea is to restrict the support of Ψ to

$$\mathcal{G}^* = \left\{ n \in \mathcal{G} : \frac{\varphi(\hat{n})}{\hat{n}} \geqslant c^* \right\},$$

where $\mathcal{G}$ and $\eta = \eta(\gamma)$ are as in Section 3. That is, we introduce $\Psi^* = \Psi 1_{\mathcal{G}^*}$ and

$$\mathcal{E}_n^* = \left\{ \alpha \in [0,1] : \exists a \in \mathbb{Z} \text{ s.t. } \begin{array}{c} a + \gamma \in \hat{n}\mathcal{I}, \\ |\hat{n}\alpha - \gamma - a| < \Psi^*(n), \\ (a, \hat{n}) \text{ is } (\gamma, \eta)\text{-shift-reduced} \end{array} \right\}.$$

<u>Case: $\Psi^*(n) \leqslant 1/2$ for large n</u>

We commence by discussing (1.18). Our modification can only reduce the left hand side of (3.5) so, by the reasoning of Proposition 3.3, it remains to justify (3.4) with $\mathcal{E}_n^*$ in place of $\mathcal{E}_n$. The upper bound is immediate from the inequality $\mu(\mathcal{E}_n^*) \leqslant \mu(\mathcal{E}_n)$, leaving us to deal with the lower bound. By (3.7), we have

$$\sum_{n \leqslant X} \mu(\mathcal{E}_n^*) \gg \mu(\mathcal{I}) \sum_{\substack{n \in \mathcal{G}^* \\ C_1 < n \leqslant X}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1 - \gamma_1\| \cdots \|\hat{n}\alpha_{k-1} - \gamma_{k-1}\|}, \tag{A.1}$$

where C_1 is a large, positive constant. We compute that

$$\begin{aligned} &\sum_{\substack{n \in \mathcal{G} \setminus \mathcal{G}^* \\ n \leqslant X}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1 - \gamma_1\| \cdots \|\hat{n}\alpha_{k-1} - \gamma_{k-1}\|} \\ &< c^* \sum_{\substack{n \in \mathcal{G} \\ n \leqslant X}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1 - \gamma_1\| \cdots \|\hat{n}\alpha_{k-1} - \gamma_{k-1}\|} \\ &\ll c^* \sum_{n \leqslant X} \psi(\hat{n}) (\log n)^{k-1}, \end{aligned}$$

where the final inequality follows from the calculations within the proof of Lemma 3.8. As c^* is small, combining this with (3.16) and (3.17) yields

$$\sum_{\substack{n\in\mathcal{G}^*\\ n\leqslant X}} \frac{\varphi(\hat{n})}{\hat{n}} \frac{\psi(\hat{n})}{\|\hat{n}\alpha_1-\gamma_1\|\cdots\|\hat{n}\alpha_{k-1}-\gamma_{k-1}\|} \gg \sum_{n\leqslant X}\psi(\hat{n})(\log n)^{k-1}.$$

Substituting this into (A.1) gives

$$\sum_{n\leqslant X}\mu(\mathcal{E}_n^*) \gg \mu(\mathcal{I})\sum_{n\leqslant X}\psi(\hat{n})(\log n)^{k-1}.$$

The upshot is that we have (3.4) with $\mathcal{E}_n^*$ in place of $\mathcal{E}_n$, which is the last remaining ingredient needed for (1.18).

Let N be large, and let us now specialise $\mathcal{I}=[0,1]$. By (3.1) and the above, we have

$$\sum_{n=1}^{\infty}\mu(\mathcal{E}_n^*)=\infty.$$

By (3.8), we have

$$\frac{\varphi_{\gamma,\eta}(\hat{n})}{\hat{n}} \gg \frac{\varphi(\hat{n})}{\hat{n}} \geqslant c^* \qquad (n\in\mathcal{G}^*,\quad n>N),$$

and note also that

$$\mu(\mathcal{E}_n^*)\leqslant\mu(\mathcal{E}_n)\ll\Psi(n)\leqslant\Phi(\hat{n}) \qquad (n>N).$$

Therefore

$$\sum_{n=1}^{\infty}\frac{\varphi_{\gamma,\eta}(\hat{n})}{\hat{n}}\Phi(\hat{n}) \gg c^*\sum_{\substack{n\in\mathcal{G}^*\\ n>N}}\mu(\mathcal{E}_n^*)=c^*\sum_{n>N}\mu(\mathcal{E}_n^*)=\infty,$$

which implies (1.17). Having established (1.17) and (1.18), we have completed the proof of the theorem in this case.

<u>Case: $\Psi^*(n)>1/2$ infinitely often</u>

Let N be large, and put

$$\mathcal{S}=\{n\in\mathbb{N}:\Psi^*(n)>1/2,\quad n>N\}.$$

The inequalities

$$\frac{\varphi_{\gamma,\eta}(\hat{n})}{\hat{n}}\gg c^*,\quad \Phi(\hat{n})\geqslant\Psi^*(n)>1/2 \qquad (n\in\mathcal{S}),$$

together with the infinitude of $\mathcal{S}$, yield

$$\sum_{n\in\mathcal{S}}\frac{\varphi_{\gamma,\eta}(\hat{n})}{\hat{n}}\Phi(\hat{n})=\infty,$$

which implies (1.17).

For $n \in \mathbb{N}$, define

$$\Psi^\dagger(n) = \begin{cases} 1/2, & \text{if } n \in \mathcal{S} \\ 0, & \text{if } n \notin \mathcal{S}. \end{cases}$$

and

$$\mathcal{E}_n^\dagger = \left\{ \alpha \in [0,1] : \exists a \in \mathbb{Z} \text{ s.t.} \quad \begin{array}{c} a + \gamma \in \hat{n}\mathcal{I}, \\ |\hat{n}\alpha - \gamma - a| < \Psi^\dagger(n), \\ (a, \hat{n}) \text{ is } (\gamma, \eta)\text{-shift-reduced} \end{array} \right\}.$$

Let $\mathcal{I}$ be a non-empty subinterval of $[0,1]$. Henceforth, our implied constants will be allowed to depend on c^* but not $\mathcal{I}$. By (3.8), we have

$$\mu(\mathcal{E}_n^\dagger) \gg \frac{\varphi_0(\hat{n})}{\hat{n}} \gg \mu(\mathcal{I}) \qquad (n \in \mathcal{S}),$$

and clearly

$$\mu(\mathcal{E}_n^\dagger \cap \mathcal{E}_m^\dagger) \leqslant \mu(\mathcal{E}_n^\dagger) \ll \mu(\mathcal{I}) \qquad (n, m \in \mathcal{S}).$$

Applying Lemmata 2.9 and 2.10, as in Proposition 3.3, furnishes

$$\mu\left(\limsup_{n\to\infty} \mathcal{A}_n\right) = 1,$$

where

$$\mathcal{A}_n = \left\{ \alpha \in [0,1] : \exists a \in \mathbb{Z} \text{ s.t.} \quad \begin{array}{c} |\hat{n}\alpha - \gamma - a| < \Psi^\dagger(n), \\ (a, \hat{n}) \text{ is } (\gamma, \eta)\text{-shift-reduced} \end{array} \right\}.$$

This implies (1.18), since $\Phi(\hat{n}) \geqslant \Psi^\dagger(n)$ for all n, and completes the proof of Theorem 1.26.

Remark A.1. In the context of Remark 1.25, our proof works for any sufficiently small η, by scaling ε_0 accordingly.

References

[1] F. Adiceam, E. Nesharim and F. Lunnon, *On the t-adic Littlewood conjecture*, Duke Math. J., to appear, DOI: 10.1215/00127094-2020-0077, arXiv:1806.04478.

[2] C. Aistleitner, *A note on the Duffin–Schaeffer conjecture with slow divergence*, Bull. Lond. Math. Soc. **46** (2014), 164–168.

[3] C. Aistleitner, T. Lachmann, M. Munsch, N. Technau and A. Zafeiropoulos, *The Duffin–Schaeffer conjecture with extra divergence*, Adv. Math. **356** (2019).

[4] D. Badziahin, *On multiplicatively badly approximable numbers*, Mathematika **59** (2013), 31–55.

[5] D. Badziahin and S. Velani, *Multiplicatively badly approximable numbers and generalised Cantor sets*, Adv. Math. **228** (2011), 2766–2796.

[6] F. Barroero and M. Widmer, *Counting lattice points and o-minimal structures*, Int. Math. Res. Not. **2014,** 4932–4957.

[7] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation*, Ann. of Math. (2), **175** (2012), 187–235.

[8] V. Beresnevich, D. Dickinson and S. Velani, *Measure theoretic laws for lim sup sets*, Mem. Amer. Math. Soc. **179** (2006).

[9] V. Beresnevich, D. Dickinson and S. Velani, *Diophantine approximation on planar curves and the distribution of rational points*, Ann. of Math. (2) **166** (2007), 367–426.

[10] V. Beresnevich, G. Harman, A. Haynes and S. Velani, *The Duffin-Schaeffer conjecture with extra divergence II*, Math. Z. **275** (2013), 127–133.

[11] V. Beresnevich, A. Haynes and S. Velani, *Sums of reciprocals of fractional parts and multiplicative Diophantine approximation*, Mem. Amer. Math. Soc. **263** (2020).

[12] V. Beresnevich, F. Ramírez and S. Velani, *Metric Diophantine Approximation: some aspects of recent work*, Dynamics and Analytic Number Theory, London Math. Soc. Lecture Note Ser. (N.S.) **437,** Cambridge University Press, 2016, pp. 1–95.

[13] V. Beresnevich and S. Velani, *An inhomogeneous transference principle and Diophantine approximation*, Proc. Lond. Math. Soc. (3) **101** (2010), 821–851.

[14] V. Beresnevich and S. Velani, *A note on three problems in metric Diophantine approximation*, Recent Trends in Ergodic Theory and Dynamical Systems, Contemp. Math. **631** (2015), 211–229.

[15] Y. Bugeaud, *Approximation by algebraic numbers*, Cambridge Tracts in Mathematics **160,** Cambridge University Press, Cambridge, 2004.

[16] Y. Bugeaud and M. Laurent, *On exponents of homogeneous and inhomogeneous Diophantine approximation*, Mosc. Math. J. **5** (2005), 747–766.

[17] Y. Bugeaud and M. Laurent, *On transfer inequalities in Diophantine approximation, II*, Math. Z. **265** (2010), 249–262.

[18] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer, 1997.

[19] S. Chow, *Bohr sets and multiplicative diophantine approximation*, Duke Math. J. **167** (2018), 1623–1642.

[20] S. Chow, A. Ghosh, L. Guan, A. Marnat and D. Simmons, *Diophantine transference inequalities: weighted, inhomogeneous, and intermediate exponents*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **XXI** (2020), 643–671.

[21] S. Chow and N. Technau, *Higher-rank Bohr sets and multiplicative diophantine approximation*, Compositio Math. **155** (2019), 2214–2233.

[22] S. Chow and L. Yang, *Effective equidistribution for multiplicative diophantine approximation on lines*, arXiv:1902.06081.

[23] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[24] H. Davenport, *On a principle of Lipschitz*, J. London Math. Soc. **1** (1951), 179–183.

[25] R. J. Duffin and A. C. Schaeffer, *Khintchine's problem in metric Diophantine approximation*, Duke Math. J. **8** (1941), 243–255.

[26] M. Einsiedler, A. Katok, and E. Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), 513–560.

[27] P. Erdős, *On the distribution of convergents of almost all real numbers*, J. Number Theory **2** (1970), 425–441.

[28] P. Erdős, *Representations of real numbers as sums and products of Liouville numbers* Michigan Math. J. **9** (1962), 59–60.

[29] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons, 2004.

[30] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications, vol. 57, American Mathematical Society, Providence, RI, 2010.

[31] P. X. Gallagher, *Metric simultaneous diophantine approximation*, J. Lond. Math. Soc. **37** (1962), 387–390.

[32] A. Ghosh and A. Marnat, *On diophantine transference principles*, Math. Proc. Camb. Phil. Soc. **166** (2019), 415–431.

[33] A. Gorodnik and P. Vishe, *Diophantine approximation for products of linear maps—logarithmic improvements*, Trans. Amer. Math. Soc. **370** (2018), 487–507.

[34] G. Harman, *Metric number theory*, London Math. Soc. Lecture Note Ser. (N.S.), vol. 18, Clarendon Press, Oxford 1998.

[35] G. Harman, *Some cases of the Duffin and Schaeffer conjecture*, Quart. J. Math., **41** (1990), 395–404.

[36] A. Haynes, A. Pollington and S. Velani, *The Duffin-Schaeffer Conjecture with extra divergence*, Math. Ann. **353** (2012), 259–273.

[37] J.-J. Huang, *Rational points near planar curves and Diophantine approximation*, Adv. Math. **274** (2015), 490–515.

[38] J.-J. Huang, *The density of rational points near hypersurfaces*, Duke Math. J. **169** (2020), 2045–2077.

[39] M. Hussain and D. Simmons, *The Hausdorff measure version of Gallagher's theorem — closing the gap and beyond*, J. Number Theory **186** (2018), 211–225.

[40] A. Ya. Khintchine, *Über eine Klasse linearer diophantischer Approximationen*, Rendiconti Circ. Mat. Palermo **50** (1926), 170–195.

[41] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), 339–360.

[42] D. Koukoulopoulos, *The distribution of prime numbers*, Graduate Studies in Mathematics, vol. 203, American Mathematical Society, Providence, RI, 2019.

[43] D. Koukoulopoulos and J. Maynard, *On the Duffin-Schaeffer conjecture*, Ann. of Math. (2) **192** (2020), 251–307.

[44] T.-H. Lê and J. Vaaler, *Sums of products of fractional parts*, Proc. Lond. Math. Soc. (3) **111** (2015), 561–590.

[45] G. Margulis, *Problems and conjectures in rigidity theory*, Mathematics: frontiers and perspectives, 2000, 161–174.

[46] M. Mukherjee and G. Karner, *Irrational numbers of constant type — a new characterization*, New York J. Math. **4** (1998), 31–34.

[47] L. G. Peck, *Simultaneous rational approximations to algebraic numbers*, Bull. Amer. Math. Soc. **67** (1961), 197–201.

[48] A. D. Pollington and R. C. Vaughan, *The k-dimensional Duffin and Schaeffer conjecture*, Mathematika (2) **37** (1990), 190–200.

[49] A. D. Pollington and S. L. Velani, *On a problem in simultaneous Diophantine approximation: Littlewood's conjecture*, Acta Math. **185** (2000), 287–306.

[50] F. Ramírez, *Counterexamples, covering systems, and zero-one laws for inhomogeneous approximation*, Int. J. Number Theory **13** (2017), 633–654.

[51] F. Ramírez, *Khintchine's theorem with random fractions*, Mathematika **66** (2020) 178–199.

[52] A. Rockett, and P. Szüsz, *Continued Fractions*. World Scientific (1992), Singapore.

[53] U. Shapira, *A solution to a problem of Cassels and Diophantine properties of cubic numbers*, Ann. of Math. (2) **173** (2011), 543–557.

[54] P. Szüsz, *Über die metrische Theorie der Diophantischen Approximation*, Acta. Math. Sci. Hungar. **9** (1958), 177–193.

[55] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Stud. Adv. Math., vol. 105, Cambridge University Press, Cambridge, 2006.
[56] T. Tao and V. Vu, *John-type theorems for generalized arithmetic progressions and iterated sumsets*, Adv. Math. **219** (2008), 428–449.
[57] J. L. Thunder, *The number of solutions of bounded height to a system of linear equations*, J. Number Theory **43** (1993), 228–250.
[58] J. Vaaler, *On the metric theory of Diophantine approximation*, Pacific J. Math. **76** (1978), 527–539.
[59] R. C. Vaughan and S. Velani, *Diophantine approximation on planar curves: the convergence theory*, Invent. Math. **166** (2006), 103–124.
[60] H. Yu, *A Fourier-analytic approach to inhomogeneous Diophantine approximation*, Acta Arith. **190** (2019), 263–292.
[61] H. Yu, *On the metric theory of inhomogeneous Diophantine approximation: An Erdős-Vaaler type result*, J. Number Theory **224** (2021), 243–273.

Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL, United Kingdom

Email address: sam.chow@warwick.ac.uk

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA

Email address: technau@math.wisc.edu