Christoph Aistleitner · Valentin Blomer · Maksym Radziwiłł

Triple correlation and long gaps in the spectrum of flat tori

Abstract. We evaluate the triple correlation of eigenvalues of the Laplacian on generic flat tori in an averaged sense. As two consequences we show that (a) the limit inferior (resp. limit superior) of the triple correlation is almost surely at most (resp. at least) Poissonian, and that (b) almost all flat tori contain infinitely many gaps in their spectrum that are at least 2.006 times longer than the average gap.

The significance of the constant 2.006 lies in the fact that there exist sequences with Poissonian pair correlation and with consecutive spacings bounded uniformly from above by 2, as we also prove in this paper. Thus our result goes beyond what can be deduced solely from the Poissonian behavior of the pair correlation.

Keywords. Billiard, flat torus, long gaps, spectrum, Berry–Tabor conjecture, Poisson statistics, pair correlation, triple correlation, Diophantine inequalities

1. Introduction

1.1. The Berry–Tabor conjecture for tori

A central objective in quantum chaos is the classification of quantum systems according to universal statistical properties. We consider the energy spectrum $\lambda_1 \leq \lambda_2 \leq \cdots$ of a bound Hamiltonian system, such as the geodesic flow on a compact Riemannian manifold, where the energy levels are the eigenvalues of the associated Laplace–Beltrami operator. According to central conjectures [2,4], one expects a fundamental dichotomy between integrable and chaotic systems. In particular, the local statistics of the sequence, normalized to have

Christoph Aistleitner: Institute of Analysis and Number Theory, Graz University of Technology, 8010 Graz, Austria; aistleitner@math.tugraz.at

Valentin Blomer: Mathematisches Institut, University of Bonn, 53177 Bonn, Germany; blomer@math.uni-bonn.de

Maksym Radziwiłł: Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, USA; maksym.radziwill@gmail.com

mean spacing 1, should imitate those of certain random matrix ensembles. In the former case, the local statistical properties are expected to coincide with those of a sequence of points coming from a Poisson process, at least in generic cases. In particular, the probability that a randomly chosen interval $[X, X + r]$ of fixed length r contains exactly k of the numbers Λ_j is conjectured to be $(k!)^{-1} r^k e^{-r}$. From a rigorous mathematical standpoint these conjectures are completely out of reach, and we content ourselves here with the simple, yet fundamental example of the following class of integrable Hamiltonian systems.

Let $\Lambda \subseteq \mathbb{R}^2$ be a lattice of rank 2. Then $\mathbb{R}^2/\Lambda$ endowed with the Euclidean metric is a flat torus and $\Delta = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$ is the Laplace–Beltrami operator on $\mathbb{R}^2/\Lambda$. The eigenvalues of the Laplacian Δ on $\mathbb{R}^2/\Lambda$ are given by $4\pi^2 \|\boldsymbol{\omega}\|_2^2$ with $\boldsymbol{\omega}$ belonging to the dual lattice $\Lambda^\star$. Thus the eigenvalues are integral values of quadratic forms,

$$q_{\boldsymbol{\alpha}}(m, n) = \alpha_1 m^2 + \alpha_2 mn + \alpha_3 n^2, \quad \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3), \tag{1.1}$$

where $\boldsymbol{\alpha}$ is always constrained by $4\alpha_1\alpha_3 > \alpha_2^2$ since $q_{\boldsymbol{\alpha}}$ is necessarily positive-definite, and $\boldsymbol{\alpha}$ is determined by the lattice Λ up to $\mathrm{GL}_2(\mathbb{Z})$-invariance. A standard fundamental domain (up to boundary) is

$$\mathscr{D} := \{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3 \mid 0 \leq \alpha_2 \leq \alpha_1 \leq \alpha_3\},$$

and for convenience we will restrict $\boldsymbol{\alpha}$ to this set. Each quadratic form $q_{\boldsymbol{\alpha}}$ has the automorphism $(m, n) \mapsto (-m, -n)$, so each positive eigenvalue occurs with multiplicity at least 2. It is therefore natural to desymmetrize the spectrum and consider only the values $q_{\boldsymbol{\alpha}}(m, n)$ with $m > 0$ or $m = 0$ and $n \geq 0$. This does not catch all multiplicities (e.g. if $\boldsymbol{\alpha} = (1, 0, 1)$), but for typical $\boldsymbol{\alpha}$ one may expect that desymmetrized spectrum is multiplicity-free.

Given $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ we denote by $0 < \Lambda_1 \leq \Lambda_2 \leq \cdots$ the ordered set of values

$$\frac{\pi}{\sqrt{4\alpha_1\alpha_3 - \alpha_2^2}} \cdot q_{\boldsymbol{\alpha}}(m, n) \tag{1.2}$$

where $m > 0$ or $m = 0$ and $n \geq 0$. Asymptotically the average spacing between the Λ_i is 1, thus we think of Λ_i as the properly re-scaled multi-set of eigenvalues of the Laplacian on $\mathbb{R}^2/\Lambda$ (desymmetrized after removing the obvious double multiplicity of eigenvalues).

1.2. Triple correlation

In this setting, a consequence of the Berry–Tabor conjecture is that for "generic" tori the distribution of the gaps $\Lambda_{i+1} - \Lambda_i$ should coincide with the distribution of nearest neighbor spacings of the Poisson process.[1] Precisely, the number of $i \leq N$ for which

---

[1] As an aside, the parallelogram rule implies that there are linear relations in the sequence of Λ_i, but they do not seem to influence the fine-scale statistical properties.

$\Lambda_{i+1} - \Lambda_i \in [0, r]$ should be asymptotically $N \int_0^r e^{-u} \, du$. The word "generic" implicitly assumes a measure, and in this case the natural measure on $\mathcal{D}$ is the hyperbolic measure

$$d\mu_{\text{hyp}}(\alpha_1, \alpha_2, \alpha_3) = \frac{d\alpha_1 \, d\alpha_2 \, d\alpha_3}{(4\alpha_1\alpha_3 - \alpha_2^2)^{3/2}}.$$

While the Berry–Tabor conjecture for generic flat tori remains wide open, some results have been obtained for *n-correlations*

$$T_n(\boldsymbol{\alpha}; I_2, \ldots, I_n; N)$$
$$:= \frac{1}{N} |\{(i_1, \ldots, i_n) \in \mathcal{J}_n(N) \mid \Lambda_{i_j}(\boldsymbol{\alpha}) - \Lambda_{i_1}(\boldsymbol{\alpha}) \in I_j \text{ for } 2 \leq j \leq n\}|,$$

where $\mathcal{J}_n(N) \subset [1, N]^n$ is the subset of tuples all of whose entries are pairwise distinct and $I_2, \ldots, I_n$ are intervals in $\mathbb{R}$. When $I_2 = \cdots = I_n = I$ we write $T_n(\boldsymbol{\alpha}; I; N)$. We expect that the *n*-correlations coincide for all $n \geq 1$ with those of a sequence generated by a Poisson process, i.e. for almost all $\boldsymbol{\alpha}$ with respect to the measure μ_{hyp} and for any fixed intervals $I_2, \ldots, I_n$ we expect to have

$$T_n(\boldsymbol{\alpha}; I_2, \ldots, I_n; N) \to \mu(I_2) \cdot \ldots \cdot \mu(I_n) \quad (N \to \infty), \tag{1.3}$$

where μ denotes Lebesgue measure.

Sarnak [14] established (1.3) for $n = 2$ for almost all $\boldsymbol{\alpha}$ with respect to the hyperbolic measure μ_{hyp} (equivalently for almost all $\boldsymbol{\alpha}$ with Lebesgue measure on $\alpha_1, \alpha_2, \alpha_3$). This was generalized in [16] to four-dimensional tori. Eskin, Margulis and Mozes [7] gave explicit Diophantine conditions on $\boldsymbol{\alpha}$ under which (1.3) holds when $n = 2$.

In contrast nothing is known about correlations of higher order, that is, for $n > 2$. (For higher-dimensional tori, VanderKam [17] has results for higher order correlations, which are based on the fact that more Diophantine variables are available.) While the pair correlation is associated to an orthogonal group of signature $(2, 2)$, no useful group structure seems to be available in the case of higher correlations. In particular, the methods of [7] are not available. Our main result sheds some light on the case $n > 2$ by establishing (1.3) for $n = 3$ in an averaged sense.

Theorem 1. *Let $R \subseteq \mathcal{D}$ be a three-dimensional rectangle of finite hyperbolic measure $\mu_{\text{hyp}}(R)$, and let $J \subset \mathbb{R}$ be a finite interval. Then*

$$\lim_{N \to \infty} \int_R T_3(\boldsymbol{\alpha}; I_1, I_2; N) \, d\mu_{\text{hyp}}(\boldsymbol{\alpha}) = \mu(I_1)\mu(I_2)\mu_{\text{hyp}}(R), \tag{1.4}$$

uniformly for all intervals $I_1, I_2 \subseteq J$.

A consequence of Theorem 1 is a sharp upper bound for $\liminf T_3(\boldsymbol{\alpha}; I; N)$ for almost all $\boldsymbol{\alpha} \in \mathcal{D}$ with respect to μ_{hyp}.

Corollary 2. *Let $I \subseteq \mathbb{R}$ be a finite interval. For almost all $\boldsymbol{\alpha} \in \mathcal{D}$ with respect to the measure $\mu_{\text{hyp}}(\boldsymbol{\alpha})$ we have*

$$\liminf_{N \to \infty} T_3(\boldsymbol{\alpha}; I; N) \leq \mu(I)^2 \quad \text{and} \quad \limsup_{N \to \infty} T_3(\boldsymbol{\alpha}; I; N) \geq \mu(I)^2.$$

Although the statistical investigation of the spectrum of a Riemann surface is a problem of mathematical physics, the main input of the proof of Theorem 1 is of Diophantine nature and consists in an involved lattice point count in a certain 6-dimensional region; see Section 3. It uses to some extent ideas of [14], but since Sarnak studies an L^2-mean of the pair correlation, he has 8 variables at his disposal, while we have only 6. Consequently, the Diophantine analysis becomes more delicate and depends, among other things, on non-trivial bounds for Kloosterman sums. In fact, our result can be viewed as a "thin subset" version of Sarnak's result. These remarks are explained in more detail in Section 1.4.

1.3. Long gaps in Poissonian sequences

Theorem 1 was originally motivated by the problem of establishing the existence of large gaps between consecutive Λ_i. This problem stands in sharp contrast with the problem of establishing small gaps between consecutive Λ_i. In the case of rectangular tori (i.e. $\alpha_2 = 0$) it is possible to obtain nearly optimal results for almost all $\boldsymbol{\alpha}$, that is, gaps of size $N^{-1+\varepsilon}$ among the first N eigenvalues (birthday paradox); see [3]. As long gaps in a Poisson process typically grow at a logarithmic scale, their investigation is a more delicate endeavour. In addition, long gaps are harder to detect from a technical point of view. While a difference of any two eigenvalues is in particular an upper bound for the smallest gap, no such approximation is available for a lower bound of long gaps.

It is therefore not surprising that comparatively little is known about large gaps between consecutive Λ_i. Using a variant of Corollary 2 (see Lemma 2) we obtain the following partial progress on this question.

Corollary 3. *Let G be the unique positive solution to the equation*

$$35 - 18G - 9G^2 + (12G - 10)\sqrt{6G - 5} = 0.$$

Then for almost all $\boldsymbol{\alpha}$ with respect to the measure μ_{hyp} we have

$$\limsup_{i \to \infty} (\Lambda_{i+1}(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha})) \geq G = 2.00636193892510\ldots.$$

Remark 1. In fact we prove a slightly stronger result: Given $\varepsilon > 0$, for almost all $\boldsymbol{\alpha}$ there exists a subsequence $N_1 < N_2 < \cdots$ along which we have $\Lambda_{i+1} - \Lambda_i > G - \varepsilon$ for a positive proportion of $i \leq N_j$ as $N_j \to \infty$.

As mentioned above, the Berry–Tabor conjecture predicts that the gaps $\Lambda_{i+1}(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha})$ are exponentially distributed (for generic $\boldsymbol{\alpha}$), which clearly would imply that the maximal gap size is unbounded. However, proving this seems to be far out of reach. As a remark pointing to this direction, note that by inclusion-exclusion it is not hard to see that if we could show (1.3) for all $n \leq 2k$ ($k \in \mathbb{N}, k \geq k_0$) and some $\boldsymbol{\alpha}$, then the corresponding torus has infinitely many gaps in its spectrum that are of length at least $k/2$. In particular, if (1.3) is known for all n, one obtains unbounded gaps.

The particular significance of Corollary 3 lies in the fact that the spacings that we obtain are strictly greater than 2. As we show in the following theorem, there exist sequences with gaps uniformly bounded by 2 and whose pair correlation is Poissonian. In other words, Theorem 1 shows that the gap structure of almost all flat tori has properties that cannot be derived from the pair correlation result alone.

It is an interesting problem in its own right to determine the minimal possible value of the largest gap of a sequence whose pair correlation is assumed to be Poissonian. To fix notation, let us denote by $\mathscr{S}$ the set of all increasing sequences (γ_i) of non-negative real numbers with mean spacing 1 and Poissonian pair correlation, i.e.

$$\frac{1}{N}\gamma_N \to 1 \quad \text{and} \quad \frac{1}{N}\sum_{\substack{\gamma_{i_1}-\gamma_{i_2}\in I \\ i_1,i_2\leq N,\, i_1\neq i_2}} 1 \to \mu(I) \quad (N\to\infty) \tag{1.5}$$

for all finite intervals $I \subset \mathbb{R}$.

Theorem 4. *Let*

$$\nu := \inf_{(\gamma_i)\in\mathscr{S}} \limsup_i (\gamma_{i+1}-\gamma_i).$$

Then $\nu \in [3/2, 2]$.

The lower bound $\nu \geq 3/2$ shows that every sequence with Poissonian pair correlation contains infinitely many gaps that are 1.5 times larger than the average gap.[2] By [7], this applies for instance to the ordered sequence of values $x^2 + \sqrt{2}\,y^2$, $x, y \in \mathbb{N}$. The lower bound is obtained using a reasoning similar to the one which leads from Corollary 2 to Corollary 3.

The problem of determining ν precisely is interesting in its own right since it touches upon the questions of how one can construct a sequence having a preassigned pair correlation distribution, and for which pair correlation distributions such a construction can exist at all and for which it cannot. We will discuss this problem a bit further in the closing remarks of the introduction (Section 1.6).

Finally, there are interesting sequences whose pair correlation is Poissonian, but whose level spacing is not, for instance the (suitably normalized) sequence of fractional parts of $\sqrt{n}$ (cf. [5,6]), which has in fact gaps of unbounded length.

1.4. Outline of the proof of Theorem 1

Roughly speaking, establishing Theorem 1 amounts to evaluating asymptotically

$$\frac{1}{N}\int_{(\alpha,\beta)\in\mathcal{R}}\int_N^{2N}\left(\sum_{\substack{|u-\alpha m^2-mn-\beta n^2|\leq\eta \\ M\leq m,n\leq 2M}} 1\right)^3 \mathrm{d}u\,\mathrm{d}\alpha\,\mathrm{d}\beta$$

---

[2]During the course of the refereeing process, the number $3/2$ was improved to $3/2 + 10^{-9}$ in [1]. We also take the opportunity to mention related correlation results for fractional parts [12,15] that were established during the refereeing process.

as $N \to \infty$, with $\eta > 0$ fixed, $M^2 = N$ and $\mathcal{R}$ a fixed rectangle. Note that for $\alpha_2 \neq 0$ we have $\Lambda_i((\alpha_1, \alpha_2, \alpha_3)) = \Lambda_i((\alpha_1/\alpha_2, 1, \alpha_3/\alpha_2))$, so without loss of generality we can assume $\alpha_2 = 1$. Expanding the third power gives rise to three conditions $|u - \alpha m_i^2 - m_i n_i - \beta n_i^2| \leq \eta$ with $i = 1, 2, 3$ and $m_i, n_i \in [M, 2M]$ say. We detect each of them using a Fourier integral. Subsequently we execute the integrations in u, α, β and make the following change of variables

$$a_1 = n_1 - n_2, \quad a_2 = n_1 + n_2, \quad a_3 = m_1 - m_3, \quad a_4 = m_1 + m_3,$$
$$b_1 = m_1 - m_2, \quad b_2 = m_1 + m_2, \quad b_3 = n_1 - n_3, \quad b_4 = n_1 + n_3.$$

Here the new variables are constrained by $a_1 + a_2 = b_3 + b_4$ and $b_1 + b_2 = a_3 + a_4$. After this change of variables and repeated integration by parts in the resulting Fourier integral, we are led to the problem of understanding asymptotically an expression that roughly looks like

$$\sum_{\substack{M \leq a_i, b_i \leq 2M \\ |\Delta_1|, |\Delta_2| \ll \Delta + M^2 \\ a_1 + a_2 = b_3 + b_4 \\ b_1 + b_2 = a_3 + a_4}} \frac{1}{\Delta + M^2}, \tag{1.6}$$

where

$$\Delta = a_1 a_2 a_3 a_4 - b_1 b_2 b_3 b_4,$$
$$\Delta_1 = a_1 a_2 b_3 a_4 + a_1 a_2 a_3 b_4 - a_1 b_2 b_3 b_4 - b_1 a_2 b_3 b_4,$$
$$\Delta_2 = a_1 b_2 a_3 a_4 + b_1 a_2 a_3 a_4 - b_1 b_2 a_3 b_4 - b_1 b_2 b_3 a_4.$$

The expected main term for (1.6) is of size M^2. It is important to note that we are over-simplifying the situation here by assuming that a_i, b_i are of size M. In practice this is not necessarily the case, but we assume it here for the sake of simplicity.

In [14], Sarnak obtains the same expression as (1.6) but without the constraints $a_1 + a_2 = b_3 + b_4$ and $b_1 + b_2 = a_3 + a_4$. As a result his expression is asymptotically of size M^4 and the counting problem is easier since there are fewer constraints. This explains our earlier remark on our work being a "thin subset" version of Sarnak's work.

When $|\Delta| > M^{4-\delta}$ we can estimate (1.6) asymptotically by applying the Lipschitz principle (that is, the Euler–Maclaurin approximation), getting a main term of size M^2. It therefore remains to show that the contribution to (1.6) of the terms with $|\Delta| \leq M^{4-\delta}$ is negligible, that is, $\ll M^{2-\varepsilon}$ for some $\varepsilon > 0$.

For $|\Delta| \in [D, 2D]$ with $D < M^{4-\delta}$ we use a substitution trick of Sarnak to see that the condition $|\Delta_1| \ll D + M^2$ implies that $|kl| \ll D + M^2$ where

$$k := a_2 a_4 - b_2 b_4, \quad l := b_2 b_3 - a_2 a_3. \tag{1.7}$$

In Sarnak's case this alone is enough to conclude via an elementary argument. In our case his elementary argument is slightly insufficient, and we need to non-trivially exploit the new condition that $a_1 + a_2 = b_3 + b_4$ and $a_3 + a_4 = b_1 + b_2$.

We proceed by

(1) solving for a_4 in terms of a_2, a_3, b_2, k, l using the conditions $a_1 + a_2 = b_3 + b_4$ and $a_3 + a_4 = b_1 + b_2$, definition (1.7), and the condition $|\Delta| \in [D, 2D]$; we refer to this as the "geometric restriction" on a_4;

(2) expressing b_3 in terms of the variables a_2, a_3, b_2, l by requiring that $a_3 \equiv -\overline{a_2}l$ (mod b_2) and

$$(b_3 =) \; \frac{a_3 a_2 + l}{b_2} \approx M;$$

(3) expressing b_4 in terms of the variables a_2, a_4, b_2, k by requiring that $a_4 \equiv \overline{a_2}k$ (mod b_2) and

$$(b_4 =) \; \frac{a_2 a_4 - k}{b_2} \approx M.$$

This leads to an upper bound for (1.6) that is roughly of the form

$$\frac{1}{D + M^2} \sum_{a_2, b_2 \approx M} \sum_{k, l \approx M^2} \sum_{\substack{a_3 \equiv -\overline{a_2}l \, (\mathrm{mod} \, b_2) \\ (a_3 a_2 + l)/b_2 \approx M}} \sum_{\substack{a_4 \equiv \overline{a_2}k \, (\mathrm{mod} \, b_2) \\ (a_2 a_4 - k)/b_2 \approx M}} \Phi_{a_2, a_3, b_2, k, l}(a_4), \qquad (1.8)$$

where Φ is a smooth function capturing the "geometric restriction" on a_4 in terms of the variables a_2, a_3, b_2, k, l that we mentioned in (1) of the preceding list. The function Φ only needs to capture a barely non-trivial piece of this "geometric restriction". It is not necessary for Φ to be as precise as to allow one to exactly reconstruct a_4 in terms of the other variables a_2, a_3, b_2, k, l.

We now apply Poisson summation on the variables a_2, a_4. The diagonal term gives a negligible contribution thanks to the presence of Φ (if Φ were replaced by 1 we would have obtained a diagonal of size matching the main term M^2 and this would not allow us to win). The off-diagonal terms give rise to sums of Kloosterman sums and we win by applying the Weil bound. It is, however, important to design Φ very carefully so as to capture only a barely non-trivial piece of our condition (1), otherwise we encounter again a counting problem with highly cuspidal regions which is hard to analyze.

We close by recalling that this discussion assumed that the variables a_i, b_i are in generic position, that is, all of size about M. However, ranges in which a_i are smaller require different treatment. For instance in some ranges we continue after (1.8) by also applying Poisson summation in k (and even l). Moreover, in the non-generic ranges the diagonal gives an acceptable contribution and we do not need to introduce the function Φ. This is welcome, since the behavior of Φ is complicated so we are happy to avoid it whenever possible.

1.5. Outline of the proof of Theorem 4

The lower bound in Theorem 4 comes from noticing that if the gaps $\gamma_{i+1} - \gamma_i$ are uniformly bounded by, say $3/2 - \varepsilon$, then the distribution function of the gaps $\gamma_{i+1} - \gamma_i$ can grow at most linearly between $1/2 - \varepsilon$ and $3/2 - \varepsilon$ because the pair correlation function

is Poissonian. However, integrating by parts we see that there can be no such distribution function unless $\varepsilon = 0$, since the average spacing of the sequence is 1 by assumption.

The construction of the upper bound is more involved and proceeds as follows. In each interval $[i, i+1]$ we first select a point γ_i uniformly at random. The resulting sequence is such that almost surely

$$\frac{1}{N} \sum_{\substack{\gamma_i - \gamma_j \leq b \\ i > j}} 1 \rightarrow \begin{cases} b^2/2 & \text{if } b \in [0, 1], \\ b - 1/2 & \text{if } b \geq 1. \end{cases} \tag{1.9}$$

and the mean spacing of γ_i is 1 by construction.

We then make a deterministic correction to the above random construction by introducing a small amount of clusters of points. Since there are few clusters and they are far apart, they will not interact with our random construction. Specifically, for each $m < \sqrt{N}$ we insert into each interval $[m^2, m^2 + 1]$ a sequence of $\lceil \sqrt{2m} \rceil$ equally spaced points. For $b \geq 1$ each such interval contribute an additional $\frac{1}{2} \cdot \lceil \sqrt{2m} \rceil^2$ to (1.9), and therefore a total of $\frac{1}{2}$ when summed over all $m < \sqrt{N}$. On the other hand, for $b \leq 1$ each interval $[m^2, m^2 + 1]$ contributes an additional

$$\frac{1}{N} \cdot \left(\lceil \sqrt{2m} \rceil (1 - b) \cdot \lceil \sqrt{2m} \rceil b + \sum_{k < b \lceil \sqrt{2m} \rceil} k \right) \sim \frac{2m}{N} \cdot \left(b - \frac{b^2}{2} \right),$$

and summing over $m < \sqrt{N}$ adds exactly the missing $b - b^2/2$ to (1.9) when $b \leq 1$.

1.6. Closing remarks

We have not been able to construct a sequence with bounded gaps whose pair correlation and triple correlation are both Poissonian; it would be interesting to know if such a sequence exists, and which bound for the maximal gap size can be achieved. More generally, it is an interesting question to ask which distribution functions can occur for the distribution function of the spacings $\gamma_{i+1} - \gamma_i$ given that the pair correlation function of γ_i is Poissonian. Concerning the general problem of generating a sequence having preassigned pair correlation behavior, there exist several results in this direction when the sequence is given by a so-called random point process; see for example [10, 11] and the references given there. However, these results only cover the purely random case. As far as we know, the question which asymptotic distributions of the pair correlation can be realized by a deterministic sequence (γ_i) has never been studied. We hope that our investigations initiate further research in this direction.

On a somewhat unrelated theme, we note that if we attempt to compute the 4-correlations, then the first problem we face is that we have to estimate the number of solutions in $M \leq x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4 \leq 2M$ to

$$-M^4 \leq \det \begin{vmatrix} x_1^2 - x_2^2 & x_1^2 - x_3^2 & x_1^2 - x_4^2 \\ y_1^2 - y_2^2 & y_1^2 - y_3^2 & y_1^2 - y_4^2 \\ x_1 y_1 - x_2 y_2 & x_1 y_1 - x_3 y_3 & x_1 y_1 - x_4 y_4 \end{vmatrix} \leq M^4, \tag{1.10}$$

asymptotically as $M \to \infty$. At present we do not know how to accomplish this, since this corresponds to a very thin region. What complicates matters is that the variety in (1.10) is highly singular. If it were smooth then the methods of [9] might offer some hope.

Any attempt to estimate the 3-correlations in an L^2 sense runs into a similar problem. More precisely, we would have to asymptotically estimate as $M \to \infty$ the number of solutions in $M \le m_1, m_2, m_3, m'_1, m'_2, m'_3, n_1, n_2, n_3, n'_1, n'_2, n'_3 \le 2M$ to

$$
-D \le \det \begin{vmatrix}
m_1^2 - m_2^2 & m_1^2 - m_3^2 & m_2'^2 - m_1'^2 & m_3'^2 - m_1'^2 \\
n_1^2 - n_2^2 & n_1^2 - n_3^2 & n_3'^2 - n_1'^2 & n_3'^2 - n_1'^2 \\
m_1 n_1 - m_2 n_2 & m_1 n_1 - m_3 n_3 & 0 & 0 \\
0 & 0 & m_2' n_2' - m_1' n_1' & m_3' n_3' - m_1' n_1'
\end{vmatrix} \le D
$$

uniformly in $M^6 \le D \le M^8$ and the case $D = M^6$ appears to be of comparable difficulty to (1.10).

1.7. Notational conventions

For the rest of the paper, we will apply the following notational conventions. Boldface letters like $\mathbf{a}$ or $\mathbf{A}$ denote vectors with components $(a_1, \ldots, a_n)$ resp. $(A_1, \ldots, A_n)$ whose dimension n will always be clear from the context. Given a vector $\mathbf{v} = (v_1, \ldots, v_n)$ and a real $M \ge 0$, the notation $\mathbf{v} \ll M$ means that there exists an absolute constant $C > 0$ such that $|v_i| \le CM$ for all $i = 1, \ldots, n$. In Section 3 we will make use of the notation $X \preccurlyeq Y$ to mean $X \ll_\varepsilon Y M^\varepsilon$ for $\varepsilon > 0$ where the meaning of ε can change from line to line.

Given a Schwartz function h we define its Fourier transform as

$$
\hat{h}(\xi_1, \ldots, \xi_n) := \int_{\mathbb{R}^n} h(x_1, \ldots, x_n) e(-x_1 \xi_1 - \cdots - x_n \xi_n) \, d\xi_1 \ldots d\xi_n,
$$

where $e(x) := e^{2\pi i x}$. For two vectors $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ we denote by $\psi(\mathbf{v}, \mathbf{w})$ the function $\psi(v_1, \ldots, v_n, w_1, \ldots, w_m)$. Moreover for $\mathbf{v} = (v_1, \ldots, v_n)$ we write

$$
\sum_{\mathbf{v}} \psi(\mathbf{v}) := \sum_{v_1, \ldots, v_n \in \mathbb{Z}} \psi(v_1, \ldots, v_n), \quad \int_{\mathbb{R}^n} \psi(\mathbf{x}) \, d\mathbf{x} := \int_{\mathbb{R}^n} \psi(x_1, \ldots, x_n) \, dx_1 \ldots dx_n,
$$

and we will also sometimes write $d(x_1, \ldots, x_n)$ in place of $dx_1 \ldots dx_n$.

2. Proof of Theorem 1: Reduction to Diophantine problems

Since $T_3(\lambda \boldsymbol{\alpha}; I_1, I_2; N) = T_3(\boldsymbol{\alpha}; I_1, I_2; N)$ by (1.2) for all $\lambda > 0$, we find for $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and $R = I \times J \times K \subseteq \mathcal{D}$ with intervals $I, J, K \subseteq \mathbb{R}_{\ge 0}$ that

$$
\int_R T_3(\boldsymbol{\alpha}; I_1, I_2; N) \, d\mu_{\text{hyp}}(\boldsymbol{\alpha})
$$

$$
= \int_J \int_{I/\alpha_2} \int_{K/\alpha_2} T_3((\alpha_1, 1, \alpha_3); I_1, I_2; N) \, d_{\text{hyp}}(\alpha_1, 1, \alpha_3) \frac{d\alpha_2}{\alpha_2} \tag{2.1}
$$

where

$$d_{\text{hyp}}(\alpha_1, 1, \alpha_3) = \frac{d\alpha_1 \, d\alpha_3}{(4\alpha_1\alpha_3 - 1)^{3/2}} \tag{2.2}$$

and I/α_2, K/α_2 are the corresponding re-scaled intervals with the interpretation $I/\alpha_2 = \mathbb{R}_{\geq 0}$ for $\alpha_2 = 0$. Since $\{\boldsymbol{\alpha} \subseteq \mathcal{D} \mid \alpha_2 = 0\}$ is a null set, to prove Theorem 1 it is enough to obtain an asymptotic formula for

$$\int_R T_3(\boldsymbol{\alpha}; I_1, I_2; N) \, d_{\text{hyp}}\boldsymbol{\alpha}$$

with $R \subseteq \mathcal{D}_0 := \{(\alpha_1, \alpha_3) \in (0, \infty)^2 \mid 1 \leq \alpha_1 \leq \alpha_3\}$ a rectangle, $\boldsymbol{\alpha} = (\alpha_1, 1, \alpha_3)$ and $d_{\text{hyp}}(\boldsymbol{\alpha})$ defined as in (2.2). This is the purpose of Proposition 5 below. As before we write $q_{\boldsymbol{\alpha}}(m, n) = \alpha_1 m^2 + mn + \alpha_3 n^2$ for $\boldsymbol{\alpha} = (\alpha_1, 1, \alpha_3)$. We also define

$$D(\boldsymbol{\alpha}) = \frac{1}{\pi}\sqrt{4\alpha_1\alpha_3 - 1},$$

so that $\{q_{\boldsymbol{\alpha}}(m, n)/D(\boldsymbol{\alpha}) \mid m > 0 \text{ or } m = 0, n \geq 0\}$ is the sequence of $\Lambda_j(\boldsymbol{\alpha})$.

Proposition 5. *Let F, W_i $(i = 1, 2)$, and V be fixed smooth functions with compact support in $\mathcal{D}_0$, $(-\infty, \infty)$ and $(0, \infty)$ respectively. For $M \geq 1$ let*

$$\mathcal{T}(M) := \frac{1}{8}\int_{\mathcal{D}_0} F(\boldsymbol{\alpha}) \sum_{\mathbf{x},\mathbf{y}}^* W_1\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_2, y_2)}{D(\boldsymbol{\alpha})}\right)$$

$$\times W_2\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_3, y_3)}{D(\boldsymbol{\alpha})}\right) V\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1)}{M^2 D(\boldsymbol{\alpha})}\right) d_{\text{hyp}}\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha} = (\alpha_1, 1, \alpha_3)$, $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$, $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{Z}^3$ and $\sum^*$ indicates summation with the conditions

$$(x_1, y_1) \neq \pm(x_2, y_2), \quad (x_1, y_1) \neq \pm(x_3, y_3), \quad (x_2, y_2) \neq \pm(x_3, y_3). \tag{2.3}$$

Then

$$\mathcal{T}(M) = \text{vol}(F)\hat{W}_1(0)\hat{W}_2(0)\hat{V}(0)M^2 + O(M^{2-1/258}),$$

where

$$\text{vol}(F) = \int_{\mathcal{D}_0} F(\boldsymbol{\alpha}) \, d_{\text{hyp}}(\boldsymbol{\alpha}), \quad \boldsymbol{\alpha} = (\alpha_1, 1, \alpha_3).$$

The factor $\frac{1}{8}$ is included to compensate for the fact that we are summing over $\mathbf{x}$ and $\mathbf{y}$ of arbitrary signs, and it therefore takes care of the desymmetrization of the spectrum mentioned in the introduction. Theorem 1 is an immediate consequence of Proposition 5: choosing $M^2 = N$ and choosing F, W_i and V to be smooth approximations to the characteristic functions of R, I_i $(i = 1, 2)$ and $[2^{-k-1}, 2^{-k}]$ $(k = 0, 1, \ldots, K = 1/\varepsilon$, say) respectively, we obtain

$$\int_R \frac{1}{N}\left|\left\{(i, j, k) \,\middle|\, \begin{matrix} \Lambda_j(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha}) \in I_1, \ \Lambda_k(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha}) \in I_2 \\ |\{i, j, k\}| = 3, \ 1 \leq \Lambda_i(\boldsymbol{\alpha}) \leq N \end{matrix}\right\}\right| d_{\text{hyp}}\boldsymbol{\alpha}$$

$$\sim \mu_{\text{hyp}}(R)\mu(I_1)\mu(I_2).$$

as $N \to \infty$. Plugging this into (2.1) we obtain (1.4), noting that the condition $N \le \Lambda_i(\boldsymbol{\alpha}) \le 2N$ can be replaced by $N \le i < 2N$, since the cardinality of the symmetric difference of $\{i \mid N \le \Lambda_i(\boldsymbol{\alpha}) \le 2N\}$ and $[N, 2N)$ is $o(N)$ as $N \to \infty$. The rate of convergence is continuous in the endpoints of I, and hence uniform as long as I varies within a fixed finite interval J.

We start the *proof of Proposition* 5 by Fourier-inverting the weight functions W_1, W_2 and V. The support of V, W_1, W_2 implies automatically $\mathbf{x}, \mathbf{y} \ll M$. We remember this by inserting a smooth redundant weight function $\psi(\mathbf{x}/M, \mathbf{y}/M)$, where ψ is a suitable smooth function that is constantly 1 on some sufficiently large fixed box in $\mathbb{R}^6$ and constantly zero outside some slightly larger box. We obtain

$$\mathcal{T}(M) = \frac{1}{8} \int_{\mathcal{D}_0} F(\boldsymbol{\alpha}) \sum_{\mathbf{x}, \mathbf{y}}^* \psi\left(\frac{\mathbf{x}}{M}, \frac{\mathbf{y}}{M}\right)$$

$$\times \int_{\mathbb{R}^3} \hat{V}(z) \hat{W}_1(u) \hat{W}_2(v) e\left(u \frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_2, y_2)}{D(\boldsymbol{\alpha})}\right)$$

$$\times e\left(v \frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_3, y_3)}{D(\boldsymbol{\alpha})}\right) e\left(\frac{z q_{\boldsymbol{\alpha}}(x_1, y_1)}{M^2 D(\boldsymbol{\alpha})}\right) d(u, v, z) \frac{d\boldsymbol{\alpha}}{\pi^3 D(\boldsymbol{\alpha})^3}. \quad (2.4)$$

We change variables $z \leftarrow z/D(\boldsymbol{\alpha})$, $u \leftarrow u/D(\boldsymbol{\alpha})$, $v \leftarrow v/D(\boldsymbol{\alpha})$ and integrate over $\boldsymbol{\alpha} = (\alpha_1, 1, \alpha_3) \in \mathcal{D}_0 \subseteq \mathbb{R}^2$. This gives

$$\mathcal{T}(M) = \frac{1}{8} \sum_{\mathbf{x}, \mathbf{y}}^* \psi\left(\frac{\mathbf{x}}{M}, \frac{\mathbf{y}}{M}\right) \int_{\mathbb{R}^3} e\left(u(x_1 y_1 - x_2 y_2) + v(x_1 y_1 - x_3 y_3) + \frac{z}{M^2} x_1 y_1\right)$$

$$\times \hat{G}\left(u(x_1^2 - x_2^2) + v(x_1^2 - x_3^2) + \frac{z x_1^2}{M^2}, u(y_1^2 - y_2^2) + v(y_1^2 - y_3^2) + \frac{z y_1^2}{M^2}; z, u, v\right)$$

$$\times d(u, v, z) \quad (2.5)$$

where

$$G(\alpha_1, \alpha_3; z, u, v) = \pi^{-3} F(\boldsymbol{\alpha}) \hat{V}(z D(\boldsymbol{\alpha})) \hat{W}_1(u D(\boldsymbol{\alpha})) \hat{W}_2(v D(\boldsymbol{\alpha}))$$

and the Fourier transform $\hat{G}$ of G is taken with respect to the first two variables α_1, α_3. Note that G is a Schwartz-class function in all variables. Therefore so is $\hat{G}$, i.e.

$$\mathcal{D}\hat{G}(U, V; z, u, v) \ll_{A, \mathcal{D}} \left((1 + |U|)(1 + |V|)(1 + |z|)(1 + |u|)(1 + |v|)\right)^{-A} \quad (2.6)$$

for all $A > 0$ and any differential operator $\mathcal{D}$.

Next we make an invertible change of integer variables as follows: each 6-tuple $(x_1, y_1, x_2, y_2, x_3, y_3)$ satisfying (2.3) is mapped bijectively to an 8-tuple $(a_1, \ldots, a_4, b_1, \ldots, b_4)$ satisfying

$$a_1 \equiv a_2 \pmod{2}, \quad b_1 \equiv b_2 \pmod{2} \quad (2.7)$$

as well as

$$(a_1, a_2, a_3, a_4) \ne (b_3, b_4, b_1, b_2), \quad (a_1, a_2, a_3, a_4) \ne (b_4, b_3, b_2, b_1),$$
$$(a_1, b_1) \ne (0, 0), \quad (a_2, b_2) \ne (0, 0), \quad (a_3, b_3) \ne (0, 0), \quad (a_4, b_4) \ne (0, 0) \quad (2.8)$$

and

$$a_1 + a_2 = b_3 + b_4, \quad a_3 + a_4 = b_1 + b_2 \tag{2.9}$$

via

$$
\begin{aligned}
a_1 &= y_1 - y_2, \quad a_2 = y_1 + y_2, \quad a_3 = x_1 - x_3, \quad a_4 = x_1 + x_3, \\
b_1 &= x_1 - x_2, \quad b_2 = x_1 + x_2, \quad b_3 = y_1 - y_3, \quad b_4 = y_1 + y_3.
\end{aligned} \tag{2.10}
$$

Note that the conditions $a_3 \equiv a_4 \pmod 2$ and $b_3 \equiv b_4 \pmod 2$ follow automatically from (2.7) and (2.9). This gives

$$\mathcal{T}(M) = \frac{1}{8} {\sum_{\mathbf{a}, \mathbf{b}}}' \tilde{\psi}\left(\frac{\mathbf{a}}{M}, \frac{\mathbf{b}}{M}\right) \mathcal{H}(\mathbf{a}, \mathbf{b}), \tag{2.11}$$

where $\sum'$ indicates conditions (2.7)–(2.9), $\tilde{\psi}$ is a smooth function such that

$$\tilde{\psi}(\mathbf{a}, \mathbf{b}) = \psi\left(\frac{b_1 + b_2}{2}, \frac{b_2 - b_1}{2}, \frac{a_4 - a_3}{2}, \frac{a_1 + a_2}{2}, \frac{a_2 - a_1}{2}, \frac{b_4 - b_3}{2}\right) = \psi(\mathbf{x}, \mathbf{y})$$

and

$$
\begin{aligned}
\mathcal{H}(\mathbf{a}, \mathbf{b}) := \int_{\mathbb{R}^3} & e\left(\frac{u}{2}(a_1 b_2 + a_2 b_1) + \frac{v}{2}(a_3 b_4 + a_4 b_3) + \frac{z}{4M^2}(a_1 + a_2)(b_1 + b_2)\right) \\
& \times \hat{G}\left(u b_1 b_2 + v a_3 a_4 + \frac{z}{4M^2}(b_1 + b_2)^2, u a_1 a_2 + v b_3 b_4 + \frac{z}{4M^2}(a_1 + a_2)^2; z, u, v\right) \\
& \hspace{10cm} \times \, \mathrm{d}(u, v, z).
\end{aligned}
$$

For notational convenience we introduce the following functions in the variables $a_1, \ldots, a_4, b_1, \ldots, b_4$. Put

$$
\begin{aligned}
P &= \max(|a_1 a_2|, |a_3 a_4|, |b_1 b_2|, |b_3 b_4|), \quad \Delta = a_1 a_2 a_3 a_4 - b_1 b_2 b_3 b_4, \\
\Delta_1 &= a_1 a_2 b_3 a_4 + a_1 a_2 a_3 b_4 - a_1 b_2 b_3 b_4 - b_1 a_2 b_3 b_4, \\
\Delta_2 &= a_1 b_2 a_3 a_4 + b_1 a_2 a_3 a_4 - b_1 b_2 a_3 b_4 - b_1 b_2 b_3 a_4.
\end{aligned}
$$

Note that (2.8) and (2.9) imply $P \neq 0$.

If $\Delta \neq 0$, we can change variables

$$u = \frac{a_3 a_4 V - b_3 b_4 U}{\Delta}, \quad v = \frac{a_1 a_2 U - b_1 b_2 V}{\Delta}$$

to see that $\mathcal{H}(\mathbf{a}, \mathbf{b})$ equals

$$
\begin{aligned}
\frac{1}{|\Delta|} \int_{\mathbb{R}^3} & \hat{G}\left(U + \frac{z(b_1 + b_2)^2}{4M^2}, V + \frac{z(a_1 + a_2)^2}{4M^2}; z, \frac{a_3 a_4 V - b_3 b_4 U}{\Delta}, \frac{a_1 a_2 U - b_1 b_2 V}{\Delta}\right) \\
& \times e\left(\frac{U \Delta_1}{2\Delta} + \frac{V \Delta_2}{2\Delta} + \frac{z}{4M^2}(a_1 + a_2)(b_1 + b_2)\right) \mathrm{d}(U, V, z). \tag{2.12}
\end{aligned}
$$

By repeated partial integration in U and V we conclude from (2.6) that

$$\mathcal{H}(\mathbf{a}, \mathbf{b}) \ll_A \frac{1}{|\Delta|} \min\left(1, \frac{|\Delta|}{P}\right)\left(\left(1 + \frac{|\Delta_1|}{|\Delta| + P}\right)\left(1 + \frac{|\Delta_2|}{|\Delta| + P}\right)\right)^{-A}$$

$$\ll \frac{1}{|\Delta| + P}\left(\left(1 + \frac{|\Delta_1|}{|\Delta| + P}\right)\left(1 + \frac{|\Delta_2|}{|\Delta| + P}\right)\right)^{-A} \quad (2.13)$$

for any $A \geq 0$ and all $\mathbf{a}, \mathbf{b} \ll M$. The final bound remains true for $\Delta = 0$, which we show now. Suppose without loss of generality that $P = |a_1 a_2| \neq 0$ (the other cases are similar). We observe that $\Delta = 0$ implies $b_1 b_2 \Delta_1 = -a_1 a_2 \Delta_2$, in particular $|\Delta_2| \leq |\Delta_1|$. In this case we change variables simply by

$$ub_1 b_2 + v a_3 a_4 = \frac{b_1 b_2}{a_1 a_2} U, \quad u a_1 a_2 + v b_3 b_4 = U$$

and recast $\mathcal{H}(\mathbf{a}, \mathbf{b})$ as

$$\frac{1}{|a_1 a_2|} \int_{\mathbb{R}^3} \widehat{G}\left(\frac{b_1 b_2}{a_1 a_2} U + \frac{z(b_1 + b_2)^2}{4M^2}, U + \frac{z(a_1 + a_2)^2}{4M^2}; z, \frac{U - b_3 b_4 v}{a_1 a_2}, v\right)$$

$$\times e\left(\frac{U(b_1 a_2 + b_2 a_1)}{2a_1 a_2} + \frac{v \Delta_1}{2a_1 a_2} + \frac{z}{4M^2}(a_1 + a_2)(b_1 + b_2)\right) \mathrm{d}(U, v, z).$$

Now integration by parts with respect to v confirms (2.13) again.

In order to evaluate asymptotically (2.11), we would like to apply a kind of Lipschitz principle and replace the sum over $\mathbf{a}, \mathbf{b}$ by an integral. This is not directly possible, because $\mathcal{H}(\mathbf{a}, \mathbf{b})$ is quite oscillatory if Δ is small. Using Diophantine techniques, we will establish in the next section the following crucial result, which basically tells us that the contribution of small values of Δ can be absorbed in an error term.

Proposition 6. *Let $0 < \delta \leq 1/2$. For any $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that*

$$\sideset{}{'}\sum_{\substack{\mathbf{a}, \mathbf{b} \ll M \\ |\Delta| \leq M^{4-\delta} \\ \Delta_1, \Delta_2 \ll (|\Delta| + P)^{1+\varepsilon'}}} \frac{1}{|\Delta| + P} \ll M^{2 - \delta/128 + \varepsilon}.$$

Let ϕ be a smooth function with support on $[1/2, \infty]$ that is 1 on $[1, \infty]$, and write

$$\Phi(\mathbf{a}, \mathbf{b}) := \tilde{\psi}\left(\frac{\mathbf{a}}{M}, \frac{\mathbf{b}}{M}\right) \phi\left(\frac{|\Delta|}{M^{4-\delta}}\right). \quad (2.14)$$

From (2.11), (2.13) and Proposition 6 we obtain

$$\mathcal{T}(M) = \frac{1}{8} \sideset{}{'}\sum_{\mathbf{a}, \mathbf{b}} \Phi(\mathbf{a}, \mathbf{b}) \mathcal{H}(\mathbf{a}, \mathbf{b}) + O(M^{2 - \delta/129}). \quad (2.15)$$

Note that condition (2.8) is now automatic since it is only violated in the case $\Delta = 0$ (which is not in the support of Φ). On the support of Φ we deduce from (2.6) and (2.12) that

$$\|\nabla(\Phi(\mathbf{a}, \mathbf{b})\mathcal{H}(\mathbf{a}, \mathbf{b}))\| \ll \frac{1}{|\Delta|} \cdot \frac{M^3}{|\Delta|} \ll M^{-5+2\delta}. \tag{2.16}$$

Using (2.9), we interpret the main term in (2.15) as a sum over $a_1, a_2, a_3, b_1, b_2, b_3$ subject only to (2.7), and in the following we apply the same interpretation for all functions in $\mathbf{a}, \mathbf{b}$, such as Φ, $\mathcal{H}$, P, Δ, Δ_1 and Δ_2. By a standard application of the Euler–Maclaurin summation formula, using (2.16), we see that

$$\frac{1}{8} \sideset{}{'}\sum_{\mathbf{a}, \mathbf{b}} \Phi(\mathbf{a}, \mathbf{b})\mathcal{H}(\mathbf{a}, \mathbf{b}) = \frac{1}{32} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \Phi(\mathbf{a}, \mathbf{b})\mathcal{H}(\mathbf{a}, \mathbf{b}) \, \mathrm{d}(a_1, a_2, a_3) \, \mathrm{d}(b_1, b_2, b_3)$$
$$+ O(M^{6-5+2\delta}).$$

The following continuous analogue of Proposition 6 allows us to remove the cut-off function ϕ in the definition (2.14).

Lemma 1. *Let $\delta \leq 1/2$. For any $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that*

$$\int \cdots \int_{\substack{\mathbf{a}, \mathbf{b} \ll M \\ |\Delta| \leq M^{4-\delta} \\ \Delta_1, \Delta_2 \ll (|\Delta|+P)^{1+\varepsilon'}}} \frac{1}{|\Delta| + P} \, \mathrm{d}(a_1, a_2, a_3) \, \mathrm{d}(b_1, b_2, b_3) \ll M^{2-\delta/4}.$$

We will prove this in the final section. Choosing $\delta = 1/2$, we have now arrived at

$$\mathcal{T}(M) = \frac{1}{32} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \tilde{\psi}\left(\frac{\mathbf{a}}{M}, \frac{\mathbf{b}}{M}\right)\mathcal{H}(\mathbf{a}, \mathbf{b}) \, \mathrm{d}(a_1, a_2, a_3) \, \mathrm{d}(b_1, b_2, b_3) + O(M^{2-1/258}). \tag{2.17}$$

At this point we invert the change of variables $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{a}, \mathbf{b})$ in (2.10), undo the integration (2.5) over α_1, α_3 and revert the Fourier inversions (2.4). This shows that

$$\frac{1}{32} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \tilde{\psi}\left(\frac{\mathbf{a}}{M}, \frac{\mathbf{b}}{M}\right)\mathcal{H}(\mathbf{a}, \mathbf{b}) \, \mathrm{d}(a_1, a_2, a_3) \, \mathrm{d}(b_1, b_2, b_3) = \mathcal{I}(M), \tag{2.18}$$

where

$$\mathcal{I}(M) := \frac{1}{8} \int_{\mathcal{D}} F(\boldsymbol{\alpha}) \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \psi\left(\frac{\mathbf{x}}{M}, \frac{\mathbf{y}}{M}\right) W_1\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_2, y_2)}{D(\boldsymbol{\alpha})}\right)$$
$$\times W_2\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1) - q_{\boldsymbol{\alpha}}(x_3, y_3)}{D(\boldsymbol{\alpha})}\right) V\left(\frac{q_{\boldsymbol{\alpha}}(x_1, y_1)}{M^2 D(\boldsymbol{\alpha})}\right) \mathrm{d}\mathbf{x} \, \mathrm{d}\mathbf{y} \, \mathrm{d}_{\mathrm{hyp}}\boldsymbol{\alpha}$$

is the continuous analogue of $\mathcal{T}(M)$. Here we can drop the function $\psi(\mathbf{x}/M, \mathbf{y}/M)$ because it is redundant. By a change of variables

$$x_j \leftarrow \left(x_j + \frac{y_j}{2\alpha_1}\right) \frac{\alpha_1^{1/2}}{(4\alpha_1\alpha_2 - 1)^{1/4}}, \quad y_j \leftarrow y_j \frac{(4\alpha_1\alpha_2 - 1)^{1/4}}{(4\alpha_1)^{1/2}}$$

the arguments of V, W_1, W_2 become independent of α, and we obtain

$$\mathcal{I}(M) = \text{vol}(F) \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} V\left(\frac{(x_1^2 + y_1^2)\pi}{M^2}\right) W_1((x_1^2 + y_1^2 - x_2^2 - y_2^2)\pi)$$
$$\times W_2((x_1^2 + y_1^2 - x_3^2 - y_3^2)\pi)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}.$$

Changing to polar coordinates, we get

$$\mathcal{I}(M) = (2\pi)^3 \text{vol}(F) \int_{[0,\infty)^3} V\left(\frac{r_1^2 \pi}{M^2}\right) W_1((r_1^2 - r_2^2)\pi) W_2((r_1^2 - r_3^2)\pi) r_1 r_2 r_3 \, \mathrm{d}\mathbf{r}.$$

Next changing $r_j \leftarrow r_j^2 \pi$, this simplifies to

$$\mathcal{I}(M) = \text{vol}(F) \int_{[0,\infty)^3} V\left(\frac{r_1}{M^2}\right) W_1(r_1 - r_2) W_2(r_1 - r_3)\, \mathrm{d}\mathbf{r}$$
$$= \text{vol}(F) M^2 \int_0^\infty V(r_1) \int_{-\infty}^{Mr_1} \int_{-\infty}^{Mr_1} W_1(r_2) W_2(r_3)\, \mathrm{d}\mathbf{r},$$

and for M sufficiently large this equals $\text{vol}(F)\hat{V}(0)\hat{W}_1(0)\hat{W}_2(0)M^2$. Combining this with (2.17) and (2.18), we complete the proof of Proposition 5.

3. Proof of Theorem 1: Diophantine analysis

In this section we prove Proposition 6. For notational convenience we introduce the notation $X \preccurlyeq Y$ to mean $X \ll_\varepsilon Y M^\varepsilon$ for $\varepsilon > 0$, where the meaning of ε can change from line to line. Next we observe that conditions (2.7)–(2.9) as well as $|\Delta|$, P and the set $\{|\Delta_1|, |\Delta_2|\}$ are invariant under the following symmetries:

- interchanging indices 1 and 2;
- interchanging indices 3 and 4;
- interchanging indices 1, 2 with 3, 4;
- interchanging a with b.

These four involutions generate a 2-subgroup of S_8 of order 16. In particular, without loss of generality we can and will assume that

$$\min(|a_3|, |a_4|, |b_3|, |b_4|) \leq \min(|a_1|, |a_2|, |b_1|, |b_2|),$$
$$\max(|a_1|, |a_2|, |b_1|, |b_2|) = |a_1|, \tag{3.1}$$
$$|b_2 b_3 - a_2 a_3| \geq |a_2 a_4 - b_2 b_4|.$$

Now we put all variables into dyadic intervals and suppose that $A_1 \leq |a_1| \leq 2A_1, \ldots,$ $A_4 \leq |a_4| \leq 2A_4$, $B_1 \leq |b_1| \leq 2B_1, \ldots, B_4 \leq |b_4| \leq 2B_4$ with $0 \leq A_1, \ldots, B_4 \ll M$. By (3.1) we have

$$\min(A_3, A_4, B_3, B_4) \ll \min(A_1, A_2, B_1, B_2), \quad \max(A_1, A_2, B_1, B_2) = A_1. \tag{3.2}$$

We write $P_0 := \max(A_1 A_2, B_1 B_2, A_3 A_4, B_3 B_4) \ll M^2$. In addition we assume that

$$D \leq |\Delta| \leq 2D, \quad 0 \leq D \leq M^{4-\delta}. \tag{3.3}$$

Let $\mathcal{N}(\mathbf{A}, \mathbf{B}, D)$ denote the number of 8-tuples $(a_1, \ldots, a_4, b_1, \ldots, b_4)$ subject to these size constraints and

$$\Delta_1, \Delta_2 \preccurlyeq D + P_0 \tag{3.4}$$

as well as conditions (2.7)–(2.9).

Using (2.9) and a standard divisor bound, we have the trivial bound

$$\mathcal{N}(\mathbf{A}, \mathbf{B}, D) \preccurlyeq P_0^2 \min(A_3, A_4) \min(B_3, B_4) \leq P_0^3. \tag{3.5}$$

First we dispose of two easy degenerate cases.

3.1. Degenerate case I: $a_1 a_2 a_3 a_4 b_1 b_2 b_3 b_4 = 0$

Let us first assume that one of the variables, say a_1, equals 0, but none of the b-variables vanishes. By a divisor argument we can choose the b-variables in $\preccurlyeq D$ ways, and then by (2.9) we have A_3 (say) choices for a_2, a_3, a_4. This gives a total number of

$$\preccurlyeq D A_3 \ll (D + P_0) M \tag{3.6}$$

choices.

Next, if any three of the variables $a_1, \ldots, a_4, b_1, \ldots, b_4$ vanish, then by (2.9) it is easy to see that we have at most

$$\ll P_0^{1/2} M^2 \ll (D + P_0)^{1/2} M^2 \tag{3.7}$$

choices for the remaining ones.

Up to symmetry, the only remaining case is that a_1 and exactly one of b_2, b_3, b_4 vanish (recall (2.8)). If $b_2 = 0$, then (3.4) gives $a_3 a_4 a_2 b_1 \preccurlyeq P_0 + D$ in non-zero variables. There are $\preccurlyeq P_0 + D$ choices by the usual divisor argument, and after choosing b_3, b_4 is determined by (2.9), which matches the contribution in (3.6). If $b_3 = 0$ (the case $b_4 = 0$ is similar), then (3.4) and the first equations in (2.9) give

$$P_0 + D \succcurlyeq a_2 a_3 a_4 b_1 - a_3 b_1 b_2 b_4 = a_2 a_3 b_1 (a_4 - b_2).$$

If $a_4 \neq b_2$, then we fix $a_2, a_3, b_1, a_4 - b_2$ in $\preccurlyeq D + P_0$ ways as well as a_4 in $\ll M$ ways, then the rest is determined by (2.9), so that we end up with a contribution as in (3.6). On the other hand, if $a_4 = b_2$, then everything is determined from a_2, a_3, a_4, and we obtain $P_0^{1/2} M^2 \ll (D + P_0)^{1/2} M^2$ solutions as in (3.7).

We summarize that the number $\mathcal{N}_0(\mathbf{A}, \mathbf{B}, D)$ of tuples where one of the variables vanishes is at most

$$\mathcal{N}_0(\mathbf{A}, \mathbf{B}, D) \preccurlyeq (D + P_0) M + (D + P_0)^{1/2} M^2. \tag{3.8}$$

From now on we focus on $\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D)$ where $a_1 a_2 a_3 a_4 b_1 b_2 b_3 b_4 \neq 0$.

3.2. Degenerate case II: $a_2a_4 = b_2b_4$

In this subsection we treat another degenerate case, namely the case

$$a_2a_4 = b_2b_4.$$

We call this contribution $\mathcal{N}_{*,0}(\mathbf{A}, \mathbf{B}, D)$. Fixing b_2, b_4 (non-zero), the values for a_2, a_4 are prescribed up to a divisor function. Moreover, using (2.9) we have

$$\frac{D}{b_2b_4} \gg d := a_1a_3 - b_1b_3 = a_1(b_2 - a_4) + a_3(b_4 - a_2) - (a_4 - b_2)(a_2 - b_4).$$

Fix a value for d and suppose initially $d \neq 0$. Then automatically $(b_2 - a_4, b_4 - a_2) \neq (0, 0)$. Choosing a suitable one of a_1 or a_3, the other one is determined, and then also b_1, b_3 from (2.9). So the total number of choices is

$$\leqslant B_2 B_4 \frac{D}{B_2 B_4}(A_1 + A_3) \ll DM \ll (D + P_0)M.$$

Now suppose that $d = 0$, so that $a_1a_3 = b_1b_3$. Put $r = (a_1, b_1)$ and $t = a_1/r$, $s = b_1/r$. Then $(s, t) = 1$ and $a_1 = rt$, $b_1 = rs$. Since $(s, t) = 1$, we must have $s \mid a_3$, say $a_3 = su$. Then $b_3 = tu$. Similarly, we can parametrize the equation $b_2b_4 = a_2a_4$ as $b_2 = xy, b_4 = zw, a_2 = xz, a_4 = yw$ with $(y, z) = 1$. Then (2.9) becomes

$$rt + xz = tu + zw, \quad su + yw = rs + xy,$$

in other words

$$t(r - u) = z(w - x), \quad s(r - u) = y(w - x).$$

We have $r = u$ if and only if $w = x$, which is equivalent to the opposite of (2.8), so that this case is excluded. On the other hand, if $r - u \neq 0 \neq w - x$, then by coprimality $s = t$, $z = y$, so that

$$a_1 = b_1, \quad a_2 = b_2, \quad a_3 = b_3, \quad a_4 = b_4.$$

Invoking also (2.9), we find in this case at most

$$\ll A_1 A_2 A_3 \ll P_0 A_3 \ll (D + P_0)M$$

solutions. We conclude that

$$\mathcal{N}_{*,0}(\mathbf{A}, \mathbf{B}, D) \leqslant (D + P_0)M. \tag{3.9}$$

From now on we assume $a_2a_4 \neq b_2b_4$.

3.3. Substituting

Solving the first inequality in (3.3) for b_1 and substituting into the bound for Δ_1 in (3.4), we obtain

$$(a_2a_4 - b_2b_4)(b_2b_3 - a_2a_3) \leqslant \left|\frac{b_2}{a_1}\right|(D + P_0) + D\left|\frac{a_2}{a_1}\right| \ll D + P_0$$

by (3.2). We write

$$k = a_2 a_4 - b_2 b_4, \quad l = b_2 b_3 - a_2 a_3. \tag{3.10}$$

By (3.1) we have $|k| \le |l|$, and by the assumption from the previous subsection we have $kl \ne 0$. We assume $K \le |k| \le 2K$, $L \le |l| \le 2L$ with

$$L \ge K \ge 1, \quad KL \preccurlyeq D + P_0. \tag{3.11}$$

Note that a_1, b_1 are determined from the other variables by (2.9). We call the corresponding contribution $\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L)$.

Let $d = (a_2, b_2)$ and $a_2 = da_2'$, $b_2 = db_2'$. Then we must have $k = dk'$, $l = dl'$ and $a_4 \equiv \overline{a_2'} k' \pmod{|b_2'|}$, $a_3 \equiv -\overline{a_2'} l' \pmod{|b_2'|}$.

For the remaining part of the argument, ρ is a parameter such that

$$0 < \rho < \delta/4 \tag{3.12}$$

with δ as in Proposition 6, and η is a parameter with

$$0 < \eta < 1/6.$$

For later purposes we take the opportunity to dispose of another somewhat degenerate case at this point. Suppose that

$$B_4 \le M^\eta, \quad K \asymp A_2 A_4 \quad \text{or} \quad B_3 \le M^\eta, \quad L \asymp A_2 A_3. \tag{3.13}$$

The second set of conditions is analogous, so we focus on the first. In this case we fix $d, a_2', b_2', a_4, b_4, l', a_3$, which then determines k and b_3. Thus we see by elementary means that

$\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L)$

$$\preccurlyeq \sum_d \sum_{a_2', b_2'} \sum_{a_4, b_4} \sum_{l'} \sum_{a_3 \equiv -\overline{a_2'} l' \pmod{|b_2'|}} 1 \ll A_2 B_2 A_4 B_4 \frac{D + P_0}{K} \left(\frac{A_3}{B_2} + 1 \right)$$

$$\ll (A_3 + B_2) B_4 (D + P_0) \ll (D + P_0) M^{1+\eta}. \tag{3.14}$$

under condition (3.13). This is sufficient for our purposes, and from now we therefore assume that (3.13) does not hold.

Let W be a smooth non-negative function with support in $[1/2, 3]$ that is 1 on $[1, 2]$. We now fix $d, a_2', a_3', k', l', a_3, a_4$; then b_3, b_4 are determined, and we obtain

$\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L)$

$$\le \sum_d \sum_{B_2 \le |b_2| \le 2B_2 \, (a_2', b_2') = 1} \sum W \left(\frac{a_2'}{A_2/d} \right) \sum_{k', l'} W \left(\frac{k'}{K/d} \right) W \left(\frac{l'}{L/d} \right)$$

$$\times \sum_{a_3 \equiv -\overline{a_2'} l' \pmod{|b_2'|}} \Phi_{a_2', b_2', -l'; A_3, B_3}(a_3) \sum_{a_4 \equiv \overline{a_2'} k' \pmod{|b_2'|}} \Phi_{a_2', b_2', k'; A_4, B_4}(a_4), \tag{3.15}$$

where

$$\Phi_{a_2', b_2', n; A, B}(x) = W \left(\frac{|x|}{A} \right) W \left(\frac{|a_2' x - n|}{B |b_2'|} \right). \tag{3.16}$$

We have

$$\partial_x^{j_1} \partial_a^{j_2} \partial_n^{j_3} \Phi_{a,b_2',n;A,B}(x) \ll_{j_1,j_2,j_3} \min\left(A, \frac{BB_2}{A_2}\right)^{-j_1} \left(\frac{B|b_2'|}{A}\right)^{-j_2} (B|b_2'|)^{-j_3} \quad (3.17)$$

for $j_1, j_2, j_3 \geq 0$, which we apply with $(n, A, B) = (k', A_3, B_3)$ or $(-l', A_4, B_4)$.

In this analysis we could have interchanged the roles of a_2, a_3, a_4 and b_2, b_3, b_4 (leaving $|k|$ and $|l|$ invariant), so that without loss of generality we may and do assume

$$A_2 \geq B_2. \qquad (3.18)$$

We now distinguish two principal cases which require different treatment of the right hand side of (3.15).

3.4. Case 1

Let us assume that

$$\min(A_1, \ldots, A_4, B_1, \ldots, B_4) \leq M^{1-\rho} \quad \text{or} \quad KL \leq (D + P_0)^{1-\rho}. \qquad (3.19)$$

3.4.1. Application of Poisson summation. Here we start with an application of Poisson summation in (3.15) with respect to the a_3, a_4-sum, getting

$$\sum_d \sum_{(a_2',b_2')=1} W\left(\frac{a_2'}{A_2/d}\right) \sum_{k',l'} W\left(\frac{k'}{K/d}\right) W\left(\frac{l'}{L/d}\right) \sum_{h_4,h_3\in\mathbb{Z}} e\left(\frac{\overline{a_2'}k'h_4 - \overline{a_2'}l'h_3}{|b_2'|}\right)$$

$$\times \frac{1}{|b_2'|^2} \int_{\mathbb{R}^2} \Phi_{a_2',b_2',k';A_4,B_4}(x)\Phi_{a_2',b_2',-l';A_3,B_3}(y)e\left(\frac{-xh_4 - yh_3}{|b_2'|}\right) d(x, y). \qquad (3.20)$$

By partial integration and (3.17) we can truncate the h_j-sum ($j = 3, 4$) at

$$h_j \preccurlyeq \frac{B_2/d}{\min(A_j, B_j B_2/A_2)} = \frac{1}{d}\left(\frac{B_2}{A_j} + \frac{A_2}{B_j}\right) =: \frac{1}{d}H_j,$$

say, at the cost of a negligible error. For future reference we note that the double integral in (3.20) is trivially bounded by

$$\min\left(A_4, \frac{B_4 B_2}{A_2}\right)\min\left(A_3, \frac{B_3 B_2}{A_2}\right).$$

3.4.2. The diagonal terms. The total contribution of the $h_3 = h_4 = 0$ term, say $\mathcal{N}_{*,\mathrm{diag}}(\mathbf{A}, \mathbf{B}, D, K, L)$, is

$$\ll \sum_d \sum_{(a_2',b_2')=1} \sum_{k',l'} \frac{1}{|b_2'|^2} \int_{\mathbb{R}^2} \Phi_{a_2',b_2',k';A_4,B_4}(x)\Phi_{a_2',b_2',l';A_3,B_3}(y) d(x, y)$$

$$\preccurlyeq A_2 B_2 KL \min\left(\frac{A_4}{B_2}, \frac{B_4}{A_2}\right)\min\left(\frac{A_3}{B_2}, \frac{B_3}{A_2}\right) \ll \min(A_4 B_3, A_3 B_4)KL$$

so that

$$\mathcal{N}_{*,\mathrm{diag}}(\mathbf{A}, \mathbf{B}, D, K, L) \ll KLM \min(A_1, \ldots, A_4, B_1, \ldots, B_4)$$

by the first inequality in (3.2). By our current assumption (3.19) we obtain

$$\mathcal{N}_{*,\mathrm{diag}}(\mathbf{A}, \mathbf{B}, D, K, L) \ll (D + P_0)^{1-\rho} M^2 + (D + P_0) M^{2-\rho}. \qquad (3.21)$$

3.4.3. The off-diagonal terms. Let us now assume $(h_3, h_4) \neq (0, 0)$ in (3.20) and call this term $\mathcal{N}_{*,\mathrm{off}}(\mathbf{A}, \mathbf{B}, D, K, L)$. We now apply Poisson summation also to a_2'. This gives

$$\sum_d \sum_{b_2'} \sum_{k',l'} \frac{1}{|b_2'|^3} W\left(\frac{k'}{K/d}\right) W\left(\frac{l'}{L/d}\right) \sum_{\substack{h_3 \preccurlyeq H_3, h_4 \preccurlyeq H_4 \\ (h_3, h_4) \neq 0}} \sum_{r_2} S(k'h_4 - l'h_3, r_2, |b_2'|)$$

$$\times \int_{\mathbb{R}^3} W\left(\frac{zd}{A_2}\right) \Phi_{z,b_2',k';A_4,B_4}(x) \Phi_{z,b_2',-l';A_3,B_3}(y) e\left(\frac{-yh_3 - xh_4 - zr_2}{|b_2'|}\right) \mathrm{d}(x, y, z). \qquad (3.22)$$

Here S denotes the usual Kloosterman sum

$$S(a, b, c) = \sum_{\substack{d \pmod c \\ (d,c)=1}} e\left(\frac{ad + b\bar{d}}{c}\right),$$

which satisfies the Weil bound $|S(a, b, c)| \le c^{1/2}(a, b, c)^{1/2}\tau(c)$ [13, 18] where τ is the divisor function.

By partial integration with respect to z and (3.17) we can truncate the r_2-sum at

$$r_2 \preccurlyeq R := \frac{B_2}{A_2} + \frac{A_4}{B_4} + \frac{A_3}{B_3}$$

at the cost of a negligible error. Again we distinguish several cases.

3.4.4. Off-diagonal case 1. Suppose that

$$B_2 B_4 \ge A_2 A_4 M^{-\eta} \quad \text{and} \quad B_2 B_3 \ge A_2 A_3 M^{-\eta}. \qquad (3.23)$$

In this case we apply the Weil bound in (3.22). We first treat the degenerate case where $r_2 = k'h_4 - l'h_3 = 0$ in which the Kloosterman sum is large (and in which case $h_3 h_4 \neq 0$). This contributes at most

$$\preccurlyeq \sum_d \frac{d}{B_2} \min\left(\frac{KH_4}{d^2}, \frac{LH_3}{d^2}\right) \frac{A_2}{d} \min\left(A_4, \frac{B_4 B_2}{A_2}\right) \min\left(A_3, \frac{B_3 B_2}{A_2}\right)$$

$$\preccurlyeq \frac{1}{B_2}\left(\left(\frac{B_2}{A_4} + \frac{A_2}{B_4}\right)\left(\frac{B_2}{A_3} + \frac{A_2}{B_3}\right)KL\right)^{1/2} A_2 \min\left(A_4, \frac{B_4 B_2}{A_2}\right) \min\left(A_3, \frac{B_3 B_2}{A_2}\right)$$

$$\le (KL)^{1/2}(B_2(B_3 B_4)^{1/2} + (A_2 A_3 B_2 B_4)^{1/2} + (A_2 A_4 B_2 B_3)^{1/2} + A_2(A_3 A_4)^{1/2})$$

$$\preccurlyeq (D + P_0)^{1/2} M^2.$$

The remaining terms contribute, by Weil's bound,

$$\preccurlyeq \sum_d \frac{d^{3/2}}{B_2^{3/2}} \frac{KL}{d^2} \left(\frac{H_3}{d} + \frac{H_4}{d} + \frac{H_3 H_4}{d^2} \right) (1+R) \frac{A_2}{d} \min\left(A_4, \frac{B_4 B_2}{A_2} \right) \min\left(A_3, \frac{B_3 B_2}{A_2} \right)$$

$$\ll \frac{KLA_2}{B_2^{3/2}} \left(\frac{B_2}{A_4} + \frac{A_2}{B_4} + \frac{B_2}{A_3} + \frac{A_2}{B_3} + \left(\frac{B_2}{A_4} + \frac{A_2}{B_4} \right)\left(\frac{B_2}{A_3} + \frac{A_2}{B_3} \right) \right)$$

$$\times \left(1 + \frac{B_2}{A_2} + \frac{A_4}{B_4} + \frac{A_3}{B_3} \right) \min\left(A_4, \frac{B_4 B_2}{A_2} \right) \min\left(A_3, \frac{B_3 B_2}{A_2} \right)$$

$$\ll KLB_2^{1/2} \left(A_4 + A_3 + A_2 + B_4 + B_3 + B_2 + \frac{B_4 A_3}{B_3} + \frac{A_2 A_3}{B_3} + \frac{A_4 B_3}{B_4} + \frac{A_2 A_4}{B_4} \right)$$

$$\preccurlyeq (D + P_0)\left(M^{3/2} + B_2^{1/2} \left(\frac{B_4 A_3}{B_3} + \frac{A_2 A_3}{B_3} + \frac{A_4 B_3}{B_4} + \frac{A_2 A_4}{B_4} \right) \right). \tag{3.24}$$

So far we have not used our assumption (3.18), but we use it now. Combining (3.18) with our current assumption (3.23), we must have $B_4 \geq A_4 M^{-\eta}$ and $B_3 \geq A_3 M^{-\eta}$, so that (3.24) is bounded by

$$\preccurlyeq (D + P_0) M^{3/2+\eta}(1 + B_2/A_2) \ll (D + P_0) M^{3/2+\eta},$$

which gives in total

$$\mathcal{N}_{*,\mathrm{off}}(\mathbf{A}, \mathbf{B}, D, K, L) \preccurlyeq (D + P_0) M^{3/2+\eta} + (D + P_0)^{1/2} M^2 \tag{3.25}$$

in our present case.

3.4.5. Off-diagonal case 2. Suppose now that

$$B_2 B_4 \leq A_2 A_4 M^{-\eta}, \quad \text{but} \quad B_2 B_3 \geq A_2 A_3 M^{-\eta}.$$

Then necessarily $A_2 A_4 \asymp K$, otherwise $k = a_2 a_4 - b_2 b_4$ has no solution. Since we are assuming that (3.13) does not hold, we have $B_4 \geq M^{\eta}$. We claim now that the present assumptions imply that only the terms $h_4 = 0$, $h_3 \neq 0$ contribute non-negligibly to (3.22). Indeed, applying Poisson summation to (3.22) with respect to k', we see that up to a negligible error the dual sum has length $\preccurlyeq B_4^{-1} + B_2/K \ll M^{-\eta}$, in other words, only the central Poisson term survives. Since

$$\sum_{k \,(\mathrm{mod}\,|b_2'|)} S(k'h_4 - l'h_3, r_2, |b_2'|) = 0 \tag{3.26}$$

unless $h_4 = 0$, we conclude that this forces $h_4 = 0$, so that $h_3 \neq 0$.

Having established our claim, we apply Weil's bound for $S(-h_3 l', r_2, |b_2'|)$ and bound $\mathcal{N}_{*,\mathrm{off}}(\mathbf{A}, \mathbf{B}, D, K, L)$ in (3.22) as above by

$$\preccurlyeq \sum_d \frac{KL/d^2}{B_2^{3/2}/d^{3/2}} \frac{H_3}{d} (1+R) \frac{A_2}{d} \min\left(A_4, \frac{B_4 B_2}{A_2} \right) \min\left(A_3, \frac{B_3 B_2}{A_2} \right)$$

$$\preccurlyeq (D + P_0)\left(M^{3/2} + B_2^{1/2} \frac{B_4 A_3}{B_3} \right)$$

$$\ll (D + P_0) M^{3/2+\eta}(1 + B_2/A_2) \ll (D + P_0) M^{3/2+\eta}. \tag{3.27}$$

3.4.6. Off-diagonal case 3. Of course, the dual situation $B_2 B_3 \leq A_2 A_3 M^{-\eta}$, but $B_2 B_4 \geq A_2 A_4 M^{-\eta}$ is handled similarly to the previous case.

3.4.7. Off-diagonal case 4. Finally, we treat the case

$$B_2 B_4 \leq A_2 A_4 M^{-\eta} \quad \text{and} \quad B_2 B_3 \leq A_2 A_3 M^{-\eta}.$$

Here we can apply Poisson in both k, l in (3.22). By the same argument as in (3.26) this forces $h_3 = h_4 = 0$, up to a negligible error, but this case is excluded in the off-diagonal contribution.

3.4.8. The final bound. Collecting (3.25) and (3.27) we have shown that

$$\mathcal{N}_{*,\mathrm{off}}(\mathbf{A}, \mathbf{B}, D, K, L) \preccurlyeq (D + P_0) M^{3/2+\eta} + (D + P_0)^{1/2} M^2.$$

Combining this with (3.21), we see that altogether

$$\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L) \preccurlyeq (D + P_0)^{1-\rho} M^2 + (D + P_0) M^{2-\rho} \tag{3.28}$$

in the present case (3.19).

3.5. Case 2

We now turn to the second case

$$\min(A_1, \ldots, A_4, B_1, \ldots, B_4) \geq M^{1-\rho}, \quad \text{and} \quad KL \geq (D + P_0)^{1-\rho}. \tag{3.29}$$

We could proceed in this case as in Case 1, but we will find that the diagonal term is too large. Therefore we need to introduce an additional constraint into our upper bound for $\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L)$. In this case we will obtain additional constraints for the variable a_4.

3.5.1. Creating an additional constraint on a_4. Solving (3.10) for b_3, b_4 and substituting this as well as (2.9) into the definition of Δ, one finds by brute force algebraic manipulation

$$\left(a_4 - \frac{1}{2} \left(b_2 + \frac{k}{a_2} \right) \right)^2 = H(a_2, b_2, a_3, k, l) - \frac{b_2}{a_2 l} \Delta, \tag{3.30}$$

where

$$H(a_2, b_2, a_3, k, l) = -\frac{a_3 b_2 k}{l} + \frac{a_3^2 k}{l} + \frac{b^2}{4} + \frac{a_3 k}{a_2} - \frac{b_2 k}{2a_2} + \frac{k^2}{4a_2^2}.$$

Solving instead (3.10) for a_3, a_4, we may exchange the roles of a_2, a_3, a_4 and b_2, b_3, b_4, so that again without loss of generality we may assume (3.18). Using the coarse bound $\sqrt{x + y} = \sqrt{x} + O(\sqrt{|y|})$, this implies

$$a_4 = G_\pm(a_2, b_2, a_3, k, l) + O(E^{1/2}), \quad E := D/L, \tag{3.31}$$

where $G_\pm(a_2, b_2, a_3, k, l) = \frac{1}{2}(b_2 + k/a_2) \pm \sqrt{H(a_2, b_2, a_3, k, l)}$. For future reference we note that

$$\partial_{a_2}^j H(a_2, b_2, a_3, k, l) \ll_j M^2 A_2^{-j}, \quad \partial_{a_3}^j H(a_2, b_2, a_3, k, l) \ll_j M^{2-j} \qquad (3.32)$$

for $j \geq 0$, where we have used the fact that (3.18) implies $K/A_2 \ll M$.

Notice that because of (3.30) and the bound $\left|\frac{b_2 \Delta}{a_2 \ell}\right| \ll E = D/L$, if we have $H(a_2, b_2, a_3, k, l) < -CD/L$ for a sufficiently large absolute constant $C > 0$ then there are no solutions for a_4. Hence we can attach a factor $W((a_4 - G_\pm(a_2, b_2, a_3, k, l))/E^{1/2})$ to the a_4-sum in (3.15). Unfortunately, $G_\pm$ may be highly oscillatory in a_2, a_3, which is problematic for the application of the Poisson summation formula in a moment. However, since we are facing a counting problem where all terms are positive we are free to "enlarge" E and $G_\pm$. By (3.29), (3.11), (3.3) and (3.12) we have

$$\frac{D}{L} \leq \frac{D}{(D + P_0)^{(1-\rho)/2}} \leq D^{1/2+\rho/2} \leq M^{(4-\delta)(1+\rho)/2} \leq M^2.$$

We choose a parameter $0 < \alpha < 1/6$ and write

$$E_0 := \left(\frac{D+1}{L}\right)^\alpha M^{2(1-\alpha)} \gg E \qquad (3.33)$$

and obtain

$$a_4 = \tilde{G}_\pm(a_2, b_2, a_3, k, l) + O(E_0^{1/2})$$

with

$$\tilde{G}_\pm(a_2, b_2, a_3, k, l) = \frac{1}{2}\left(b_2 + \frac{k}{a_2}\right) \pm \sqrt{H(a_2, b_2, a_3, k, l) + E_0}.$$

This gives

$$\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L) \leq \sum_d \sum_{(a_2', b_2')=1} W\left(\frac{a_2'}{A_2/d}\right) \sum_{k',l'} W\left(\frac{k'}{K/d}\right) W\left(\frac{l'}{L/d}\right)$$

$$\times \sum_{a_3 \equiv -\overline{a_2'}l' \pmod{|b_2'|}} \Phi_{a_2', b_2', -l'; A_3, B_3}(a_3) \sum_{a_4 \equiv \overline{a_2'}k' \pmod{|b_2'|}} \tilde{\Phi}_{d, a_2', b_2', a_3, k', l'}(a_4) \qquad (3.34)$$

with Φ as in (3.16) and

$$\tilde{\Phi}_{d, a_2', b_2', a_3, k', l'}(x) = W\left(\frac{x}{A_4}\right) W\left(\frac{a_2' x - k'}{B_4 b_2'}\right) \tilde{W}\left(\frac{x - \tilde{G}_\pm(a_2' d, b_2' d, a_3, k' d, l' d)}{E_0^{1/2}}\right)$$

for a suitable smooth weight function $\tilde{W}$. From (3.32) we obtain

$$\partial_{a_2}^j \tilde{G}_\pm(a_2, b_2, a_3, k, l) \ll_j \frac{M^2}{A_2^j E_0^{1/2}} + \frac{M^{2j}}{A_2^j E_0^{(2j-1)/2}},$$

$$\partial_{a_3}^j \tilde{G}_\pm(a_2, b_2, a_3, k, l) \ll_j \frac{M^{2-j}}{E_0^{1/2}} + \frac{M^j}{E_0^{(2j-1)/2}} \qquad (3.35)$$

for $j \geq 1$. We see now why it is useful to have E_0 not too small.

3.5.2. Application of Poisson summation. We now apply Poisson summation in a_3, a_4 to (3.34). This gives an expression very similar to (3.20) except that $\Phi_{a_2', b_2', k'; A_4, B_4}(x)$ is replaced with $\tilde{\Phi}_{d, a_2', b_2', y, k', l'}(x)$. More precisely, we get

$$
\sum_d \sum_{(a_2', b_2') = 1} W\left(\frac{a_2'}{A_2/d}\right) \sum_{k', l'} W\left(\frac{k'}{K/d}\right) W\left(\frac{l'}{L/d}\right) \sum_{h_4, h_3 \in \mathbb{Z}} e\left(\frac{\overline{a_2'}k'h_4 - \overline{a_2'}l'h_3}{|b_2'|}\right)
$$
$$
\times \frac{1}{|b_2'|^2} \int_{\mathbb{R}^2} \tilde{\Phi}_{d, a_2', b_2', y, k', l'}(x) \Phi_{a_2', b_2', -l'; A_3, B_3}(y) e\left(\frac{-xh_4 - yh_3}{|b_2'|}\right) d(x, y). \quad (3.36)
$$

In this case, it follows from (3.35) and (3.17) and integration by parts with respect to x and y in

$$
\int_{\mathbb{R}^2} \tilde{\Phi}_{d, a_2', b_2', y, k', l'}(x) \Phi_{a_2', b_2', -l'; A_3, B_3}(y) e\left(\frac{-xh_4 - yh_3}{|b_2'|}\right) d(x, y)
$$

that we can truncate the h_j-sum ($j = 3, 4$) at

$$
\begin{aligned}
h_3 &\preccurlyeq \frac{B_2/d}{\min(A_3, B_3 B_2/A_2, E_0/M)} = \frac{1}{d}\left(\frac{B_2}{A_3} + \frac{A_2}{B_3} + \frac{B_2 M}{E_0}\right) =: \frac{1}{d}\tilde{H}_3, \\
h_4 &\preccurlyeq \frac{B_2/d}{\min(A_4, B_4 B_2/A_2, E_0^{1/2})} = \frac{1}{d}\left(\frac{B_2}{A_4} + \frac{A_2}{B_4} + \frac{B_2}{E_0^{1/2}}\right) =: \frac{1}{d}\tilde{H}_4.
\end{aligned} \quad (3.37)
$$

As before, we denote by $\mathcal{N}_{*,\mathrm{diag}}(\mathbf{A}, \mathbf{B}, D, K, L)$ the contribution of the diagonal terms $(h_3, h_4) = (0, 0)$ and by $\mathcal{N}_{*,\mathrm{off}}(\mathbf{A}, \mathbf{B}, D, K, L)$ the contribution of the off-diagonal terms $(h_3, h_4) \neq (0, 0)$.

3.5.3. The diagonal terms. The introduction of $\tilde{\Phi}$ shortens the a_4-sum a bit and by the same argument as in the treatment of diagonal terms in the previous Case 1 together with (3.33) and (3.11), the central term $h_3 = h_4 = 0$ in (3.36) is bounded by

$$
\begin{aligned}
\mathcal{N}_{*,\mathrm{diag}}(\mathbf{A}, \mathbf{B}, D, K, L) &\ll A_2 B_2 \min\left(\frac{\sqrt{E_0}}{B_2}, \frac{A_4}{B_2}, \frac{B_4}{A_2}\right) \min\left(\frac{A_3}{B_2}, \frac{B_3}{A_2}\right) KL \\
&\leq M^{1-\alpha}(D+1)^{\alpha/2} B_3 K L^{1-\alpha/2} \preccurlyeq M^{2-\alpha}(D+1)^{\alpha/2}(D+P_0)^{1-\alpha/4} \\
&\leq M^{2-\alpha}(D+1)^{\alpha/4}(D+P_0). \quad (3.38)
\end{aligned}
$$

3.5.4. The off-diagonal terms. It remains to obtain an acceptable bound for the contribution of $(h_3, h_4) \neq (0, 0)$.

Because of condition (3.29) all variables have roughly the same size, so this case is much less delicate. As in (3.22) of Case 1 we apply in (3.36) Poisson summation in the a_2'-variable. By (3.35) the dual variable can be truncated at

$$
r_2 \preccurlyeq \tilde{R} := \frac{B_2}{A_2} + \frac{A_4}{B_4} + \frac{A_3}{B_3} + \frac{B_2 M^2}{A_2 E_0}.
$$

We now run the same computation as already done in Case 1. The terms with $r_2 = kh_4 - lh_3 = 0$ contribute

$$\ll \sum_d \frac{d}{B_2} \min\left(\frac{K\tilde{H}_4}{d^2}, \frac{L\tilde{H}_3}{d^2}\right) \frac{A_2}{d} \min\left(A_4, \frac{B_4 B_2}{A_2}\right) \min\left(A_3, \frac{B_3 B_2}{A_2}\right)$$

$$\preccurlyeq M^2 (KL)^{1/2} (H_3 H_4)^{1/2} \ll M^2 (D + P_0)^{1/2}\left(M^\rho + \frac{M^2}{E_0}\right)$$

by (3.18) and (3.29). Similarly, the remaining terms contribute

$$\preccurlyeq \sum_d \frac{d^{3/2}}{B_2^{3/2}} \frac{KL}{d^2}\left(\frac{\tilde{H}_3}{d} + \frac{\tilde{H}_4}{d} + \frac{\tilde{H}_3 \tilde{H}_4}{d^2}\right)(1 + \tilde{R})\frac{A_2}{d} \min\left(A_4, \frac{B_4 B_2}{A_2}\right) \min\left(A_3, \frac{B_3 B_2}{A_2}\right)$$

$$\preccurlyeq (D + P_0) M^{3/2}(\tilde{H}_3 + \tilde{H}_4 + \tilde{H}_3 \tilde{H}_4)(1 + \tilde{R}) \preccurlyeq (D + P_0) M^{3/2}\left(M^\rho + \frac{M^2}{E_0}\right)^3.$$

We have shown

$$\mathcal{N}_{*,\text{off}}(\mathbf{A}, \mathbf{B}, D, K, L) \preccurlyeq M^2 (D + P_0)^{1/2}\left(M^\rho + \frac{M^2}{E_0}\right) + (D + P_0) M^{3/2}\left(M^\rho + \frac{M^2}{E_0}\right)^3.$$

By (3.33) we have

$$M^2/E_0 \leq L^\alpha M^{2-2(1-\alpha)} \ll M^{4\alpha}$$

and we recall $M \ll P_0^{1/(2-2\rho)} \leq (D + P_0)^{1/(2-2\rho)}$. Hence in total we obtain

$$\mathcal{N}_{*,\text{off}}(\mathbf{A}, \mathbf{B}, D, K, L)$$
$$\preccurlyeq (D + P_0) M^{3/2 + \max(3\rho, 12\alpha)} + (D + P_0)^{\frac{1}{2} + \frac{\max(\rho, 4\alpha)}{2 - 2\rho}} M^2. \quad (3.39)$$

3.5.5. The final bound. Combining (3.38) and (3.39), we conclude under the present assumption (3.29) that

$$\mathcal{N}_*(\mathbf{A}, \mathbf{B}, D, K, L) \preccurlyeq M^{2-\alpha}(D + 1)^{\alpha/4}(D + P_0)$$
$$+ (D + P_0) M^{3/2 + \max(3\rho, 12\alpha)} + (D + P_0)^{\frac{1}{2} + \frac{\max(\rho, 4\alpha)}{2 - 2\rho}} M^2 \quad (3.40)$$

for a parameter $0 < \alpha < 1/6$ that is at our disposal.

3.6. The endgame

We summarize our findings. Suppose that

$$\max(3\rho, 12\alpha) \leq \frac{1}{2} - \rho, \quad \frac{\max(\rho, 4\alpha)}{2 - 2\rho} \leq \frac{1}{2} - \rho. \quad (3.41)$$

Combining (3.5), (3.8), (3.9), (3.14), (3.28) and (3.40) and summing over $(\log M)^2$ values of K, L, we obtain

$$\mathcal{N}(\mathbf{A}, \mathbf{B}, D) \preccurlyeq (D + P_0) \min\left(P_0^2, M^{2-\rho} + (D + P_0)^{-\rho} M^2 + M^{2-\alpha}(D + 1)^{\alpha/4}\right)$$

in all cases, where $0 < \rho, \alpha \leq 1/6$ are subject to (3.12) and (3.41), but otherwise at our disposal. By (3.3), we can simplify

$$\mathcal{N}(\mathbf{A}, \mathbf{B}, D) \preccurlyeq (D + P_0)M^2 \min\left((P_0/M)^2, M^{-\rho} + P_0^{-\rho} + M^{-\alpha\delta/4}\right).$$

Since $\min((P_0/M)^2, P_0^{-\rho}) \leq M^{-2\rho/(2+\rho)}$, we obtain finally

$$\mathcal{N}(\mathbf{A}, \mathbf{B}, D) \preccurlyeq (D + P_0)M^2(M^{-2\rho/(2+\rho)} + M^{-\alpha\delta/4}).$$

Choosing $\rho = \delta/4 \leq 1/8$, $\alpha = 1/32$, both conditions (3.12) and (3.41) are satisfied, and Proposition 6 follows.

4. Conclusion of the proof of Theorem 1: The continuous case

It is now a simple matter to prove Lemma 1, which follows the strategy of the previous section, but without any arithmetic input. We symmetrize the 6-dimensional integration region and consider all 8 variables $a_1, \ldots, a_4, b_1, \ldots, b_4$ subject to (2.9). We put all variables into dyadic intervals $A_j \leq |a_j| \leq 2A_j$, $B_j \leq |b_j| \leq 2B_j$, $D \leq |\Delta| \leq 2D$, $\Delta_1, \Delta_2 \preccurlyeq \Delta + P_0$ where the meaning of P_0 is the same as in the previous section. The numbers A_j, B_j, D run through logarithmically many positive and negative powers of 2 and are bounded by $M^{-100} \leq A_j, B_j \ll M$, $M^{-100} \leq D \ll M^{4-\delta}$. (If one of these is $\ll M^{-100}$, the coarsest trivial estimates suffice.) The same substitution as in Section 3.3 yields

$$kl := (a_2a_4 - b_2b_4)(b_2b_3 - a_2a_3) \preccurlyeq D + P_0.$$

We also put k, l into logarithmically many dyadic intervals $K \leq |k| \leq 2K$, $L \leq |l| \leq 2L$, $KL \preccurlyeq D + P_0$, and assume $K \leq L$. We also have relation (3.31), so that the region of integration for a_4 is of length $E^{1/2} = (D/L)^{1/2}$. The total volume of this region can be computed, of course, just as the central Poisson term. We integrate over $a_4, b_4, b_3, a_3, b_2, a_2$ in this order and see that it is bounded by

$$\preccurlyeq A_2 B_2 \frac{L}{A_2} B_3 \frac{K}{B_2} \sqrt{E} = KL^{1/2}D^{1/2}B_3$$
$$\ll (D + P_0)^{3/4}D^{1/2}M \leq (D + P_0)D^{1/4}M \ll (D + P_0)M^{2-\delta/4}.$$

Lemma 1 follows.

5. Proof of Corollary 2

The main purpose of this section is to deduce Corollary 2 from Theorem 1. The following general lemma implies Corollary 2, but it is stated in such generality that it can also be used in the proof of Corollary 3 in Section 6. We prove only the lim inf part of the claim since the proof for the lim sup is identical (up to reversing inequalities).

Lemma 2. *Let σ be a measure with bounded support in $\mathbb{R}$ and let γ be a real number. Then for almost all $\boldsymbol{\alpha}$ with respect to μ_{hyp} we have*

$$\liminf_{N \to \infty} \int_\gamma^\infty T_3(\boldsymbol{\alpha}; [\gamma, x]; N) \, d\sigma(x) \le \int_\gamma^\infty (x - \gamma)^2 \, d\sigma(x).$$

Remark 2. The lemma implies a bound for $T_3(\boldsymbol{\alpha}; I; N)$ as in the conclusion of Corollary 2, by writing $I = [a, b]$, picking $\gamma = a$, and setting $\sigma = \delta_b$ with δ_b denoting the Dirac measure centered on b. Then the right-hand side of Lemma 2 becomes $(b - a)^2 = \mu(I)^2$, and thus Corollary 2 follows from the lemma.

Proof of Lemma 2. Let $\varepsilon > 0$ be arbitrary. Let $B = B(\varepsilon, \gamma, \sigma)$ denote the set of $\boldsymbol{\alpha}$ such that for all $N \ge N_0(\varepsilon, \gamma, \sigma)$ we have

$$\int_\gamma^\infty T_3(\boldsymbol{\alpha}; [\gamma, x]; N) \, d\sigma(x) > \int_\gamma^\infty (x - \gamma)^2 \, d\sigma(x) + \varepsilon. \tag{5.1}$$

We need to show that $\mu_{\text{hyp}}(B) = 0$. Suppose the contrary, i.e. $\mu_{\text{hyp}}(B) > 0$ for some $\varepsilon > 0$. For some small $0 < \varepsilon_1 < 1/2$ to be determined later we choose a (finite) rectangle R with $\mu_{\text{hyp}}(R) > 0$ and $\mu_{\text{hyp}}(B \cap R) \ge (1 - \varepsilon_1)\mu_{\text{hyp}}(R)$. Such an R exists by the Lebesgue density theorem. By Theorem 1 we have

$$\int_R \left(\int_\gamma^\infty T_3(\boldsymbol{\alpha}; [\gamma, x]; N) \, d\sigma(x) \right) d_{\text{hyp}}\boldsymbol{\alpha} = \int_\gamma^\infty \int_R T_3(\boldsymbol{\alpha}; [\gamma, x]; N) \, d_{\text{hyp}}\boldsymbol{\alpha} \, d\sigma(x)$$

$$\le \int_\gamma^\infty (x - \gamma)^2 \, d\sigma(x) \, \mu_{\text{hyp}}(R) + \varepsilon_1$$

for $N \ge N_1 = N_1(\varepsilon_1, \gamma, \sigma, R)$, so that for all such N there exists a set $S_N \subseteq R$ such that $\mu_{\text{hyp}}(S_N) \ge 2\varepsilon_1 \mu_{\text{hyp}}(R)$ and

$$\int_\gamma^\infty T_3(\boldsymbol{\alpha}; [\gamma, x]; N) \, d\sigma(x) \le \frac{1}{1 - 2\varepsilon_1} \left(\int_\gamma^\infty (x - \gamma)^2 \, d\sigma(x) + \varepsilon_1 \right) \tag{5.2}$$

for all $\boldsymbol{\alpha} \in S_N$ and $N \ge N_1$. Let

$$S := \limsup S_N = \bigcap_{K \ge 1} \bigcup_{N \ge K} S_N \subseteq R.$$

The set S is the subset of those $\boldsymbol{\alpha} \in R$ which belong to infinitely many S_N. It follows from Lebesgue's dominated convergence theorem that $\mu_{\text{hyp}}(S) \ge 2\varepsilon_1 \mu_{\text{hyp}}(R)$, so that $\mu_{\text{hyp}}(S \cap B) \ge \mu_{\text{hyp}}(S) - \mu_{\text{hyp}}(R \setminus (B \cap R)) \ge \varepsilon_1 \mu_{\text{hyp}}(R)$ and in particular $S \cap B \ne \emptyset$. On the other hand, if

$$\int_\gamma^\infty (x - \gamma)^2 \, d\sigma(x) + \varepsilon \ge \frac{1}{1 - 2\varepsilon_1} \left(\int_\gamma^x (x - \gamma)^2 d\sigma(x) + \varepsilon_1 \right)$$

then $S \cap B = \emptyset$ by (5.1), and this can be achieved by choosing ε_1 sufficiently small. This contradiction shows $\mu_{\text{hyp}}(B) = 0$. ∎

6. Proof of Corollary 3

In this section, we show how Corollary 3 can be deduced from Corollary 2. We assume that Theorem 1 holds, and that consequently we may apply Lemma 2. For $b \geq 0$ let

$$f_{N,\boldsymbol{\alpha}}(b) := \frac{1}{N}|\{1 \leq i \leq N - 1 \mid \Lambda_{i+1}(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha}) \leq b\}|.$$

This is a piecewise constant non-decreasing function. Let $\varepsilon > 0$ be given. We will show that there is an infinite sequence of indices N for which $f_{N,\boldsymbol{\alpha}}(G - \varepsilon) < 1$, with G as in the statement of Corollary 3, thus establishing the existence of infinitely many pairs of consecutive elements of the sequence which are more than $G - \varepsilon$ apart. Suppose to the contrary that $f_{N,\boldsymbol{\alpha}}(G - \varepsilon) = 1$ for all large enough $N > N_0(\boldsymbol{\alpha}, \varepsilon)$. Then, by partial summation,

$$\Lambda_{N+1}(\boldsymbol{\alpha}) - \Lambda_1(\boldsymbol{\alpha}) = \sum_{i=1}^{N-1}(\Lambda_{i+1}(\boldsymbol{\alpha}) - \Lambda_i(\boldsymbol{\alpha})) = (N-1)\left((G-\varepsilon) - \int_0^{G-\varepsilon} f_{N,\boldsymbol{\alpha}}(x)\,\mathrm{d}x\right).$$

The left hand side is $(1 + o_{\boldsymbol{\alpha}}(1))N$ since the $\Lambda_i(\boldsymbol{\alpha})$ are normalized so that the average gap between them is $1 + o(1)$. In other words,

$$G \geq 1 + (1 - 10^{-6})\varepsilon + \int_0^{G-\varepsilon} f_{N,\boldsymbol{\alpha}}(x)\,\mathrm{d}x \tag{6.1}$$

for every $N \geq N_0 = N_0(\boldsymbol{\alpha}, \varepsilon)$. We now record three elementary inequalities:

- For all indices i we have

$$\mathbf{1}_{\Lambda_{i+1}-\Lambda_i \leq b} \geq \sum_{\substack{j>i \\ \Lambda_j - \Lambda_i \leq b}} 1 - \sum_{\substack{k>j>i \\ \Lambda_j - \Lambda_i \leq b \\ \Lambda_k - \Lambda_i \leq b}} 1. \tag{6.2}$$

- For all indices i we have

$$\mathbf{1}_{\Lambda_{i+1}-\Lambda_i \leq b} \geq \frac{2}{3} \sum_{\substack{j>i \\ \Lambda_j - \Lambda_i \leq b}} 1 - \frac{1}{3} \sum_{\substack{k>j>i \\ \Lambda_j - \Lambda_i \leq b \\ \Lambda_k - \Lambda_i \leq b}} 1. \tag{6.3}$$

- For all $i > N_0(\boldsymbol{\alpha}, \varepsilon)$,

$$\mathbf{1}_{\Lambda_{i+1}-\Lambda_i \leq b} = 1 - \mathbf{1}_{\Lambda_{i+1}-\Lambda_i \in (b, G-\varepsilon]} \geq 1 - \sum_{\substack{j>i \\ \Lambda_j - \Lambda_i \in (b, G-\varepsilon]}} 1. \tag{6.4}$$

To prove the first two inequalities notice that if there are exactly $\ell \geq 1$ values Λ_j in the interval $(\Lambda_i, \Lambda_i + c]$ then inequalities (6.2) and (6.3) amount to, respectively,

$$1 \geq \ell - \binom{\ell}{2} \quad \text{and} \quad 1 \geq \frac{2}{3}\ell - \frac{1}{3}\binom{\ell}{2},$$

which are easily seen to hold. Since moreover (6.2) and (6.3) are trivially true when there are no Λ_j's in $(\Lambda_i, \Lambda_i + c]$, we conclude that (6.2) and (6.3) always hold. The final inequality (6.4) follows from our assumption that $\Lambda_{i+1} - \Lambda_i \leq G - \varepsilon$ for all $N > N_0(\boldsymbol{\alpha}, \varepsilon)$ and from the trivial upper bound

$$\mathbf{1}_{\Lambda_{i+1} - \Lambda_i \in I} \leq \sum_{\substack{j > i \\ \Lambda_j - \Lambda_i \in I}} 1,$$

which holds for all intervals $I \subset \mathbb{R}$.

Summing each of (6.2), (6.3) and (6.4) over i gives, respectively,

$$f_{N,\boldsymbol{\alpha}}(b) \geq T_2(\boldsymbol{\alpha}; [0, b]; N) - \tfrac{1}{2} T_3(\boldsymbol{\alpha}; [0, b]; N),$$
$$f_{N,\boldsymbol{\alpha}}(b) \geq \tfrac{2}{3} \cdot T_2(\boldsymbol{\alpha}; [0, b]; N) - \tfrac{1}{6} T_3(\boldsymbol{\alpha}; [0, b]; N),$$
$$f_{N,\boldsymbol{\alpha}}(b) \geq 1 - T_2(\boldsymbol{\alpha}; [b, G - \varepsilon]; N),$$

where each of these inequalities holds for all possible values of b. Let $C = \sqrt{6G - 5} - 1 \ (\approx 1.65)$. Using the above inequalities we see that

$$\int_0^{G-\varepsilon} f_{N,\boldsymbol{\alpha}}(b) \, \mathrm{d}b \geq \int_0^1 \left(T_2(\boldsymbol{\alpha}; [0, b]; N) - \tfrac{1}{2} T_3(\boldsymbol{\alpha}; [0, b]; N) \right) \mathrm{d}b$$
$$+ \int_1^C \left(\tfrac{2}{3} T_2(\boldsymbol{\alpha}; [0, b]; N) - \tfrac{1}{6} T_3(\boldsymbol{\alpha}; [0, b]; N) \right) \mathrm{d}b$$
$$+ \int_C^{G-\varepsilon} \left(1 - T_2(\boldsymbol{\alpha}; [b, G - \varepsilon]; N) \right) \mathrm{d}b. \tag{6.5}$$

We evaluate the integrals using Sarnak's pair correlation result [14] (where it is easy to see that the convergence is uniform in intervals $[0, b]$ for $b \leq 1$, say), and using Lemma 2 with $\gamma = 0$ and σ a measure equal to $\tfrac{1}{2}\mathrm{d}x$ on $0 \leq x \leq 1$ resp. $\tfrac{1}{6}\mathrm{d}x$ on $1 \leq x \leq C$ and vanishing otherwise. In this way we find that for almost all $\boldsymbol{\alpha}$ there exists a subsequence $N_1 < N_2 < \cdots$ along which the right hand side of (6.5) is at least

$$\int_0^1 (b - b^2/2) \, \mathrm{d}b + \int_1^C \left(\tfrac{2}{3}b - \tfrac{1}{6}b^2 \right) \mathrm{d}b + \int_C^{G-\varepsilon} (1 - (G - \varepsilon - b)) \, \mathrm{d}b - 10^{-6}\varepsilon.$$

Inserting this lower bound into (6.1) gives

$$G \geq 1 + (1 - 10^{-6})\varepsilon + \tfrac{1}{3} + \left(3G - \tfrac{1}{9}\sqrt{6G - 5} \cdot (3G + 5) - \tfrac{43}{18} \right)$$
$$+ \left(G\varepsilon - G^2/2 - \varepsilon^2/2 + \sqrt{6G - 5} \cdot (G - \varepsilon) - 3G + 3 \right) - 10^{-6}\varepsilon.$$

Using the definition of G this simplifies to

$$0 \geq \left(1 - 2 \cdot 10^{-6} + G - \sqrt{6G - 5} \right)\varepsilon - \varepsilon^2/2.$$

Since $2 \leq G \leq 2.1$, this is a contradiction for all ε sufficiently small. This proves Corollary 3.

7. Proof of Theorem 4

7.1. The lower bound

Let

$$f_N(x) := \frac{1}{N}|\{i \le N - 1 \mid \gamma_{i+1} - \gamma_i \le x\}| \quad \text{and} \quad T_2(x; N) := \frac{1}{N} \sum_{\substack{\gamma_i - \gamma_j \in [0,x] \\ 1 \le j < i \le N}} 1.$$

Suppose without loss of generality that the gaps $\gamma_{i+1} - \gamma_i$ are uniformly bounded by some $G > 0$ for all sufficiently large i. Then by integration by parts (as in the previous section), we obtain

$$G \ge 1 + \int_0^G f_N(x)\,dx + o(1) \tag{7.1}$$

as $N \to \infty$. Notice now that for all i large enough,

$$\mathbf{1}_{\gamma_{i+1} - \gamma_i \le x} = 1 - \mathbf{1}_{\gamma_{i+1} - \gamma_i \in (x,G]} \ge 1 - \sum_{\substack{\gamma_i - \gamma_j \in (x,G] \\ 1 \le j < i \le N}} 1.$$

Therefore, summing over $i < N$ and dividing by N we get

$$f_N(x) \ge 1 - (T_2(G; N) - T_2(x; N)) + o(1) \tag{7.2}$$

uniformly in $0 \le x \le G$ as $N \to \infty$. Given $\varepsilon > 0$ there exists an $N_0(\varepsilon)$ such that for all $N > N_0(\varepsilon)$,

$$T_2(G; N) - T_2(x; N) \le (G - x) + \varepsilon \tag{7.3}$$

for all $0 \le x \le G$.[3] In particular, fixing $\varepsilon > 0$, combining $f_N(x) \ge 0$ and (7.1)–(7.3) and taking N to infinity, we conclude that

$$G \ge 1 + \int_{G-1}^G (1 - G + x)\,dx + \varepsilon.$$

Squeezing ε to zero this implies that $G \ge 3/2$.

7.2. The upper bound

We define a semi-random sequence $(\gamma_i)_{i \ge 1} = (\gamma_i(\omega))_{i \ge 1}$ in the following way. Let $X_1, X_2, \ldots$ be a sequence of independent random variables defined on a probability space $(\Omega, \Sigma, \mathbb{P})$ such that X_i has uniform distribution on the interval $[i, i + 1]$ for $i \ge 1$. We may assume that none of the X_i has an integer value. Furthermore, for every $m \ge 1$ we define numbers

$$y_i^{(m)} = m^2 + \frac{i}{\lceil \sqrt{2m} \rceil}, \quad 0 \le i \le \lceil \sqrt{2m} \rceil.$$

Note that the numbers $y_i^{(m)}, 0 \le i \le \lceil \sqrt{2m} \rceil$, are all contained in the interval $[m^2, m^2 + 1]$.

---

[3]To see this, it is enough to split $[0, G]$ into a disjoint union of intervals $I_1, I_2, \ldots$ of length ε^2 and find an N_0 large enough so that $T_{2,N}(I_k)$ is less than $\mu(I_k) + \varepsilon^2$ for all k.

Thus these point sets do not intersect or overlap for different values of m. Our sequence will consist of the random numbers $(X_i)_i$ from above, as well as of the deterministic numbers $(y_i^{(m)})_{i,m}$. To prevent the random components of the sequence from mixing with the deterministic ones (which would make it unnecessarily complicated to sort the sequence in increasing order), we discard those random numbers whose index is a perfect square, since their values lie in the same range as the values of some of the deterministic numbers. More precisely, we define $(\gamma_i)_{i \geq 1} = (\gamma_i(\omega))_{i \geq 1}$ as the sequence which contains

- all numbers X_i, $i \geq 1$, for which i is not a perfect square, as well as
- all numbers $y_i^{(m)}$ for $m \geq 1$ and $0 \leq i < \lceil \sqrt{2m} \rceil$,

sorted in increasing order. Note that the gaps between consecutive elements of (γ_i) are uniformly bounded by 2.

We split the index set $\mathbb{N}$ into two classes $\mathcal{N}_1$ and $\mathcal{N}_2$. The first class $\mathcal{N}_1$ contains all those indices i for which the value of γ_i comes from one of the random numbers $(X_j)_{j \geq 1}$. Furthermore, we set $\mathcal{N}_2 = \mathbb{N} \setminus \mathcal{N}_1$, that is, $\mathcal{N}_2$ contains those indices i for which γ_i comes from one of the clusters of deterministic points. The set $\mathcal{N}_2$ decomposes into classes C_m, $m \geq 1$, such that C_m contains those indices $i \in \mathcal{N}_2$ for which $\gamma_i \in [m^2, m^2 + 1]$. It is easily seen that

$$|\{i \in \mathcal{N}_2 \mid \gamma_i \leq N\}| \leq \sum_{m^2 \leq N} (\sqrt{2m} + 1) \ll N^{3/4}.$$

Consequently,

$$|\{i \in \mathbb{N} \mid \gamma_i \leq N\}| = N + O(N^{3/4}),$$

which implies

$$\gamma_i / i \to 1 \quad \text{as } i \to \infty.$$

In other words, the average spacing of (γ_i) is 1, as required. We claim that the pair correlation of this sequence is Poissonian, $\mathbb{P}$-almost surely.

Assume that $b > 0$ is fixed. We split $T_2(b; N)$ into different parts in the following way. We write

$$T_2(b; N) = \frac{1}{N} |\{(i_1, i_2) \mid 1 \leq i_1 < i_2 \leq N, \gamma_{i_2} - \gamma_{i_1} \leq b\}|$$

$$=: \frac{1}{N} \left(\sum_1 + \sum_2 + \sum_3 + \sum_{m=1}^{\infty} \sum_{4,m} \right),$$

where all sums are taken over indices $1 \leq i_1 < i_2 \leq N$ such that $\gamma_{i_2} - \gamma_{i_1} \leq b$, subject to the following additional restrictions:

$$\sum_1 : i_1, i_2 \text{ are both in } \mathcal{N}_1;$$

$$\sum_2 : \text{exactly one of } i_1, i_2 \text{ is in } \mathcal{N}_1;$$

$$\sum_3 : i_1 \text{ and } i_2 \text{ are both in } \mathcal{N}_2, \text{ but not in the same set } C_m \text{ for some } m;$$

$$\sum_{4,m} : i_1 \text{ and } i_2 \text{ are both in } \mathcal{N}_2 \text{ and both in the same set } C_m.$$

Recall that for $\gamma_i \in \mathcal{N}_1$ we have $\gamma_i = X_i = i + Y_i$ where the Y_i are independent and uniformly distributed on $[0, 1]$. Therefore,

$$\frac{1}{N} \sum_1 = \sum_{\ell \geq 1} \left(\frac{1}{N} \sum_{\substack{k+\ell < N \\ \gamma_{k+\ell} - \gamma_k \leq b}} 1 \right) = \sum_{\ell \geq 1} \left(\frac{1}{N} \sum_{\substack{k+\ell < N \\ Y_{k+\ell} - Y_k \leq b - \ell}} 1 \right) \to \sum_{\ell \leq b+1} \mathbb{P}(X - Y \leq b - \ell)$$

$\mathbb{P}$-almost surely as $N \to \infty$ by the strong law of large numbers, where X, Y are uniformly distributed on $[0, 1]$. Since X, Y are independent, we have

$$\mathbb{P}(X - Y \leq x) = \int_{-\infty}^{x} h_{X-Y}(v) \, dv, \quad h_{X-Y}(v) := \int_{\mathbb{R}} \mathbf{1}_{-1 \leq y \leq 0} \cdot \mathbf{1}_{0 \leq v - y < 1} \, dy.$$

Therefore,

$$\sum_{\ell \geq 1} \mathbb{P}(X - Y \leq b - \ell) = \int_{-\infty}^{b} \sum_{\ell \geq 1} h_{X-Y}(x - \ell) \, dx.$$

We now notice that

$$\sum_{\ell \geq 1} h_{X-Y}(x - \ell) = \int_{\mathbb{R}} \mathbf{1}_{-1 \leq y \leq 0} \left(\sum_{\ell \geq 1} \mathbf{1}_{0 \leq x - \ell - y < 1} \right) dy$$

$$= \int_{\mathbb{R}} \mathbf{1}_{-1 \leq y \leq 0} \cdot \mathbf{1}_{x - y \geq 1} \, dy = \begin{cases} 1 & \text{if } x \geq 1, \\ x & \text{if } 0 \leq x \leq 1, \\ 0 & \text{if } x \leq 0. \end{cases}$$

Combining the above two equations we conclude that

$$\sum_{\ell \geq 1} \mathbb{P}(X - Y \leq b - \ell) = \begin{cases} 0, & b \leq 0, \\ b^2/2, & b \in [0, 1], \\ b - 1/2, & b \geq 1. \end{cases} \tag{7.4}$$

Furthermore, we have

$$\frac{1}{N} \sum_2 \ll \frac{|\{i \in \mathcal{N}_2 \mid i \leq N\}|}{N} \ll \frac{N^{3/4}}{N} \to 0, \tag{7.5}$$

since for each $i \in \mathcal{N}_2$ there are at most $2\lceil b \rceil$ indices $j \in \mathcal{N}_1$ such that $|\gamma_i - \gamma_j| \leq b$. Finally, we have

$$\frac{1}{N} \sum_3 \to 0 \quad \text{as } N \to \infty, \tag{7.6}$$

since b is assumed to be fixed and the gaps between C_m and C_{m+1} grow to infinity as $m \to \infty$. Thus it remains to calculate the contribution of the sums $\sum_{4,m}$. So let m be given, and assume that $C_m \subset \{1, \ldots, N\}$. Consider

$$\sum_{4,m} = |\{(i_1, i_2) \in C_m^2 \mid i_1 < i_2, \gamma_{i_2} - \gamma_{i_1} \leq b\}|.$$

If $b \geq 1$ then by construction all such pairs of indices are within distance b of each other, and thus for $b \geq 1$ we have

$$\sum\nolimits_{4,m} = |\{(i_1, i_2) \in C_m^2 \mid i_1 < i_2\}| \sim |C_m|^2/2 \sim m.$$

The maximal m for which $C_m \subset \{1, \ldots, N\}$ is of order $\sqrt{N}(1 + o(1))$. There may also be one set C_m which is only partially contained in $\{1, \ldots, N\}$, but its contribution is of order at most $|C_m|^2 \ll m \ll \sqrt{N}$ and thus negligible. So overall in the case $b \geq 1$ we have

$$\sum_{m=1}^{\infty} \sum\nolimits_{4,m} \sim \sum_{1 \leq m \leq \sqrt{N}} m \sim \frac{N}{2}.$$

Thus, together with (7.4)–(7.6), for $b \geq 1$ we obtain

$$T_2(b; N) \to b - \frac{1}{2} + \frac{1}{2} = b \quad \text{as } N \to \infty, \; \mathbb{P}\text{-almost surely},$$

which coincides with Poissonian behavior.

It remains to consider the case $b \in [0, 1]$. So fix b in this range. Let a set C_m be given, and assume again that it is contained in $\{1, \ldots, N\}$. The points γ_i, $i \in C_m$, are positioned at the values

$$m^2, \; m^2 + \frac{1}{\lceil \sqrt{2m} \rceil}, \; m^2 + \frac{2}{\lceil \sqrt{2m} \rceil}, \; \ldots, \; m^2 + 1.$$

We could give an explicit formula for the number of pairs of indices $i_1 < i_2$ for which $\gamma_{i_2} - \gamma_{i_1} \leq b$, but it is sufficient to know that the cardinality of this set of pairs is of order

$$\sim |C_m|^2 \left((1 - b)b + \int_0^b y \, dy \right) \sim 2m(b - b^2/2),$$

which is easily verified (for example, by looking at the problem in terms of counting lattice points of a square lattice contained in a certain polygon). Thus for $b \in [0, 1]$ we have

$$\sum_{m=1}^{\infty} \sum\nolimits_{4,m} \sim \sum_{1 \leq m \leq \sqrt{N}} 2m(b - b^2/2) \sim N(b - b^2/2).$$

As a consequence, again in combination with (7.4), (7.5) and (7.6), we obtain

$$T_2(b; N) \to b^2/2 + b - b^2/2 = b \quad \text{as } N \to \infty, \; \mathbb{P}\text{-almost surely}.$$

This gives the desired result for $b \in [0, 1]$. Finally, we note that by continuity it is sufficient to consider all b contained in a countable dense subset of $[0, \infty)$, such as $\mathbb{Q} \cap [0, \infty)$. For each individual b there is an exceptional set of $\mathbb{P}$-measure zero, and since the union of countably many sets of measure zero also has measure zero, we conclude that there is an element $\omega \in \Omega$ for which $T_2(b; N) \to b$ holds for all $b \geq 0$. This completes the proof of Theorem 4.

References

[1] Altman, D., Chase, Z.: On the smallest gap in a sequence with Poisson pair correlations. arXiv:2210.17466 (2022)

[2] Berry, M., Tabor, M.: Level clustering in the regular spectrum. Proc. Roy. Soc. London A **356**, 375–394 (1977)   Zbl 1119.81395

[3] Blomer, V., Bourgain, J., Radziwiłł, M., Rudnick, Z.: Small gaps in the spectrum of the rectangular billiard. Ann. Sci. École Norm. Sup. (4) **50**, 1283–1300 (2017)   Zbl 1405.35120   MR 3720030

[4] Bohigas, O., Giannoni, M.-J., Schmit, C.: Spectral fluctuations of classically chaotic quantum systems. In: Quantum chaos and statistical nuclear physics, Lecture Notes in Physics 263, Springer, Berlin, 18–40 (1986)

[5] El-Baz, D., Marklof, J., Vinogradov, I.: The two-point correlation function of the fractional parts of $\sqrt{n}$ is Poisson. Proc. Amer. Math. Soc. **143**, 2815–2828 (2015)   Zbl 1318.11091   MR 3336607

[6] Elkies, N. D., McMullen, C. T.: Gaps in $\sqrt{n}$ mod 1 and ergodic theory. Duke Math. J. **123**, 95–139 (2004)   Zbl 1063.11020   MR 2060024

[7] Eskin, A., Margulis, G., Mozes, S.: Quadratic forms of signature $(2, 2)$ and eigenvalue spacings on rectangular 2-tori. Ann. of Math. (2) **161**, 679–725 (2005)   Zbl 1075.11049   MR 2153398

[8] Gradshteyn, I. S., Ryzhik, I. M.: Table of integrals, series, and products. 7th ed., Elsevier/Academic Press, Amsterdam (2007)   Zbl 1208.65001   MR 2360010

[9] Huang, J.-J.: The density of rational points near hypersurfaces. Duke Math. J. **169**, 2045–2077 (2020)   Zbl 1458.11115   MR 4132580

[10] Kuna, T., Lebowitz, J. L., Speer, E. R.: Realizability of point processes. J. Statist. Phys. **129**, 417–439 (2007)   Zbl 1130.60056   MR 2351408

[11] Kuna, T., Lebowitz, J. L., Speer, E. R.: Necessary and sufficient conditions for realizability of point processes. Ann. Appl. Probab. **21**, 1253–1281 (2011)   Zbl 1235.60052   MR 2857448

[12] Lutsko, C.: Long-range correlations of sequences modulo 1. J. Number Theory **234**, 333–348 (2022)   Zbl 1487.11074   MR 4370539

[13] Salié, H.: Über die Kloostermanschen Summen $S(u, v; q)$. Math. Z. **34**, 91–109 (1932)   Zbl 0002.12801   MR 1545243

[14] Sarnak, P.: Values at integers of binary quadratic forms. In: Harmonic analysis and number theory (Montreal, PQ, 1996), CMS Conference Proceedings 21, American Mathematical Society, Providence, RI, 181–203 (1997)   Zbl 0911.11032   MR 1472786

[15] Technau, N., Walker, A.: On the triple correlations of fractional parts of $n^2\alpha$. Canad. J. Math. **74**, 1347–1384 (2022)   Zbl 07613387   MR 4504666

[16] Vanderkam, J. M.: Pair correlation of four-dimensional flat tori. Duke Math. J. **97**, 413–438 (1999)   Zbl 0965.11017   MR 1682221

[17] VanderKam, J. M.: Correlations of eigenvalues on multi-dimensional flat tori. Comm. Math. Phys. **210**, 203–223 (2000)   Zbl 0963.58014   MR 1748175

[18] Weil, A.: On some exponential sums. Proc. Nat. Acad. Sci. U.S.A. **34**, 204–207 (1948)   Zbl 0032.26102   MR 27006