SPECTRAL QUASI CORRELATIONS AND PHASE-TRANSITIONS FOR THE NODAL LENGTH OF ARITHMETIC RANDOM WAVES

ANDREA SARTORI

Abstract. We study the nodal length of Arithmetic Random Waves at small scales: we show that there exists a phase-transition for the distribution of the nodal length at a logarithmic power above Planck-scale. Furthermore, we give strong evidence for the existence of an intermediate phase between Arithmetic and Berry’s random waves. These results are based on the study of small sums of lattice points lying on the same circle, called spectral quasi correlations. We show that, for generic integers representable as the sum of two squares, there are no spectral quasi correlations.

1. Introduction

1.1. The Random Wave Model and the nodal length of Laplace eigenfunctions. Given a compact Riemannian surface $(M,g)$ without boundary, let $\Delta_g$ be the Laplace-Beltrami operator on $M$. There exists an orthonormal basis for $L^2(M,d\text{Vol})$ consisting of eigenfunctions $\{f_{\lambda_i}\}$

$$\Delta_g f_{\lambda_i} + \lambda_i f_{\lambda_i} = 0$$

with $0 = \lambda_1 < \lambda_2 \leq \ldots$ listed with multiplicity, and $\lambda_i \to \infty$. One of the main characteristics of an eigenfunction $f_{\lambda}$ is its nodal set

$$\mathcal{Z}(f_{\lambda}) = \{x \in M : f_{\lambda}(x) = 0\}.$$ 

It is known that $\mathcal{Z}(f_{\lambda})$ is the union of smooth curves outside a finite set of points [12] and Yau conjectured that its volume, the nodal length, satisfies

$$c\sqrt{\lambda} \leq \mathcal{L}(f) := \mathcal{H}(\mathcal{Z}(f_{\lambda})) \leq C\sqrt{\lambda}$$

(1.1)

for some constants $c, C > 0$ which depend on $M$ only, where $\mathcal{H}(\cdot)$ is the Hausdorff measure. Donnelly and Fefferman [17] showed that Yau’s conjecture holds for any real-analytic manifold (of any dimension), and recently, Logunov and Malinnikova [28, 29, 30] proved the lower-bound in the smooth case and gave a polynomial upper-bound.

Berry [5, 6] conjectured that “generic” Laplace eigenfunctions $f_{\lambda}$ can be modelled, in balls of radius slightly larger than $O(\lambda^{-1/2})$, the Planck-scale, by monochromatic plane waves, an isotropic Gaussian field with the spectral measure the Lebesgue measure on the unit circle. This Gaussian field is also known as Berry’s Random Waves (BRW). In particular, Berry’s model suggests that “generic” Laplace eigenfunctions change their behaviour when restricted to sufficiently small balls, we are interested in exploring how this affects their nodal length.

1.2. Phase-transitions for the nodal length of Arithmetic Random Waves. We study random Laplace eigenfunctions on the flat two dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, also known as Arithmetic Random Waves (ARW). These are Gaussian random fields satisfying

$$\Delta f_n + 4\pi^2 n f_n = 0$$

1
where $\Delta$ is the flat Laplacian and the eigenvalue $n \in S := \{ n = a^2 + b^2 : a, b \in \mathbb{Z} \}$, the set of integers representable as the sum of two squares. Explicitly, $f_n$ can be defined as

$$ f_n(x) = \frac{1}{\sqrt{N}} \sum_{\xi \in \mathbb{Z}^2} a_\xi e(\langle \xi, x \rangle) \tag{1.2} $$

where $e(\cdot) = e^{2\pi i \cdot}$, $a_\xi$ are i.i.d. standard complex Gaussian random variables save for $\overline{a_\xi} = a_{-\xi}$ so that $f_n$ is real valued, and the normalisation constant $N = r_2(n)$, the number of lattice points on the circle of radius $\sqrt{n}$, in (1.2) implies that $\mathbb{E}[|f_n|^2] = 1$. Up to rescaling $\mathbb{T}^2$, ARW can equivalently be defined, via Kolmogorov’s Theorem, as the centred, stationary, Gaussian random field with spectral measure

$$ \mu_n = \frac{1}{N} \sum_{|\xi|^2 = n} \delta_{\xi/\sqrt{n}} \tag{1.3} $$

where $\delta_{\xi/\sqrt{n}}$ is the Dirac distribution at $\xi/\sqrt{n}$.

The study of the nodal length of the ARW was initiated by Oravecz, Rudnick and Wigman [37]; Rudnick and Wigman [38] found the expectation of $\mathcal{L}(f_n)$ and gave an upper bound for the variance. Subsequently, Krishnapur, Kurlberg and Wigman [25] proved that

$$ \text{Var}[\mathcal{L}(f_n)] = 1 + \hat{\mu}_n(4) \frac{n}{512} \frac{N}{N^2} (1 + o_{n \to \infty}(1)) \tag{1.4} $$

where $\hat{\mu}_n(4)$ is the fourth Fourier coefficient of the measure $\mu_n$. Notably, the accumulation points of the sequence $\hat{\mu}_n(4)$ contain the interval $[0, 1]$, [13, 27, 39]. Finally, a non-universal, non-central limit law for $\mathcal{L}(f_n)$ was found by Marinucci, Peccati, Rossi and Wigman [31].

Berry [6] showed that

$$ \text{Var}[\mathcal{L}(f_\mu)] = \frac{1}{256\pi} \log n (1 + o_{n \to \infty}(1)) $$

where $f_\mu$ are BRW (in a square of side 1), $\mu$ is the Lebesgue measure on the unit circle and, for the sake of consistency, $n$ represent the eigenvalue. Therefore, the asymptotic expansion (1.4) shows that the total nodal length of the ARW behaves differently than the nodal length of the BRW. However, since the spectral measure $\mu_n$ converges to $\mu$ for almost all $n \in S$ [18, 24], the field $f_n$ generically behaves like BRW in balls of radius $O(n^{-1/2})$. This suggests the existence of a phase-transition in the behaviour of the nodal length of “generic” ARW and, in analogy with the study of the ARW at small scales in [22], we expect the said phase-transition to happen at some logarithmic power above Planck-scale.

Investigating the nodal length of ARW at small scales, Benatar, Marinucci and Wigman [4], suggested that there actually exits a intermediate phase between ARW and BRW: letting

$$ \mathcal{L}(f_n, s) = \text{Vol}\{x \in B(s) : f_n(x) = 0\} $$

where $B(s)$ is the ball of radius $s$ centred at the origin, it is expected that there exists some exponent $A_0 > 0$ such that the law of $\mathcal{L}(f_n, s)$ agrees with the law of $\mathcal{L}(f_\mu)$ for $s > (\log n)^{A_0}/n^{1/2}$ and behaves differently for $s < (\log n)^{A_0}/n^{1/2}$. In this direction, Benatar, Marinucci and Wigman [4] found that, for a density one subsequence of $n \in S$, the asymptotic expansion of the variance of $\mathcal{L}(f_n, s)$ agrees, once appropriately rescaled, with (1.4) and they deduced that $\mathcal{L}(f_n, s)$ fully correlates with $\mathcal{L}(f_\mu)$, provided that $s > n^{-1/2+}$. Moreover, they related $\mathcal{L}(f_n, s)$ to the notion of spectral quasi-correlations, while $\mathcal{L}(f_\mu)$ is related to the notion of spectral correlations [8, 25]. We are now going to describe spectral
correlations and quasi-correlations and we will give the details of such relations in section 1.4 below.

1.3. Statements of main results. The set of spectral correlations is

$$S(l, n) := \{(\xi_1, ..., \xi_l) : \xi_1 + ... + \xi_l = 0 \ | \xi_i|^2 = n\}$$

where \(l\) is a positive integer, the length of the correlations, \(n \in S\) and \(\xi_i\) are the lattice points on the circle of radius \(\sqrt{n}\). While the set of spectral quasi-correlations is

$$Q(l, n, K) := \{(\xi_1, ..., \xi_l) : 0 < |\xi_1 + \xi_2 + \xi_3 + ... + \xi_l| \leq K \ |\xi_i|^2 = n\}$$

where \(K > 0\) is some parameter. Importantly, \(Q(2l, n, K)\) excludes the set of “diagonal” solutions \(\xi_1 = -\xi_2, ..., \xi_{2l-1} = -\xi_l\) which is contained in \(S(2l, n)\). To study ARW, we are interested in the largest \(K\) such that \(Q(l, n, K) = \emptyset\) as \(n \to \infty\).

Harman and Lewis [23] showed that there are infinitely many primes of the form \(p = a^2 + b^2\) with \(|b| \leq p^c\) for some small constant \(0 < c < 0.119\). For the such primes, there are two lattice points, \(\xi_\pm = (\pm a, b)\), with \(|\xi_+ + \xi_-| = b\), thus \(Q(2, p, p^c) \neq \emptyset\). Moreover, if there exist infinitely many primes of the form \(p = m^2 + 1\), as it was conjectured by Landau, then \(Q(2, p, O(1)) \neq \emptyset\). However, if we consider a generic integer \(n \in S\), we can prove the following:

**Theorem 1.1.** Let \(l \geq 2\) be an integer, \(\epsilon > 0\) and define \(c(l) = c(l, \epsilon)\) recursively as follows: \(c(2) = \log 2 + \epsilon, c(3) = 3\log 2/2 + \epsilon\) and \(c(l) = l\log 2/2 + c([l/2]) + \epsilon\), where \([l/2]\) represent the largest integer smaller than \(l/2\). Then for almost all \(n \in S\), we have

$$Q(l, n, n^{1/2}/(\log n)^{c(l)}) = \emptyset.$$ 

It is also relevant to us to investigate when \(Q(l, n, K) \neq \emptyset\). Erdős-Hall [18, Theorem 3] showed that, for almost all \(n \in S\), \(Q(2, n, n^{1/2}\log n^{\log 3/2+c+\epsilon}) \neq \emptyset\). Therefore, by the triangle inequality, we have

$$Q(2l, n, l \cdot n^{1/2}/(\log n)^{\log 3/2+c+\epsilon}) \neq \emptyset$$

for almost all \(n \in S\) and, in section 5.2 below, we will show that also odd length quasi-correlations can be small, in the appropriate sense.

In order to understand the size of \(Q(l, n, K)\), we study a random model for a generic integer \(n \in S\), see also [22, Remark 3.3]. In this model, the angles of the Gaussian primes dividing \(n\) are represented by i.i.d. uniform random variables on \([0, 2\pi)\) so that the lattice points \(\xi_i\) are random variables taking values on the circle of radius \(\sqrt{n}\), more details will be given in sections 2.1 and 4 below. We then define the random sums \(X_\iota := (\xi_{i_1} + ... + \xi_{i_l})/n^{1/2}\) where \(\iota = (i_1, ..., i_l)\) for \(1 \leq i_j \leq N\) and prove the following:

**Theorem 1.2.** Let \(l \geq 2\), \(n \in S\), \(0 < \alpha < 1\) be some parameter which may depend on \(n\) and \(X_\iota\) for \(\iota = (i_1, ..., i_l)\) be as above. Then we have

$$\mathbb{E}[^\#\{X_\iota : |X_\iota| \leq \alpha\}] \approx N(\alpha + O_l(\alpha^2))N^l(1 + o_{N \to \infty}(1))$$

where \(A \lessapprox B\) if there exist two constants \(c, C > 0\) such that \(cA \leq B \leq CA\) and the constants implied in the notation depend on \(l\) only.

We observe that, taking \(\alpha = N^{-l}\) in Theorem 1.2, we expect \(Q(l, n, O(n^{1/2}/N)) \neq \emptyset\). Since, for almost all \(n \in S\), \(N \sim (\log n)^{\log 2/2+\epsilon}\), see Lemma 2.3 below, and \(c(l) \leq 2l\), Theorem 1.2 suggests that Theorem 1.1 gives the right order of growth, in \(l\), for the constant \(c(l)\).

Thanks to Theorem 1.1, following similar techniques to [4], we are able to prove the following upper bound for the phase-transitions:
Theorem 1.3. Let \( A = \frac{29}{6} \log 2 = 3.3512... \) and \( \epsilon > 0 \). There exists a density one subsequence of \( n \in S \) such that the following holds:

1. \( N(n) \to \infty \) and the set of accumulation points of \( \{ \hat{\mu}_n(4) \} \) contains \([0,1]\).

2. Uniformly for \( s > (\log n)^{A+\epsilon} \cdot n^{-1/2} \), we have
   \[
   \text{Var}(\mathcal{L}(f_n, s)) = \frac{1 + \hat{\mu}_n(4)}{512} (\pi s^2)^2 \frac{n}{N^2} (1 + o_{N \to \infty}(1)).
   \]

3. We have
   \[
   \sup_{s > (\log n)^{A+\epsilon}/n^{1/2}} \left| \text{Corr}(\mathcal{L}(f_n, s), \mathcal{L}(f_n)) - 1 \right| \to 0
   \]
   where, \( \text{Corr}(X,Y) = \text{Cov}(X,Y)/(\text{Var}(X))^{1/2}(\text{Var}(Y))^{1/2} \).

Given a sequence of \( n \in S \) that satisfies the conclusion of Theorem 1.1 and (1.9), part (2) and part (3) of Theorem 1.3 follow directly using the techniques in [4]. However, \emph{a priori}, for any said sequence, \( \hat{\mu}_n(4) \) might have only one accumulation point. To rule this out, we explicitly construct sequences of \( n \in S \), satisfying (1.9) and the conclusion of Theorem 1.1, for which we can control the distribution of lattice points on \( \sqrt{n} S^1 \). Benatar, Marinucci and Wigman’s argument relies on density estimates, thus our result seems to be the first to give explicit examples of such sequences in the literature.

Finally, we show that there exists some \( B > 0 \) such that \( \mathcal{L}(f_n, s) \) behaves like the nodal length of BRW for \( s < (\log n)^{B} \cdot n^{-1/2} \). This provides a lower bound for the phase-transitions and shows that the behaviour of the nodal length changes at some logarithmic power above the Planck scale.

Theorem 1.4. Let \( B = \frac{1}{81} \log \frac{\pi}{2} = 0.0053... \), \( \epsilon > 0 \) and \( R > 1 \). Moreover, let \( F_{n}(\cdot) = f_{n}(R \cdot /\sqrt{n}) \) and \( F_{\mu}(\cdot) = f_{\mu}(R \cdot) \), where \( f_{\mu} \) is the BRW. There exists a density one subsequence of \( n \in S \) such that for all \( R \leq \log n^{B-\epsilon} \) and all fixed \( t \in (-\infty, \infty) \), we have
\[
|\mathbb{E}[\exp(it\mathcal{L}(F_{n}))] - \mathbb{E}[\exp(it\mathcal{L}(F_{\mu}))]| \to 0 \quad n \to \infty.
\]

As mentioned in section 1.2, \( \text{Var}[\mathcal{L}(f_{\mu})] \) was computed by Berry [6] and Wigman [42] found the variance for the nodal length of random spherical harmonics on the two dimensional sphere. The law, for random spherical harmonics, was discovered by Marinucci, Rossi and Wigman [32]. Subsequently, Nourdin, Peccati and Rossi [36] found the law of \( \mathcal{L}(f_{\mu}) \) to be Normal, once appropriately normalised.

1.4. Intermediate phase. Let \( f_{n} \) be as in (1.2) and \( s > 0 \) be some parameter. Thanks to the Kac-Rice formula, moments of \( \mathcal{L}(f_{n}, s) \) can be expressed in terms of the \emph{restricted} moments of covariance function
\[
r_{n}(x, y) = \int_{S^1} e(\langle x - y, \lambda \rangle) d\mu_{n}(\lambda)
\]
where \( S^1 \subset \mathbb{R}^2 \) is the unit circle. That is, for \( l \geq 2 \) we are interested in asymptotically evaluating
\[
\int_{B(s)} r(x)^{l} dx = \frac{1}{N^l} \sum_{\xi_1, \ldots, \xi_l} \int_{B(s)} e(\langle \xi_1 + \ldots + \xi_l, x \rangle) dx
\]
where \( B(s) \) is the ball centred at 0 of radius \( s \). Separating the terms with \( \xi_1 + \ldots + \xi_l = 0 \), we obtain
\[
\int_{B(s)} r(x)^{l} dx = \pi s^{2} \frac{\#S(l, n)}{N^l} + \frac{2\pi s^{2}}{N^l} \sum_{|\xi_1 + \ldots + \xi_l| > 0} \frac{J_{l}(s|\xi_1 + \ldots + \xi_l|)}{s|\xi_1 + \ldots + \xi_l|}
\]
where $S(l, n)$ is as in (1.5) and $J_1(\cdot)$ is the Bessel function of the first kind.

Spectral correlations have been studied by Bombieri and Bourgain [8] who showed that

$$\# S(2l, n) = \frac{(2l)!}{l!} N^l(1 + o(1))$$  \hspace{1cm} (1.9)

for almost all $n \in S$, while $S(l, n) = \emptyset$ for $l$ odd by congruence obstruction modulo 2, see also [25] and section 2.1 below for a more detailed discussion. Since $J_1(T) \ll T^{1/2}$ for $T$ large enough, the second term in (1.8) would asymptotically vanish if $Q(l, n, s^{-1}) = \emptyset$, which is in particular the case if $s = O(1)$. Hence, the phase-transition for $\mathcal{L}(f_n, s)$ can be compared to the change in the asymptotic law of the second term in (1.8).

In light of Theorem 1.2, we expect that $A_0 = 2 \log 2 = 1.3862\ldots$. Indeed, thanks to the calculations in [4], we need to control the asymptotic in (1.8) only for $l = 2, 4, 6$. Thus, given some $l \geq 2$, Theorem 1.2, with $\alpha = N^{-l/2}$, suggests that there are at most $o(N^{l/2})$ tuples $(\xi_1, \ldots, \xi_l)$ such that $s|\xi_1 + \ldots + \xi_l| < N^{l/6}$, where $s = N^{2l/3}/n^{l/2}$. For the remaining $l$-tuples, we have that $(J_1(s|\xi_1 + \ldots + \xi_l|)/s|\xi_1 + \ldots + \xi_l|) = o(N^{l/2})$. Thus, using the bound $J_1(T)/T = O(1)$ for the former set of tuples and the bound $J_1(T)/T \ll T^{-3/2}$ for the latter, the second term in (1.8) is negligible compared to the first one. Taking $l = 6$ and bearing in mind that $N \asymp \log(n)^{\log 2 \pm 3\epsilon}$, we obtain $s = (\log(n))^{2\log 2 + 3\epsilon}/n^{l/2}$.

### 1.5. Related work.

Bourgain and Rudnick [10] first studied length 2 quasi-correlations and showed that $Q(2, n, n^{1/2-\epsilon}) = \emptyset$ for a density one subset of $n \in S$. Subsequently, Granville and Wigman [22] showed that

$$\# \{ n \in S : n \leq X \text{ and } Q(2, n, K) \neq \emptyset \} = C\sqrt{X}K\left( (2 \log K)^{1/2} + O(1) \right),$$

for some explicit $C > 0$, this implies that

$$Q(2, n, n^{1/2}/\Psi(n) \log n) = \emptyset$$  \hspace{1cm} (1.10)

for almost all $n \in S$ and any function $\Psi(n) \to \infty$ as $n \to \infty$. Theorem 1.1 refines (1.10) to $Q(2, n, n^{1/2}/\log n^{2+2\epsilon}) = \emptyset$. This is also directly related to the question of estimating the number of lattice points $\xi_i$ on an arc of length $n^\delta$ for $\delta > 0$. Cilleruelo and Cordoba [14] showed that there are at most $O(1)$ such lattice points if $\delta < 1/4$ and Cilleruelo and Granville [15] conjectured that this remains true for every $\delta < 1/2$. Theorem 1.1 implies that there are at most 2 lattice points on any arc of length at most $n^{1/2}/\log n^{2+2\epsilon}$ on a generic circle of radius $\sqrt{n}$. Furthermore, Benatar, Marinucci and Wigman [4] showed that $Q(l, n, n^{1/2-\epsilon}) = \emptyset$ for almost all $n \in S$. Theorem 1.1 not only refines their bound, but also gives an explicit dependence of $K$ on $l$ which is essential in the study of toral eigenfunctions at small scales.

The proof of Theorem 1.4 relies on the quantitative convergence of the spectral measure $\mu_n$ to $\mu$, given by Erdős-Hall [18] and Kátaï- Környei [24], see also Theorem 2.7 below, and a recent result of Beliaev-Maffucci [2]. In this scenario, the Continuous Mapping Theorem suggests that convergence of the spectral measure, in the weak sense, implies convergence in distribution of the nodal length. This principle has already been rigorously implemented by Granville and Wigman [21] for trigonometric polynomials and by Todino [41] for spherical harmonics in the two dimensional sphere. The author was also recently communicated that Dierickx, Nourdin, Peccati and Rossi [16] showed that the said principle applies in a quite general scenario: They showed, from the appropriate convergence of covariances, the convergence to $\mathcal{L}(f_\mu)$, in mean square and distribution, of the nodal length of Gaussian monochromatic random waves on Riemann surfaces without conjugate points. As a consequence of their method, with the notation of Theorem 1.4, they found the
variance and the distribution of $\mathcal{L}(F_n)$ to agree with $\mathcal{L}(F_\mu)$ for $B = \frac{1}{18} \log(\pi/2)$; thus giving a sharper, value for $B$ in our Theorem 1.4.

Finally, in light of the techniques in [21], it is conceivable that the value of $B$ can be increased further, to maybe $B = \log(\pi/2)/2 = 0.2257\ldots$, using Crofton’s formula and Hurwitz’s theorem. Nevertheless, since this new value for $B$ would still be far from our expected $A_0$, we opted for a short and, in our view, elegant proof based on the stability of the nodal set, as in [34], and a quantitative version of the Continuous Mapping Theorem.

1.6. Notation. Let $u \to \infty$ be some parameter, we say that the quantity $X = X(u)$ and $Y = Y(u)$ satisfy $X \ll Y, X \gg Y$ if there exists some constant $C$, independent of $u$, such that $X \leq CY$ and $X \geq CY$ respectively. If $X \ll Y$ and $Y \ll X$, we write $X \asymp Y$. We also write $O(X)$ for some quantity bounded in absolute value by a constant times $X$ and $X = o(Y)$ if $X/Y \to 0$ as $u \to \infty$, in particular we denote by $o(1)$ any function that tends to 0 (arbitrarily slowly) as $u \to \infty$. We denote by $B(s)$ the (open) ball centred at 0 and by $\overline{B}(s)$ the closure of $B(s)$. When the specific radius is unimportant, we simply write the ball as $B$ and $\frac{1}{2}B$ for the concentric ball with half the radius. Moreover, for a positive integer $n$, we denote by $\omega(n)$ the number of its prime factors without multiplicity. Furthermore, given some $k \geq 0$ and some $k$-times differentiable function $f : B \to \mathbb{R}$, we denote by $\|f\|_{C^k(B)} = \sum_{m=0}^{k} \sup_{x \in B} |f^{(m)}(x)|$, where $f^{(m)}(x)$ is the $m$-th derivative. Finally, we denote by $\Omega$ an abstract probability space where every random object is defined.

2. Number theoretic background

2.1. An equivalent formulation of Theorem 1.1. Given $n \in S$, we can express the representations $\xi_i$ of $n$ as products of prime ideal in $\mathbb{Z}[i]$. Let $n = 2^{\alpha_2} \prod_k P_k^{\alpha_k} \prod_v Q_v^{\beta_v}$, where $P_k$ and $Q_v$ are primes $p \equiv 1 \pmod{4}$ and $q \equiv 3 \pmod{4}$ and $\alpha$’s and $\beta$’s are positive integers. Let $(n) \subset \mathbb{Z}[i]$ be the ideal generated by $n$, then, by unique factorisation of ideals in $\mathbb{Z}[i]$, bearing in mind that primes $p \equiv 1 \pmod{4}$ split and primes $q \equiv 3 \pmod{4}$ are inert, we have

$$(n) = Z^{2\alpha_2} \prod_k (P_k \overline{P_k})^{\alpha_k} \prod_v Q_v^{\beta_v},$$

where $Z$ is an ideal above 2, $P_k$ an ideal above $p_k$ and $Q_v$ an ideal above $q_v$. Thus, if $(n) = (x + iy)(x - iy)$ for some $x, y \in \mathbb{Z}$, then

$$(x + iy) = Z^{\alpha_2} \prod_k P_k^{\gamma_{kj}} \prod_v Q_v^{\beta_v - \gamma_{kj}} \prod_v Q_v^{\beta_v/2}.$$  \hspace{1cm} (2.1)

for some $0 \leq \gamma_{kj} \leq \alpha_k$. It follows that the $\beta$’s must be even and the representations $\xi_i$ of $n$ are in one to one correspondence with ideal of the form (2.1). Therefore, taking into account the symmetries $\xi \to -\xi$ and $x + iy \to y + ix$, we have

$$N(n) = 4 \prod_k (\alpha_k + 1) \ll \exp \left( O \left( \frac{\log n}{\log \log n} \right) \right).$$  \hspace{1cm} (2.2)

where the inequality follows from the divisor bound. Moreover, by (2.1), we see that the factor $Z^{\alpha_2} \prod_v Q_v^{\beta_v/2}$ is common to every representation. This has the effect of rotating the lattice points, but it does not affect their spacial distribution. Thus, it does not effect the set of solutions to (1.6). Hence, we can restate Theorem 1.1 as follows.
**Theorem 2.1.** Let \( S' := \{ n \in S : p \mid n \text{ then } p \equiv 1 \pmod{4} \} \), \( l \geq 2 \) be an integer, \( \epsilon > 0 \) and \( c(l, \epsilon) = c(l) \) be as in Theorem 1.1. Then there exists a density one subset of \( n \in S' \) such that

\[
Q(l, n, n^{1/2}/(\log n)^{c(l)}) = \emptyset.
\]

**2.2. Spectral correlations.** Recall that \( S(l, n) = \{(\xi_1, \ldots, \xi_l) : \xi_1 + \ldots + \xi_l = 0\} \) and that, by congruence obstruction modulo \( 2 \), \( S(l, n) = \emptyset \) if \( l \) is odd. If \( l \) is even, we have the “diagonal” solutions given by \( \xi_1 = -\xi_2, \ldots, \xi_{l-1} = -\xi_l \), thus \( S(l, n) \gg N^{1/2} \). For \( l = 2 \) the only solutions are \( \xi_1 = -\xi_2 \), thus \( S(2, n) = N \). For \( l = 4 \) Zygmund [43] observed that the only solutions are \( \xi_1 = -\xi_2 \) and \( \xi_3 = -\xi_4 \), therefore

\[
S(4, n) = 3N^2 + O(N).
\]

For \( l = 6 \) Bourgain [25, Theorem 2.2] showed that \( S(6, n) = o(N^4) \). Subsequently, Bombieri and Bourgain [8] gave the bound

\[
S(6, n) \ll N^{7/2}.
\]

Finally, using the deep work of Evertse-Schlickewei-Schmidt [19] on additive relations in multiplicative subgroups of \( \mathbb{C}^* \) of bounded rank, see [8, Theorem 5] and [9, Lemma 5], we can explicitly construct sub-sequences of \( n \in S \) which satisfy (1.9):

**Lemma 2.2.** Let \( n = \prod_{i=1}^l p_i^{\alpha_i} \in S' \) and \( l \geq 2 \) be an even integer. If \( \sum_i \log(\alpha_i + 1)/r \to \infty \), then the number of solutions to (1.9) is

\[
\frac{l!}{2(l/2) \cdot (l/2)!} N^{l/2} + O(N^\gamma)
\]

for some \( 0 < \gamma < 1/2 \).

**2.3. Lattice points and geometry of numbers.** In this section we collect some facts which will be used thorough the rest of the article. By Landau’s Theorem, see for example [20, Theorem 14.2], there exists some explicit constant \( c > 0 \) such that

\[
\#\{n \in S' : n \leq X\} = c \frac{X}{\sqrt{\log X}} (1 + o(1))
\]

(2.4)

Thanks to (a weak version of) the Erdös-Kac Theorem, see for example [40, Part III Chapter 3], we have

**Lemma 2.3** (Erdös-Kac). Let \( \epsilon > 0 \), then, for a density one subset of \( n \in S' \), we have

\[
\frac{1}{2} \log \log n(1-\epsilon) \leq \#\{p \mid n : p \equiv 1 \pmod{4}\} \leq \frac{1}{2} \log \log n(1+\epsilon).
\]

where the primes are counted without multiplicity. In particular, via (2.2), we have

\[
N(n) = N \asymp (\log n)^{\frac{3}{2} - \epsilon}.
\]

As another consequence of (a slightly stronger version of) the Erdös-Kac Theorem, we can also control the size of the prime in the factorisation of a generic integer \( n \in S' \), see for example [40, Part III Chapter 3, Theorem 8 and Theorem 9] for a standard derivation of the following fact from the Erdös-Kac Theorem.

**Lemma 2.4.** Let \( n \in S' \) and let \( p_1 < p_2 < \ldots < p_r \) be its prime factors. Then for a density one subset of \( n \in S' \) we have

\[
\sup_{\log \log \log n < k < r} \left| \frac{2^{-1} \log \log p_k - k}{\sqrt{k \log k}} \right| \leq 3/2.
\]
In particular, by Lemma 2.3, we have

\[ p_r \geq \exp((\log n)^{1/3}) \]

**Remark 2.5.** Lemma 2.4 is not sharp, in particular the constant 1/3 can be replace with a larger constant, but it will suffice for our purposes.

We will also need the following result of Kubilius [26] about Gaussian primes, which are primes \( P \subset \mathbb{Z}[i] \) such that \( |P|^2 = p \) with \( p \equiv 1 \pmod{4} \).

**Lemma 2.6 (Kubilius).** Let \( \theta_1, \theta_2 \in [0, 2\pi] \). Then, the number of Gaussian primes in the sector \( \arg(P) \in [\theta_1, \theta_2] \) such that \( |P|^2 \leq X \) is

\[
\frac{2}{\pi} (\theta_1 - \theta_2) \int_2^X \frac{dx}{\log x} + O(X \exp(-c\sqrt{\log X})).
\]

Finally, we will need the following result about the distribution of lattice points for generic \( n \). Recall the spectral measure \( \mu_n \) in (1.3) and \( \mu \), the Lebesgue measure on the interval \([0, 1]\), then have the following theorem, see [18, 24].

**Theorem 2.7 (Erdős-Hall).** Let \( \kappa = \frac{1}{2} \log \frac{\pi}{2} \) and \( \epsilon > 0 \). Then, for a density one subset of \( n \in S \), we have

\[
\sup_{0 < a < b < 1} |\mu_n(a, b) - \mu(a, b)| \leq (\log n)^{-\kappa + \epsilon}.
\]

### 3. Proof of Theorem 1.1

The argument in this section is inspired by [8, Theorem 14]. As discussed in section 2.1 it is enough to prove Theorem 2.1. To ease the exposition we divide the proof into two parts: \( n \) square-free and \( n \) not square-free. We begin by proving Theorem 2.1 in the square-free case.

#### 3.1. Proof of Theorem 2.1 for the square-free case.

During the proof of Theorem 2.1 we will need the following direct consequence of the Landau’s Theorem (2.4):

**Lemma 3.1.** Let \( X > 1 \) be some large parameter and \( S' \) be as in section 2.1, then we have

\[
\sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{1}{n \log(X/n)} \ll (\log X)^{-1/2}.
\]

**Proof.** Using the expansion \( \log(X/n)^{-1} = (\log X)^{-1} \left(1 + O(\log(n)/\log X)\right) \) valid for \( n \leq X/100 \) say, we have

\[
\sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{1}{n \log(X/n)} \ll \frac{1}{\log X} \sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{1}{n} + \frac{1}{(\log X)^2} \sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{\log n}{n}
\]

(3.1)

By partial summation using (2.4), we have

\[
\sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{1}{n} \ll (\log X)^{1/2}
\]

(3.2)

\[
\sum_{\substack{n \in S' \\cap \mathbb{N} \leq X/100}} \frac{\log n}{n} \ll (\log X)^{3/2}
\]

Thus, the lemma follows by inserting (3.2) into (3.1). \( \square \)
Let $X > 1$ be some (large) parameter, for a positive integer $n$, which we assume to be always square-free, let $\omega(n)$ be the number of its prime divisors without multiplicity. Moreover, given some integer $l \geq 2$ and $\epsilon > 0$, we define

$$\Phi(n, l) = \Phi(n) = (\log n)^{c(l)} \quad (3.3)$$

where $c(2, \epsilon) = c(2) = \log 2 + \epsilon$, $c(3, \epsilon) = c(3) = 3 \log 2 + \epsilon$ and $c(l, \epsilon) = c(l) = l \log 2 / 2 + c([l/2]) + \epsilon$ and $[l/2]$ represent the largest integer smaller than $l/2$. Finally, let $F_r(X, l, \epsilon) = F_r(l)$ be the set of $n \in S'$ such that $\omega(n) = r$ and

$$Q(l, n, n^{1/2}/\Phi(n)) \neq \emptyset,$$

but for all divisors $d|n$ and $d \neq n$, we have

$$Q(l, d, d^{1/2}/\Phi(d)) = \emptyset.$$

We are going to prove following bound:

**Proposition 3.2.** Let $l \geq 2$ be a positive integer and $\epsilon > 0$, then for a density one subset of $n \in S'$, we have

$$|F_r(X, l, \epsilon)| \ll_l \frac{X}{(\log X)^{1/2 + r/2}},$$

uniformly for all $r \geq 1$.

**Proof.** First, by Lemma 2.3 and Lemma 2.4, we may assume that for all $\epsilon > 0$ and all $n \in F_r(l)$ the following hold:

$$r \leq \frac{1}{2} \log \log X(1 + \epsilon/4) \quad p_r \gg \exp((\log X)^{1/3}) \quad (3.4)$$

where $p_r$ is the largest prime divisor of $n$. We are now going to prove the Proposition by induction on $l$.

**Base case, $l = 2$.** Let $n \in F_r(2)$, then there exist two lattice points $\xi_1 \neq \xi_2$ such that $|\xi_1|^2 = |\xi_2|^2 = n$ and

$$0 < |\xi_1 - \xi_2| < n^{1/2}/\Phi(n). \quad (3.5)$$

Let $p_1 < p_2 < ... < p_r$ be the prime factors of $n$. With the same notation as in (2.1) and bearing in mind that $n$ is square-free, we may write

$$\xi_i = \prod_{k=1}^r P_k^{\gamma_{ki}} P_k^{1-\gamma_{ki}}$$

where $\gamma_{ki} \in \{0, 1\}$ and $P_k$ is the Gaussian prime above $p_k$ (which we select by insisting that $\text{Arg}(P_k) \in (0, \pi/2)$). Fixing $P_1, ..., P_{r-1}$, we can rewrite (3.5), in view of the fact that $n \in F_r(2)$, as

$$0 < |P_r a - \overline{P_r} b| \leq n^{1/2}/\Phi(n) \quad (3.6)$$

for some $a, b \in \mathbb{Z}[i]$ such that $|a| = |b| = (n/p_r)^{1/2}$. Thus, dividing both sides of (3.6) by $n^{1/2}$, we deduce

$$\text{Arg}(P_r) \in \left[\theta - \frac{1}{\Phi(n)}, \theta + \frac{1}{\Phi(n)}\right] := I_{a,b}(n) = I(n). \quad (3.7)$$

for some $\theta = \theta(a, b)$. Thus, given $p_1 < ... < p_{r-1}$, there are at most $2^{2(r-1)}$ choices of for $a$ and $b$, once these are fixed we obtain that $p_r$ is a prime of modulus $\exp((\log X)^{1/3}) \ll p_r \leq$
\[ X/ \prod_{k} p_k \] and the argument of the Gaussian prime \( P_r \) above it satisfies (3.7). Hence, we have the following key bound

\[ \#\{n \in \mathcal{F}_r(l) : n \leq X \text{ and } \mathcal{O}(l, n, n^{1/2}/\Phi(n)) \neq \emptyset\} \leq \sum_{p_1 < \cdots < p_r} 2^{2r} \sum_{|P_r|^2 \leq X/\prod_{k} p_k} \sum_{|P_r|^2 \gg \exp((\log X)^{1/3})} 1. \] (3.8)

Since \( \Phi(n) \gg \Phi(X) \), letting \( Y = X/\prod_{k} p_k \) and bearing in mind that \( \int_{2}^{Y} (\log x)^{-1} dx = Y ((\log Y)^{-1} + O((\log Y)^{-2})) \), Lemma 2.6 gives

\[ \sum_{|P_r|^2 \leq X/\prod_{k} p_k} 1 \ll \frac{Y}{\log Y} + Y \exp(-c\sqrt{\log Y}). \] (3.9)

We first consider the first term on the right hand side of (3.9): let \( Y = |P_r|^2 \gg \exp((\log X)^{1/3}) := Z \), then

\[ \sum_{p_1 < \cdots < p_r} \frac{Y}{\log Y} \ll X \sum_{p_1 < p_2 < \cdots < p_r} \prod_{k} p_k \log(X/\prod_{k} p_k) \leq \sum_{n \in S'} \frac{1}{n \log(X/n)} \] (3.10)

where we have extended the inner sum by lifting the restriction on the number of prime factors of \( n \). Hence, combining Lemma 3.1 and (3.10), we have

\[ \frac{1}{\Phi(X)} \sum_{p_1, \ldots, p_{r-1}} \frac{Y}{\log Y} \ll \frac{X}{(\log X)^{1/2}\Phi(X)} \] (3.11)

We now consider the second term on the right hand side of (3.9):

\[ \sum_{p_1, \ldots, p_{r-1}} Y \exp(-c\sqrt{\log Y}) \leq X \sum_{\prod_{k} p_k \leq X/Z} \exp(-c \log(X/\prod_{k} p_k)^{1/2}) \]

\[ \frac{1}{\prod_{k} p_k} \leq X \exp(-c \log(Z)^{1/2}) \sum_{n \in S'} \frac{1}{n} \]

\[ \ll X \exp(-c \log(X)^{1/6}) \] (3.12)

which is smaller than (3.10). Hence the base case follows from inserting (3.9) into (3.8), using (3.11) and (3.12) and noticing that (3.4) gives \( 2^{2r}/\Phi(X) \ll (\log X)^{-c/2} \).

**Induction step.** Let us assume that the Proposition holds for all \( l' < l \), then, repeating the argument in the base case and maintaining the same notation, we find the following: given \( n \in \mathcal{F}_r(l) \) with prime factorisation \( p_1 < \cdots < p_r \), let \( P_i \) be the Gaussian prime above \( p_i \), then

\[ 0 < |P_r a - \overline{P_r} b| \leq n^{1/2}/\Phi(n, l) \] (3.13)

for some non-zero \( a, b \in \mathbb{Z}[i] \) depending \( P_1, \ldots, P_{r-1} \). Now, we claim that for all but at most \( O(X/(\log X)^{1/2+c/2}) \) integers \( n \in S' \) up to \( X \), we have

\[ |a| \geq \left( \frac{n}{p_r} \right)^{1/2} \cdot \frac{1}{\Phi(n/p_r, [l/2])} \quad \text{or} \quad |b| \geq \left( \frac{n}{p_r} \right)^{1/2} \cdot \frac{1}{\Phi(n/p_r, [l/2])} \] (3.14)
Indeed, since either $a$ or $b$ is the sum of at most $\lfloor l/2 \rfloor$ lattice points on the circle of radius $(n/p_r)^{1/2}$, the induction hypothesis, with $2\epsilon$ instead of $\epsilon$, gives that the number of exceptions is at most
\[
\sum_{p_r \leq X^{2 \epsilon}} F_{r-1} \left( \frac{X}{p_r}, \lfloor l/2 \rfloor \right) \ll X \sum_{p \leq X^{2 \epsilon}} \frac{1}{p^{(\log X/p)^{1/2+\epsilon}}}
\ll \frac{X}{(\log X)^{1/2+\epsilon}} \left( \sum_{p \leq X} \frac{1}{p} + \frac{1}{\log X} \sum_{p \leq X} \log p \right)
\ll \frac{X \log \log X}{(\log X)^{1/2+\epsilon}} \ll \frac{X}{(\log X)^{1/2+\epsilon/2}}
\]
where, in the second but last line, we have used the expansion $(\log(X/p))^{1/2+\epsilon} = \log X^{1/2+\epsilon}(1 + O((\log p/\log X))$. Thus, we have proved the claim and we may assume that (3.14) holds.

Without loss of generality we might assume that $|a| \geq |b|$, so that, inserting (3.14) into (3.13), we obtain
\[
0 < \left| P_r - \overline{P_r} \right| \leq p_r^{1/2} \Phi(n/p_r, \lfloor l/2 \rfloor) / \Phi(n, l) \leq p_r^{1/2} \Phi(n, \lfloor l/2 \rfloor) / \Phi(n, l) =: p_r^{1/2} \Psi(n).
\]
Following step by step the computations in the base case with $\Phi$ substituted by $\Psi$, bearing in mind the number of exceptions in claim (3.14), we deduce that
\[
F_r(l) \ll 2^{1/r} \frac{X}{(\log X)^{1/2}} \cdot \frac{1}{\Psi(X)} + O \left( \frac{X}{(\log X)^{1/2+\epsilon/2}} \right)
\]
which, in light of (3.3) and (3.4), concludes the induction. □

We are finally ready to prove Theorem 2.1:

**Proof of Theorem 2.1 for the square-free case.** Let $l \geq 2$ and $X > 1$ be given, if $n \in S'$ less than $X$ is such that $Q(l, n, n^{1/2}/\Phi(n)) \neq \emptyset$ then $n \in F_r(l)$ for some $r \geq 1$. Moreover, by Lemma 2.3, we may assume that $r \leq \log \log X$. Hence, summing the bound in Proposition 3.2 over the all possible values of $r$, we obtain the Theorem. □

3.2. **Proof of Theorem 1.1 for the non square-free case.** We are now going to prove Theorem 2.1 in the non-square free case. We need the following standard lemma:

**Lemma 3.3.** Let $n \in S'$ and let $\omega(n)$ be the number of prime factors of $n$ without multiplicity and $\Omega(n)$ be the number of prime factors of $n$ with multiplicity, then
\[
\sum_{n \leq X} \left| \Omega(n) - \omega(n) \right| \ll \frac{X}{\sqrt{\log X}}
\]

**Proof.** Using Landau’s Theorem (2.4) and following similar calculations as in the proof of Proposition 3.2, we obtain
\[
\sum_{n \leq X} \left| \Omega(n) - \omega(n) \right| \leq X \sum_{p \text{ prime}} \sum_{i \geq 2} \frac{1}{p^{\lfloor \log(X/p) \rfloor^{1/2}}} \ll \frac{X}{\sqrt{\log X}} \sum_{p} \frac{1}{p^2} \ll \frac{X}{\sqrt{\log X}}
\]

We are now ready to begin the proof of Theorem 2.1 in the non-square free case.
Proof of Theorem 2.1 for \( n \) not square-free. First, let us define \( \tilde{F}_r(l) \) to be \( F_r(l) \) but \( \omega(n) \) is substituted by \( \Omega(n) \). Observe that, by Lemma 3.3 and with the same notation, we have, for a density of subset of \( n \in S' \), \( \Omega(n) - \omega(n) \leq \log \log n \). Thus, we may assume that \( r \leq 2 \log \log n \). Thus, as we have seen in the proof of Theorem 2.1 for the square-free case, it is enough to show that

\[
\tilde{F}_r(l) \ll X \log(X)^{1/2+\epsilon/2}.
\] (3.15)

To prove (3.15), we may assume, by again Lemma 3.3, that at most \( O(\log \log \log n) \) of the prime divisors of \( n \) have multiplicity. Therefore, given \( n \in S' \) and letting \( p_1 < \ldots < p_r \) be its prime factors, by Corollary 2.4 we may assume that there exists some \( s \geq \log \log n / 10 \) such that

\[
p_s \gg \exp((\log X)^{1/3})
\]

and moreover \((p_s)^2\) does not divide \( n \). Fixing \( s \), the proof now proceeds step by step as the proof of Proposition 3.2 in the square-free case. Finally, summing the bound in Proposition 3.2 over the \( \log \log n \) choices for \( s \) gives (3.15), up to changing the values of \( \epsilon \).

4. Random model for lattice points, proof of Theorem 1.2

As discussed in section 2.1, we may assume that a generic integer \( n \in S' \) has \( \omega(n) \approx \log \log n \) prime factors, most of which are not repeated by Lemma 3.3, and the distribution of the angle of Gaussian prime is uniform in \([0, 2\pi)\). Thus, we may model representations of a generic integer \( n \in S' \) as

\[
\xi_i = \exp\left(2\pi i \sum_{k=1}^{\omega(n)} \eta_{ik} \theta_k\right)
\] (4.1)

where \( \theta_k \)'s are i.i.d random variables uniformly distributed on \([0, 1)\) and \( \eta_{ik} \in \{-1, 1\} \) are deterministic, see also [22, Remark 3.3]. This gives \( 2^{\omega(n)} \approx N \) representations of \( n \). Given \( l \geq 2 \), we define the random variables

\[
X_i = \xi_{i_1} + \ldots + \xi_{i_l}
\]

for \( 1 \leq i_j \leq N \). Before proving Theorem 1.2, we need two preliminary results.

4.1. Distribution and independence of the \( \xi_i \). In this section we show that the random \( \xi_i \) as in (4.1) have the same distribution and are “generically” independent.

Lemma 4.1. Let \( \xi_i \) be as in (4.1) and let \( r \in \mathbb{Z} \) be some integer, then

\[
\mathbb{E}[\xi_i^r] = \begin{cases} 
1 & r = 0 \\
0 & r \neq 0
\end{cases}.
\]

Proof. Since the \( \theta_k \) are independent, we have

\[
\mathbb{E}[\xi_i^r] = \prod_k \int_0^1 e(r \eta_{ik} \theta_k) \, d\theta_k
\]

which gives the lemma as \( \int e(r \eta_{ik} \theta_k) \, d\theta_k = 0 \) unless \( r = 0 \). □

Lemma 4.2. Let \( l \geq 2 \) and let \( \xi_1, \ldots, \xi_l \) be as in (4.1). Then \( \xi_1, \ldots, \xi_l \) are independent for all but \( o(\omega(n) \to \infty (2^{\omega(n)!}) \) choices of \( l \)-tuples \( (\xi_1, \ldots, \xi_l) \).
Proof. Since $|\xi_i| \leq 1$ surely, the joint distribution of $(\xi_1, \ldots, \xi_l)$ is fully determined by its moments. Thus, it is enough to prove that given $l$ integers $m_1, \ldots, m_l$, we have

$$E \left[ \prod_{i=1}^{l} \xi_i^{m_i} \right] = \begin{cases} 1 & m_1 = m_2 = \ldots = m_l = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (4.2)$$

Observe that

$$E \left[ \prod_{i=1}^{l} \xi_i^{m_i} \right] = E \left[ \exp \left( 2\pi i \sum_{k} (\sum_{i=1}^{l} m_i \eta_{ik}) \phi_k \right) \right], \quad (4.3)$$

integrating, we obtain that $m_1 \eta_{1k} + \ldots + m_l \eta_{lk} = 0$ for all $k$. Therefore $(m_1, \ldots, m_l)$ is in the kernel of the matrix $\{\eta_{ik}\}$ for $1 \leq i \leq l$ and $1 \leq k \leq \omega(n)$. Hence, it is enough to prove the following claim:

**Claim 4.3.** Consider the $l \times \omega(n)$ random matrix with entry $\eta_{ik} = 1$ with probability $1/2$ and $\eta_{ik} = -1$ with probability $1/2$. Then with probability greater than $1 - o_{\omega(n) \to \infty}(1)$, the matrix $\{\eta_{ik}\}$ for $1 \leq i \leq l$ and $1 \leq k \leq \omega(n)$ has rank $l$.

Proof. Consider the first two rows $\{\eta_{ik}\}$ and $\eta_{2k}$, outside a set $\Omega_1$ of probability at most $2 \cdot 2^{-\omega(n)}$ there exists $k_1$ and $k_2$ such that $\eta_{1k_1} = \eta_{2k_1}$ and $\eta_{1k_2} \neq \eta_{2k_2}$. Therefore, applying row and columns operations, we reduce the matrix as

$$\begin{bmatrix} 1 & 1 & \ldots \\ -1 & 1 & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix} \to \begin{bmatrix} 1 & 1 & \ldots \\ 0 & 2 & \ldots \\ 0 & 0 & \ldots \\ 0 & 0 & \ldots \end{bmatrix}. \quad (4.3)$$

Consider now rows $\{\eta_{3k}\}$ and $\{\eta_{4k}\}$. By the above, we have $\eta_{3k} = \eta_{4k} = 0$ for $k = 1, 2$, and for $k \geq 3$ the entries are $\{-2, 0, 2\}$ with probability $1/4, 1/2$ and $1/4$ respectively. Therefore, outside a set $\Omega_2$ of probability at most $2 \cdot 2^{-\omega(n)+2}$, there exists some $k_3, k_4 \geq 3$ such that $\eta_{3k_3} = \eta_{4k_3} \neq 0$ and $\eta_{3k_4} \neq \eta_{4k_4}$. Therefore, applying row and columns operations, we have one of the following matrices

$$\begin{bmatrix} 1 & 1 & \ldots \\ 0 & 2 & \ldots \\ 0 & 0 & 2 & \ldots \\ 0 & 0 & 0 & 2 & \ldots \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 & \ldots \\ 0 & 2 & \ldots \\ 0 & 0 & 2 & 2 & \ldots \\ 0 & 0 & 2 & -2 & \ldots \end{bmatrix} \to \begin{bmatrix} 1 & 1 & \ldots \\ 0 & 2 & \ldots \\ 0 & 0 & 2 & 2 & \ldots \\ 0 & 0 & 4 & \ldots \end{bmatrix}. \quad (4.4)$$

Now consider rows $\{\eta_{5k}\}$ and $\{\eta_{6k}\}$. By the above, we have $\eta_{5k} = \eta_{6k} = 0$ for $k \leq 4$, and for $k \geq 5$ the entries are $\{-4, -2, 0, 2, 4\}$ with probability $1/8, 2/9, 1/3, 2/9$ and $1/8$ respectively. Since $l$ is fixed, we can repeat the above argument to find subsets $\Omega_1, \ldots, \Omega_{[l/2]}$ such that $P(\Omega_i) = o(1)$ and outside $\bigcup_i \Omega_i$, we have $l$ linearly independent columns, as required.

Having proved the claim, the Lemma follows.

4.2. **Proof of Theorem 1.2.** We are finally ready to prove Theorem 1.2.

Proof of Theorem 1.2. Since $|\xi_i| \leq 1$ surely, we also have $|\Re(\xi_i)| = |2^{-1}(\xi_i + \overline{\xi_i})| \leq 1$ and $|\Im(\xi_i)| = |2^{-1}(\xi_i - \overline{\xi_i})| \leq 1$. Therefore, by Levi’s continuity Theorem [7, Theorem 26.2], the distribution of $\xi_i$ is fully determined by its integer moments. Then Lemma 4.1 implies that each $\xi_i$ have the same distribution and in particular they are uniformly distributed.
on the unit circle $S^1 \subset \mathbb{R}^2$. Given $l \geq 2$ suppose that $\xi_1, \ldots, \xi_l$ are i.i.d. random variables and identify $S^1$ with the unit interval $[-1/2, 1/2)$, then the random variable

$$X_1 = \xi_1 + \ldots + \xi_l$$

has Irwin–Hall distribution. In particular, the density function of $X_1$ is piece-wise analytic and depends only on $l$. Therefore, given any $\alpha > 0$, by Taylor’s expansion, we have

$$P(|X_i| \leq \alpha) \approx l \alpha (1 + O(\alpha)). \quad (4.4)$$

By Lemma 4.2 we have that $\xi_1, \ldots, \xi_l$ are independent for all but $o_{\omega(n) \rightarrow \infty}(2^{\omega(n)}l)$ choices of $\xi_1, \ldots, \xi_l$. Thus, bearing in mind (4.4), we obtain

$$\mathbb{E}[\#\{X_i : |X_i| \leq \alpha\}] \approx l 2^{\omega(n)}l(1 + O(\alpha^2))(1 + o_{\omega(n) \rightarrow \infty}(1))$$

which concludes the proof.

5. Proof of Theorem 1.3

The proof of Theorem 1.3 essentially follows the proof of the main Theorem in [4] and, for the sake of completeness, we summarise here the main steps. The main difference is that we explicitly construct sequences of $n \in S$ for which we can control the distribution of lattice points on $\sqrt{n}S^1$. This is the content of the next section.

5.1. Limit points and spectral correlations. In this section, we prove the following proposition:

**Proposition 5.1.** Let $w \in [0,1]$ and $l \geq 2$ be an integer. Then, there exists a sub-sequence of integers $n \in S'$ such that $N \rightarrow \infty$ as $n \rightarrow \infty$, $\hat{\mu}_n(4) \rightarrow w$, (1.9) hold and $\mathcal{Q}(l, n, c(n)n^{1/2}) = \emptyset$ for any function $c(n) \rightarrow 0$ arbitrarily slowly.

In order to prove Proposition 5.1, we need a two preliminary results. The first is Lemma 2.2 in section 2.1. The second is a standard tool to control the size of exponential sums, see [35].

**Lemma 5.2 (Remez’ inequality).** Let $F(t) = \sum_{i=1}^{J} a_\eta e(\eta_i \cdot t)$ where $t \in \mathbb{R}$, $J \in \mathbb{N}$, $a_\eta \in \mathbb{C}$ and suppose that $\eta_i \in \mathbb{R}$ are distinct. Then, for any interval $B \subset \mathbb{R}$ and any sub-interval $I \subset B$, we have

$$\sup_I |F| > \left( \frac{|I|}{|B|} \right)^{J-1} \sup_B |F|.$$  

for some explicit $C > 0$ independent of $F$.

We are now ready to prove Proposition 5.1.

**Proof of Proposition 5.1.** We pick integers of the form $n = p^m \cdot q$ for some primes $p, q \equiv 1 \pmod{4}$ and $m \geq 1$ to be chosen later. Let $\theta_p, \theta_q$ be the angle of the Gaussian prime lying above $p$ and $q$ respectively. Observe that, by Lemma 2.2 and (2.2), if $m \rightarrow \infty$ as $n \rightarrow \infty$, then both (1.9) and $N \rightarrow \infty$ are satisfied. So we assume that $m$ is a sufficiently slow growing function of $n$ to be specified later. The rest of the proof relies on two claims.

**Claim 5.3.** Let $\epsilon > 0$. Then, there exist $\delta_1 = \delta_1(\epsilon) > 0$ and some interval $I = I(\epsilon, m) \subset [0, 2\pi)$ such that for all $\theta_q \in (0, \delta_1)$ and $\theta_p \in I$, we have

$$|\hat{\mu}_n(4) - w| \leq \epsilon. \quad (5.1)$$

Proof of Claim 5.3. Define the de-symmetrize probability measure on $S^1$ to be $d\nu_n(\theta) = d\mu_n(\theta/4)$. Then $\hat{\nu}_n(1) = \hat{\nu}_n(1)$ and, by convolution properties of the Fourier transform, we also have

$$\nu_n = \nu_{p^m} \ast \nu_q.$$  \hspace{1cm} (5.2)

A direct computation shows that

$$\hat{\nu}_{p^m}(1) = \frac{1}{m + 1} \sum_{j=0}^{m} \cos((m - 2j)\theta_p) = \frac{\sin((m + 1)\theta_p)}{(m + 1) \sin \theta_p}$$

Thus, using (5.2) and properties of the Fourier Transform, we deduce

$$\hat{\nu}_n(1) = \hat{\nu}_{p^m}(1) \ast \hat{\nu}_{q}(1) = \frac{\sin((m + 1)\theta_p) \sin(2\theta_q)}{(m + 1) \sin \theta_p \cdot 2 \sin \theta_q}. \hspace{1cm} (5.3)$$

Observe that, since the function $\sin(2x)/2 \sin(x)$ tends to 1 as $x \to 0$ and it is decreasing in a small neighbourhood to the right of $x = 0$, we can find some small $\delta_1 = \delta_1(\epsilon) > 0$ such that for all $x \in (0, \delta_1)$, we have $|\sin(2x)/2 \sin(x) - 1| \leq \epsilon/2$. With this choice of $\delta_1$, bearing in mind that $|\sin(m\theta)/(m \sin \theta)| \leq 1$, equation (5.3) becomes

$$\left| \hat{\nu}_n(1) - \frac{\sin((m + 1)\theta_p)}{(m + 1) \sin \theta_p} \right| \leq \epsilon/2 \hspace{1cm} (5.4)$$

The claim follows by (5.4) and the continuity of $\sin((m + 1)\theta)/(m + 1) \sin \theta)$.

Before stating the next claim, we introduce some notation. Observe that are only finitely many (depending on $m$ and $l$) sums 0 $\neq \sum_i \xi_i$, let us label them as $S_1, S_2, \ldots$. By (2.1), we can write

$$\xi_i = \sqrt{n} e(a_i \theta_p + \theta_q + b_i \pi/2) \hspace{1cm} (5.5)$$

for some $|a_i| \leq m$ and $b_i \in \{0, 1, 2, 3\}$. Thus, collecting terms with equal $a_i$, we can write

$$S_k = \sqrt{n} \sum_{j=-m}^{m} c_{jk} e(j \theta_p) := \sqrt{n} \cdot F_k(\theta_p) \hspace{1cm} (5.6)$$

for some $c_{ik} \in \mathbb{C}$ with $|c_{ik}| \leq 2l$.

Claim 5.4. Let $F_k(\theta)$ be as (5.6). With the notation of Claim 5.3, there exists some $q \asymp n^{1/m+1}$ with $\theta_q \in (0, \delta_1)$ and some $\delta_2 = \delta_2(\epsilon, m) > 0$ such that $\min_k \max_f |F_k(\theta)| \geq \delta_2$.

Proof. First, we prove that there exists some $q \asymp n^{1/m+1}$ with $\theta_q \in (0, \delta_1)$ such that $\max_{[0,1]} |F_k(\theta)| \geq \delta_3$ for some $\delta_3 = \delta_3(\epsilon, m, l)$ independent of $k$. Bounding the infinity norm by the $L^2$-norm, we have

$$\max_{[0,1]} |F_k(\theta)|^2 \geq \int |F_k(\theta)|^2 d\theta = \sum_{j} |c_{jk}|^2. \hspace{1cm} (5.7)$$

Since $F_k(\theta)$ is not identically zero, there exists some $j_1 = j_1(k)$ such that $|c_{j_1k}| = |c_{j_1}| > 0$. By (5.5), we have

$$c_{j_1} = \alpha e(\theta_q) + \beta e(-\theta_q)$$

for some $\alpha = \alpha(k), \beta = \beta(k) \in \mathbb{Z}[i]$ with $|\alpha| \leq l$ and $|\beta| \leq l$. If $|\alpha| \neq |\beta|$, then $|c_{j_1}| \geq ||\alpha| - |\beta|| \geq 1$. If $|\alpha| = |\beta|$, then, writing $\alpha = re(\psi_1)$ and $\beta = re(\psi_2)$, we have

$$|c_{j_1}| = r |e(2\theta_q + \psi_1 - \psi_2) + 1| = \sqrt{2r}|1 + \cos(2\pi(2\theta_q + \psi_1 - \psi_2))|. $$
Since there are at most $O(l^2)$ choices for $\psi_1$ and at most $O(l^2)$ for $\psi_2$, there exists some $\theta_0 = \theta_0(l) \in (0, \delta_1)$ such that
\[ |c_{j_1}| \geq \delta_3 \]
for some $\delta_3 = \delta_3(\epsilon, m, l) > 0$ independent of $k$. Via Lemma 2.6, we choose $q < n^{1/m+1}$ such that $\theta_1 = \theta_0 + O(\log n^{-100})$. Thus, bearing in mind that the continuity of $F_k$ depends only on $m$ and $l$, together with (5.7), we have $\max_{[0,1]} |F_k(\theta)| \geq \delta_4$, for some $\delta_4 = \delta_4(\epsilon, m, l)$. Therefore, applying Lemma 5.2 with $B = [0,1]$ and $\Omega = I$, we deduce that
\[ \max_I |F_k(\theta)| \geq C^{-m} |I|^{-m} \delta_4 = \delta_2 \]
uniformly for every $k$ and for some absolute constant $C > 0$, as required. □

To conclude the proof, we need to choose a sequence of primes $p, q \equiv 1 \pmod{4}$ and a function $m = m(n)$ such that $\hat{\mu}_{\omega}(4) \to \omega$ and $Q(l, n, c(n)n^{1/2}) = \emptyset$. Let $\epsilon > 0$ and $I$ and $q$ be given by Claim 5.3 and Claim 5.4 respectively. Via Claim 5.4, let $\theta_1 \in I$ be such that $\min_k |F_k(\theta)| \geq \delta_2$. Then, via Lemma 2.6, choose $p \approx n^{1/(m+1)}$ such that $\theta_p = \theta_1 + O(\log n^{-100})$. Finally, bearing in mind that the continuity of $F_k$ depends only on $m$ and $l$, choose $m$ slow enough such that $\min_k |F_k(\theta)| \geq \delta_2/2$ and $\delta_2/2 > 3c(n)$. With this choices of $p, q$ and $m$, we have
\[ |\hat{\mu}(4) - \omega| \leq \epsilon \quad \min_k |S_k| > c(n) \]
as required. □

5.2. Small 3 spectral quasi-correlations. In this section, we construct a sequence of $n \in S'$ such that $Q(l, n, \exp(-c\sqrt{\log n})) \neq \emptyset$, for some absolute constant $c > 0$. We pick integers of the form $n = p_1 \cdot p_2 \cdot p_3$, where $p_i \equiv 1 \pmod{4}$. Let $\theta_{p_1}$ be the angle of the Gaussian prime above $p_1$, then, by Lemma 2.6, we can choose $p_1 \approx n^{1/3}$ so that
\[ \theta_1 = 0 + O\left(\exp(-c_1 \sqrt{\log n})\right) \]
\[ \theta_2 = \frac{\pi}{3} + O\left(\exp(-c_1 \sqrt{\log n})\right) \]
\[ \theta_3 = -\frac{\pi}{3} + O\left(\exp(-c_1 \sqrt{\log n})\right) \]
for some $c_1 > 0$. Thus, we have the representations
\[ \xi_1 = \exp(i(\theta_{p_1} + \theta_{p_2} + \theta_{p_3})) = 1 + O(\exp(-c\sqrt{\log n})) \]
\[ \xi_2 = \exp(i(\theta_{p_1} + \theta_{p_2} - \theta_{p_3})) = \exp(i2\pi/3) + O(\exp(-c\sqrt{\log n})) \]
\[ \xi_3 = \exp(i(-\theta_{p_1} + \theta_{p_2} - \theta_{p_3})) = \exp(-i2\pi/3) + O(\exp(-c\sqrt{\log n})). \]
Hence,
\[ |\xi_1 + \xi_2 + \xi_3| n^{-1/2} \ll \exp(-c\sqrt{\log n}). \]

5.3. Kac-Rice premises. We have the following formula for the variance of $\mathcal{L}(f_n, s)$, see [4, Lemma 3.1, Lemma 3.4, Lemma 3.5 and and page 16].

Lemma 5.5. Let $s > 0$ and write $r = r_n(\cdot)$ as in (1.7). Then, we have
\[ \text{Var}(\mathcal{L}(f_n, s)) = \frac{n}{2} \int_{B(s) \times B(s)} (L_2(x - y) + e(x - y)) \, dx \, dy + O\left(\int_{B(s) \times B(s)} r(x - y)^6 \, dx \, dy\right) \]
where
\[
8L_2(x) = r^2 + \text{Tr}(X) + \frac{\text{Tr}(Y^2)}{4} + \frac{3}{4}r^4 - \frac{\text{Tr}(XY^2)}{8} - \frac{\text{Tr}(X^2)}{16} + \frac{\text{Tr}(Y^4)}{128}
\]
\[
+ \frac{\text{Tr}(Y^2)^2}{256} - \frac{\text{Tr}(X) \text{Tr}(Y^2)}{16} + \frac{r^2 \text{Tr}(X)}{2} + \frac{r^2 \text{Tr}(Y^2)}{8}
\]
\[
|\epsilon(x)| \ll r^6 + \text{Tr}(X^3) + \text{Tr}(Y^6)
\]

and
\[
X = -\frac{2}{n(1-r^2)}(\nabla r)^4 \nabla r \quad \quad Y = -\frac{2}{n} \left( H + \frac{r}{1-r^2} (\nabla r)^4 \nabla r \right)
\]

where \(H\) is the Hessian of \(r\), that is \(H_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} r\).

To evaluate the integrals in Lemma 5.5, we need the following lemma.

**Lemma 5.6.** Let \(A = \frac{20}{9} \log 2\), \(\epsilon > 0\), \(n \in S\) and \(s > n^{-1/2} (\log n)^{A + \epsilon / 2}\). Suppose that \(Q(l, n, n^{1/2}/(\log n)^{2 + \epsilon}) = \emptyset\) for \(l = 2, 4, 6\), and \(N = (\log n)^{\frac{2}{3} + \epsilon / 4}\), then
\[
\int_{B(s) \times B(s)} r(x - y)^2 dxdy = \frac{(\pi s^2)^2}{N^2} \left( 1 + o \left( \frac{1}{N^2} \right) \right)
\]
\[
\int_{B(s) \times B(s)} r(x - y)^4 dxdy = \frac{3(\pi s^2)^2}{N^2} \left( 1 + o \left( \frac{1}{N^2} \right) \right)
\]
\[
\int_{B(s) \times B(s)} r(x - y)^6 dxdy = o \left( \frac{s^4}{N^2} \right)
\]

**Proof.** Let \(l\) be either 2, 4 or 6. A direct computation gives
\[
\int_{B(s) \times B(s)} r(x - y)^l dxdy = \frac{1}{N^l} \int_{B(s) \times B(s)} \sum_{\xi_1, \ldots, \xi_l} e(\langle \xi_1 + \ldots + \xi_l, x - y \rangle) dxdy
\]
\[
= \frac{(\pi s^2)^2}{N^l} \sum_{\xi_1 + \ldots + \xi_l = 0} 1 + O \left( \frac{1}{N^l} \sum_{\xi_1 + \ldots + \xi_l > 0} \left| \int_{B(s)} e(\langle \xi_1 + \ldots + \xi_l, x \rangle) dx \right|^2 \right)
\]

(5.8)

The first term on the right hand side of (5.8) is equal to \((\pi s^2)^2/N\) if \(l = 2\), \(3(\pi s^2)^2/N^2\) if \(l = 4\) and if \(l = 6\) we use (2.3) to see that it is bounded by \(O(s^4/N^{5/2})\). Thus, we are left with bounding the second term on the right of (5.8). Carrying out the integral gives
\[
\frac{1}{N^l} \sum_{\xi_1 + \ldots + \xi_l > 0} \left| \int_{B(s)} e(\langle \xi_1 + \ldots + \xi_l, x \rangle) dx \right|^2 = \frac{(\pi s^2)^2}{N^l} \sum_{\xi_1 + \ldots + \xi_l > 0} \frac{J_1(s|\xi_1 + \ldots + \xi_l|)^2}{s|\xi_1 + \ldots + \xi_l|} \quad (5.9)
\]

By assumption \(Q(l, n, n^{1/2}/(\log n)^{9\log 2/2+\epsilon}) = \emptyset\), thus \(s|\xi_1 + \ldots + \xi_l| \geq (\log n)^{\frac{1}{4} \log 2 - \epsilon / 2}\), bearing in mind that \(J_1(T)/T \ll T^{-3/2}\) and \(N = (\log n)^{\frac{1}{2} + \epsilon / 4}\), we obtain
\[
\text{RHS}(5.9) \ll \frac{1}{(\log n)^{\frac{1}{2} + 3\epsilon / 2}} = o(N^{-2})
\]
as required. \(\square\)

Using Lemma 5.5 and following similar calculations to Lemma [4, Lemma 3.4], we obtain the following lemma.
Lemma 5.7. Under the assumptions of Lemma 5.6, we have

\[ \int_{B(s) \times B(s)} \Tr X(x-y)\, dx dy = (\pi s^2)^2 \left( -\frac{2}{N} - \frac{2}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} \Tr Y(x-y)^2\, dx dy = (\pi s^2)^2 \left( \frac{4}{N} - \frac{4}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} \Tr X(x-y)Y(x-y)\, dx dy = (\pi s^2)^2 \left( -\frac{4}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} \Tr X(x-y)^2\, dx dy = (\pi s^2)^2 \left( \frac{8}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]

\[ \int_{B(s) \times B(s)} \Tr Y(x-y)^4\, dx dy = (\pi s^2)^2 \left( \frac{2(11 + \hat{\mu}(4))^2}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} (\Tr Y(x-y)^2)^2\, dx dy = (\pi s^2)^2 \left( \frac{4(7 + \hat{\mu}(4))^2}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} \Tr X(x-y) \Tr Y(x-y)^2\, dx dy = (\pi s^2)^2 \left( -\frac{8}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} r(x-y)^2 \Tr X(x-y)\, dx dy = (\pi s^2)^2 \left( -\frac{4}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} r(x-y)^2 \Tr Y(x-y)^2\, dx dy = (\pi s^2)^2 \left( \frac{8}{N^2} + o\left( \frac{1}{N^2} \right) \right) \]
\[ \int_{B(s) \times B(s)} \Tr Y(x-y)^3\, dx dy = o\left( \frac{s^4}{N^2} \right) \]
\[ \int_{B(s) \times B(s)} \Tr Y(x-y)^6\, dx dy = o\left( \frac{s^4}{N^2} \right) . \]

5.4. Concluding the proof of Theorem 1.3. We are finally ready to prove Theorem 1.3.

Proof of Theorem 1.3. By Proposition 5.1 and Theorem 1.1, there exists a density one subsequence of integers \( n \in S \) such that both (1) and Lemma 5.7 hold. For such sequence we can evaluate the variance using Lemma 5.5 (and following identical calculation to the proof of [4, Theorem 1.1]), so (2) follows. To prove (3) we again resort to calculations in [4]: for all \( s > 0 \) we have

\[ \Cov(\mathcal{L}(f_n), \mathcal{L}(f_n)) = (\pi \cdot s^2)^2 \Var(\mathcal{L}(f_n)). \]  

(5.10)

Hence, combining (5.10), part (2) and (1.4) we obtain (3).

\[ \square \]

6. Proof of Theorem 1.4

To prove Theorem 1.4, we use the stability of the nodal set under small perturbations, as in [34]. To prove Theorem 1.4 will need a series of results.
6.1. **Stability of the nodal set.** The following deterministic lemma, inspired by [3, Lemma 4.7], will be our main tool in studying small perturbations of the nodal set of $f_n$.

**Lemma 6.1.** Let $h, \vartheta : B(1) \to \mathbb{R}$ be two smooth functions and assume the following:

1. For some $\beta > 0$ we have
   \[
   \min_{y \in B(1)} \max \{|h|, |\nabla h|\} > \beta.
   \]
2. For some $M > 0$ we have
   \[
   ||h||_{C^2}, ||h + \vartheta||_{C^2} < M.
   \]
3. For some $\tau > 0$ we have
   \[
   ||\vartheta||_{C^2} < \tau.
   \]

Then, provided that $\tau \leq \beta^2 \cdot (16M)^{-1}$, we have

\[
\mathcal{L}(h + \vartheta) = \mathcal{L}(h) \left(1 + O \left(\frac{M^3}{\beta^4}\right)\right).
\]

**Proof.** Let $\gamma_h$ be a connected component of $h^{-1}(0)$, fix some $z_0 \in \gamma_h$, and let $N(z_0) = \nabla h/|\nabla h|$ be a unit normal vector of $\gamma_h$ at $z_0$. By assumption (1), $|\nabla h(z_0)| > \beta$, thus, bearing in mind that all second derivatives of $h$ are bounded, we can find some $0 < r_0 = r_0(\beta, M) \leq \beta/4M$ such that

\[
N(z_0) \cdot \nabla h(z) > \beta/2
\]
for all $z$ in a $r_0$-neighbourhood of $z_0$. Now, consider the function

\[
\zeta(r) = h(z_0 + rN(z_0)) + \vartheta(z_0 + rN(z_0)).
\]

Bearing in mind assumption (3) and using (6.1), we have

\[
\zeta'(r) > \beta/2 - \tau > \beta/4
\]
for $\tau < \beta/4$ and uniformly for all $|r| < r_0$. Now, suppose that $0 < \zeta(0) = \vartheta(z_0) \leq \tau$ and that $4\tau/\beta \leq r_0$ then, bearing in mind (6.2), the mean value theorem implies that there exist a unique (negative) $r = r(z_0)$ with $|r| \leq 4\tau/\beta \leq r_0$, that is $\tau \leq \beta^2 \cdot (16M)^{-1}$, such that $\zeta(r(z_0)) = 0$. Arguing similarly in the case $\zeta(0) < 0$, and taking $r = 0$ if $\zeta(0) = 0$, we find that the map

\[
z \to z + r(z)N(z)
\]
is an injection of $\gamma_h$ into $\gamma_{h + \vartheta}$ (where $\gamma_{h + \vartheta}$ is a connected component of $(h + \vartheta)^{-1}(0)$), provided $\tau \leq \beta^2 \cdot (16M)^{-1}$.

**Claim 6.2.** Via the implicit function theorem, parametrize $\gamma_h$ in some neighbourhood $U$ around $z_0$ as $C(t) = (t, q(t))$ for some smooth function $q : U \to \mathbb{R}$, then

\[
\frac{d}{dt}r(C(t)) = \nabla r C' \ll \tau M^2/\beta^4
\]

**Proof.** By (6.3), we have the following system:

\[
(h + \vartheta)(C(t) + r(C(t))N(C(t))) = 0 \quad \text{and} \quad h(C(t)) = 0
\]
Taking the derivative with respect to $t$, we obtain

\[
\nabla(h + \vartheta)(C + rN) \cdot [C' + (\nabla r C') N + r(N'C')] = 0 \quad \text{and} \quad \nabla h(C) C' = 0
\]
By the implicit function theorem $|C'| \leq M/\beta$, thus, bearing in mind assumption (2) and the definition of $N$, $|NC'| \leq M^2/\beta^2$. Moreover, by construction $r \leq 4\tau/\beta$, $|N| \leq 1$, and by assumption (2) $|\nabla(h + \vartheta)| \leq M$, therefore we can re-write (6.4) as

$$\nabla(h + \vartheta)(C + rN) \cdot (\nabla rC')N = -\nabla(h + \vartheta)(C + rN) \cdot C' + O \left( \frac{M^3}{\beta^3} \right) \quad (6.5)$$

$$\nabla h(C)C' = 0.$$ 

Note that, by assumptions (2) and (3), for $r < 1$ say, we have

$$\nabla(h + \vartheta)(C + rN) = \nabla h(C) + O \left( \frac{M}{\beta} \cdot \tau \right). \quad (6.6)$$

Thus, using the expansion (6.6) on the right hand side of (6.5) and subtracting the second equation from the first, we have

$$\nabla(h + \vartheta)(C + rN) \cdot (\nabla rC') \ll \frac{M^3}{\beta^3}. \quad (6.7)$$

As $\nabla(h + \vartheta) \geq \beta/2$, the claim follows from (6.7). \hfill \square

Finally, fix some $z_0 \in \gamma_h$ and let $U$ be as in Claim 6.2, moreover let $V$ be the image of $U$ under the map (6.3) so that $V$ is parametrised by $C(t) + r(C(t)) \cdot N(C(t)) =: \tilde{C}(t)$. Now, Claim 6.2 implies that $|C'(t)| = |	ilde{C}'(t)| + O(\tau M^3 \beta^{-4})$, thus

$$\mathcal{L}(h|U) = \int_U |C'(t)|dt = \int_V |\tilde{C}'(t)|dt + O \left( \frac{M^3}{\beta^4} \right)$$

$$= \mathcal{L}((h + \vartheta)|V) + O \left( \frac{M^3}{\beta^4} \right). \quad (6.8)$$

Summing (6.8) over the zero set of $h$ we obtain the required result. \hfill \square

6.2. **Quantifying** $M$. By [33, Lemma 3.12] and [2, Corollary 2.2], we have the following result:

**Lemma 6.3.** Let $f_n$ be as in (1.2), $R > 1$ and $k \geq 0$, and let $F_n(y) = f_n(y, \sqrt{n})$ for $y \in B(1)$. Then we have

$$\mathbb{P} \left( \|F_n\|_{C^k(B(1))} \gg R^{k+1} \log R \right) \leq e^{-\mathcal{C}(\log R)^2}$$

for some $\mathcal{C} > 0$.

6.3. **Quantifying** $\beta$.

**Lemma 6.4.** Let $f_n$ be as in (1.2), $R > 1$ and $F_n(y) = f_n(y, \sqrt{n})$ for $y \in B(1)$. Suppose that $n \in S$ satisfies the conclusion of Theorem 2.7, then

$$\mathbb{P} \left( \min_{y \in B(1)} \max \{|F_n(y)|, |\nabla F_n(y)|\} \leq R^{-4}(\log R)^{-3} \right) \leq (\log R)^{-1}$$

**Proof.** Let $\epsilon_1 > 0$ be some (small) parameter. Differentiating $r_n(y, y) = 1$, we see that $F_n(y)$ and $\nabla F_n(y)$ are independent random variables. Therefore, bearing in mind that $F_n$ is a stationary field, we have

$$\mathbb{P} \left( |F_n(y)| \leq \epsilon_1 \text{ and } |\nabla F_n(y)| \leq \epsilon_1 \right) = \mathbb{P} \left( |F_n(y)| \leq \epsilon_1 \right) \mathbb{P} \left( |\nabla F_n(y)| \leq \epsilon_1 \right)$$

$$= \mathbb{P} \left( |F_n(0)| \leq \epsilon_1 \right) \mathbb{P} \left( |\nabla F_n(0)| \leq \epsilon_1 \right) = \frac{1}{\sqrt{2\pi}} \int_{-\epsilon_1}^{\epsilon_1} e^{t^2/2} dt \cdot \mathbb{P} \left( |\nabla F_n(0)| \leq \epsilon_1 \right)$$

$$\leq \epsilon_1 \mathbb{P} \left( |\nabla F_n(0)| \leq \epsilon_1 \right) \quad (6.9)$$
The covariance matrix of the Gaussian vector $\nabla F_n(0)$ is given by
\[
C(0) = \frac{4\pi^2 R^2}{N} \left( \sum_{\xi,\xi'} (\xi_1^2 \xi'_2 - \xi_1 \xi_2 \xi'_1 \xi'_2) \right)
\]
with determinant
\[
\det C(0) = \frac{16\pi^4 R^4}{N^2} \sum_{\xi,\xi'} (\xi_1^2 (\xi'_2)^2 - \xi_1 \xi_2 \xi'_1 \xi'_2) = \frac{16\pi^4 R^4}{N^2} \sum_{\xi,\xi'} \sin^2(\theta_\xi - \theta_{\xi'})^2
\]
where $\xi = e^{2\pi i \theta_\xi}$. Using Theorem 2.7 to pass from the sum to the integral, and the identity $\sin(\cdot)^2 = (1 - \cos(2\cdot))/2$, we have
\[
\det C(0) = 8\pi^4 R^4 + \int_0^{2\pi} \int_0^{2\pi} \cos(2(x-y))dxdy + o(1) = 8\pi^4 R^4 + o(1)
\]
Hence,
\[
\mathbb{P}( |\nabla F_n(0)| \leq \epsilon_1) = \frac{1}{4\sqrt{2\pi} R^2} \int_{[-\epsilon_1,\epsilon_1]^2} \exp \left( -\frac{1}{2} \sum_{\xi} \xi_1^2 x_1^2 - 2\xi_1 \xi_2 x_1 x_2 + \xi_2^2 x_2^2 \right) dx_1 dx_2
\]
(6.10)
Inserting (6.10) into (6.9), we obtain
\[
\mathbb{P}( |F_n(y)| \leq \epsilon_1 \text{ and } |\nabla F_n(y)| \leq \epsilon_1) \leq \epsilon_1^3 R^2. \quad (6.11)
\]
Now, consider an $\eta$-net on $B(1)$ and denote by $y_i$ the points of the net. By Lemma 6.3, we know that $||F_n||_{C^2(B(1))} \ll R^3 \log R$ outside an event of probability at most $O(e^{-C(\log R)^2})$. Thus, since every point $y \in B(1)$ is at distance at most $\eta$ from a point $y_i$ on the net, we have
\[
F_n(y) = F_n(y_i) + O(\eta R^3 \log R) \quad \nabla F_n(y) = \nabla F_n(y_i) + O(\eta R^3 \log R),
\]
for some $y_i$. Therefore, if $|F_n(y)| \leq \beta$ and $|\nabla F_n(y)| \leq \beta$, then, taking $\eta = c\beta (R^3 \log R)^{-1}$ for some sufficiently small $c > 0$, also $|F_n(y_i)| \leq \beta/2$ and $|\nabla F_n(y_i)| \leq \beta/2$, which has probability at most $O(\beta^2/R^2)$ by (6.11). Taking the union bound over the net, which has $O(\eta^{-2})$ points, we deduce that
\[
\mathbb{P} \left( \min_{y \in B(1)} \max \{|F_n(y)|, |\nabla F_n(y)|\} \leq \beta \right) \ll \beta^3 R^2 \eta^{-2} + e^{-C(\log R)^2} \ll \beta R^4 \log R^2 + e^{-C(\log R)^2}.
\]
Hence, taking $\beta = R^{-4}(\log R)^{-3}$ we deduce the lemma.

6.4. Quantifying $\tau$. To quantify $\tau$ we use the following recent result [2, Theorem 5.5]:

**Lemma 6.5** (Beliaev-Maffucci). Let $R > 1$, $\epsilon > 0$, $n \in S$, $F_n(y) = f(Ry/\sqrt{n})$ and $F_n(y) = f_n(Ry)$ for $y \in B(1)$. Suppose that $n$ satisfies the conclusion of Theorem 2.7, then there exists a coupling such that the field $\tilde{F} = F_n - F_n$ satisfies
\[
\mathbb{P}( ||\tilde{F}||_{C^2(B(1))} \gg R^2 \log R \cdot (\log n)^{-2\epsilon/3+\epsilon} ) \ll (\log R)^{-1}.
\]
6.5. **Proof of Theorem 1.4.** We are finally ready to prove Theorem 1.4.

**Proof of Theorem 1.4.** Take a subsequence of $n \in S$ such that the conclusion of Theorem 2.7 holds (and $N \to \infty$). Let $R > 1$, rescale $f_n$ as $F_n(y) = f_n(Ry/\sqrt{n})$ and $f_\mu$ as $F_\mu(y) = f_\mu(Ry)$ for $y \in B(1)$. By Lemma 6.4, Lemma 6.3 and Lemma 6.5, outside an event of probability at most $O((\log R)^{-1})$, we have the following bounds:

1. $$\min_{y \in B(1)} \max |F_n|, |\nabla F_n| \geq R^{-4}(\log R)^{-3} := \beta$$

2. $$\|F_n\|_{C^2} \leq 2R^3 \log R := M$$

3. $$\|F_\mu - F_n\|_{C^2} \leq R^2 \log R (\log n)^{-2\kappa/3+\epsilon} := \tau$$

Lemma 6.1, provided that $\tau \ll \beta^2/M = R^{-11}(\log R)^{-5}$, implies that

$$\mathcal{L}(F_\mu) = \mathcal{L}(F_n) \left(1 + O \left(\frac{M^3}{\beta^4}\right)\right) = \mathcal{L}(F_n) \left(1 + O \left(R^{27} \log R^{16} \cdot (\log n)^{-2\kappa/3+\epsilon}\right)\right)$$

outside an event of probability at most $O(\log R^{-1})$. The Kac-Rice formula [1, Theorem 6.3] implies that $\mathbb{E}[\mathcal{L}(F_n)] \ll R$. Thus, outside and event of probability at most $O(\log R^{-1})$, $\mathcal{L}(F_n) \ll R \log R$ which implies

$$|\exp(it \mathcal{L}(F_n)) - \exp(it \mathcal{L}(F_\mu))| \ll t|\mathcal{L}(F_n) - \mathcal{L}(F_\mu)| \leq tR^{28}(\log R)^{17}(\log n)^{-2\kappa/3+\epsilon}. \quad (6.12)$$

Taking $R = (\log n)^{\frac{2\kappa}{3\epsilon}}$ and $n$ large enough depending on $t$, (6.12) implies the Theorem.

ACKNOWLEDGEMENT

The author would like to thank Igor Wigman for pointing out the question considered here and for the many discussions, Zeev Rudnick for valuable comments that helped improving the presentation of the article, as well as Oleksiy Klurman for useful conversations. The author would also like to thank the anonymous referees for pointing our an error in the previous draft of the article and their valuable comments that greatly helped to improve the presentation. This work was supported by the Engineering and Physical Sciences Research Council [EP/L015234/1]. The EPSRC Centre for Doctoral Training in Geometry and Number Theory (The London School of Geometry and Number Theory), University College London.

REFERENCES

[1] J.M. Azais, M. Wschebor. *Level sets and extrema of random processes and fields*. John Wiley and Sons Inc., Hoboken NJ, (2009).

[2] D. Beliaev, R. Maffucci. *Coupling of stationary fields with applications to Arithmetic Waves*, Arxiv preprint, https://arxiv.org/pdf/1912.09470.pdf (2019).

[3] D. Beliaev, I. Wigman. *Volume distribution of nodal domains of random band-limited functions*. Probab. Theory Related Fields 172, no. 1-2, 453–492 (2018).

[4] J. Benatar, D. Marinucci and I. Wigman *Planck-scale distribution of nodal length of arithmetic random waves*, J. Anal. Math (to appear), Preprint arXiv:1710.06153 (2017).

[5] M. Berry. *Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature*. J. Phys. A: Math. Gen. 35 , 3025-3038 (2002).

[6] M. Berry. *Semiclassical mechanics of regular and irregular motion. Chaotic behavior of deterministic systems* Les Houches, 171-271, (1981).
SPECTRAL QUASI CORRELATIONS AND ARITHMETIC RANDOM WAVES

[7] P. Billingsley, *Probability and measure*, Third edition, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, Inc., New York, (1995).

[8] E. Bombieri, J. Bourgain, A problem on sums of two squares, IMRN 11, 3343-3407 (2015).

[9] J. Bourgain, On toral eigenfunctions and the random wave model, Israel J. Math. 201, no. 2, 611-630, (2014).

[10] J. Bourgain, Z. Rudnick On the geometry of the nodal lines of eigenfunctions of the two-dimensional torus. Ann. Henri Poincaré 12, no. 6, 1027-1053, (2011).

[11] J. Bourgain, Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces and nodal sets, Geom. Funct. Anal. 22, no. 4, 878-937, (2012).

[12] S. Y. Cheng, Eigenfunctions and nodal sets. Comm. Math. Helv. 51 , 43-55 (1976).

[13] J. Cilleruelo, The distribution of the lattice points on circles, J. Number Theory 43, no. 2, 198-202, (1993).

[14] J. Cilleruelo, A. Córdoba, Trigonometric polynomials and lattice points, Proc. Amer. Math. Soc. 115 (4) , 899-905, (1992).

[15] J. Cilleruelo, A. Granville, Lattice points on circles, squares in arithmetic progressions and sums of squares, in Additive Combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc, Providence, RI, 241-262, (2007).

[16] G. Dierckx, I. Nourdin, G. Peccati, M. Rossi, Small Scale CLTs for the Nodal Length of Monochromatic Waves, Arxiv preprint: https://arxiv.org/abs/2005.06577 (2020).

[17] H. Donnelly, C. Fefferman Nodal sets of eigenfunctions on Riemannian manifolds. Invent. Math. 93 , 161-183. (1988).

[18] P. Erdős, R.R. Hall. On The Angular Distribution Of Gaussian Integers With Fixed Norm. Discrete Mathematics 200, 87-94, (1999).

[19] J. Evertse, H. Schlickewei, W. Schmidt, Linear equations with variables which lie in a multiplicative group, Annals of Math (2), 807-836, (2002).

[20] J. Friedlander, H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications, 57. American Mathematical Society, (2010).

[21] A. Granville, I. Wigman, The distribution of the zeros of random trigonometric polynomials, Amer. J. Math, 133, no. 2, 295-357, (2011).

[22] A. Granville, I. Wigman, Planck-scale mass equidistribution of toral Laplace eigenfunctions, Comm. Math. Phys., 355(2), 767-802, (2017).

[23] G. Harman and P. Lewis, Gaussian primes in narrow sectors, Mathematika, 119-135, (2001).

[24] I. Kátai, I. Környei, On the distribution of lattice points on circles , Ann. Univ. Sci. Budapest. Eötvös Sect. Math., 87-91, (1977).

[25] M. Krishnapur, P. Kurlberg, I. Wigman, Igor Nodal length fluctuations for arithmetic random waves. Ann. of Math. (2) 177, no. 2, 699-737. (2013).

[26] I. Kubilius. The distribution of Gaussian primes in sectors and contours, (in Russian) Leningrad Gos. Univ. Uc. Zap. Ser. Mat. Nauk, 40-52, (1950).

[27] P. Kurlberg , I. Wigman On Probability Measures Arising From Lattice Points On Circles. Mathematische Annalen (2016).

[28] A. Logunov, E. Malinnikova, Nodal sets of Laplace eigenfunctions: estimates of the Hausdorff measure in dimensions two and three, 50 years with Hardy spaces, 333-344, (2018).

[29] A. Logunov, Nodal sets of Laplace eigenfunctions: proof of Nadirashvili’s conjecture and of the lower bound in Yau’s conjecture, Ann. of Math. (2) 187 , no. 1, 241-262, (2018).

[30] A. Logunov, Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure, Ann. of Math. (2) 187, no. 1, 221-239, (2018).

[31] D. Marinucci, G. Peccati, M. Rossi, I. Wigman, Non-universality of nodal length distribution for arithmetic random waves. Geom. Funct. Anal. 26, 926-960, (2016).

[32] D. Marinucci, M. Rossi, I. Wigman The asymptotic equivalence of the sample trispectrum and the nodal length for random spherical harmonics, Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, 56, 1, 374-390 (2020).

[33] S. Muirhead, H. Vanneuville, The sharp phase transition for level set percolation of smooth planar gaussian fields, Ann. Inst. Henri Poincaré Probab. Stat. 56, no. 2, 1358-1390, (2020).

[34] F. Nazarov, M.Sodin, Asymptotic laws for the spatial distribution and the number of connected components of zero sets of Gaussian random functions. Zh. Mat. Fiz. Anal. Geom. 12 , no. 3, 205-278, (2016).
[35] F. Nazarov, Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type. Algebra i Analiz 5, no. 4, 3–66, (1993).

[36] I. Nourdin, G. Peccati, M. Rossi. Nodal statistics of planar random waves, Communications in Mathematical Physics, 369, 1, 99–151 (2019).

[37] F. Oravecz, Z. Rudnick and I. Wigman, The Leray measure of nodal sets for random eigenfunctions on the torus, Annales de l’Institut Fourier 57, (2007).

[38] Z. Rudnick, and I. Wigman. “On The Volume Of Nodal Sets For Eigenfunctions Of The Laplacian On The Torus”. Annales Henri Poincaré 9.1, 109-130, (2008).

[39] A. Sartori, On the fractal structure of attainable probability measures, Bull. Pol. Acad. Sci. Math. 66, no. 2, 123-133, (2018).

[40] G. Tenenbaum, Introduction to analytic and probabilistic number theory, Third edition, Graduate Studies in Mathematics, 163, American Mathematical Society, Providence, RI, (2015).

[41] A. P. Todino, Nodal Lengths in Shrinking Domains for Random Eigenfunctions on $S^2$, Bernoulli 26, no. 4, 3081-3110, (2020).

[42] I. Wigman, Fluctuations of the nodal length of random spherical harmonics, Comm. Math. Phys. 298, no. 3, 787-831, (2010).

[43] A. Zygmund, On Fourier coefficients and transforms of functions of two variables. Studia Math. 50, 189-201, (1974).

Departement of Mathematics, King’s College London, Strand, London WC2R 2LS, England, UK
Email address: andrea.sartori.16@ucl.ac.uk