Annals of Mathematics **177** (2013), 699–737
http://dx.doi.org/10.4007/annals.2013.177.2.8

Nodal length fluctuations for arithmetic random waves

By Manjunath Krishnapur, Pär Kurlberg, and Igor Wigman

Abstract

Using the spectral multiplicities of the standard torus, we endow the Laplace eigenspaces with Gaussian probability measures. This induces a notion of random Gaussian Laplace eigenfunctions on the torus ("arithmetic random waves"). We study the distribution of the nodal length of random eigenfunctions for large eigenvalues, and our primary result is that the asymptotics for the variance is *nonuniversal*. Our result is intimately related to the arithmetic of lattice points lying on a circle with radius corresponding to the energy.

1. Introduction

The purpose of this paper is to investigate the variance of the fluctuations of nodal lengths of random Laplace eigenfunctions on the standard 2-torus $\mathbb{T} := \mathbb{R}^2/\mathbb{Z}^2$. The nodal set of a function f is simply the zero set of f, and if $f : \mathbb{T} \to \mathbb{R}$ is a Laplace eigenfunction, i.e., if f is nonconstant and

$$\Delta f + Ef = 0, \quad E > 0, \tag{1}$$

then the nodal set of f consists of a union of smooth curves outside a finite set of singular points (see [12]). Hence $\mathrm{length}(f^{-1}(0))$, the *nodal length of* f, is well defined.

A fundamental conjecture by Yau [29], [30] asserts that for any smooth compact Riemannian manifold M, there exist constants $c_2(M) \ge c_1(M) > 0$ such that

$$c_1(M) \cdot \sqrt{E} \le \mathrm{Vol}(f^{-1}(0)) \le c_2(M) \cdot \sqrt{E} \tag{2}$$

M.K. was supported in part by DST and UGC through DSA-SAP-PhaseIV.

P.K. was supported in part by grants from the Göran Gustafsson Foundation, the Knut and Alice Wallenberg foundation, and the Swedish Research Council.

I.W. was supported in part by grant KAW 2005.0098 from the Knut and Alice Wallenberg Foundation and an EPSRC grant EP/J004529/1 under the First Grant Scheme.

for any Laplace eigenfunction f on M with eigenvalue E. By the work of Donnelly and Fefferman [15] and Brüning and Gromes [9], [10], Yau's conjecture is known to be true for manifolds with real analytic metrics and, in particular, for $M = \mathbb{T}$.

For $\mathbb{T} = \mathbb{R}^2/\mathbb{Z}^2$, the sequence of eigenvalues, or energy levels, are related to integers expressible as a sum of two integer squares; if we define $S := \{n : n = a^2 + b^2,\ a, b \in \mathbb{Z}\}$, the eigenvalues are all of the form

$$E_n := 4\pi^2 n, \quad n \in S. \tag{3}$$

For $n \in S$, let

$$\Lambda_n := \left\{\lambda \in \mathbb{Z}^2 : \|\lambda\|^2 = n\right\}$$

denote the corresponding frequency set. Using the standard notation $e(z) := \exp(2\pi i z)$, the $\mathbb{C}$-eigenspace $\mathcal{E}_n$ corresponding to E_n is spanned by the L^2-orthonormal set of functions $\{e(\langle\lambda, x\rangle)\}_{\lambda\in\Lambda_n}$. The dimension of $\mathcal{E}_n$, denoted by

$$\mathcal{N}_n := \dim \mathcal{E}_n = r_2(n) = |\Lambda_n|,$$

is equal to the number $r_2(n)$ of different ways n may be expressed as a sum of two squares.

1.1. *Arithmetic random waves.* The set Λ_n can be identified with the set of lattice points lying on a circle with radius $\sqrt{n}$, and its properties are intimately related to representations of integers by the quadratic form x^2+y^2. The frequency set is thus of arithmetic nature. A particular consequence is that the sequence of spectral multiplicities $\{\mathcal{N}_n\}_{n\geq 1}$ is *unbounded*. It is thus natural to consider properties of "generic," or "random," eigenfunctions $f_n \in \mathcal{E}_n$, and our primary interest is the high energy asymptotics of the distribution of their nodal length $\mathcal{L}(f_n)$ as n tends to infinity in such a way that $\mathcal{N}_n \to \infty$. More precisely, let $f_n : \mathbb{T} \to \mathbb{R}$ be the random Gaussian field of (real valued) $\mathcal{E}_n$-functions with eigenvalue E_n, i.e.,

$$f_n(x) = \frac{1}{\sqrt{2\mathcal{N}_n}} \sum_{\lambda\in\Lambda_n} a_\lambda e(\langle\lambda, x\rangle), \tag{4}$$

where $a_\lambda = b_\lambda + ic_\lambda$ are independent standard complex Gaussian random variables, save for the relations $a_{-\lambda} = \overline{a}_\lambda$. This just means that $b_\lambda, c_\lambda \sim \mathcal{N}(0, 1)$ are standard real Gaussians satisfying the relation $b_{-\lambda} = b_\lambda$, $c_{-\lambda} = -c_\lambda$ and otherwise independent. Our object of study is the random variable

$$\mathcal{L}_n := \mathcal{L}(f_n) = \text{length}(f_n^{-1}(0)),$$

henceforth called the *nodal length* of f_n.

1.2. *Prior work on this model.* In this setting, Rudnick and Wigman [24] computed the expected nodal length of f_n to be $\mathbb{E}[\mathcal{L}_n] = \frac{1}{2\sqrt{2}} \cdot \sqrt{E_n}$, in agreement with Yau's conjecture

$$\mathrm{Var}(\mathcal{L}_n) = O\left(\frac{E_n}{\sqrt{\mathcal{N}_n}}\right) \tag{5}$$

for the variance, and conjectured that the stronger bound

$$\mathrm{Var}(\mathcal{L}_n) = O\left(\frac{E_n}{\mathcal{N}_n}\right) \tag{6}$$

holds. A nice consequence of (5) is that $\mathcal{L}(f_n)$ concentrates around its mean. More precisely, there is a sequence $\delta_n \to 0$ such that

$$\mathbf{P}\left((1-\delta_n)\frac{\sqrt{E_n}}{2\sqrt{2}} \le \mathcal{L}(f_n) \le (1+\delta_n)\frac{\sqrt{E_n}}{2\sqrt{2}}\right) \to 1 \quad \text{as } \mathcal{N}_n \to \infty.$$

In this paper we shall determine the leading order asymptotic of $\mathrm{Var}(\mathcal{L}_n)$ as $\mathcal{N}_n \to \infty$. As consequence we improve on the conjectured bound (6) and obtain the *sharp* bounds

$$\frac{E_n}{\mathcal{N}_n^2} \ll \mathrm{Var}(\mathcal{L}_n) \ll \frac{E_n}{\mathcal{N}_n^2}.$$

It turns out that the asymptotic behaviour of the variance is *nonuniversal* in the sense that it depends on the angular distribution of the points in the frequency set Λ_n. In the proof, a leading order sum involving many terms of size $E_n/\mathcal{N}_n$ surprisingly *cancels perfectly*, and the variance is therefore much smaller than expected! We may say that $\mathbb{T}$ exhibits "arithmetic Berry cancellation" (cf. Section 1.6.2).

1.3. *Our results.* In order to describe our results we shall need some further notation. The set Λ_n induces a discrete probability measure μ_n on the circle $\mathcal{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ by defining

$$\mu_n := \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \delta_{\frac{\lambda}{\sqrt{n}}}, \tag{7}$$

where δ_x is the Dirac delta measure supported at x. As usual, the Fourier transform of μ_n is, for any $k \in \mathbb{Z}$, given by $\widehat{\mu}_n(k) := \int_{\mathcal{S}^1} z^{-k} d\mu_n(z)$. For $n \in S$, we define

$$c_n := \frac{1 + \widehat{\mu}_n(4)^2}{512}; \tag{8}$$

it is then easy to see that c_n is real and that $c_n \in [1/512, 1/256]$. (Since Λ_n is invariant under the transformations $z \to \overline{z}$ and $z \to i \cdot z$, the same holds for μ_n; hence, $\widehat{\mu}_n(4) \in \mathbb{R}$. Further, since μ_n is a probability measure, $|\widehat{\mu}_n(4)| \le 1$, and consequently $\widehat{\mu}_n(4)^2 \in [0, 1]$.)

We can now formulate our principal result.

THEOREM 1.1. *If $(n_i)_{i\geq 1}$ is any sequence of elements in S such that $\mathcal{N}_{n_i} \to \infty$, then*

$$\operatorname{Var}(\mathcal{L}_{n_i}) = c_{n_i} \cdot \frac{E_{n_i}}{\mathcal{N}_{n_i}^2}(1 + o(1)). \tag{9}$$

Further, given any $c \in [1/512, 1/256]$, there exists a sequence $(n_i)_{i\geq 1}$ of elements in S such that as $i \to \infty$, we have $\mathcal{N}_{n_i} \to \infty$ together with $c_{n_i} \to c$ so that

$$\operatorname{Var}(\mathcal{L}_{n_i}) = c \cdot \frac{E_{n_i}}{\mathcal{N}_{n_i}^2}(1 + o(1)).$$

1.4. *Attainable measures.* The second part of the theorem, in light of the first one, amounts to the following. Given any $\alpha \in [0, 1]$, there exists a sequence $(n_i)_{i\geq 1}$ such that $\widehat{\mu}_{n_i}(4)^2 \to \alpha$. We briefly describe the measures μ_n giving rise to the extremal points, as well as intermediate values attainable by c_n. (See Section 7 for full details, in particular the precise notions of generic and thin used below.)

It is well known that the lattice points Λ_n are equidistributed on $\mathcal{S}^1$ along generic subsequences of energy levels; see, e.g., [16, Prop. 6]. Thus, for $(n_i)_{i\geq 1}$ a *generic* sequence of elements in S, the variance is *minimal* in the limit since $\widehat{\mu}_{n_i}(4) \to 0$, and thus $c_{n_i} \to 1/512$. It is also worthwhile mentioning that the nodal length variance of f_n for such generic sequences differs by an order of magnitude from the corresponding quantity for superposition of random planar waves with same wavelength and directions chosen uniformly on the unit circle. (This is especially noteworthy in light of the fact that both have the same *scaling* properties. See the last paragraph of Section 1.6.1 for a more detailed explanation.)

As for the maximum, Cilleruelo [13] has shown that there are thin sequences $(n_i)_{i\geq 1}$, with $\mathcal{N}_{n_i} \to \infty$, such that μ_{n_i} converges weakly to an atomic probability measure supported at the four symmetric points ± 1, $\pm i$; hence, $\widehat{\mu}_{n_i}(4) \to 1$ and $c_{n_i} \to 1/256$. For the intermediate values, we construct thin sequences $(n_i)_{i\geq 1}$ of elements in S such that μ_{n_i} converges weakly to the uniform probability measure supported on a union of four arcs.

More precisely, for $a \in [0, \frac{\pi}{4}]$, define a probability measure ν_a on $\mathcal{S}^1$ by

$$\nu_a := \left(\frac{1}{4} \sum_{k=0}^{3} \delta_{i^k} \right) \star \tilde{\nu}_a, \tag{10}$$

where $\star$ stands for convolutions of measures and $\tilde{\nu}_a$ is the uniform measure on $[-a, a]$ (identifying $\mathcal{S}^1 \cong \mathbb{R}/2\pi\mathbb{Z}$). More explicitly,

$$\tilde{\nu}_a(f) = \frac{1}{2a} \int_{-a}^{a} f\left(e^{i\theta}\right) d\theta, \qquad \nu_a(f) = \frac{1}{8a} \sum_{k=0}^{3} \int_{-a+k\frac{\pi}{2}}^{a+k\frac{\pi}{2}} f\left(e^{i\theta}\right) d\theta.$$

For $a = 0$, we shall use the notational convention that $\nu_0 = \frac{1}{4} \sum_{k=0}^{3} \delta_{i^k}$.

Proposition 1.2. *For every $a \in [0, \frac{\pi}{4}]$, there exists a sequence E_{n_i} of energy levels such that $\mu_{n_i} \Rightarrow \nu_a$ with ν_a as in (10). In particular, for every $b \in [0, 1]$, there exists a sequence E_{n_i} of energy levels such that $c_{n_i} \to \frac{1+b}{512}$.*

Note that the second statement follows from the first because the values of $\widehat{\nu}_a(4)$ range over the whole of $[0, 1]$ as a ranges over $[0, \frac{\pi}{4}]$. (In fact, an easy computation shows that $\widehat{\nu}_a(4) = \frac{\sin(4a)}{4a}$.) Further, the extremal values $b = 0$ and $b = 1$ are attained by $\nu = \nu_{\frac{\pi}{4}}$, the uniform measure on $\mathcal{S}^1$, and $\nu = \nu_0$, the atomic symmetrized measure. The proof of Proposition 1.2 will be given in Section 7.

1.5. *Independence of eigenbasis choice and the covariance function.* The random field (4) is centered, Gaussian and stationary in the sense that for any $x_1, \ldots, x_k \in \mathbb{T}$ and $y \in \mathbb{T}$, the random vector

$$(f_n(x_1 + y), \ldots, f_n(x_k + y)) \in \mathbb{R}^k$$

is a mean zero multivariate Gaussian, whose distribution does not depend on y. The covariance function[1] $r(x) = r_n(x) := \mathbb{E}[f_n(y) f_n(x + y)]$ thus depends only on x, and we may express it explicitly as

$$r_n(x) = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda} e\left(\langle \lambda, x \rangle\right) = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda} \cos\left(2\pi \langle \lambda, x \rangle\right). \tag{11}$$

Though the normalizing factor in the definition (4) of f_n has no bearing on the nodal length, it is convenient to work with, and we have chosen to have $r_n(0) = 1$ or, equivalently, for every $x \in \mathbb{T}$, $\mathbb{E}[f_n(x)^2] = 1$.

The covariance function *determines* the distribution of a centered Gaussian random field and, in principle, one may express any aspect of the geometry of f_n in terms of r_n only (cf. Kolmogorov's Theorem, [14, Ch. 3.3]). This important fact also shows that we would get the same random field in (4) had we chosen a different orthonormal basis of $\mathcal{E}_n$ in the Gaussian linear combination.

1.6. *Background and results in related models.* The question of distribution of various local quantities such as the nodal length, or the total curvature of nodal lines in different settings, has been extensively studied. It is widely believed [2] that for *generic chaotic* billiards, one can model the nodal lines for eigenfunctions of eigenvalue of order $\approx E$ with nodal lines of planar monochromatic random waves of wavenumber $\sqrt{E}$. (This is called Berry's Random Wave Model or RWM, see (12) for the definition.) Berry [3] found that the expected nodal length (per unit area) for the RWM is of size approximately $\sqrt{E}$, and he argued that the variance should be of order $\log E$.

---

[1] The covariance function is widely referred to as the 2-point function in the physics literature.

The 2-dimensional unit sphere $\mathcal{S}^2$ is another manifold with degenerate Laplace spectrum. Here the Laplace eigenvalues are all the numbers $E = m(m+1)$ with $m \geq 0$ an integer, and the corresponding eigenspace is the space of degree m spherical harmonics; its dimension is $2m+1$. One may define the random field of degree m spherical harmonics similarly to the torus (4) with the plane waves (exponentials) replaced by any L^2-orthonormal basis $\{\eta_{m;1}, \dots, \eta_{m;2m+1}\}$ of real valued spherical harmonics of degree m:

$$f_m^{\mathcal{S}^2}(x) = \frac{1}{\sqrt{2m+1}} \sum_{k=1}^{2m+1} a_k \eta_{m;k}(x),$$

with a_k the independent and identically distributed standard Gaussian. Setting $\mathcal{L}\left(f_m^{\mathcal{S}^2}\right)$ to be the nodal length of $f_m^{\mathcal{S}^2}$, Berard [1] computed the expected nodal length

$$\mathbb{E}\left[\mathcal{L}\left(f_m^{\mathcal{S}^2}\right)\right] = \sqrt{2}\pi \cdot \sqrt{E}.$$

Wigman [27] found that the nodal length variance is asymptotic to

$$\mathrm{Var}\left(\mathcal{L}\left(f_m^{\mathcal{S}^2}\right)\right) \sim c \log m,$$

which is consistent with Berry's prediction for the RWM.

1.6.1. *Comparing the random wave model to the torus and the sphere.* The logarithmic variance is much smaller than one would expect. Taking into account that the wavelength for either the sphere or the RWM scales as $\frac{1}{\sqrt{E}}$, one may rescale them to unit wavelength to argue that the nodal length variance should be proportional to $\sqrt{E}$. However, a computation reveals that the coefficient in front of the expected leading term $\sqrt{E}$ surprisingly *vanishes* due to "Berry's Cancellation Phenomenon" — the leading term for nodal length variance is in fact logarithmic. A similar cancellation phenomenon is responsible for the variance (9) in our situation being of order of magnitude $\frac{E_n}{\mathcal{N}_n^2}$, rather than of order $\frac{E_n}{\mathcal{N}_n}$ as was originally conjectured in [24].

As already remarked, in general, defining a centered (or mean zero) Gaussian random field f on an arbitrary domain T is equivalent to specifying its covariance function $r_f(x,y) := \mathbb{E}[f(x)f(y)]$ on $T \times T$. For the planar random waves (RWM), the covariance function is

$$r_{\mathrm{RWM}}(x,y) = J_0(\sqrt{E}\|x-y\|), \tag{12}$$

with J_0 the standard Bessel function, and

$$r_{\mathcal{S}^2}(x,y) = P_m(\cos d(x,y))$$

for the degree m spherical harmonics, where P_m are the usual Legendre polynomials, and d is the (spherical) distance. The latter scales as

$$P_m(\cos(\psi/m)) \approx J_0(d)$$

uniformly for $\psi \in [0, m \cdot \frac{\pi}{2}]$. As the corresponding eigenvalues are $m(m+1)$, this is consistent with the RWM scaling. The covariance function $r_n(x)$ for our ensemble f_n of random toral eigenfunctions given by (4) is of arithmetic flavour. It is given by the summation (11) over lattice points Λ_n lying on a circle.

The equidistribution of Λ_n along generic sequences of energy levels on the torus mentioned earlier implies that for any *fixed* $y \in \mathbb{R}^2$, one may approximate

$$r_n(y/(2\pi\sqrt{n_i})) \approx \int\limits_{\mathcal{S}^1} \cos(\langle y, z\rangle) dz = J_0(\|y\|)$$

for a generic sequence $\{n_i\} \subseteq S$. Although it is the same scaling limit as before, the latter holds for y of fixed size only, and by no means uniformly for $y \in [0, n]^2$. In particular, as opposed to the other cases, no "intermediate range asymptotic" for $r_n(x)$ is known, i.e., for $x \cdot \sqrt{n} \to \infty$. It is remarkable that even in this case, in spite of the fact that the covariance function for f_n has the same scaling limit as the RWM and random spherical harmonics random fields, the nodal length variance (9) of random arithmetic waves is of different order of magnitude compared to the other cases.

1.6.2. *Berry's cancellation phenomenon and the* 2*-point correlation function.* In order to evaluate moments of the nodal length of a random field f, we exploit a suitable Kac-Rice type formula (see Section 2.1). For the variance, it means that we need to understand the fluctuations of the so called 2-point correlation function (defined on the domain of f) around its scaled asymptotic value at infinity. For both $\mathbb{R}^2$ (RWM) and $\mathcal{S}^2$, the main contribution for the variance comes from the "intermediate range" (i.e., a few wavelengths away from the origin), where the asymptotic behaviour of the covariance function and its first two derivatives translates to asymptotics for the 2-point correlation function. No analogous asymptotics is known for the torus. The cancellation phenomenon amounts to the fact that the leading term in the intermediate range asymptotics for the 2-point correlation function is purely oscilliatory and its contribution to the integral is negligible. Then, the main contribution comes from the second term in the asymptotics.

As a substitute for pointwise asymptotics for the 2-point correlation function, we use an arithmetic formula (see (33)) valid outside a suitably defined "singular set" (arithmetic in nature; its analogue for $\mathbb{R}^2$ and sphere is a neighbourhood of the origin, with radius of order wavelength). Although the arithmetic formula does not give the pointwise behaviour of the 2-point correlation function, its arithmetic structure is exploited for averaging over the torus and is essential for evaluating the variance. The "arithmetic Berry's cancellation" amounts to the Fourier expansion of the highest magnitude term of the 2-point correlation function vanishing at the origin, an artifact of the seemingly unrelated trigonometric identity

$$4\cos(\theta/2)^4 = 1 + 2\cos\theta + \cos(\theta)^2;$$

see Section 4.2 for more details.

1.7. *Some other related results.* For a generic compact manifold $\mathcal{M}$ with no spectral degeneracies, one can also consider random Gaussian linear combinations of eigenfunctions with different eigenvalues (sometimes referred to as "random wave on $\mathcal{M}$"). Berard [1] and Zelditch [31] found that, given a spectral parameter E, the expected nodal length for random Gaussian superpositions of eigenfunctions with eigenvalues lying either about E or below[2] E is of order $\sqrt{E}$, consistent with Berry's RWM. The subtle question of the nodal length variance in this generic setup is to be addressed in [23].

Some other generic results concerning random waves with spectral parameter E are the following. Toth and Wigman [26] found that the expected number of open nodal lines, i.e., the connected component of the zero set that intersect the boundary, of the random wave with spectral parameter E on a generic surface with boundary is again of order $\sqrt{E}$. Moreover, Nicolaescu [21] evaluated the expected number of critical points to be of order E; the latter is also an upper bound for the number of nodal domains.

For other related or relevant results, we refer the interested reader to the recent survey [28]. Also, for recent interesting results and conjectures on nonlocal quantities, such as *nodal domains* (i.e., the connected components of the complement $\mathcal{M} \setminus f^{-1}(0)$ of the nodal line), see [4], [20], [6], [8], [7].

1.8. *Outline of the paper.* The paper is organized as follows. The proof of Theorem 1.1, assuming certain preparatory results is given in Section 2; this proof relies on an arithmetic formula (16) in Proposition 2.1, whose proof is commenced in Section 4. The proof of the formula (16) is based on studying the behaviour of the so-called 2-point correlation function introduced in Section 3; its subtle asymptotic analysis is given in Sections 4, 5 and 8, with the latter containing a certain technical computation essential for understanding the asymptotic properties.

Section 6 is dedicated to the proof of Theorem 2.2, an arithmetic bound, due to Bourgain, needed for the admissability of the error term in (16). In Section 7, sequences of energy levels with corresponding discrete probability measures (7) converging to the measures ν_a as in (10) are constructed to prove the attainability of the latter.

[2] Called the short or long energy window random combinations respectively.

1.9. *Acknowledgements.* The authors would like to thank Zeév Rudnick for suggesting the problem as well as having many inspiring and fruitful discussions. In addition, we would like to thank Mikhail Sodin, Peter Sarnak and Jean Bourgain for stimulating and fruitful discussions on various aspects of the problem solved in this manuscript. Part of the research presented was conducted during, or as a result of, the visit of I.W. to the Institute for Advanced Study, Princeton, and he would like to acknowledge the friendly environment, and the generous financial support. We are especially grateful to Jean Bourgain for his kind permission to include Theorem 2.2 and its proof.

2. Proof of Theorem 1.1

2.1. *Kac-Rice formulas.* Moments of the nodal length for smooth random fields can be computed using the Kac-Rice formulas [14]. To state them, we need some notation. For $f = f_n$, we define its first and second correlations as follows:

$$
\begin{aligned}
K_1 &= \mathbb{E}\left[\|\nabla f(y)\| \mid f(y)=0\right], \\
\tilde{K}_2(x) &= \mathbb{E}\left[\|\nabla f(y)\| \cdot \|\nabla f(y+x)\| \mid f(y)=f(y+x)=0\right], \\
K_2(x) &= \frac{2}{E_n}\tilde{K}_2(x).
\end{aligned}
\tag{13}
$$

Observe that K_1 and K_2 are independent of y because f_n is stationary. (For general smooth Gaussian fields, they become $K_1(y)$ and $\tilde{K}_2(x,y)$.) They are called the first and second correlations of the nodal set $f_n^{-1}(0)$. K_2 is just a scaled version of the second correlation. As we are dealing with Gaussians, it is possible to write analytical expressions for these as Gaussian integrals in terms of r_n and its derivatives (see (29), (25) and (24)).

Then, the Kac-Rice formulas say that

$$
\mathbb{E}[\mathcal{L}_n] = \int_{\mathbb{T}} K_1 dy = K_1, \qquad \mathbb{E}[\mathcal{L}_n^2] = \int_{\mathbb{T}} \tilde{K}_2(x)dx.
$$

The first of these formulas gives $\mathbb{E}[\mathcal{L}_n] = \frac{1}{2\sqrt{2}} \cdot \sqrt{E_n}$, as was quoted earlier. Using this and the second gives

$$
\operatorname{Var}(\mathcal{L}_n) = \frac{E_n}{2}\int_{\mathbb{T}} \left(K_2(x) - \frac{1}{4}\right) dx. \tag{14}
$$

Full justification of their validity in our context may be found in [24]. For the Kac-Rice formulas in general, consult [14].

It is instructive to intuitively understand the function $\tilde{K}_2(x)$ in the following way. Let $x \in \mathbb{T}$, and take a small positive number $\varepsilon > 0$. We define the random variables

$$
\mathcal{L}_n^{x,\varepsilon} = \operatorname{length}\left(f_n^{-1}(0) \cap B(x,\varepsilon)\right),
$$

where $B(x,\varepsilon)$ is the disk of radius ε centered at x; $\mathcal{L}_n^{x,\varepsilon}$ measures the nodal length of f_n inside the corresponding disk. Then we have

$$\tilde{K}_2(x) = \frac{2}{E_n} \lim_{\varepsilon\to 0} \frac{1}{\pi^2\varepsilon^4} \mathbb{E}\left[\mathcal{L}_n^{x,\varepsilon}\mathcal{L}_n^{0,\varepsilon}\right].$$

2.2. *Computing the variance — the Gaussian integral.* To understand $\operatorname{Var}(\mathcal{L}_n)$, we need to understand the integral in (14). The function K_2 may be implicitly expressed in terms of the covariance function r_n of f_n as a Gaussian expectation of a 4-variate centered Gaussian $(\nabla f_n(0), \nabla f_n(x))$ conditioned on $f_n(0) = f_n(x) = 0$, with 4×4 covariance matrix $\Omega_n(x)$ depending on r and its derivatives (see (29), (25) and (24)). In our case, the covariance function is the arithmetic function (11).

In order to study the asymptotic behaviour of the integral above, we use some ideas from [22], [24] and divide the torus into a singular set B and the nonsingular complement $\mathbb{T}\setminus B$; only the latter is convenient to work with, so it is essential make the former as small as possible. We improve the analysis of the earlier paper [24] on both B and $\mathbb{T}\setminus B$.

A better upper bound for the measure of B is proved using the sixth moment of r rather than the fourth one. As an artifact of the definition of B, one has a lower bound for the values $|r(x)|$ on B; using a Chebyshev-like inequality on the sixth moment of $r(x)$, we will bound the measure of B so that its contribution to the variance is negligible.

On $\mathbb{T}\setminus B$, where the main contribution comes from, we establish a precise asymptotic expression for the 2-point correlation function compared to a partial upper bound as in [24] (see Proposition 4.5). Here the (scaled) covariance matrix $\Omega_n(x)$, defined by (25), is a perturbation of the identity matrix. In order to understand its contribution to the integral, we expand $K_2(x)$ as a function of Ω_n into a 4-variate Taylor polynomial around $\Omega_n = I_4$, the identity matrix. In principle, this could be performed using brute force; we choose to work with Berry's elegant method [3].

The computation above culminates in the "arithmetic formula" given in the following proposition. It is of arithmetic essence and at the heart of the variance being nonuniversal, but the derivation itself involves no arithmetic. Before stating the proposition, we define

$$\mathcal{R}_k(n) := \int_{\mathbb{T}} |r_n(x)|^k dx. \tag{15}$$

PROPOSITION 2.1. *The nodal length variance is given by the asymptotic formula*

$$\operatorname{Var}(\mathcal{L}_n) = c_n \cdot \frac{E_n}{\mathcal{N}_n^2} + O\left(E_n \cdot \mathcal{R}_5(n)\right), \tag{16}$$

where we used the notation (8) *for* c_n.

The reason we refer to (16) as "arithmetic" is that both the main term and the error term in (16) are of arithmetic nature: c_n is related to the distribution of lattice points Λ_n on the circle (see Section 8), and $\mathcal{R}_5$ is controlled in terms of arithmetics of spectrum correlation.

The proof of Proposition 2.1 will commence in Section 4. It is lengthy and quite technical, so it may be omitted on a first reading of the paper.

In case of random spherical harmonics or the random wave model, one arrives at analogous propositions for the variance. The proof essentially ends there as the error term may be checked to be smaller than the main term.

2.3. *Computing the variance — the arithmetic part.* In our setting however, the main obstacle is in proving the admissability of the error term. We will control various error terms in terms of the moments $\mathcal{R}_k(n) := \int_{\mathbb{T}} |r_n(x)|^k dx$ (cf. (15)). The even moments are naturally related to the spectral correlations; for example, it is straightforward to check that

$$\mathcal{R}_6(n) = \frac{1}{\mathcal{N}_n^6}|S_6(n)|, \tag{17}$$

where S_6 is the 6-correlation set of frequencies

$$S_6(n) = \left\{(\lambda_1, \dots, \lambda_6) \in \Lambda_n : \sum_{i=1}^{6} \lambda_i = 0\right\}. \tag{18}$$

Since for any choice of $\lambda_1, \dots, \lambda_4 \in \Lambda_n$ there are at most four possible choices for $\lambda_5, \lambda_6 \in \Lambda_n$, it follows that $|S_6(n)| = O\left(\mathcal{N}_n^4\right)$ or, equivalently,

$$\mathcal{R}_6(n) = O\left(\frac{1}{\mathcal{N}_n^2}\right). \tag{19}$$

The latter bound is not quite sufficient for our purposes, but the following result, due to J. Bourgain, is sufficiently strong for our purposes.

THEOREM 2.2. *As $\mathcal{N}_n \to \infty$, we have the following estimate*:

$$|S_6(n)| = o\left(\mathcal{N}_n^4\right). \tag{20}$$

Consequently, $\mathcal{R}_6(n) = o\left(\frac{1}{\mathcal{N}_n^2}\right)$.

Theorem 2.2 will be proven in Section 6. We give a brief preview of the proof of Theorem 2.2. To refute the possibility that $|S_6(n)| \gg \mathcal{N}_n^4$, we invoke some techniques from additive combinatorics (see Section 6). Utilizing a notion of "additive energy" defined in Section 6, a certain set A related to the sum set of Λ_n is shown to contain a large subset A_1 with "bounded doubling." Using a suitable version of Freiman's Theorem, this implies that A_1 is essentially a *generalized arithmetic progression* (GAP; see Theorem 6.3), i.e., is contained inside a slightly larger GAP. This then leads to a contradiction via an application of Chang's result [11] on the number of representations of a complex number as a product of elements inside a GAP.

We note that the bound $|S_6(n)| = o(\mathcal{N}_n^4)$ is quite far from the truth; Bombieri and Bourgain [5] have recently obtained an exponent savings. We further remark that a trivial lower bound is that $|S_6(n)| \gg \mathcal{N}_n^3$; taking any $\lambda_1, \lambda_2, \lambda_3 \in \Lambda_n$ and letting $\lambda_4 = -\lambda_1, \lambda_5 = -\lambda_2$, and $\lambda_6 = -\lambda_3$ yields $|\Lambda_n|^3 = \mathcal{N}_n^3$ solutions of "diagonal type." We believe that essentially all solutions arise like this, and we conjecture that for every $\varepsilon > 0$,

$$|S_6(n)| = O_\varepsilon\left(\mathcal{N}_n^{3+\varepsilon}\right).$$

(Possibly the even stronger bound $|S_6(n)| = O\left(\mathcal{N}_n^3\right)$ holds.)

2.4. *Proof of Theorem 1.1 assuming the preparatory results.* Given Propositions 1.2 and 2.1 and Theorem 2.2, it is now straightforward to deduce Theorem 1.1, our main result. Recall that $\mathcal{R}_k$ are the moments (15) of r_n. Using the Cauchy-Schwarz inequality on $|r(x)|^5 = r^2(x) \cdot |r(x)|^3$ together with the bound $\mathcal{R}_4 = O\left(\frac{1}{\mathcal{N}^2}\right)$ (which follows from the same argument that yielded (19)), and the bound $\mathcal{R}_6(n) = o\left(\frac{1}{\mathcal{N}_n^2}\right)$ from Theorem 2.2, we obtain

$$\mathcal{R}_5(n) = o\left(\frac{1}{\mathcal{N}_n^2}\right). \tag{21}$$

Now, using (16) together with (21), we obtain (9).

Finally, using the second part of Proposition 1.2 and the definition of c_n (see (7)), we find that any $c \in [1/512, 1/256]$ is attainable as a limit.

3. The 2-point correlation function of f_n

In this section we use the Kac-Rice formula (14) that expresses $\mathrm{Var}(\mathcal{L}_n)$ as an integral of the (scaled) 2-point correlation K_2 defined in (13). For this we will need to study some aspects of the random field f_n first. From this point on we fix n and will usually suppress the n-dependency with no further note.

3.1. *Joint distribution of values and gradients.* In order to study the variance, we shall need to study the random vector

$$W = W_{n;x} = (u_1, u_2, v_1, v_2) = (f_n(0), f_n(x), \nabla f_n(0), \nabla f_n(x)) \in \mathbb{R}^6.$$

Since W is a linear transformation of the standard Gaussian $a = (a_\lambda) \in \mathbb{R}^{\mathcal{N}_n}$, its distribution is also centered (or mean zero) Gaussian, and by the stationarity, 0 and x may be replaced by any y and $y + x$.

Let

$$D = D_{1\times 2}(x) = \nabla r(x) = \frac{2\pi i}{\mathcal{N}_n} \sum_{\lambda \in \Lambda} e\left(\langle \lambda, x\rangle\right) \cdot \lambda \tag{22}$$

(cf. (11)). The vector W is centered Gaussian with covariance matrix (cf. [24, §5.1])

$$\Sigma = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix},$$

where

$$A(x) = \begin{pmatrix} 1 & r(x) \\ r(x) & 1 \end{pmatrix},$$

$$B(x) = \begin{pmatrix} 0 & D(x) \\ -D(x) & 0 \end{pmatrix}$$

and

$$C(x) = \begin{pmatrix} \frac{E}{2} I_2 & -H(x) \\ -H(x) & \frac{E}{2} I_2 \end{pmatrix},$$

where $H_{2\times 2}(x)$ is the Hessian

(23) $$H(x) = \left(\frac{\partial^2 r}{\partial x_i \partial x_j} \right) = -\frac{4\pi^2}{\mathcal{N}_n} \sum_{\lambda\in\Lambda} e\left(\langle \lambda, x\rangle\right) \cdot (\lambda^t \lambda),$$

by (11). (Note that λ is a row vector so that $\lambda^t\lambda$ is a 2×2 matrix.)

The covariance matrix of $(\nabla f(0), \nabla f(x))$, conditioned on $f(0) = f(x) = 0$ is

(24) $$\tilde{\Omega}_{4\times 4} = C - B^t A^{-1} B = \begin{pmatrix} \frac{E}{2} I_2 & -H \\ -H & \frac{E}{2} I_2 \end{pmatrix} - \frac{1}{1-r^2} \begin{pmatrix} D^t D & rD^t D \\ rD^t D & D^t D \end{pmatrix},$$

where we write $r = r(x)$ for brevity. Thus, $K_2(x) = \mathbb{E}[\|\tilde{V}_1\| \cdot \|\tilde{V}_2\|]$, where $\tilde{V}_i$ are 2-dimensional random vectors with $(\tilde{V}_1, \tilde{V}_2)$ having Gaussian distribution with zero mean and covariance matrix $\tilde{\Omega}(x)$.

3.2. *The scaled* 2*-point correlation function.* It is more convenient to work with the *scaled* covariance matrix

(25) $$\Omega(x) = \Omega_n(x) = \frac{2}{E_n} \tilde{\Omega}_n(x).$$

Then, the scaled 2-point correlation function defined in (13) may be written as

$$K_2(x) = \frac{1}{2\pi\sqrt{1 - r_n(x)^2}} \mathbb{E}[\|V_1\| \cdot \|V_2\|],$$

where V_1, V_2 are centered Gaussians with covariance matrix $\Omega(x)$.

At the origin $x = 0$ the matrix $\Omega(x)$ is singular and hence corresponds to a covariance matrix of a degenerate Gaussian. However for almost all $x \in \mathbb{T}$, $\Omega(x)$ is nonsingular; see [24, Prop. A.1]. We claim that $\Omega(x)$ is a small perturbation of the 4×4 identity matrix I_4, at least, for "generic" x. To quantify the latter statement, write

$$\Omega(x) = I + \begin{pmatrix} X & Y \\ Y & X \end{pmatrix}, \tag{26}$$

where

$$X = -\frac{2}{E_n(1-r^2)} D^t D, \quad Y = -\frac{2}{E_n}\left(H + \frac{r}{1-r^2} D^t D\right), \tag{27}$$

and both $X = X_n(x)$ and $Y = Y_n(x)$ are small for "typical" x.

With these computations, we may rewrite the Kac-Rice formula (14) as follows.

PROPOSITION 3.1 (Cf. [24, Prop. 5.2]). *The nodal length variance is given by*

$$\mathrm{Var}(\mathcal{L}_n) = \frac{E_n}{2} \int\limits_{\mathbb{T}} \left(K_2(x) - \frac{1}{4}\right) dx, \tag{28}$$

where K_2 is the scaled 2-point correlation function given by

$$K_2(x) = K_{2;n}(x) = \frac{1}{2\pi\sqrt{1 - r_n(x)^2}} \mathbb{E}[\|V_1\| \cdot \|V_2\|]; \tag{29}$$

here $V_1, V_2 \in \mathbb{R}^2$ are centered Gaussians with covariance matrix given by (26), with X and Y as in (27).

We shall need the following lemma later.

LEMMA 3.2. *The matrices X_n and Y_n are uniformly bounded (entry-wise); i.e.,*

$$X_n(x), Y_n(x) = O(1), \tag{30}$$

where the constant involved in the 'O'-notation is universal. In particular,

$$K_{2;n}(x) \ll \frac{1}{\sqrt{1 - r_n(x)^2}}. \tag{31}$$

Proof. To prove that (30) holds it is sufficient to show that the diagonal entries of X are uniformly bounded, by (26). (The nondiagonal entries of a covariance matrix are dominated by the diagonal ones, by the Cauchy-Schwarz inequality.) For the latter, it is sufficient to notice that the diagonal entries of Ω are positive, and the diagonal entries of X are ≤ 0 (recall (27)).

To prove that the bound in (31) holds, we use (29), the Cauchy-Schwarz inequality and (30). $\square$

4. **Proof of Proposition 2.1**

To find the asymptotics of the integral (28), we will study the pointwise asymptotic behaviour of K_2. Even though we will only be able to determine the precise asymptotics outside the so-called *singular set*, already used in [22] and [24], we will prove that the exceptional singular set is *small*, so that its contribution is negligible (see Lemma 4.4). To quantify the last statement, we will control the contribution using a Chebyshev-like inequality, so that the corresponding error term will naturally involve the moments (15) of the covariance function (11). We improve upon the analysis of [22] by working with the sixth moment $\mathcal{R}_6(n)$ rather than $\mathcal{R}_4(n)$.

4.1. *The singular set.* For $r(x)$ bounded away from 1 we may expand the $\frac{1}{\sqrt{1-r^2}}$ factor in (29) and related expressions into the Taylor series around $r = 0$. Since the moments of r are "small" (by Theorem 2.2, say), a Chebyshev-like inequality implies that $r(x)$ is small outside a small set. This is the main idea behind the notion of the singular set to follow. We use a slightly stronger definition in order to endow the exceptional set with a structure as a union of squares, necessary in order to find its contribution to the integral (28). The following definitions are borrowed directly from [22, §6.1].

Definition 4.1. A point $x \in \mathbb{T}$ is a positive singular point if there is a set of frequencies $\Lambda_x \subseteq \Lambda$ with density

$$\frac{|\Lambda_x|}{|\Lambda|} > \frac{7}{8}$$

for which $\cos(2\pi\langle\lambda, x\rangle) > 3/4$ for all $\lambda \in \Lambda_x$. Similarly we define a negative singular point to be a point x where there is a set $\tilde{\Lambda}_x \subseteq \Lambda$ of density $> \frac{7}{8}$ for which $\cos(2\pi\langle\lambda, x\rangle) < -3/4$ for all $\lambda \in \tilde{\Lambda}_x$.

Let $M \approx \sqrt{E_n}$ be a large integer. We decompose the torus $\mathbb{T}$ as a union of M^2 closed squares $I_{\vec{k}}$ of side length $1/M$ centered at $\vec{k}/M$, $\vec{k} \in \mathbb{Z}^2$. The squares have disjoint interiors.

Definition 4.2. A square $I_{\vec{k}}$ is a positive (resp. negative) singular square if it contains a positive (resp. negative) singular point.

Definition 4.3. The singular set $B = B_n$ is the union of all singular squares.

Note that, by the definition, each singular square contains a singular point; however, points in B are not necessarily all singular. Let $y \in B$ be a point lying in a positive singular cube, x be the corresponding positive singular point lying in the same singular cube and $\Lambda_x \subseteq \Lambda$ the frequency set prescribed by the definition of a positive singular point. It is easy to see that

$$|\cos(2\pi\langle\lambda, y\rangle) - \cos(2\pi\langle\lambda, x\rangle)| \ll \frac{\sqrt{E_n}}{M},$$

where the implied constant is absolute, so that one may choose $M \approx \sqrt{E_n}$ for which the latter expression is $\le \frac{1}{4}$; it will then imply

$$\cos(2\pi\langle\lambda, y\rangle) \ge \frac{1}{2}$$

for every $\lambda \in \Lambda_x$. We then conclude that

$$\begin{aligned} r(y) &= \frac{1}{|\Lambda|}\sum_{\lambda\in\Lambda_x}\cos(2\pi\langle\lambda, y\rangle) + \frac{1}{|\Lambda|}\sum_{\lambda\in\Lambda\setminus\Lambda_x}\cos(2\pi\langle\lambda, y\rangle) \\ &\ge \frac{1}{|\Lambda|}\sum_{\lambda\in\Lambda_x}\frac{1}{2} - \frac{1}{|\Lambda|}\sum_{\lambda\in\Lambda\setminus\Lambda_x} 1 \ge \frac{7}{16} - \frac{1}{8} = \frac{5}{16} \end{aligned}$$

and, similarly, if y is lying in a negative square, then $r(y) \le -\frac{5}{16}$. Hence we have $|r(y)| \ge \frac{5}{16}$ on all of B. We then write

$$\mathcal{R}_6(n) \ge \mathrm{meas}(B)\cdot\left(\frac{5}{16}\right)^4$$

to obtain the Chebyshev-type inequality

$$\mathrm{meas}(B) \ll \mathcal{R}_6(n). \tag{32}$$

It was proven ([22, §6.5]) that if S is any singular square, then its contribution to the integral (28) is

$$\ll \int_S |K_2(x)|dx \ll \frac{1}{M\sqrt{E_n}}.$$

Since the number of the singular cubes is

$$\ll M^2\,\mathrm{meas}(B),$$

the total contribution of B to (28) is bounded by

$$\int_B |K_2(x)|dx \ll M^2\,\mathrm{meas}(B)\cdot\frac{1}{M\sqrt{E_n}} = \mathrm{meas}(B)\frac{M}{\sqrt{E_n}} \ll \mathcal{R}_6(n),$$

by (32) and $M \approx \sqrt{E_n}$. The latter is summarized in the following lemma.

LEMMA 4.4 (cf. [22, §6.3]). *The contribution of the singular set to* (28) *is bounded by*

$$\int_B |K_2(x)|dx = O\left(\mathcal{R}_6(n)\right).$$

Lemma 4.4 bounds the contribution of the singular set to the integral in (28). The main contribution comes from the nonsingular set. In order to evaluate it, we will need a precise point-wise estimate for $K_2(x)$ in this range. This is given by the following proposition, up to admissible error terms. (To verify the admissibility, see Lemmas 4.6 and 5.4.)

PROPOSITION 4.5 ("Intermediate range" asymptotics for K_2). *For $x \in \mathbb{T} \setminus B$, we have*

$$K_2(x) = \frac{1}{4} + L_2(x) + \varepsilon(x), \tag{33}$$

where the main term $L_2(x)$ is given by

$$L_2(x) = \frac{1}{4}\Bigg(\frac{r^2}{2} + \frac{\operatorname{tr} X}{2} + \frac{\operatorname{tr}(Y^2)}{8} + \frac{3}{8}r^4 - \frac{\operatorname{tr}(XY^2)}{16} - \frac{\operatorname{tr}(X^2)}{32} + \frac{\operatorname{tr}(Y^4)}{256} + \frac{\operatorname{tr}(Y^2)^2}{512} - \frac{\operatorname{tr} X \operatorname{tr}(Y^2)}{32} + \frac{1}{4}r^2 \operatorname{tr} X + \frac{1}{16}r^2 \operatorname{tr}(Y^2)\Bigg) \tag{34}$$

with $X = X_n(x)$, $Y = Y_n(x)$ and $r = r_n(x)$. The error term $\varepsilon(x)$ is bounded by

$$|\varepsilon(x)| = O\left(r(x)^6 + \operatorname{tr}(X^3) + \operatorname{tr}(Y^6)\right). \tag{35}$$

Proposition 4.5 will be proved in Section 5. Assuming it, we arrive at the proof of Proposition 2.1.

Proof of Proposition 2.1 assuming Proposition 4.5. To express the nodal length variance, we invoke Proposition 3.1. Since the contribution of $K_2(x)$ to the integral (28) on B is

$$O\left(E_n \cdot \mathcal{R}_6(n)\right),$$

by Lemma 4.4, we have

$$\begin{aligned}\operatorname{Var}(\mathcal{L}_n) &= \frac{E_n}{2}\left[\int\limits_{\mathbb{T}\setminus B} \left(K_2(x) - \frac{1}{4}\right) dx\right] + O\left(E_n \cdot \mathcal{R}_6(n)\right) \\ &= \frac{E_n}{2} \int\limits_{\mathbb{T}\setminus B} L_2(x)dx + O\left(E_n \cdot \int\limits_{\mathbb{T}\setminus B} |\varepsilon(x)|dx\right) + O\left(E_n \cdot \mathcal{R}_6(n)\right),\end{aligned} \tag{36}$$

by Proposition 4.5. Note that

$$\int_{\mathbb{T}\setminus B} |\varepsilon(x)|dx \le \int_{\mathbb{T}} |\varepsilon(x)|dx = O\left(E_n \cdot \mathcal{R}_6(n)\right),$$

by (35) and Lemma 4.6 to follow (see parts 10-11), so that (36) is

$$\operatorname{Var}(\mathcal{L}_n) = \frac{E_n}{2} \int\limits_{\mathbb{T}\setminus B} L_2(x)dx + O\left(E_n \cdot \mathcal{R}_6(n)\right). \tag{37}$$

We further note that, since $L_2(x)$ is uniformly bounded thanks to Lemma 3.2,

$$\frac{E_n}{2}\int\limits_{B} L_2(x)dx = O\left(E_n \cdot \operatorname{meas}(B)\right) = O\left(E_n \cdot \mathcal{R}_6(n)\right),$$

so that we may rewrite (37) as

$$\operatorname{Var}(\mathcal{L}_n) = \frac{E_n}{2}\int\limits_{\mathbb{T}} L_2(x)dx + O\left(E_n \cdot \mathcal{R}_6(n)\right), \tag{38}$$

the upshot being that we are now able to use the definition (34) of L_2 and integrate the right-hand side of (34) term by term, as in Lemma 4.6 (where the domain of integration is the whole of torus $\mathbb{T}$ rather than $\mathbb{T}\setminus B$). We then perform the term-wise integration of (34) to obtain (with Lemma 4.6)

$$\begin{aligned}
4\cdot\int\limits_{\mathbb{T}} L_2(x)dx &= \frac{1}{\mathcal{N}_n}\left(\frac{1}{2}-\frac{1}{2}\cdot 2+\frac{1}{8}\cdot 4\right)\\
&\quad+\frac{1}{\mathcal{N}_n^2}\Bigg(-\frac{1}{2}\cdot 2-\frac{1}{8}\cdot 4+\frac{3}{8}\cdot 3+\frac{1}{16}\cdot 4-\frac{1}{32}\cdot 8+\frac{1}{256}\cdot 2(11+\widehat{\mu}_n(4)^2)\\
&\quad+\frac{1}{512}\cdot 4(7+\widehat{\mu}_n(4)^2)+\frac{1}{32}\cdot 8-\frac{1}{4}\cdot 2+\frac{1}{16}\cdot 8\Bigg)+O\left(\mathcal{R}_5(n)\right)\\
&= \frac{1}{\mathcal{N}_n^2}\cdot\frac{1+\widehat{\mu}_n(4)^2}{64}+O\left(\mathcal{R}_5(n)\right).
\end{aligned}$$

Collecting all the constants encountered and bearing in mind (38) yields (16), which is the statement of the present proposition. □

4.2. *Some remarks on arithmetic Berry cancellation.* While the constant term $\frac{1}{4}$ cancels out with the expectation squared, the leading nonconstant term of the scaled 2-point correlation function (i.e., the leading term of $K_2(x)-\frac{1}{4}$) is

$$\frac{1}{8}\left(r^2+\operatorname{tr}X+\frac{\operatorname{tr}(Y^2)}{4}\right)\approx\frac{1}{8}\left(r^2-\frac{2}{E_n}DD^t+\frac{1}{E_n^2}\operatorname{tr}H^2\right),$$

where we neglected some lower-order terms. Denote the expression in parenthesis as

$$v(x) := r^2-\frac{2}{E_n}DD^t+\frac{1}{E_n^2}\operatorname{tr}H^2. \tag{39}$$

We may substitute (11), (22) and (23) into (39) to rewrite $v(x)$ as

$$\begin{aligned}
v(x) &= \frac{1}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n} e\left(\langle\lambda_1+\lambda_2,x\rangle\right)+\frac{2}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\frac{\lambda_1\lambda_2^t}{E_n/4\pi^2}e\left(\langle\lambda_1+\lambda_2,x\rangle\right)\\
&\quad+\frac{1}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\frac{(\lambda_1\lambda_2^t)^2}{(E_n/4\pi^2)^2}e\left(\langle\lambda_1+\lambda_2,x\rangle\right)\\
&= \frac{1}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\left(1+2\frac{\lambda_1\lambda_2^t}{n}+\frac{(\lambda_1\lambda_2^t)^2}{n^2}\right)e\left(\langle\lambda_1+\lambda_2,x\rangle\right),
\end{aligned}$$

on recalling (3).

Note that

$$\frac{\lambda_1\lambda_2^t}{n} = \cos\theta(\lambda_1,\lambda_2),$$

where $\theta(\cdot,\cdot)$ is the angle between two vectors in $\mathbb{R}^2$. Thus we may write, up to lower order terms,

$$\begin{aligned} v(x) &= \frac{1}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\left(1+2\cos\theta(\lambda_1,\lambda_2)+\cos\left(\theta(\lambda_1,\lambda_2)\right)^2\right)e\left(\langle\lambda_1+\lambda_2,x\rangle\right) \\ &= \frac{4}{\mathcal{N}^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\cos\left(\frac{\theta(\lambda_1,\lambda_2)}{2}\right)^4 e\left(\langle\lambda_1+\lambda_2,x\rangle\right), \end{aligned}$$

by the usual trigonometric identities. Upon integrating (28), all the summands vanish except for $\lambda_1+\lambda_2=0$; the corresponding angle θ is given by

$$\theta = \theta(\lambda_1,\lambda_2) = \pi,$$

so that $\cos(\theta/2)=0$. Thus the arithmetic cancellation phenomenon in the length variance amounts to $\cos(\theta/2)^4$ vanishing at $\theta=\pi$.

4.3. *Integrating matrix elements.* We may obtain an asymptotic expression for the nodal length variance upon using (28) with Proposition 4.5, provided that we are able to integrate the expressions on the right-hand side of (33), term-wise. This is done in Lemma 4.6. We choose to control the various error terms encountered in terms of the moments of r, $\mathcal{R}_k$ (recall the notation (15)). It will turn out that we will be able to control the error terms in terms of $\mathcal{R}_5$ (and $\mathcal{R}_6\le\mathcal{R}_5$), admissible thanks to Theorem 2.2 via a simple Cauchy-Schwarz argument (see the proof of Theorem 1.1 in Section 2.4). The proof of Lemma 4.6 is left to Section 5.1.

Lemma 4.6. *As $\mathcal{N}_n\to\infty$, we have the following estimates:*

1. $\int\limits_{\mathbb{T}}\operatorname{tr}X(x)dx = -\frac{2}{\mathcal{N}_n}-\frac{2}{\mathcal{N}_n^2}+O\left(\mathcal{R}_6(n)\right).$
2. $\int\limits_{\mathbb{T}}\operatorname{tr}(Y(x)^2)dx = \frac{4}{\mathcal{N}_n}-\frac{4}{\mathcal{N}_n^2}+O\left(\mathcal{R}_6(n)\right).$
3. $\int\limits_{\mathbb{T}}\operatorname{tr}(X(x)Y(x)^2)dx = -\frac{4}{\mathcal{N}_n^2}+O\left(\mathcal{R}_5(n)\right).$
4. $\int\limits_{\mathbb{T}}\operatorname{tr}(X(x)^2)dx = \frac{8}{\mathcal{N}_n^2}+O\left(\mathcal{R}_6(n)\right).$
5. $\int\limits_{\mathbb{T}}\operatorname{tr}(Y(x)^4)dx = \frac{2}{\mathcal{N}_n^2}(11+\widehat{\mu}_n(4)^2)+O\left(\mathcal{R}_6(n)\right).$
6. $\int\limits_{\mathbb{T}}\operatorname{tr}(Y(x)^2)^2dx = \frac{4}{\mathcal{N}_n^2}(3+\widehat{\mu}_n(4)^2)+O\left(\mathcal{R}_6(n)\right).$
7. $\int\limits_{\mathbb{T}}\operatorname{tr}X(x)\operatorname{tr}(Y(x)^2)dx = -\frac{8}{\mathcal{N}_n^2}+O\left(\mathcal{R}_6(n)\right).$
8. $\int\limits_{\mathbb{T}}r(x)^2\operatorname{tr}X(x)dx = -\frac{2}{\mathcal{N}_n^2}+O\left(\mathcal{R}_6(n)\right).$

9. $\int_{\mathbb{T}} r(x)^2 \operatorname{tr}(Y(x)^2)dx = \frac{8}{\mathcal{N}_n^2} + O\left(\mathcal{R}_6(n)\right).$
10. $\int_{\mathbb{T}} \operatorname{tr}(X(x)^3)dx = O(\mathcal{R}_6(n)).$
11. $\int_{\mathbb{T}} \operatorname{tr}(Y(x)^6)dx = O\left(\mathcal{R}_6(n)\right).$

5. **Asymptotics for the 2-point correlation function**

The ultimate goal of this section is to prove Proposition 4.5. To establish the desired asymptotics for (29), one needs to understand the behaviour of $\mathbb{E}[\|V_1\| \cdot \|V_2\|]$ where (V_1, V_2) is a centered Gaussian with covariance Ω_n, the latter being a small perturbation of the indentity matrix, given by (26), where both X and Y are small. That is, we expand $F(X,Y) = \mathbb{E}[\|V_1\| \cdot \|V_2\|]$ into a Taylor polynomial of the entries of X, Y, about $X = Y = 0$.

The degree of the required Taylor polynomial in each of the variables is determined according to its (average) order of magnitude and the admissible error term. In principle, one may compute the polynomial by brute force, computing each derivative separately, but this approach results in a long and tedious computation. In this manuscript we employ Berry's method [3] in order to compute the nodal length fluctuations for the random monochromatic planar waves. The following lemma provides the Taylor approximation of $F(X,Y)$ for perturbed standard Gaussian.

Lemma 5.1. *Let $\Delta \in M_4(\mathbb{R})$ be a positive definite matrix such that*

$$\Delta = I + \begin{pmatrix} X & Y \\ Y & X \end{pmatrix},$$

where $X, Y \in M_2(\mathbb{R})$ are symmetric, $\operatorname{rank}(X) = 1$. *Define*

$$F(X,Y) = \mathbb{E}\left[\|W_1\| \cdot \|W_2\|\right],$$

where $(W_1, W_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ is centered Gaussian with covariance Δ. Then

$$F(X,Y) = \frac{\pi}{2}\Bigg(1 + \frac{\operatorname{tr} X}{2} + \frac{\operatorname{tr}(Y^2)}{8} - \frac{\operatorname{tr}(XY^2)}{16} - \frac{\operatorname{tr}(X^2)}{32} + \frac{\operatorname{tr}(Y^4)}{256} + \frac{\operatorname{tr}(Y^2)^2}{512} - \frac{\operatorname{tr} X \operatorname{tr}(Y^2)}{32}\Bigg) + O\left(\operatorname{tr}(X^3) + \operatorname{tr}(Y^6)\right).$$

Proof of Proposition 4.5 assuming Lemma 5.1. Note that since D^tD is a rank 1 matrix, it satisfies

$$\operatorname{tr}(D^tD) = DD^t.$$

A straightforward application of Lemma 5.1 with X and Y given by (27) yields (for $x \in \mathbb{T} \setminus B$, r is bounded away from ± 1, so we may write $\frac{1}{\sqrt{1-r^2}} = 1 + \frac{1}{2}r^2 + \frac{3}{8}r^4 + O(r^6)$) the following:

$$
\begin{aligned}
K_2(x) &= \frac{1}{2\pi\sqrt{1-r^2}} \cdot F(X,Y) \\
&= \frac{1}{4\sqrt{1-r^2}}\Bigg(1+\frac{\operatorname{tr} X}{2}+\frac{\operatorname{tr}(Y^2)}{8}-\frac{\operatorname{tr}(XY^2)}{16}-\frac{\operatorname{tr}(X^2)}{32} \\
&\quad +\frac{\operatorname{tr}(Y^4)}{256}+\frac{\operatorname{tr}(Y^2)^2}{512}-\frac{\operatorname{tr} X \operatorname{tr}(Y^2)}{32}\Bigg)+O\left(\operatorname{tr}(X^3)+\operatorname{tr}(Y^6)\right) \\
&= \frac{1}{4}\Bigg(1+\frac{r^2}{2}+\frac{\operatorname{tr} X}{2}+\frac{\operatorname{tr}(Y^2)}{8}+\frac{3}{8}r^4-\frac{\operatorname{tr}(XY^2)}{16}-\frac{\operatorname{tr}(X^2)}{32} \\
&\quad +\frac{\operatorname{tr}(Y^4)}{256}+\frac{\operatorname{tr}(Y^2)^2}{512}-\frac{\operatorname{tr} X \operatorname{tr}(Y^2)}{32}+\frac{1}{4}r^2 \operatorname{tr} X+\frac{1}{16}r^2 \operatorname{tr}(Y^2)\Bigg) \\
&\quad +O\left(r^6+\operatorname{tr}(X^3)+\operatorname{tr}(Y^6)\right). \qquad \square
\end{aligned}
$$

To present the proof of Lemma 5.1, we need some notation.

Notation 5.2. For a matrix A and a number a, we write $A = O(a)$ if the corresponding inequality holds entry-wise.

Notation 5.3. For $t \in 0$, we denote $m(t) := \min\{t, 1\}$, and for $t, s \in \mathbb{R}$,

$$m(t,s) := m(t) \cdot m(s).$$

Proof of Lemma 5.1. Following Berry (see [3, eq. (24)]),

$$\sqrt{\alpha} = \frac{1}{\sqrt{2\pi}} \int_0^\infty (1 - e^{-\frac{\alpha t}{2}}) \frac{dt}{t^{3/2}},$$

and we have

$$\mathbb{E}[\|W_1\| \cdot \|W_2\|] = \frac{1}{2\pi} \iint_{\mathbb{R}_+^2} [f(0,0) - f(t,0) - f(0,s) + f(t,s)] \frac{dtds}{(ts)^{3/2}}, \tag{40}$$

where

$$f^{X,Y}(t,s) = f(x,y) := \mathbb{E}\left[\exp\left(-\frac{1}{2}\left(\|W_1\|^2 + \|W_2\|^2\right)\right)\right] = \frac{1}{\det(I+M)}, \tag{41}$$

with

$$M = \begin{pmatrix} \sqrt{t}I & 0 \\ 0 & \sqrt{s}I \end{pmatrix} \Delta \begin{pmatrix} \sqrt{t}I & 0 \\ 0 & \sqrt{s}I \end{pmatrix}.$$

Now by the well-known formula for the determinant of a block matrix (see, e.g., [14, p. 210]), we have

$$\det(I+M) = \det\left((1+t)I + tX\right) \cdot \det\left((1+s)I + sX - stY((1+t)I+tX)^{-1}Y\right),$$

so that

$$
\begin{aligned}
\det(I+M)^{-1/2} &= \det\left((1+t)I + tX\right)^{-1/2} \times \\
&\quad \times \det\left((1+s)I + sX - stY((1+t)I+tX)^{-1}Y\right)^{-1/2}.
\end{aligned} \tag{42}
$$

Now we compute each of the two factors of the right-hand side of (42), up to the admissible error terms X^3 and Y^6, as in the formulation of Lemma 5.1:

$$\det(I+A)^{-1/2} = 1 - \frac{1}{2}\operatorname{tr} A + \frac{1}{4}\operatorname{tr}(A^2) + \frac{1}{8}(\operatorname{tr} A)^2 + O(A^3), \tag{43}$$

so that the first factor in the right-hand side of (42) is

$$\begin{aligned}\det\left((1+t)I+tX\right)^{-1/2} &= \frac{1}{1+t}\det\left(I+\frac{t}{1+t}X\right)^{-1/2} \\ &= \frac{1}{1+t}\cdot\left(1-\frac{t}{2(1+t)}\operatorname{tr} X+\frac{t^2}{4(1+t)^2}\operatorname{tr}(X^2)+\frac{t^2}{8(1+t)^2}(\operatorname{tr} X)^2+O(X^3)\right).\end{aligned} \tag{44}$$

To compute the second factor in the right-hand side of (42), we write

$$(I+A)^{-1} = I - A + O(A^2),$$

and we have

$$\begin{aligned}&\frac{1}{1+s}\det\left(I+\frac{s}{1+s}X-\frac{st}{(1+s)(1+t)}Y\left(I+\frac{t}{1+t}X\right)^{-1}Y\right)^{-1/2} \\
&= \frac{1}{1+s}\det\Bigg(I+\frac{s}{1+s}X-\frac{st}{(1+s)(1+t)}Y^2 \\
&\qquad\qquad +\frac{st^2}{(1+s)(1+t)^2}YXY+O(YX^2Y)\Bigg)^{-1/2} \\
&= \frac{1}{1+s}\Bigg(1-\frac{1}{2}\frac{s}{1+s}\operatorname{tr} X+\frac{1}{2}\frac{st}{(1+s)(1+t)}\operatorname{tr}(Y^2) \\
&\qquad -\frac{1}{2}\frac{st^2}{(1+s)(1+t)^2}\operatorname{tr}(YXY)+\frac{1}{4}\frac{s^2}{(1+s)^2}\operatorname{tr}(X^2) \\
&\qquad +\frac{1}{4}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^4)-\frac{1}{2}\frac{s^2t}{(1+s)^2(1+t)}\operatorname{tr}(XY^2) \\
&\qquad +\frac{1}{8}\frac{s^2}{(1+s)^2}\operatorname{tr}(X)^2+\frac{1}{8}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^2)^2 \\
&\qquad -\frac{1}{4}\frac{s^2t}{(1+s)^2(1+t)}\operatorname{tr} X\operatorname{tr}(Y^2)+O(\operatorname{tr}(X^3)+\operatorname{tr}(Y^6))\Bigg)\end{aligned} \tag{45}$$

upon using (43) with

$$A = \frac{s}{1+s}X-\frac{st}{(1+s)(1+t)}Y^2+\frac{st^2}{(1+s)(1+t)^2}YXY+O(YX^2Y).$$

Cross multiplying (44) and (45) and substituting into (42) and finally into (41), we obtain an asymptotic expression for $f^{X,Y}(t,s)$ of the form

(46)

$$
\begin{aligned}
f^{X,Y}(t,s) = \frac{1}{(1+t)(1+s)}\Bigg(&1 - \frac{1}{2}\frac{s}{1+s}\operatorname{tr} X + \frac{1}{2}\frac{st}{(1+s)(1+t)}\operatorname{tr}(Y^2) \\
&- \frac{1}{2}\frac{st^2}{(1+s)(1+t)^2}\operatorname{tr}(YXY) + \frac{1}{4}\frac{s^2}{(1+s)^2}\operatorname{tr}(X^2) \\
&+ \frac{1}{4}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^4) - \frac{1}{2}\frac{s^2t}{(1+s)^2(1+t)}\operatorname{tr}(XY^2) \\
&+ \frac{1}{8}\frac{s^2}{(1+s)^2}(\operatorname{tr} X)^2 + \frac{1}{8}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^2)^2 \\
&- \frac{1}{4}\frac{s^2t}{(1+s)^2(1+t)}\operatorname{tr} X\operatorname{tr}(Y^2) - \frac{t}{2(1+t)}\operatorname{tr} X \\
&+ \frac{t^2}{4(1+t)^2}\operatorname{tr}(X^2) + \frac{t^2}{8(1+t)^2}(\operatorname{tr} X)^2 \\
&+ \frac{1}{4}\frac{ts}{(1+t)(1+s)}(\operatorname{tr} X)^2 - \frac{1}{4}\frac{st^2}{(1+s)(1+t)^2}\operatorname{tr} X\operatorname{tr}(Y^2) \\
&+ O((\operatorname{tr}(X^3) + \operatorname{tr}(Y^6))\Bigg) \\
= \frac{1}{(1+t)(1+s)}\Bigg(&1 - \frac{1}{2}\left(\frac{s}{1+s} + \frac{t}{1+t}\right)\operatorname{tr} X + \frac{1}{2}\frac{st}{(1+s)(1+t)}\operatorname{tr}(Y^2) \\
&- \frac{1}{2}\left(\frac{st^2}{(1+s)(1+t)^2} + \frac{s^2t}{(1+s)^2(1+t)}\right)\operatorname{tr}(XY^2) \\
&+ \left(\frac{3}{8}\frac{s^2}{(1+s)^2} + \frac{3}{8}\frac{t^2}{(1+t)^2} + \frac{1}{4}\frac{ts}{(1+t)(1+s)}\right)\operatorname{tr}(X^2) \\
&+ \frac{1}{4}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^4) + \frac{1}{8}\frac{s^2t^2}{(1+s)^2(1+t)^2}\operatorname{tr}(Y^2)^2 \\
&- \frac{1}{4}\left(\frac{s^2t}{(1+s)^2(1+t)} + \frac{st^2}{(1+s)(1+t)^2}\right)\operatorname{tr} X\operatorname{tr}(Y^2) \\
&+ O(\operatorname{tr}(X^3) + \operatorname{tr}(Y^6))\Bigg),
\end{aligned}
$$

where we used

$$\operatorname{tr}(YXY) = \operatorname{tr}(XY^2),\ \operatorname{tr}(X^2) = (\operatorname{tr} X)^2, \tag{47}$$

the latter thanks to $\operatorname{rk} X = 1$.

It is important to identify (46) as the Taylor expansion of $f^{X,Y}(t,s)$ for fixed t, s, as a function of X, Y around $X = Y = 0$ (i.e., a Taylor polynomial in terms of the entries of X and Y). In the next step we perform the integration in (40) by integrating the various terms in (46). The main problem is that the error term does not depend on t,s, so that its integral against $\frac{1}{t^{3/2}s^{3/2}}$ is divergent at the origin. We improve the error term in the following way. Define

$$g^{X,Y}(t,s) := f(0,0) - f(t,0) - f(0,s) + f(t,s), \tag{48}$$

so that under the new notation, (40) is

$$\mathbb{E}[\|W_1\| \cdot \|W_2\|] = \frac{1}{2\pi} \iint\limits_{\mathbb{R}_+^2} g^{X,Y}(t,s) \frac{dtds}{(ts)^{3/2}}. \tag{49}$$

It is evident that for every X,Y the function $g^{X,Y}$ vanishes if either $t=0$ or $s=0$, so that for $t,s \geq 0$,

$$g^{X,Y}(t,s) = O_{X,Y}(ts).$$

We now substitute (46) into (48) in order to expand $g^{X,Y}(t,s)$ into a Taylor polynomial around $X=Y=0$; by the latter observation, the remainder term may be improved from $O\left(\operatorname{tr}(X^3) + \operatorname{tr}(Y^6)\right)$ to

$$O\left(m(t,s)(\operatorname{tr}(X^3) + \operatorname{tr}(Y^6))\right)$$

(recall Notation 5.3). To compute the contribution of each of the summands in (46), we notice that each summand splits into a product $\phi(t)\psi(s)$ for some ϕ and ψ (which are read off directly, for example, for the constant term $\frac{1}{(1+t)(1+s)}$, $\phi(t) = \frac{1}{1+t}$ and $\psi(s) = \frac{1}{1+s}$), so that the corresponding term of $g^{X,Y}(t,s)$ in (48) is

$$\phi(t)\psi(s) - \phi(t)\psi(0) - \phi(0)\psi(s) + \phi(0)\psi(0) = (\phi(t)-\phi(0))(\psi(s)-\psi(0)).$$

Therefore, the corresponding term in the integral (49) splits as well. We then finally obtain

$$\begin{aligned} g^{X,Y}(t,s) = {} & \frac{ts}{(1+t)(1+s)} + \frac{1}{2}\left(\frac{t}{1+t}\frac{s}{(1+s)^2} + \frac{t}{(1+t)^2}\frac{s}{1+s}\right)\operatorname{tr} X \\ & + \frac{1}{2}\frac{t}{(1+t)^2}\frac{s}{(1+s)^2}\operatorname{tr}(Y^2) \\ & - \frac{1}{2}\left(\frac{t^2}{(1+t)^3}\frac{s}{(1+s)^2} + \frac{t}{(1+t)^2}\frac{s^2}{(1+s)^3}\right)\operatorname{tr}(XY^2) \\ & - \left(\frac{3}{8}\frac{t}{1+t}\frac{s^2}{(1+s)^3} + \frac{3}{8}\frac{t^2}{(1+t)^3}\frac{s}{1+s} - \frac{1}{4}\frac{t}{(1+t)^2}\frac{s}{(1+s)^2}\right)\operatorname{tr}(X^2) \\ & + \frac{1}{4}\frac{t^2}{(1+t)^3}\frac{s^2}{(1+s)^3}\operatorname{tr}(Y^4) + \frac{1}{8}\frac{t^2}{(1+t)^3}\frac{s^2}{(1+s)^3}\operatorname{tr}(Y^2)^2 \\ & - \frac{1}{4}\left(\frac{t^2}{(1+t)^3}\frac{s}{(1+s)^2} + \frac{t}{(1+t)^2}\frac{s^2}{(1+s)^3}\right)\operatorname{tr} X \operatorname{tr}(Y^2) \\ & + O\left(m(t,s)(\operatorname{tr}(X^3) + \operatorname{tr}(Y^6))\right). \end{aligned} \tag{50}$$

Note that rather than improving the error term in the last step, we may incorporate the improvement into more precise versions of (44) and (45) and then carry the improved error term along; it would result in the same formula (50).

Inserting (50) into (49) yields

$$\mathbb{E}[\|W_1\|\cdot\|W_2\|] = \frac{\pi}{2}\Bigg(1+\frac{\operatorname{tr} X}{2}+\frac{\operatorname{tr}(Y^2)}{8}-\frac{\operatorname{tr}(XY^2)}{16}-\frac{\operatorname{tr}(X^2)}{32} + \frac{\operatorname{tr}(Y^4)}{256}+\frac{\operatorname{tr}(Y^2)^2}{512}-\frac{\operatorname{tr} X\operatorname{tr}(Y^2)}{32}\Bigg)+O\left(\operatorname{tr}(X^3)+\operatorname{tr}(Y^6)\right),$$

using the elementary integrals

$$\int_0^\infty\left(\frac{t}{1+t}\right)\frac{dt}{t^{3/2}}=\pi;\ \int_0^\infty\frac{dt}{(1+t)^2\sqrt{t}}=\frac{\pi}{2};\ \int_0^\infty\frac{\sqrt{t}dt}{(1+t)^3}=\frac{\pi}{8},$$

which is the statement of the present lemma. □

5.1. *Proof of Lemma 4.6.* To prove Lemma 4.6 we will need the following lemma (which establishes the asymptotics for some expressions involved in X and Y; see (27)), whose proof is relegated to Section 8.

LEMMA 5.4. *We have the following estimates*:

1. $\int_{\mathbb{T}} r(x)^2dx=\frac{1}{\mathcal{N}_n}$. $\int_{\mathbb{T}} r(x)^4dx=\frac{3}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
2. $\int_{\mathbb{T}} D(x)D(x)^tdx=\frac{E_n}{\mathcal{N}_n}$. $\int_{\mathbb{T}}\left(D(x)D(x)^t\right)^2dx=2\cdot\frac{E_n^2}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
3. $\int_{\mathbb{T}} r(x)^2D(x)D(x)^tdx=\frac{E_n}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
4. $\int_{\mathbb{T}}\operatorname{tr}(H(x)^2)dx=\frac{E_n^2}{\mathcal{N}_n}$. $\int_{\mathbb{T}} r(x)^2\operatorname{tr}(H(x)^2)dx=2\cdot\frac{E_n^2}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
5. $\int_{\mathbb{T}}\operatorname{tr}(H(x)^4)dx=\frac{E_n^4}{8\mathcal{N}_n^2}(11+\widehat{\mu}_n(4)^2)+O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right)$.
   $\int_{\mathbb{T}}\operatorname{tr}(H(x)^2)^2dx=\frac{E_n^4}{4\mathcal{N}_n^2}(7+\widehat{\mu}_n(4)^2)+O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right)$.
6. $\int_{\mathbb{T}} D(x)D(x)^t\operatorname{tr}(H(x)^2)dx=\frac{E_n^3}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
7. $\int_{\mathbb{T}} r(x)D(x)H(x)D(x)^tdx=-\frac{1}{2}\cdot\frac{E_n^2}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.
8. $\int_{\mathbb{T}} D(x)H(x)^2D(x)^tdx=\frac{1}{2}\cdot\frac{E_n^3}{\mathcal{N}_n^2}\left(1+O\left(\frac{1}{\mathcal{N}_n}\right)\right)$.

9. $\int_{\mathbb{T}} (D(x)D(x)^t)^3 dx = O\left(E_n^3 \mathcal{R}_6(n)\right).$

10. $\int_{\mathbb{T}} r(x)^4 D(x)D(x)^t dx = O\left(E_n \mathcal{R}_6(n)\right).$

11. $\int_{\mathbb{T}} \operatorname{tr}(H^6) dx = O\left(E_n^6 \mathcal{R}_6(n)\right).$

Proof of Lemma 4.6 assuming Lemma 5.4. In this proof we will suppress the dependence on x (and n), i.e., use the shortcuts $r = r_n(x)$, $X = X_n(x)$, $Y = Y_n(x)$, $D = D_n(x)$, $H = H_n(x)$. We have

$$\int_{\mathbb{T}} \operatorname{tr} X dx = \int_{\mathbb{T}\setminus B} \operatorname{tr} X dx + O(\operatorname{meas}(B))$$

by the uniform boundedness (30) of X. On $\mathbb{T} \setminus B$ we use the approximation

$$\frac{1}{1-r^2} = 1 + r^2 + O(r^4),$$

and since $\operatorname{meas}(B)$ is small (32), we have

$$\begin{aligned}\int_{\mathbb{T}} \operatorname{tr} X dx &= -\frac{2}{E_n}\left(\int_{\mathbb{T}} DD^t dx + \int_{\mathbb{T}} r^2 DD^t dx\right) + O\left(\mathcal{R}_6(n)\right)\\ &= -\frac{2}{\mathcal{N}_n} - \frac{2}{\mathcal{N}_n^2} + O\left(\mathcal{R}_6(n)\right)\end{aligned}$$

by parts 10, 2 and 3 of Lemma 5.4. Arguing in a similar fashion, we obtain

$$\int_{\mathbb{T}} \operatorname{tr}(Y^2) dx \sim \frac{4}{E_n^2} \int_{\mathbb{T}} \left[\operatorname{tr}(H^2) + 2rDHD^t\right] dx = \frac{4}{\mathcal{N}_n} - \frac{4}{\mathcal{N}_n^2} + O\left(\mathcal{R}_6(n)\right),$$

$$\int_{\mathbb{T}} \operatorname{tr}(XY^2) dx \sim -\frac{8}{E_n^3} \int_{\mathbb{T}} DH^2D^t dx = -\frac{4}{\mathcal{N}_n^2} + O(\mathcal{R}_5(n)),$$

$$\int_{\mathbb{T}} \operatorname{tr}(Y^4) dx \sim \frac{16}{E_n^4} \int_{\mathbb{T}} \operatorname{tr}(H^4) + O\left(\mathcal{R}_6(n)\right),$$

$$\int_{\mathbb{T}} \operatorname{tr}(Y^2)^2 dx = \frac{16}{E_n^4} \int_{\mathbb{T}} \operatorname{tr}(H^2)^2 dx + O\left(\mathcal{R}_6(n)\right) = \frac{4}{\mathcal{N}_n^2}(7 + \widehat{\mu}_n(4)^2) + O\left(\mathcal{R}_6(n)\right).$$

This shows parts 1, 2, 3, 5 and 6, parts 4, 7, 8 and 9 being similar.

To see part 10, we notice that as X is uniformly bound (30) and $\operatorname{meas}(B)$ is small (32), it is sufficient to bound the contribution on $\mathbb{T} \setminus B$ only, so that we may assume that r is bounded away from ± 1:

$$\int_{\mathbb{T}} \operatorname{tr}(X^3) dx \ll \frac{1}{E^3} \int_{\mathbb{T}} (DD^t)^3 dx + \mathcal{R}_6(n). \tag{51}$$

Part 10 of Lemma 4.6 then follows upon applying part 9 of Lemma 5.4 with (51). The proof for part 11 is very similar, using part 11 of Lemma 5.4, and we omit it here. □

6. Proof of Theorem 2.2

We begin by recalling some needed results from additive combinatorics. An additive set is a finite and nonempty subset of an ambient (additive) abelian group Z. Given an additive set A, we define $E(A,A)$, the *additive energy of* A, by

$$E(A,A) := \left|\left\{(y_1,y_2,y_3,y_4) \in A^4 : y_1+y_2 = y_3+y_4\right\}\right|.$$

We shall use the following "large energy version" of the Balog-Szemeredi-Gowers theorem (see [25, Chs. 2.4–5]).

THEOREM 6.1 (BSG). *Let A be an additive set, and let $K \geq 1$. There exists an absolute constant C with the following property. If $E(A,A) \geq |A|^3/K$, then there exists a subset $A_1 \subseteq A$ satisfying*

$$|A_1| > K^{-C}|A| \tag{52}$$

and

$$|A_1 + A_1| < K^{C}|A|. \tag{53}$$

Remark 6.2. Theorem 6.1 can easily be deduced from Proposition 2.26 and Theorem 2.31 of [25] as follows. By Theorem 2.31, $E(A,A) \geq |A|^3/K$ implies that there exist subsets $A_1 \subseteq A$, $A_2 \subseteq A$ with $|A_1| > K^{-C'}|A|, |A_2| > K^{-C'}|A|$ satisfying $d(A_1,A_2) \leq C' \log K$ (where $d(A_1,A_2)$ denotes the Rusza distance between A_1, A_2, and C' is an absolute constant.) By Proposition 2.26, $d(A_1,A_2) \leq C' \log K$ implies that $|A_1 + A_1| \leq K^{C''}|A_1|$ for some C'' only depending on C'. Taking $C = \max(C', C'')$, the result follows.

If G is a (torsion free) abelian group, a Generalized Arithmetic Progression (GAP) of dimension d is a subset $P \subseteq G$ of the form

$$P = \left\{\xi_0 + \sum_{k=1}^{d} j_k \xi_k : 0 \leq j_k < J_k \text{ for } k = 1,\dots,d\right\}, \tag{54}$$

with $\xi_0,\dots,\xi_d \in G$. A GAP P is called *proper* if $|P| = \prod_{k=1}^{d} J_k$ (i.e., all elements in the sum $\xi_0 + \sum_{k=1}^{d} j_k\xi_k$ are distinct). It is easy to see that a GAP has "bounded doubling," i.e., that $|A+A|/|A|$ is "small." A surprising converse is Freiman's celebrated structure theorem — an additive set with small doubling is essentially a proper GAP.

THEOREM 6.3 ([25, Th. 5.33]). *Let A be an additive set in a torsion free group G such that $|A+A| \le K|A|$. Then there exists a proper generalized arithmetic progression P, of rank at most $K-1$, that contains A such that*

$$|P| \le \exp\left(O\left(K^{O(1)}\right)\right)|A|.$$

If $A \subset \mathbb{C}$ and $z \in \mathbb{C}$, let $r_2(z, A)$ denote the number of representations of z as a product of two elements from A. The following result by Chang shows that $r_2(z, A)$ is quite small when A is a GAP.

PROPOSITION 6.4 ([11, Prop. 3]). *Let $P \subseteq \mathbb{C}$ be a* GAP *of the form* (54), *where $\xi_0, \ldots, \xi_d \in \mathbb{C}$. Then, for all $z \in \mathbb{C}$,*

$$r_2(z, P) < \exp\left(C_d \frac{\log J}{\log\log J}\right), \tag{55}$$

where $J = \max_{1\le k<d} J_k$ and the constant C_d only depends on the dimension d of P.

Proof of Theorem 2.2. Assume that $|S_n| = o(\mathcal{N}_n^4)$ does not hold, i.e., that there exists some $\delta > 0$ such that

$$|S_6(n)| > \delta \mathcal{N}_n^4 \tag{56}$$

for $\mathcal{N}_n$ arbitrarily large. Using sum-product type estimates, we will show that this leads to a contradiction.

To simplify the notation, let $S = S_6(n)$ and $N = \mathcal{N}_n$. From this point on in this proof we assume that δ is fixed; we will write $F \lesssim G$ for some expressions F, G (resp. $F \gtrsim G$), if there exists a constant C (which may depend on δ only), such that $F \le C \cdot G$ (resp. $F \ge C \cdot G$).

Define

$$A = A_n := (\Lambda_n + \Lambda_n) \setminus \{0\}. \tag{57}$$

Note that A then consists of elements that have two (or exactly one for elements of the form 2λ, $\lambda \in \Lambda_n$) representations as sums of elements of Λ_n. Also also note that A is symmetric around the origin. Thus

$$|A| = N^2/2 + O(N), \tag{58}$$

and (56) implies

$$|\{(y_1, y_2) \in A \times A : y_1 + y_2 \in A\}| \gtrsim N^4. \tag{59}$$

(Note that the number of solutions to $\sum_{i=1}^6 \lambda_i = 0$ with the additional constraint that one of $\lambda_1 + \lambda_2, \lambda_3 + \lambda_4, \lambda_5 + \lambda_6$ equals zero is $O(N^3)$; this follows immediately on noting that $\lambda_i + \lambda_j = z$ has at most four solutions if $z \neq 0$.)

Letting $\mathbb{1}_A$ denote the characteristic function of the set $A \subseteq \mathbb{Z}^2$, we have

$$|\{(y_1, y_2) \in A \times A : y_1 + y_2 \in A\}| = \langle \mathbb{1}_A \star \mathbb{1}_A, \mathbb{1}_A \rangle, \tag{60}$$

where we understand both the inner product and the convolution as defined on $L^2(\mathbb{Z}^2)$. Together with (59) and the Cauchy-Schwarz inequality, the observation (60) yields

$$N^4 \lesssim \langle \mathbb{1}_A \star \mathbb{1}_A, \mathbb{1}_A \rangle \le \|\mathbb{1}_A \star \mathbb{1}_A\|_2 \cdot |A|^{1/2} \sim \frac{1}{\sqrt{2}} \|\mathbb{1}_A \star \mathbb{1}_A\|_2 \cdot N.$$

We may hence estimate the additive energy of A as

$$\begin{aligned} E(A, A) &= |\{(y_1, y_2, y_3, y_4) \in A^4 : y_1 + y_2 = y_3 + y_4\}| \\ &= \|\mathbb{1}_A \star \mathbb{1}_A\|_2^2 \gtrsim N^6 \gtrsim |A|^3. \end{aligned}$$

We now apply Theorem 6.1 on A with $K = K(\delta)$ constant to construct a large subset $A_1 \subseteq A$ having the "bounded doubling" property (53) and, in addition, (52). Together with (58), the latter implies

$$|A_1| \gtrsim N^2. \tag{61}$$

Hence, by applying Theorem 6.3 with $G = \mathbb{Z}^2$ and $A_1 \subset G$, there exists a proper GAP

$$P = \left\{ \xi_0 + \sum_{k=1}^{d} j_k \xi_k : 0 \le j_k < J_k \text{ for } k = 1, \dots, d \right\}, \tag{62}$$

as in (54), of bounded dimension (depending on δ only)

$$d(\delta) = d(K(\delta)), \tag{63}$$

so that $A \subseteq P$ and

$$|P| \le \exp\left(O\left(K^{O(1)} \right) \right) |A_1| \lesssim |A_1|.$$

We then have

$$|A_1| = |P \cap A_1| \le |P \cap A| \le \sum_{x \in \Lambda_n} |(P - x) \cap \Lambda_n|,$$

where for the latter inequality we used the definition (57) of A. Hence, by (61),

$$N^2 \lesssim \sum_{x \in \Lambda_n} |(P - x) \cap \Lambda_n|,$$

and therefore (the length of summation being N)

$$|(P - x) \cap \Lambda_n| \gtrsim N$$

for some $x \in \Lambda_n$. Replacing P by $P - x$ if necessary, we may assume that

$$|P \cap \Lambda_n| \gtrsim N. \tag{64}$$

Using Chang's Proposition 6.4 the latter leads to a contradiction as follows. If $P = \left\{\xi_0 + \sum_{k=1}^{d} j_k\xi_k : 0 \le j_k < J_k \text{ for } k = 1, \dots, d\right\}$, then $P \cup \overline{P}$ is contained in a GAP, of dimension $2d+1$, of the form

$$P' = \left\{ \xi_0 + j_0(\overline{\xi_0} - \xi_0) + \sum_{k=1}^{d} j_k\xi_k + \sum_{k=d+1}^{2d} j_k\overline{\xi_{k-d}} \right\},$$

where $0 \le j_0 < 2$, $0 \le j_k < J_k$ for $k = 1, \dots, d$ and $0 \le j_k < J_{k-d}$ for $k = d+1, \dots, 2d$.

Considering P' as a subset of $\mathbb{Z} + i\mathbb{Z}$, it is clear (since for every $z \in \Lambda_n$, $z \cdot \overline{z} = n$) that

$$r_2(n, P') \ge |P' \cap \Lambda_n| \gtrsim N, \tag{65}$$

by (64). On the other hand, Proposition 6.4 applied on P' implies that

$$r_2(n, P') < \exp\left(C_{2d+1}\frac{\log J}{\log\log J}\right),$$

where

$$J = \max_{1\le k\le d} J_k \le |P| \lesssim N^2,$$

with J_k as in (62), and thus

$$r_2(n, P') < \exp\left(C_{2d+1}\frac{\log N}{\log\log N}\right).$$

Combined with (65) the latter estimate implies

$$N \lesssim \exp\left(C_{2d+1}\frac{\log N}{\log\log N}\right)$$

or, taking logarithm of both sides,

$$\log N \le C\frac{\log N}{\log\log N}$$

for some $C = C(\delta)$ that may depend on δ only (by (63)). This is clearly impossible for N arbitrarily large, and the desired contradiction concludes the proof. $\square$

7. Probability measures on $\mathcal{S}^1$ arising from Λ_n

Recalling that $S = \{n \in \mathbb{Z} : n = a^2 + b^2,\ a, b \in \mathbb{Z}\}$, define

$$S(x) := \{n \in S : n \le x\},$$

and for a subset $S' \subseteq S$, similarly define $S'(x) := \{n \in S' : n \le x\}$. We say that a set $S' \subseteq S$ has *asymptotic density* $s \in [0,1]$ if $\lim_{x\to\infty} \frac{S'(x)}{S(x)} = s$. Further, we say that a subsequence $(n_i)_{i\ge1}$ of elements in S is *thin* if the subset $\{n_i\}_{i\ge1} \subset S$ has asymptotic density zero.

It is known [19] that as $x \to \infty$, $S(x) \sim \frac{cx}{\sqrt{\log x}}$, where $c > 0$ is known as the Landau-Ramanujan constant. In particular, $\mathcal{N}_n$ grows as $\sim c\sqrt{\log x}$ on average for $n \le x$. Moreover, a straightforward modification of an Erdös-Kac type argument to the set S shows that

$$|\{n \in S(x) : \log \mathcal{N}_n \gg \log\log n\}| = |S(x)| \cdot (1 + o(1))$$

as $x \to \infty$, and consequently there exists a density one subset $S' \subset S$ such that $\mathcal{N}_n \to \infty$ if $n \in S'$ and $n \to \infty$.

Further, the lattice points Λ_n are equidistributed on $\mathcal{S}^1$ along *generic* subsequences of energy levels (see, e.g., [16, Prop. 6]) in the following sense.[3] There exists a density 1 subsequence $S'' \subseteq S$ so that $\mu_n \underset{n \in S''}{\Rightarrow} \nu$, where ν is the uniform probability measure $d\nu(\theta) = \frac{1}{2\pi} d\theta$ on $\mathcal{S}^1 \cong \mathbb{R}/2\pi\mathbb{Z}$. (As usual, the notation $\upsilon_i \Rightarrow \upsilon$ stands for weak convergence of probability measures on $\mathcal{S}^1$, i.e., that $\int f d\upsilon_i \to \int f d\upsilon$ for every continuous bounded test function f.) In particular, for a generic sequence of elements $n \in S$, $\mathcal{N}_n \to \infty$ and the points in Λ_n are equidistributed in $\mathcal{S}^1$.

In the other direction, Cilleruelo [13] has shown that there are thin sequences $(E_{n_i})_{i \ge 1}$ with $\mathcal{N}_{n_i} \to \infty$ such that μ_{n_i} converges to the atomic probability measure supported at the 4 symmetric points ± 1, $\pm i$:

$$\mu_{n_i} \Rightarrow \nu_0 := \frac{1}{4} \sum_{k=0}^{3} \delta_{i^k}. \tag{66}$$

7.1. *Some number theoretic prerequisites on Gaussian integers.* Before proceeding with the proof of Proposition 1.2, we begin with some number theoretic preliminaries on μ_n (see, e.g., [13]).

To describe μ_n we recall some basic facts about Gaussian integers. Given a prime $p \equiv 1 \mod 4$, the equation $x^2 + y^2 = p$ has exactly eight solutions in integers x, y, and there is a unique solution satisfying $0 \le y_p \le x_p$. We can hence attach an angle $\theta_p \in [0, \pi/4]$ to each such p by writing $x_p + iy_p = \sqrt{p}e^{i\theta_p}$. On the other hand, given a prime $q \equiv 3 \mod 4$, the equation $x^2 + y^2 = q$ has no solutions, whereas $x^2 + y^2 = 2$ has exactly four solutions. Moreover, the following holds for the ring of Gaussian integers: the units are given by i^k for $k \in \{0, 1, 2, 3\}$, the set of Gaussian primes are, up to units, given by $1+i$, primes $q \in \mathbb{Z}^+$ with $q \equiv 3 \mod 4$, and to each prime $p \equiv 1 \mod 4$ there correspond two Gaussian primes, namely $x_p + iy_p = \sqrt{p}e^{i\theta_p}$ and $x_p - iy_p = \sqrt{p}e^{-i\theta_p}$.

[3]Proposition 6 in [16] implies that all the exponential sums are $o(1)$ for a density one sequence of energy levels. The equidistribution follows from the Weyl's criterion.

The elements of Λ_n can then be parametrized as follows. Let

$$n = 2^{e_2} \cdot \prod_{p^{e_p}||n} p_i^{e_p} \cdot \prod_{q_i^{2e_q}||n} q_i^{2e_q},$$

where p_i and q_i are all the primes satisfying $p_i \equiv 1 \mod 4$ and $q \equiv 3 \mod 4$. Each pair (x, y) arises as follows. With $z = x + iy$, we have

$$z = x + iy = i^k \cdot (1+i)^{e_2} \cdot \prod_{p^{e_p}||n} (\sqrt{p}^{e_p} e^{i(e_p - 2l_p)\theta_p}) \cdot \prod_{q_i^{2e_q}||n} q_i^{e_q},$$

where $k \in \{0, 1, 2, 3\}$ and $0 \le l_p \le e_p$ for each $p|n$.

We can now describe μ_n as convolutions over prime powers. Define

$$\mu_1 := \frac{1}{4} \sum_{k=0}^{3} \delta_{i^k}$$

($\mu_1 = \nu_0$ as in (66)), $\tilde{\mu}_{2^{e_2}} := \delta_{((1+i)/\sqrt{2})^{e_2}}$, and

$$\tilde{\mu}_{p^{e_p}} := \frac{1}{e_p + 1} \sum_{l_p=0}^{e_p} \delta_{e^{i(e_p - 2l_p)\theta_p}},$$

(the "desymmetrized" version of μ_n). Then

$$\mu_n = \mu_1 \star (\star_{p|n} \tilde{\mu}_{p^{e_p}}),$$

where the convolution of two measures μ, μ' on $\mathcal{S}^1$ is given by

$$(\mu \star \mu')(z) = \int_{S^1} \mu(w) \mu'(z/w) dw.$$

7.2. *Proof of Proposition 1.2.*

Proof. That ν_0 and $\nu_{\pi/4}$ arise as weak limits of $(\mu_{n_i})_{i \ge 1}$ was already noted in the introduction of Section 7. To show that the same is true for ν_a for $a \in (0, \pi/4)$, we argue as follows. An easy consequence of Gaussian primes being equidistributed in sectors (cf. [17], [18]) is that there exists an infinite sequence of primes $p_1 < p_2 < \cdots$ such that $\theta_{p_j} \to 0$, where each $p_j \equiv 1 \mod 4$. (Also see [13].) To proceed we construct a sequence of integers n_j such that

$$\tilde{\mu}_{n_k} \to \tilde{\nu}_a; \tag{67}$$

this will immediately imply that $\mu_{n_k} \to \nu_a$ since μ_{n_k} and ν_a are the symmetrized versions of $\tilde{\mu}_{n_k}$ and $\tilde{\nu}_a$ respectively.

Thus, let $e_k = [a/\theta_{p_k}]$, where $[a/\theta_{p_k}]$ denotes the integer part of a/θ_{p_k}, and define

$$n_k = p_k^{e_k}.$$

Then, for f any continuous function on S^1,

$$\int_{S^1} f(\theta) d\tilde{\mu}_{n_k}(\theta) = \frac{1}{e_k + 1} \sum_{l=0}^{e_k} f(e^{i(e_k - 2l)\theta_{p_k}}). \tag{68}$$

The latter is the $2\theta_{p_k}$-spaced Riemann sum for the integral

$$\frac{1}{2\alpha_k}\int_{-\alpha_k}^{\alpha_k} f(e^{i\theta})d\theta$$

with

$$\alpha_k = \theta_{p_k}\cdot[a/\theta_{p_k}] = a - \theta_{p_k}\cdot\{a/\theta_{p_k}\},$$

where $\{\cdot\}$ is the fractional part of a real number.

Note that since $\theta_{p_k}\to 0$ (so that the Riemann sum spacing vanishes),

$$|\theta_{p_k}\cdot\{a/\theta_{p_k}\}| \le \theta_{p_k}\to 0,$$

and so $\alpha_k\to a$. Therefore, as $k\to\infty$, we have

$$\int_{S^1} f(\theta)d\tilde{\mu}_{n_k}(\theta) = \frac{1}{e_k+1}\sum_{l=0}^{e_k} f(e^{i(e_k-2l)\theta_{p_k}}) \to \frac{1}{2a}\int_{-a}^{a} f(e^{i\theta})d\theta$$

for any continuous test function f. Thus all ν_a are attainable as limiting measures, and the proof of the first statement is concluded.

To see that the Fourier coefficient $\widehat{\nu}_a(4)$, for $a\in[0,\frac{\pi}{4}]$, attains all values in $[0,1]$, it is sufficient to notice that

$$\widehat{\nu}_0(4) = \widehat{\mu_1}(4) = 1, \quad \widehat{\nu_{\pi/4}}(4) = 0,$$

and clearly the function $a\mapsto\widehat{\nu}_a(4)$ is continuous. Therefore, by the intermediate value theorem, given any value $b\in[0,1]$, there exists a number $a=a(b)$ so that $\nu_a(4) = b$. Since ν_a is attainable for all a, the second statement of Proposition 1.2 follows. □

8. **Proof of Lemma 5.4**

For a probability measure μ on $\mathcal{S}^1$, we define

$$B_4(\mu) := \int_{z_1,z_2\in\mathcal{S}^1} \langle z_1, z_2\rangle^4 d\mu(z_1)d\mu(z_2),$$

the fourth moment of cosine of the angle between two random points on $\mathcal{S}^1$ drawn independently according to μ. For instance, if $\mu=\mu_n$ are the atomic measures in (7), then

$$B_4(\mu_n) = \frac{1}{\mathcal{N}_n^2 n^4}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\langle\lambda_1,\,\lambda_2\rangle^4 = \frac{1}{\mathcal{N}_n^2}\sum_{\lambda_1,\lambda_2\in\Lambda_n}\cos(\theta(\lambda_1,\lambda_2))^4 \tag{69}$$

is the fourth moment of cosine of the angle $\theta(\lambda_1,\lambda_2)$ between two random uniformly and independently drawn Λ_n-points λ_1,λ_2. While the expression $B_4(n)$ comes up naturally from some of the expressions evaluated in Lemma 5.4, it is simply related to $\widehat{\mu}_n(4)$ as in the following lemma.

LEMMA 8.1. *For any probability measure μ on $\mathcal{S}^1$, invariant with respect to $x \mapsto ix$ and $x \mapsto \bar{x}$, we have*

$$B_4(\mu) = \frac{3}{8} + \frac{1}{8}\widehat{\mu}(4)^2.$$

Proof. For $z_1, z_2 \in \mathcal{S}^1$, use

$$\langle z_1, z_2 \rangle = \frac{z_1\bar{z_2} + \bar{z_1}z_2}{2}$$

together with the binomial formula for $\langle z_1, z_2 \rangle^4$ or, alternatively, the standard identity

$$\cos(\theta)^4 = \frac{3}{8} + \frac{1}{8}\cos(4\theta) + \frac{1}{2}\cos(2\theta),$$

to rewrite $B_4(\mu)$ as

$$B_4(\mu) = \frac{3}{8} + \int\limits_{z_1, z_2 \in \mathcal{S}^1} \left(\frac{1}{8}\Re\left(z_1^4\bar{z_2}^4\right) + \frac{1}{2}\Re\left(z_1^2\bar{z_2}^2\right)\right) d\mu(z_1)d\mu(z_2).$$

The statement of the present lemma follows upon noting that $\int\limits_{\mathcal{S}} z^4 d\mu(z) = \int\limits_{\mathcal{S}} \bar{z}^4 d\mu(z) = \widehat{\mu}(4) \in \mathbb{R}$ and $\int\limits_{\mathcal{S}} z^2 d\mu(z) = \int\limits_{\mathcal{S}} \bar{z}^2 d\mu(z)$ vanish by the symmetry assumptions. □

Proof of Lemma 5.4. In order to evaluate the integrals we will use (11), (22) and (23), and the orthogonality relations of the exponentials

$$\int\limits_{\mathbb{T}} e\left(\langle \lambda, x \rangle\right) dx = \begin{cases} 1 & \lambda = 0, \\ 0 & \text{otherwise}. \end{cases} \tag{70}$$

Most of the computations are similar in nature, and we will only show a few examples in detail, omitting the rest.

The statement of part 1 of the present lemma concerning the second moment of r is evident in light of (11) and (70). For the fourth moment, we have

$$\int\limits_{\mathbb{T}} r(x)^4 dx = \frac{1}{\mathcal{N}_n^4}|S_4(n)|, \tag{71}$$

where

$$S_4(n) = \left\{(\lambda_1, \ldots, \lambda_4) \in \Lambda_n : \sum_{i=1}^{4} \lambda_i = 0\right\}$$

is the length-4 correlation set of frequencies (cf. (18)). Note that since two circles may have at most two intersections (i.e., circles of radius $\sqrt{n}$ centered at 0 and $\lambda_1 + \lambda_2$), $(\lambda_1, \dots, \lambda_4) \in S_4(n)$ implies that either of the following holds:

$$(\lambda_1 = -\lambda_2 \text{ and } \lambda_3 = -\lambda_4) \text{ or}$$
$$(\lambda_1 = -\lambda_3 \text{ and } \lambda_2 = -\lambda_4) \text{ or} \tag{72}$$
$$(\lambda_1 = -\lambda_4 \text{ and } \lambda_2 = -\lambda_3).$$

Conversely, every tuple of either of the forms above is lying inside $S_4(4)$. In particular,

$$|S_4(n)| = 3\mathcal{N}_n^2 \left(1 + O\left(\frac{1}{\mathcal{N}_n}\right)\right),$$

the error term being an artifact of the existence of degenerate tuples of the form

$$(\pm\lambda, \pm\lambda, \pm\lambda, \pm\lambda) \in S_4(4)$$

(with precisely two plus and two minus signs). Part 1 of the present lemma then follows upon substituting the latter into (71). We will use the fine structure (72) of $S_4(n)$ in the course of the proof of most of the other statements of the present lemma.

Now we turn to part 2 of the present lemma. While the first statement is clear from (22) and (70), to show the other statement, we invoke the fine structure (72) of $S_4(n)$. We have

$$\int_{\mathbb{T}} (D(x)D(x)^t)^2 dx = \frac{(2\pi)^4}{\mathcal{N}_n^4} \sum_{(\lambda_1,\dots,\lambda_4)\in S_4(n)} \lambda_1\lambda_2^t\lambda_3\lambda_4^t$$
$$= \frac{(2\pi)^4}{\mathcal{N}_n^4}\left[\mathcal{N}_n^2 n^2 + \sum_{\lambda_1,\lambda_2\in\Lambda_n} \langle\lambda_1,\lambda_2\rangle^2 + \sum_{\lambda_1,\lambda_2\in\Lambda_n} \lambda_1\lambda_2^t\lambda_2\lambda_1 + O\left(\mathcal{N}n^2\right)\right].$$

The result of the present computation then follows upon making the simple observations

$$\sum_{\lambda\in\Lambda_n} \langle\lambda,\xi\rangle^2 = \frac{1}{2}\mathcal{N}_n n\|\xi\|^2$$

for every $\xi \in \mathbb{R}^2$ (see [22, Lemma 5.2]) and

$$\sum_{\lambda\in\Lambda_n} \lambda^t\lambda = \frac{1}{2}\mathcal{N}_n n \cdot I_2.$$

To show part 3, we note

$$\int_{\mathbb{T}} r(x)^2 D(x)D(x)^t dx = -\frac{(2\pi)^2}{\mathcal{N}^4} \sum_{(\lambda_1,\dots,\lambda_4)\in S_4(n)} \lambda_3\lambda_4^t,$$

and only tuples with $\lambda_3 = -\lambda_4$ contribute to the latter summation, i.e., those of the first type in (72). The computation for part 4 is very similar to what we encountered before, using (23) with (70) for the first statement and exploiting the fine structure (72) of $S_4(n)$ for the second one:

$$\int_{\mathbb{T}} \operatorname{tr}(H(x)^2)dx = \frac{(4\pi^2)^2}{\mathcal{N}^2} \sum_{(\lambda_1,\dots,\lambda_4)\in S_4(n)} \lambda_3^t\lambda_3\lambda_4^t\lambda_4.$$

To compute the integrals in part 5, we exploit Lemma 8.1. Similarly to the previous computations, by (23) and (70), we have

$$\begin{aligned}
&\int_{\mathbb{T}} \operatorname{tr}(H(x)^4)dx = \frac{(4\pi^2)^4}{\mathcal{N}_n^4} \sum_{(\lambda_1,\dots,\lambda_4)\in S_4(n)} \operatorname{tr}(\lambda_1^t\lambda_1\lambda_2^t\lambda_2\lambda_3^t\lambda_3\lambda_4^t\lambda_4)\\
&= \frac{(4\pi^2)^4}{\mathcal{N}_n^4}\Bigg[\sum_{\lambda_1,\lambda_2\in\Lambda_n} \operatorname{tr}\left(\lambda_1^t\lambda_1\lambda_1^t\lambda_1\lambda_2^t\lambda_2\lambda_2^t\lambda_2\right) + \sum_{\lambda_1,\lambda_2\in\Lambda_n} \operatorname{tr}\left(\lambda_1^t\lambda_1\lambda_2^t\lambda_2\lambda_1^t\lambda_1\lambda_2^t\lambda_2\right)\\
&\qquad + \sum_{\lambda_1,\lambda_2\in\Lambda_n} \operatorname{tr}\left(\lambda_1^t\lambda_1\lambda_2^t\lambda_2\lambda_2^t\lambda_2\lambda_1^t\lambda_1\right)\Bigg] + O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right)\\
&= \frac{(4\pi^2)^4}{\mathcal{N}_n^4}\Bigg[n^2\sum_{\lambda_1,\lambda_2\in\Lambda_n}\langle\lambda_1,\lambda_2\rangle^2 + \sum_{\lambda_1,\lambda_2\in\Lambda_n}\langle\lambda_1,\lambda_2\rangle^4\\
&\qquad + n^2\sum_{\lambda_1,\lambda_2\in\Lambda_n}\langle\lambda_1,\lambda_2\rangle^2\Bigg] + O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right)\\
&= \frac{E_n^4}{\mathcal{N}_n^2}\left(1+B_4(n)\right) + O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right),
\end{aligned}$$

where we used sums as above and the definition (69) of $B_4(n)$. Using Lemma 8.1, we may then rewrite the latter expression in terms of $\widehat{\mu}_n(4)$, as in the statement of the present lemma.

A similar computation shows that the second integral in part 5 is given by

$$\int_{\mathbb{T}} \operatorname{tr}(H(x)^2)^2dx = \frac{E_n^4}{\mathcal{N}_n^2}\left(1+2B_4(n)\right) + O\left(\frac{E_n^4}{\mathcal{N}_n^3}\right),$$

and using Lemma 8.1 again yields the result given. Evaluating the integrals for parts 6–8 of the present lemma is straightforward and very similar to the above computations, and we omit it here.

We now prove part 9 of the present lemma. By symmetry, we have

$$\int_{\mathbb{T}} (D(x)D(x)^t)^3dx = \int_{\mathbb{T}} \|D(x)\|^6dx \ll \int_{\mathbb{T}} \left(\frac{\partial r}{\partial x_1}\right)^6 dx, \tag{73}$$

whence

$$\begin{aligned}
\int_{\mathbb{T}} \left(\frac{\partial r}{\partial x_1}\right)^6 dx &= \frac{(2\pi)^6}{\mathcal{N}_n^6}\int_{\mathbb{T}}\left(\sum_{\lambda=(\lambda^1,\lambda^2)\in\Lambda}\lambda^1\sin(2\pi\langle\lambda,x\rangle)\right)^6 dx\\
&= \frac{(2\pi)^6}{\mathcal{N}_n^6}\sum_{\substack{\lambda_1,\dots,\lambda_6\in\Lambda\\ \sum_{i=1}^6\lambda_i=0}}\lambda_1^1\cdot\ldots\cdot\lambda_6^1 \ll \frac{E^3}{\mathcal{N}_n^6}\cdot|S_6(n)| = E^3\int_{\mathbb{T}} r(x)^6dx,
\end{aligned} \tag{74}$$

by (17), where we used the uniform bound

$$|\lambda_i^1| \leq \sqrt{n} \ll \sqrt{E}.$$

The present statement of Lemma 5.4 then follows upon substituting (74) into (73). The proofs for the last parts 10 and 11 of the present lemma are very similar, and we omit them here. □

References

[1] P. Bérard, Volume des ensembles nodaux des fonctions propres du laplacien, in *Bony-Sjöstrand-Meyer Seminar*, 1984–1985, École Polytech., Palaiseau, 1985, pp. Exp. No. 14 , 10. MR 0819780. Zbl 0589.58033.

[2] M. V. Berry, Regular and irregular semiclassical wavefunctions, *J. Phys. A* **10** (1977), 2083–2091. MR 0489542. Zbl 0377.70014. http://dx.doi.org/10.1088/0305-4470/10/12/016.

[3] M. V. Berry, Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature, *J. Phys. A* **35** (2002), 3025–3038. MR 1913853. Zbl 1044.81047. http://dx.doi.org/10.1088/0305-4470/35/13/301.

[4] E. Bogomolny and C. Schmit, Percolation model for nodal domains of chaotic wave functions, *Phys. Rev. Lett.* **88** (2002), 114102. http://dx.doi.org/10.1103/PhysRevLett.88.114102.

[5] E. Bombieri and J. Bourgain, A problem on sums of two squares, to be published.

[6] J. Bourgain and Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces, *C. R. Math. Acad. Sci. Paris* **347** (2009), 1249–1253. MR 2561032. Zbl 1192.58017. http://dx.doi.org/10.1016/j.crma.2009.08.008.

[7] J. Bourgain and Z. Rudnick, On the geometry of the nodal lines of eigenfunctions of the two-dimensional torus, *Ann. Henri Poincaré* **12** (2011), 1027–1053. MR 2823207. Zbl 1227.58008. http://dx.doi.org/10.1007/s00023-011-0098-z.

[8] J. Bourgain and Z. Rudnick, On the nodal sets of toral eigenfunctions, *Invent. Math.* **185** (2011), 199–237. MR 2810800. Zbl 1223.58025. http://dx.doi.org/10.1007/s00222-010-0307-2.

[9] J. Brüning, Über Knoten von Eigenfunktionen des Laplace-Beltrami-Operators, *Math. Z.* **158** (1978), 15–21. MR 0478247. Zbl 0349.58012.

[10] J. Brüning and D. Gromes, Über die Länge der Knotenlinien schwingender Membranen, *Math. Z.* **124** (1972), 79–82. MR 0287202. Zbl 0226.35078.

[11] M. Chang, Factorization in generalized arithmetic progressions and applications to the Erdős-Szemerédi sum-product problems, *Geom. Funct. Anal.* **13** (2003), 720–736. MR 2006555. Zbl 1029.11006. http://dx.doi.org/10.1007/s00039-003-0428-5.

[12] S. Y. Cheng, Eigenfunctions and nodal sets, *Comment. Math. Helv.* **51** (1976), 43–55. MR 0397805. Zbl 0334.35022.

[13] J. Cilleruelo, The distribution of the lattice points on circles, *J. Number Theory* **43** (1993), 198–202. MR 1207499. Zbl 0777.11036. http://dx.doi.org/10.1006/jnth.1993.1017.

[14] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes. Sample Function Properties and their Applications*, Dover Publications, Mineola, NY, 2004, reprint of the 1967 original. MR 2108670. Zbl 0162.21102.

[15] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions on Riemannian manifolds, *Invent. Math.* **93** (1988), 161–183. MR 0943927. Zbl 0659.58047. http://dx.doi.org/10.1007/BF01393691.

[16] L. Fainsilber, P. Kurlberg, and B. Wennberg, Lattice points on circles and discrete velocity models for the Boltzmann equation, *SIAM J. Math. Anal.* **37** (2006), 1903–1922. MR 2213399. Zbl 1141.11046. http://dx.doi.org/10.1137/040618916.

[17] E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, *Math. Z.* **1** (1918), 357–376. MR 1544302. JFM 46.0258.01. http://dx.doi.org/10.1007/BF01465095.

[18] E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, *Math. Z.* **6** (1920), 11–51. MR 1544392. JFM 47.0152.01. http://dx.doi.org/10.1007/BF01202991.

[19] E. Landau, Uber die Einteilung der positiven Zahlen nach vier Klassen nach der Mindestzahl der zu ihrer addition Zusammensetzung erforderlichen Quadrate, *Arch. Math, und Phys.* **III** (1908).

[20] F. Nazarov and M. Sodin, On the number of nodal domains of random spherical harmonics, *Amer. J. Math.* **131** (2009), 1337–1357. MR 2555843. Zbl 1186.60022. http://dx.doi.org/10.1353/ajm.0.0070.

[21] L. I. Nicolaescu, Critical sets of random smooth functions on compact manifolds. arXiv 1101.5990.

[22] F. Oravecz, Z. Rudnick, and I. Wigman, The Leray measure of nodal sets for random eigenfunctions on the torus, *Ann. Inst. Fourier* (*Grenoble*) **58** (2008), 299–335. MR 2401223. Zbl 1153.35058. Available at http://aif.cedram.org/item?id=AIF_2008__58_1_299_0.

[23] I. Polterovich and I. Wigman, Universality of length distribution of nodal lines of random waves on generic surfaces, in progress.

[24] Z. Rudnick and I. Wigman, On the volume of nodal sets for eigenfunctions of the Laplacian on the torus, *Ann. Henri Poincaré* **9** (2008), 109–130. MR 2389892. Zbl 1142.60029. http://dx.doi.org/10.1007/s00023-007-0352-6.

[25] T. Tao and V. H. Vu, *Additive Combinatorics*, *Cambridge Stud. Adv. Math.* **105**, Cambridge Univ. Press, Cambridge, 2006. MR 2289012. Zbl 1127.11002. http://dx.doi.org/10.1017/CBO9780511755149.

[26] J. A. Toth and I. Wigman, Counting open nodal lines of random waves on planar domains, *Int. Math. Res. Not.* **2009**, no. 18, 3337–3365. MR 2535002. Zbl 1184.35125. http://dx.doi.org/10.1093/imrn/rnp052.

[27] I. Wigman, Fluctuations of the nodal length of random spherical harmonics, *Comm. Math. Phys.* **298** (2010), 787–831. MR 2670928. Zbl 1213.33019. http://dx.doi.org/10.1007/s00220-010-1078-8.

[28] I. Wigman, On the nodal lines of random and deterministic Laplace eigenfunctions, in *Spectral Geometry* (Proc. Internatl. Conf., Dartmouth College, July

19–23, 2010), *Proc. Sympos. Pure Math.* **84**, Amer. Math. Soc., Providence, RI, 2012. arXiv 1103.0150.

[29] S. T. Yau, Survey on partial differential equations in differential geometry, in *Seminar on Differential Geometry, Ann. of Math. Stud.* **102**, Princeton Univ. Press, Princeton, N.J., 1982, pp. 3–71. MR 0645729. Zbl 0478.53001.

[30] S.-T. Yau, Open problems in geometry, in *Differential Geometry: Partial Differential Equations on Manifolds* (Los Angeles, CA, 1990), *Proc. Sympos. Pure Math.* **54**, Amer. Math. Soc., Providence, RI, 1993, pp. 1–28. MR 1216573. Zbl 0801.53001.

[31] S. Zelditch, Real and complex zeros of Riemannian random waves, in *Spectral Analysis in Geometry and Number Theory* (International conference on the occasion of Toshikazu Sunada's 60th birthday), *Contemp. Math.* **484**, Amer. Math. Soc., Providence, RI, 2009, pp. 321–342. MR 1500155. Zbl 1176.58021. http://dx.doi.org/10.1090/conm/484/09482.

(Received: November 20, 2011)

Indian Institute of Science, Bangalore, India
E-mail: manju@math.iisc.ernet.in
http://math.iisc.ernet.in/~manju

Royal Institute of Technology, Stockholm, Sweden
E-mail: kurlberg@math.kth.se
http://www.math.kth.se/~kurlberg/

Cardiff University, Wales, UK
Current address: King's College London, UK
E-mail: igor.wigman@kcl.ac.uk
https://www.kcl.ac.uk/nms/depts/mathematics/people/atoz/wigmani.aspx