RANDOM EIGENFUNCTIONS ON FLAT TORI: UNIVERSALITY FOR THE NUMBER OF INTERSECTIONS

MEI-CHU CHANG, HOI NGUYEN, OANH NGUYEN, VAN VU

Abstract. We show that several statistics of the number of intersections between random eigenfunctions of general eigenvalues with a given smooth curve in flat tori are universal under various families of randomness.

Contents

2010 *Mathematics Subject Classification.* 15A52,11B25, 60C05, 60G50.
Key words and phrases. arithmetic random waves, universality phenomenon, arithmetic progressions.
M. C. Chang is partially supported by NSF grant DMS 1600154, the author thanks the Mathematics Department of University of California at Berkeley for its hospitality. H. Nguyen is partially supported by NSF grant DMS 1600782.

1. Introduction

Let $\mathcal{M}$ be a smooth Riemannian manifold. Let F be a real-valued eigenfunction of the Laplacian on $\mathcal{M}$ with eigenvalues λ^2,

$$-\Delta F = \lambda^2 F.$$

The nodal set N_F is defined to be

$$N_F := \{x \in \mathcal{M}, F(x) = 0\}.$$

The study of N_F is extremely important in analysis and differential geometry. In this note we are simply interested in the case when $\mathcal{M}$ is the flat tori $\mathbf{T}^d = \mathbf{R}^d/\mathbf{Z}^d$ with $d \geq 2$; more specifically we will be focusing on the intersection set of N_F with a given reference curve.

Let $\mathcal{C} \subset \mathcal{M}$ be a curve assumed to have unit length with the arc-length parametrization $\gamma : [0,1] \to \mathcal{M}$. The nodal intersection between F and $\mathcal{C}$ is defined as

$$\mathcal{Z}(F) := \#\{x : F(x) = 0\} \cap \mathcal{C}.$$

1.1. Deterministic results in $\mathbf{T}^2$.

It is known that all eigenvalues λ^2 have the form $4\pi^2 m, m \in \mathbf{Z}^+$. Let $\mathcal{E}_\lambda$ be the collection of $\mu = (\mu_1, \mu_2) \in \mathbf{Z}^2$ such that

$$\mu_1^2 + \mu_2^2 = m.$$

Denote $N = N_m = \#\mathcal{E}_\lambda$, that is $N = r_2(m)$. Note that in this case, if $m = m_1^2 m_2$ with $m_1 = 2^r \prod_{q_k \equiv 3 \mod 4} q_k^{b_k}$ and $m_2 = 2^c \prod_{p_j \equiv 1 \mod 4} p_j^{a_j}$ $(c = 0, 1)$ then (see, for example, [25])

$$N = \prod_j (a_j + 1).$$

The toral eigenfunctions $f(x) = e^{2\pi i \langle \mu, x\rangle}, \mu \in \mathcal{E}_\lambda$ form an orthonormal basis in the eigenspace corresponding to λ^2. We first introduce several deterministic results by Bourgain and Rudnick from [3, 5, 6].

Theorem 1.1. *Let $\mathcal{C} \subset \mathbf{T}^2$ be a real analytic curve with nowhere vanishing curvature, then*

$$\mathcal{Z}(F) \leq c\lambda.$$

The constant c depends on the curve $\mathcal{C}$. This bound can be achieved from [26] once we have

$$\int_{\mathcal{C}} |F|^2 d\gamma \gg e^{-c\lambda} \int_{\mathbf{T}^2} |F(x)|^2 dx.$$

This type of restriction result was obtained in [5] in the stronger form

$$\int_{\mathcal{C}} |F|^2 d\gamma \gg \int_{\mathcal{M}} |F(x)|^2 dx. \tag{1}$$

Henceforth the bound of Theorem 1.1 follows immediately.

The lower bound for $\mathcal{Z}(F)$ is also of special interest. Let B_λ denote the maximal number of lattice points which lie on an arc of size $\sqrt{\lambda}$ on the circle $|x| = \lambda$

$$B_\lambda = \max_{|x|=\lambda} \#\{\mu \in \mathcal{E} : |x - \mu| \leq \sqrt{\lambda}\}.$$

Theorem 1.2. [6] *If $\mathcal{C} \subset \mathbf{T}^2$ is smooth with nowhere vanishing curvature, then*

$$\mathcal{Z}(F) \gg \frac{\lambda}{B_\lambda^{5/2}}.$$

In particularly, as one can show that $B_\lambda \ll \log \lambda$ (see [6]), we have

Theorem 1.3.

$$\mathcal{Z}(F) \gg \lambda^{1-o(1)}.$$

According to a conjecture of [8], $B_\lambda = O(1)$ uniformly. This is known to hold for almost all λ^2, see for instance [4, Lemma 5]; we also refer the reader to Lemma 5.2 of Section 5 for a similar result (with a relatively short proof). In view of Theorem 1.2, the following was conjectured in [6]

Conjecture 1.4. *If $\mathcal{C} \subset \mathbf{T}^2$ is smooth with non-zero curvature, then*

$$\mathcal{Z}(F) \gg \lambda.$$

1.2. **Arithmetic random wave model.** We next introduce a probabilistic setting first studied by Rudnick and Wigman [23]. Consider the random gaussian function

$$F(t) = \frac{1}{\sqrt{N}} \sum_{\mu \in \mathcal{E}_\lambda} \varepsilon_\mu e^{2\pi i \langle \mu, \gamma(t) \rangle},$$

where ε_μ are iid complex standard gaussian with a saving

$$\varepsilon_{-\mu} = \bar{\varepsilon}_\mu.$$

The random function F is called *arithmetic random wave* [1, 17], whose distribution is invariant under rotation by the gaussian property of the coefficients.

We now introduce the main result of [23].

Theorem 1.5. *Let $\mathcal{C} \subset \mathbf{T}^2$ be a smooth curve on the torus, with nowhere vanishing curvature and of total length one. Then*

(1) *the expected number of nodal intersections is precisely*

$$\mathbf{E}_{\mathbf{g}}\mathcal{Z} = \sqrt{2m},$$

(2) *the variance is bounded*

$$\mathrm{Var}_{\mathbf{g}}(\mathcal{Z}) \ll \frac{m}{N}.$$

(3) *Furthermore, let $\{m\}$ be a sequence such that $N_m \to \infty$ and $\{\hat{\tau}_m(4)\}$ do not accumulate at ± 1, then*

$$\mathrm{Var}_{\mathbf{g}}(\mathcal{Z}) = \frac{m}{N}\int_{\mathcal{C}}\int_{\mathcal{C}} 4\left(\frac{1}{N}\left\langle \frac{\mu}{|\mu|}, \dot{\gamma}(t_1)\right\rangle^2 \left\langle \frac{\mu}{|\mu|}, \dot{\gamma}(t_2)\right\rangle^2 - 1\right) dt_1 dt_2 + O(\frac{m}{N^{3/2}}).$$

Here the subscript $\mathbf{g}$ is used to emphasize standard gaussian randomness, and τ_m is the probability measure on the unit circle $S^1 \subset \mathbf{R}^2$ associated to $\mathcal{E}_\lambda$,

$$\tau_m = \frac{1}{N}\sum_{\mu \in \mathcal{E}} \delta_{\mu/\sqrt{m}}.$$

A simple consequence of (1) and (2) is that Conjecture (1.4) holds for the random wave F asymptotically almost surely. In fact, the statement of (2) and (3) show that the variance is much smaller than m, indicating a large number of cancellations in the formula of the variance.

1.3. **Partial results in $\mathbf{T}^3$.** Bourgain and Rudnick [3, 5, 6] also considered the intersection $\mathcal{Z}$ between N and a smooth *hypersurface* σ for general $\mathbf{T}^d$. For $\mathbf{T}^3$, they obtained an analog of Theorem 1.1 for the L^2 restriction over $\mathcal{Z}$. However, we are not aware of similar deterministic results regarding the intersection with a smooth curve as in $\mathbf{T}^2$. On the probabilistic side, Rudnick, Wigman and Yesha [24] have recently extended Theorem 1.5 to $\mathbf{T}^3$. Here, for $\lambda^2 = 4\pi^2 m$ with $m \neq 0, 4, 7 \mod 8$, let $\mathcal{E}_\lambda$ be the collection of $\mu = (\mu_1, \mu_2, \mu_3) \in \mathbf{Z}^3$ such that $\mu_1^2 + \mu_2^2 + \mu_3^2 = m$. Again denote $N = N_m = \#\mathcal{E}_\lambda$.

Consider the random gaussian function

$$F(t) = \frac{1}{\sqrt{N}}\sum_{\mu \in \mathcal{E}_\lambda} \varepsilon_\mu e^{2\pi i \langle \mu, \gamma(t)\rangle},$$

where ε_μ are iid complex standard gaussian again with the saving

$$\varepsilon_{-\mu} = \bar{\varepsilon}_\mu.$$

Rudnick, Wigman and Yesha showed the following result.

Theorem 1.6. *Let $\mathcal{C} \subset \mathbf{T}^3$ be a smooth curve on the torus of total length one with nowhere zero curvature. Assume further that either $\mathcal{C}$ has nowhere-vanishing torsion or $\mathcal{C}$ is planar. Then*

- *The expected number of nodal intersections is precisely*

$$\mathbf{E}_{\mathbf{g}}\mathcal{Z} = \frac{2}{\sqrt{3}}\sqrt{m}.$$

- *There exists $c > 0$ such that*
$$\mathrm{Var}_{\mathbf{g}}(\mathcal{Z}) \ll \frac{m}{N^c}.$$

The proof of Theorem 1.5 and Theorem 1.6 are based on Kac-Rice formula. Let us sketch the computation of expectation for $d \geq 2$ that
$$\mathbf{E}_{\mathbf{g}}\mathcal{Z} = \frac{2}{\sqrt{d}}\sqrt{m}. \tag{2}$$
We follow the proof of [24, Lemma 2.3]. Let $r(t_1, t_2) = \mathbf{E}(F(t_1)F(t_2))$. Denote $K_1(t)$ be the gaussian expectation (first intensity)
$$K_1(t) := \frac{1}{\sqrt{2\pi}}\mathbf{E}(|F'(t)|\big|F(t) = 0).$$
By the Kac-Rice formula
$$\mathbf{E}\mathcal{Z} = \int_0^1 K_1(t)dt.$$
Let Γ be the covariance matrix of $(F(t), F'(t))$,
$$\Gamma(t) = \begin{pmatrix} r(t,t) & r_1(t,t) \\ r_2(t,t) & r_{12}(t,t) \end{pmatrix},$$
where $r_1 = \partial r/\partial t_1, r_2 = \partial r/\partial t_2, r_{12} = \partial^2 r/\partial t_1 \partial t_2$. It is not hard to show that $\Gamma(t) = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$, where $\alpha = r_{12}(t,t) = \frac{4}{d}\pi^2 m$. It thus follows
$$K_1(t) = \frac{1}{\pi}\sqrt{\alpha} = \frac{2}{\sqrt{d}}\sqrt{m}.$$
For the variance, denote $K_2(t)$ to be
$$K_2(t) := \phi_{t_1,t_2}(0,0)\mathbf{E}\Big(|F'(t_1)F'(t_1)|F(t_1) = 0, F(t_2) = 0\Big),$$
where ϕ_{t_1,t_2} is the density function of the random gaussian vector $(F(t_1), F(t_2))$. It is known that if the covariance matrix $\Sigma(t_1, t_2)$ of the vectors $(F(t_1), F(t_2), F'(t_1), F'(t_2))$ is non-singular for all $(t_1, t_2) \in A \times B$, then
$$\mathbf{E}(\mathcal{Z}\restriction_A \mathcal{Z}\restriction_B) - \mathbf{E}(\mathcal{Z}\restriction_A)\mathbf{E}(\mathcal{Z}\restriction_B) = \int_{A\times B} K_2(t_1,t_2)dt_1dt_2.$$

The main problem here is that the matrix $\Sigma(t_1, t_2)$ is not always non-singular in $[0, 1]^2$. Roughly speaking, to overcome this highly technical obstacle, Rudnick and Wigman [23] and Rudnick, Wigman and Yesha [24] divide $[0, 1]$ into subintervals I_i of length of order $1/\sqrt{m}$ each, and then show that Kac-Rice's formula is available locally on most of the cells $I_i \times I_j$. We refer the reader to [23, 24] for more detailed treatment of these issues.

1.4. **More general random waves and our main results.** Motivated by Conjecture 1.4, and by the universality phenomenon in probability, we are interested in the behavior of $\mathcal{Z}(F)$ for other random eigenfunctions F beside the gaussian arithmetic random waves as above. More specifically, consider the random function
$$F(t) = \frac{1}{\sqrt{N}}\sum_{\mu\in\mathcal{E}_\lambda} \varepsilon_\mu e^{2\pi i\langle\mu,\gamma(t)\rangle}, \tag{3}$$

where $\varepsilon_\mu = \varepsilon_{1,\mu} + i\varepsilon_{2,\mu}$, where $\varepsilon_{1,\mu}, \varepsilon_{2,\mu}, \mu \in \mathcal{E}_\lambda$ are iid random variables with the saving constraint $\varepsilon_{-\mu} = \bar{\varepsilon}_\mu$ so that $F(t)$ is real-valued as in the gaussian case.

We denote by $\mathbf{P}_{\varepsilon_\mu}, \mathbf{E}_{\varepsilon_\mu}$, and $\mathrm{Var}_{\varepsilon_\mu}$ the probability, expectation, and variance with respect to the random variables $(\varepsilon_\mu)_{\mu\in\mathcal{E}_\lambda}$.

We are interested in the following problem.

Question 1.7. *Are the statistics such as $\mathbf{E}_{\varepsilon_\mu}\mathcal{Z}(F)$ and $\mathrm{Var}_{\varepsilon_\mu}(\mathcal{Z}(F))$ with respect to the randomness of the random variables ε_μ universal?*

Note that we can write $F(t)$ as

$$F(t) = \frac{1}{\sqrt{N}} \sum_{\mu\in\mathcal{E}_\lambda} \varepsilon_\mu e^{2\pi i\langle \mu, \gamma(t)\rangle} = \frac{1}{\sqrt{N}} \sum_{\mu\in\mathcal{E}_\lambda} \varepsilon_{1,\mu} \cos(2\pi\langle \mu, \gamma(t)\rangle) + \varepsilon_{2,\mu} \sin(2\pi\langle \mu, \gamma(t)\rangle). \quad (4)$$

We now restrict to $\mathbf{T}^2$ by assuming several necessary properties of the curves and distributions.

Assumption on the reference curve.

Condition 1. *We will suppose the following.*

(i) (Non-degeneracy) The curve $\gamma(t) : [0, 1] \to \mathbf{T}^2$ has unit length with arc-length parametrization. More specifically, there exists a positive constant c such that $\|\gamma'(t)\| > c$ and $\|\gamma''(t)\| > c$ for all t.
(ii) (Analyticity) The function $\gamma(t)$ extends analytically to $t \in [0, 1] \times [-\varepsilon, \varepsilon]$ for some small constant ε.
(iii) (Large diameter) For any constant $c_0 > 0$, there exists a constant $\alpha > 0$ such that for any interval $I \subset [0, 1]$ of length c_0/λ, the segment $\{\gamma(t), t \in I\}$ cannot be contained in a ball of radius $N^{-\alpha}/\lambda$.

We will need Condition (1) (iii) when the random variables are not continuous.

Assumption on the distribution. We will assume ε_μ to have mean zero, variance one with the following properties.

Condition 2. *There is a fixed number K such that either*

(i) (Continuous distribution) ε_μ is absolutely continuous with density function p bounded $\|p\|_\infty \le K$.
(ii) (Mixed distribution) there exist positive constants c_1, c_2, c_3 such that $\mathbf{P}(c_1 \le |\varepsilon_\mu - \varepsilon'_\mu| \le c_2) \ge c_3$ where ε'_μ is an independent copy of ε_μ and one of the following holds
- *either $|\varepsilon_\mu| > 1/K$ with probability one*
- *or $\varepsilon_\mu 1_{|\varepsilon_\mu|\le 1/K}$ is continuous with density bounded above by K.*

The assumption that ε_μ stays away from zero (for discrete distribution) is necessary because otherwise the random function $F(t)$ might be vanishing with positive probability. One representative example of our consideration is Bernoulli random variable which takes values ± 1 with probability 1/2. We now state our main result for $\mathbf{T}^2$.

Theorem 1.8 (general distributions in $\mathbf{T}^2$). *With γ as above, assume that $\varepsilon_{1,\mu}, \varepsilon_{2,\mu}, \mu \in \mathcal{E}_\lambda$ are iid random variables satisfying Condition* (2). *Then, for almost all m we have*

- $\mathbf{E}_{\varepsilon_\mu} \mathcal{Z} = \mathbf{E}_{\mathbf{g}} \mathcal{Z} + O(\lambda/N^c)$,
- *More generally, for any fixed k, $\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}^k = \mathbf{E}_{\mathbf{g}} \mathcal{Z}^k + O(\lambda^k/N^c)$,*

where the subscript $\mathbf{g}$ stands for the distribution in which the $\varepsilon_{1,\mu}$ and $\varepsilon_{2,\mu}$ are independent standard gaussian. Here the implicit constants depend on the curve γ and k but not on N and λ. In particularly, with γ and λ as in Theorem 1.5

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z} = \sqrt{2m} + O(\lambda/N^c) \text{ and } \mathrm{Var}_{\varepsilon_\mu}(\mathcal{Z}) \ll \frac{\lambda^2}{N^c}.$$

To prove Theorem 1.8, we will need to show that the set $\mathcal{E}_\lambda$ satisfies the following assumption which is later proven to be satisfied in Section 6.

Assumption 1.9. *There exists a constant $\varepsilon_0 > 0$ such that the following holds. For any vector $r \in \mathbf{R}^2$ with $|r| = \frac{1}{2\pi\lambda}$, the set $\{\langle r, \mu\rangle, \mu \in \mathcal{E}_\lambda\}$ can not be covered by less than $O(N^{\varepsilon_0})$ intervals of length N^{-1} in $[-1, 1]$.*

Theorem 1.8 is stated for almost all m mainly because of the deterministic Lemma 5.2 of Section 5, which in turn is needed for the verification of one of our probabilistic conditions of the universality framework. We also need to pass to almost all m for a brief verification of Assumption 1.9 for $\mathcal{E}_\lambda$ (Section 6).

Now we turn to $\mathbf{T}^d, d \geq 3$. While in this setting the cardinality N of $\mathcal{E}_\lambda$ is relatively large compared to λ, the situation is difficult by different reasons. Consider the following example from [24].

Example 1.10. *Let $F_0(x, y)$ be an eigenfunction on $\mathbf{T}^2$ with eigenvalue $4\pi^2 m$, and S_0 a curved segment length one contained in the nodal set, admitting an arc-length parameterization $\gamma_0 : [0, 1] \to S_0$ with curvature $\kappa_0(t) = |\gamma_0''(t)| > 0$. For $n > 0$, let $F_n(x, y, z) = F_0(x, y) \cos(2\pi nz)$, which is an eigenfunction on $\mathbf{T}^3$ with eigenvalue $4\pi^2(m + n^2)$. Let $\mathcal{C}$ be the curve $\gamma(t) = (\gamma_0(t/\sqrt{2}), t/\sqrt{2})$. Standard computation shows that the curvature $\kappa(t) = \kappa_0(t/\sqrt{2})/2 > 0$ and the torsion $\tau(t) = \pm\kappa_0(t/\sqrt{2})/2$ is non-zero. Note that $\mathcal{C}$ is contained in the nodal set of F_n for all n. Thus we can have a non-trivial curve contained in the nodal set for arbitrary large λ.*

This example shows that the study of universality for discrete distributions in $\mathbf{T}^d, d \geq 3$ can be highly complex (at least if we only assume γ to have non-vanishing curvature and torsion) as there is no deterministic upper bound for $\mathcal{Z}(F)$. If we are not careful with the choice of discrete distributions, our random function F from (3) might be one of the F_n in Example 1.10 with non-zero probability, and hence $\mathbf{E}\mathcal{Z}(F)$ is infinite. To avoid such type of singularity, in what follows we will assume that the random variables ε_μ satisfy Condition (2)(i). Note that this also holds for $d = 2$.

Theorem 1.11 (continuous distributions in $\mathbf{T}^d$, $d \geq 2$)**.** *Assume that $\varepsilon_{1,\mu}, \varepsilon_{2,\mu}, \mu \in \mathcal{E}_\lambda$ are independent random variables satisfying Condition* (2)(i)*. Assume furthermore that the curve γ extends analytically to the strip $[0,1] \times [-\lambda^{-1}, \lambda^{-1}]$. Then for any fixed k we have*

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}^k = \mathbf{E}_{\mathbf{g}} \mathcal{Z}^k + O(\lambda^k / N^c).$$

In particularly for $\mathbf{T}^3$, with γ and λ as in Theorem 1.6

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z} = \frac{2}{\sqrt{3}} \sqrt{m} + O(\lambda / N^c) \text{ and } \mathrm{Var}_{\varepsilon_\mu}(\mathcal{Z}) \ll \frac{\lambda^2}{N^c}.$$

The rest of the note is organized as follows. We first introduce in Section 2 a general scheme from [27] and [22] to prove our universality result, a sketch of proof for these results will be discussed in Section 9. In the next phase, we prove Theorem 1.11 for smooth distributions in Section 3. The proof of Theorem 1.8 will be carried out throughout Section 4, Section 5, and Section 6 to check various regulatory conditions.

Notation. We consider λ as an asymptotic parameter going to infinity and allow all other quantities to depend on λ unless they are explicitly declared to be fixed or constant. We write $X = O(Y)$, $X \ll Y$, or $Y \gg X$ if $|X| \leq CY$ for some fixed C; this C can depend on other fixed quantities such as the the parameter K of Condition 1 and the curvatures of γ. All the norms in this note, if not specified, will be the usual ℓ_2-norm.

2. Supporting lemmas: general universality results

Generally speaking, our starting point uses the techniques developed by T.Tao and V. Vu from [27], and subsequently by Y. Do, O. Nguyen and V. Vu [10] and by O. Nguyen and V. Vu [22].

Let

$$H(x) = \sum_{\mu \in \mathcal{E}} \xi_\mu f_\mu(x),$$

where x belongs to some set $\mathcal{B} \subset \mathbf{R}$.

Assumption 2.1. *Consider the following conditions.*

(1) *Analyticity: H has an analytic continuation on the set $\mathcal{B} + B(0,1)$ on the complex plane, which is also denoted by H.*

(2) *Anti-concentration: For any constants A and c, there exists a constant C such that for every $x \in \mathcal{B}$, with probability at least $1 - CN^{-A}$, there exists $x' \in B(x, 1/100)$ such that $|H(x)| \geq \exp(-N^c)$.*

(3) *Boundedness: For any constants A and c, there exists a constant C such that for every $x \in \mathcal{B}$,*

$$\mathbf{P}\left(|H(z)| \leq \exp(N^c) \text{ for all } z \in B(x,1)\right) \geq 1 - CN^{-A}.$$

(4) Contribution of tail events: For any $k \geq 1$, there exist constants $A, c > 0$ such that for any $x \in \mathcal{B}$ and any event $\mathcal{A}$ with probability at most N^{-A}, we have

$$\mathbf{E}\mathcal{Z}^k_{B(x,1)}\mathbf{1}_{\mathcal{A}} = O_{k,A,c}(N^{-c}),$$

where $\mathcal{Z}_{B(x,1)}$ is the number of roots of H in the complex ball $B(x,1)$.

(5) Delocalization: There exists a constant $c > 0$ such that for every $z \in \mathcal{B} + B(0,1)$ and every $\mu \in \mathcal{E}$,

$$\frac{|f_\mu(z)|}{\sqrt{\sum_\mu f_\mu^2(z)}} \leq N^{-c},$$

(6) Derivative growth: For any constant $c > 0$, there exists a constant C such that for any real number $x \in \mathcal{B} + [-1,1]$,

$$\sum_\mu |f'_\mu(x)|^2 \leq C\left(N^c \sum_\mu |f_\mu(x)|^2\right), \tag{5}$$

as well as

$$\sup_{z\in B(x,1)} |f''_\mu(z)|^2 \leq C\left(N^c \sum_\mu |f_\mu(x)|^2\right). \tag{6}$$

Note that the last three conditions are deterministic, which are effective for trigonometric functions. Now we state the main result from [22].

Theorem 2.2 (Local universality, real roots)**.** *Let $H(x) = \sum_\mu \xi_\mu f_\mu(x)$, with $H(x)$ be a random function with f_μ satisfying Assumption 2.1. Let k be an integer constant. There exists a constant $c > 0$ such that the following holds. For any real numbers $x_1, \dots, x_k$ in $\mathcal{B}$, and for every smooth function G supported on $\prod_{j=1}^k [x_j - c, x_j + c]$ with $|\nabla^a G(z)| \leq 1$ for $0 \leq a \leq 2k$ we have*

$$\mathbf{E}_{\varepsilon_\mu} \sum_{i_1,\dots,i_k} G(\zeta_{i_1}, \dots, \zeta_{i_k}) - \mathbf{E}_{\mathbf{g}} \sum_{i_1,\dots,i_k} G(\zeta_{i_1}, \dots, \zeta_{i_k}) = O(N^{-c}), \tag{7}$$

where the ζ_i are the roots of H, the sums run over all possible assignments of $i_1, \dots, i_k$ which are not necessarily distinct.

Remark 2.3. *By induction on k, the above theorem still holds if in (7), the $i_1, \dots, i_k$ are required to be distinct.*

We will provide a sketch of the proof of this theorem in Section 9.

Now we consider F from (3). Set the scaled function $H : [0, \lambda] \to \mathbf{R}$ to be

$$\begin{aligned} H(x) := F\left(\frac{x}{\lambda}\right) &= \frac{1}{\sqrt{N}} \sum_{\mu\in\mathcal{E}_\lambda} \varepsilon_{\mu,1} \cos\left(2\pi\left\langle \mu, \gamma\left(\frac{x}{\lambda}\right)\right\rangle\right) + \varepsilon_{\mu,2} \sin\left(2\pi\left\langle \mu, \gamma\left(\frac{x}{\lambda}\right)\right\rangle\right) \\ &:= \frac{1}{\sqrt{N}} \sum_\mu \varepsilon_{\mu,1} g_\mu(x) + \varepsilon_{\mu,2} h_\mu(x). \end{aligned} \tag{8}$$

Our main contributions are the following results.

Theorem 2.4. *Under the assumptions of Theorem 1.11, let $\mathcal{B}_1 = [0, \lambda]$, then the function H in* (8) *satisfies the assumption (with $\mathcal{B} = \mathcal{B}_1$) and hence the conclusion of Theorem 2.2.*

Theorem 2.5. *Under the assumptions of Theorem 1.8, let $\mathcal{B}_2 = [0, \lambda] \setminus \cup_{\varphi \in \mathcal{D}} (\lambda S_\varphi)$ where $\mathcal{D}$ is the set of directions*

$$\mathcal{D} = \left\{ \frac{\mu_1 - \mu_2}{\|\mu_1 - \mu_2\|}, \mu_1 \neq \mu_2, \mu_1, \mu_2 \in \mathcal{E}_\lambda \right\}.$$

and

$$S_\varphi := \{t \in [0, 1], \angle(\gamma'(t), \varphi) < N^{-3}\}.$$

Then the function H in (8) *satisfies the assumption (with $\mathcal{B} = \mathcal{B}_2$) and hence the conclusion of Theorem 2.2.*

As a consequence, we have the following.

Theorem 2.6. *Let H be the function in* (8). *Under the assumptions of Theorem 1.11 (respectively Theorem 1.8), for any $k \geq 1$, there exists a constant $c > 0$ such that for any intervals $I_1, \dots, I_k \subset [0, \lambda]$ each intersects $\mathcal{B}_1$ (respectively $\mathcal{B}_2$) and has length $O(1)$, we have*

$$\mathbf{E}_{\varepsilon_\mu} \prod_{j=1}^k \mathcal{Z}_j = \mathbf{E}_{\mathbf{g}} \prod_{j=1}^k \mathcal{Z}_j + O_k(N^{-c})$$

where $\mathcal{Z}_j$ is the number of roots of H in I_j.

To prove Theorems 2.4 and 2.5, it suffices to verify all of the conditions of Assumption 2.1 for $H(x)$. We will do that in Section 3 for Theorem 2.4 and Sections 4 and 5 for Theorem 2.5.

The deduction of Theorem 2.6 from Theorems 2.4 and 2.5 is given in Section 7. Theorems 1.8 and 1.11 will be concluded from Theorem 2.6 in Section 8.

3. Proof of Theorem 2.4: the smooth case

Because of Condition (i), we have the following anti-concentration bound.

Fact 3.1. *For any $t \in I$, and any $\delta > 0$*

$$\mathbf{P}(|F(t)| \leq \delta) = O(\delta).$$

Our claim is that with very high probability all of the conditions from Assumption 2.1 hold for the function H given in (8). Note that Condition (1) follows from our assumption on the analyticity of the curve γ.

3.1. **Verification of Condition** (2)**.** For Condition (2), if suffices to establish the bound for any $\mu_0 \in \mathcal{E}_\lambda$ and $x_0 \in J_i$. Again, as either $|\cos(2\pi\langle\mu_0, \gamma(x_0/\lambda)\rangle)|$ or $|\sin(2\pi\langle\mu_0, \gamma(x_0/\lambda)\rangle)|$ has order $\Theta(1)$, by the continuity of ε_μ, we have for any $\delta > 0$,

$$\begin{aligned} \mathbf{P}\left(|H(x_0)| \geq \delta\right) &\geq \inf_a \mathbf{P}\Big(\left|\varepsilon_{1,\mu_0} \cos(2\pi \langle\mu_0, \gamma(x_0/\lambda)\rangle) + \varepsilon_{2,\mu_0} \sin(2\pi\langle\mu_0, \gamma(x_0/\lambda)\rangle) + a\right| \geq N\delta\Big) \\ &\geq 1 - O\left(N\delta\right). \end{aligned} \tag{9}$$

Let $\delta = e^{-N^c}$, we obtain the desired estimate.

3.2. **Verification of Condition** (3)**.** For every $z \in [0, \lambda] \times [-1, 1]$, let $x = \mathrm{Re}(z)$. Since $\left\langle \mu, \gamma\left(\frac{x}{\lambda}\right)\right\rangle$ is real, we have

$$\left|\mathrm{Im}\left\langle \mu, \gamma\left(\frac{z}{\lambda}\right)\right\rangle\right| \le \left|\left\langle \mu, \gamma\left(\frac{z}{\lambda}\right) - \gamma\left(\frac{x}{\lambda}\right)\right\rangle\right| = O(1), \tag{10}$$

and so

$$\left|\exp\left(i2\pi\left\langle \mu, \gamma\left(\frac{z}{\lambda}\right)\right\rangle\right)\right| = \exp\left(-2\pi\mathrm{Im}\left\langle \mu, \gamma\left(\frac{z}{\lambda}\right)\right\rangle\right) = O(1). \tag{11}$$

Thus,

$$|H(z)| = O(1)\sum_\mu |\varepsilon_\mu|.$$

By Markov's inequality, for any $M > 0$,

$$\mathbf{P}\left(|H(z)| \ge M \text{ for some } z \in [0, T] \times [-1, 1]\right) \le \mathbf{P}\left(\sum_\mu |\varepsilon_\mu| = \Omega(M)\right) \le O\left(\frac{N}{M}\right). \tag{12}$$

Setting $M = e^{N^c}$, Condition (3) then follows. We remark that this condition holds even when ε_μ has discrete distribution.

3.3. **Verification of Condition** (4)**.** Let $K = \max_{z\in B(x,1)}|H(z)|$. By Jensen's inequality,

$$\mathcal{Z}_{B(x,1/2)} = O(1)\log\frac{K}{|H(x)|}.$$

Thus,

$$\mathbf{E}\mathcal{Z}^k_{B(x,1/2)}\mathbf{1}_\mathcal{A} \ll \mathbf{E}|\log K|^k\mathbf{1}_\mathcal{A} + \mathbf{E}|\log|H(x)||^k\mathbf{1}_\mathcal{A}.$$

By Hölder's inequality,

$$\mathbf{E}|\log K|^k\mathbf{1}_\mathcal{A} \le \left(\mathbf{E}|\log K|^{2k}\right)^{1/2}\mathbf{P}(\mathcal{A})^{1/2}.$$

By the bound (12), we obtain $\mathbf{E}|\log K|^k = O_k(N)$ which yields

$$\mathbf{E}|\log K|^k\mathbf{1}_\mathcal{A} = O_k\left(N^{-(A-1)/2}\right).$$

We argue similarly for $\mathbf{E}|\log|H(x)||^k\mathbf{1}_\mathcal{A}$ using (9) (which is valid for all $\delta > 0$). Letting $A = 2$, for example, we obtain the desired statement.

3.4. **Verification of Conditions** (5) **and** (6) **for** g_μ, h_μ**.** For Condition (5), note that for any $x \in (0, 1)$ we have $\sum_\mu |g_\mu(x)|^2 + |h_\mu(x)|^2 = N$, and so

$$\frac{|g_\mu(x)| + |h_\mu(x)|}{\sqrt{\sum_\mu g_\mu(x)^2 + h_\mu(x)^2}} = O\left(\frac{1}{\sqrt{N}}\right).$$

For (5) of Condition (6), we have

$$g'_\mu(x) = \frac{2\pi}{\lambda}\left\langle \mu, \gamma'(\frac{x}{\lambda})\right\rangle\cos\left(2\pi\left\langle \mu, \gamma\left(\frac{x}{\lambda}\right)\right\rangle\right).$$

Thus

$$\sum_{\mu}|g_{\mu}'(x)|^2+\sum_{\mu}|h_{\mu}'(x)|^2\ll\sum_{\mu}\frac{1}{\lambda^2}\left\langle\mu,\gamma'(\frac{x}{\lambda})\right\rangle^2\ll N$$

where the implicit constant depends on $\max_{x\in[0,\lambda]}|\gamma'(\frac{x}{\lambda})|$. This proves (5). Finally, (6) of Condition (6) is proven similarly using the same argument together with (11).

In the remaining sections we will prove Theorem 2.5. As we already seen, for this it suffices to verify Condition (2) and Condition (4) of Assumption 2.1 only.

4. Proof of Theorem 2.5: verification of Condition (2)

As the continuous case has been treated in Section 3, here we will assume

- there exist positive constants c_1, c_2, c_3 and K such that
$$\mathbf{P}(c_1\le|\varepsilon_\mu-\varepsilon_\mu'|\le c_2)\ge c_3$$
- with probability one
$$|\varepsilon_\mu|>1/K$$

Recall that $N=|\mathcal{E}_\lambda|$. Without scaling, we will show the following which implies Condition (2).

Theorem 4.1. *Let $A>0$ be a fixed constant, then there exists a constant $C=C(A)$ such that the following holds for $F(t)$ from (3): for any interval $I\subset[0,1]$ of length c_0/λ, for any $t_1,t_2\in I$ with $\|\gamma(t_1)-\gamma(t_2)\|=\frac{N^{-\alpha}}{\gamma}$, we have*

$$\mathbf{P}(|F(t_1)|\le N^{-C})\le N^{-A}\quad\text{or}\quad\mathbf{P}(|F(t_2)|\le N^{-C})\le N^{-A}.$$

Note that by Condition (1)(iii), for any interval I of length c_0/λ, there exist $t_1,t_2\in I$ with $\|\gamma(t_1)-\gamma(t_2)\|=\frac{N^{-\alpha}}{\gamma}$.

It is clear that Condition (2) follows immediately where the sub-exponential lower bound can be replaced by polynomial bounds. To prove Theorem 4.1 we will rely on two results on additive structures. We say a set $S\subset\mathbb{C}$ is *δ-separated* if for any $s_1,s_2\in S$, $|s_1-s_2|\ge\delta$, and S is *ε-close* to a set P if for all $s\in S$, there exists $p\in P$ such that $|s-p|\le\varepsilon$.

Define a *generalized arithmetic progression* (or GAP) to be a finite subset Q of $\mathbf{C}$ of the form

$$Q=\{g_0+a_1g_1+\dots+a_rg_r : a_i\in\mathbf{Z}, |a_i|\le N_i\text{ for all }i=1,\dots,r\}$$

where $r\ge0$ is a natural number (the *rank* of the GAP), $N_1,\dots,N_r>0$ are positive integers (the *dimensions* of the GAP), and $g_0,g_1,\dots,g_r\in\mathbf{C}$ are complex numbers (the *generators* of the GAP). We refer to the quantity $\prod_{i=1}^r(2N_i+1)$ as the *volume* $\mathrm{vol}(Q)$ of Q; this is an upper bound for the cardinality $|Q|$ of Q. When $g_0=0$, we say that Q is *symmetric*. When $\sum_i a_ig_i$ are all distinct, we say that Q is *proper*.

Let ξ be a real random variable, and let $V = \{v_1, ..., v_n\}$ be a multi-set in $\mathbf{R}^d$. For any $r > 0$, we define the small ball probability as

$$\rho_{r,\xi}(V) := \sup_{x\in\mathbf{R}^d} \mathbf{P}\left(v_1\xi_1 + \cdots + v_n\xi_n \in B(x,r)\right)$$

where $\xi_1, ..., \xi_n$ are iid copies of ξ, and $B(x,r)$ denotes the closed disk of radius r centered at x in $\mathbf{R}^d$.

Theorem 4.2. [20, Theorem 2.9] *Let $A > 0$ and $1/2 > \varepsilon_0 > 0$ be constants. Let $\beta > 0$ be a parameter that may depend on n. Suppose that $V = \{v_1, \dots, v_n\}$ is a (multi-) subset of $\mathbf{R}^d$ such that $\sum_{i=1}^n \|v_i\|^2 = 1$ and that V has large small ball probability*

$$\rho := \rho_{\beta,\xi}(V) \geq n^{-A},$$

where ξ is a real random variable satisfying Condition 2. Then the following holds: for any number $n^{\varepsilon_0} \leq n' \leq n$, there exists a proper symmetric GAP $Q = \{\sum_{i=1}^r x_i g_i : |x_i| \leq L_i\}$ such that

- *At least $n - n'$ elements of V are $O(\beta)$-close to Q.*
- *Q has constant rank $d \leq r = O(1)$, and cardinality*

$$|Q| = O(\rho^{-1} n'^{(-r+d)/2}).$$

For Theorem 4.1, first fix $t \in I$, and let $x = \gamma(t)$. Set $\beta = N^{-C}$, with C sufficiently large to be chosen, and assume that

$$\mathbf{P}\left(\left|\sum_{\mu\in\mathcal{E}_\lambda} \varepsilon_{1,\mu}\cos(2\pi\langle\mu,x\rangle) + \varepsilon_{2,\mu}\sin(2\pi\langle\mu,x\rangle)\right| \leq \beta\right) \geq N^{-A}. \tag{13}$$

We will choose ε_0 to be the constant in Assumption 1.9. Then by Theorem 4.2 (applied to the sequences $\{\cos(2\pi\langle\mu,x\rangle), \mu \in \mathcal{E}_\lambda\}$ and $\{\sin(2\pi\langle\mu,x\rangle), \mu \in \mathcal{E}_\lambda\}$ separately with $N' = N^{\varepsilon_0}$), there exist proper GAPs $P_1, P_2 \subset \mathbf{R}$ and $|\mathcal{E}_\lambda| - 2N'$ indices $\mu \in \mathcal{E}_\lambda$ such that with $z_\mu(t) = \cos(2\pi\langle\mu,x\rangle) + i\sin(2\pi\langle\mu,x\rangle) = \exp(2\pi i\langle\mu,\gamma(t)\rangle)$,

$$\mathrm{dist}(z_\mu(t), P_1 + iP_2) \leq 2\beta$$

and such that the cardinalities of P_1 and P_2 are $O\left(N^{O_A(1)}\right)$ and the ranks are $O(1)$. The properness implies that the dimensions of the GAPs P_1 and P_2 are bounded by $O\left(N^{O_A(1)}\right)$.

For short, we denote the complex GAP $P_1 + iP_2$ by $P(t)$.

Now assume for contradiction that (13) holds for both $t = t_1$ and $t = t_2$. By applying the above process to t_1 and t_2, we obtain two GAPs $P(t_1)$ and $P(t_2)$ which are 2β-close to the points $z_\mu(t_1)$ and $z_\mu(t_2)$ respectively for at least $N - 4N^{\varepsilon_0}$ indices μ.

Since the $z_\mu(t_1)$ and $z_\mu(t_2)$ have magnitude 1, the product set $P(t_1)\bar{P}(t_2) = \{p_1\bar{p_2}, p_1 \in P_1(t), p_2 \in P_2(t)\}$ will $O(\beta)$-approximate the points $z_\mu = z_\mu(t_1)\bar{z}_\mu(t_2) = \exp(2\pi\langle\mu,\gamma(t_1) - \gamma(t_2)\rangle)$ for at least $N - 4N^{\varepsilon_0}$ indices μ. Let $\mathcal{S}$ be the collection of these points z_μ.

By definition, $P = P(t_1)\bar{P}(t_2)$ is another GAP whose rank is $O(1)$ and dimensions are of order $O\left(N^{O_A(1)}\right)$.

Now we look at the set $\mathcal{S}$. On one hand, $\mathcal{S}$ is "stable" under multiplication in the sense that $|z_{\mu_1} z_{\mu_2}| = 1$ for all μ_1, μ_2. On the other hand, as z_μ can be well approximated by elements of a GAP of small size, the collection of sums $z_{\mu_1} + z_{\mu_2}$ can also be approximated by another GAP of small size. Roughly speaking, in line of the "sum-product" phenomenon in additive combinatorics [12], this is only possible if the GAP sizes are extremely small. Rigorously, we will need the following continuous analog of a result by the first author [9].

Theorem 4.3. *Let $P = \{g_0 + \sum_{i=1}^r n_i g_i : |n_i| < M\}$ be a generalized arithmetic progression of rank r on the complex plane. Then there exists an (explicit) constant C_r with the following property. Let $0 < \delta < 1$ and $\varepsilon < M^{-C_r}\delta^{C_r}$ and let $S \subset P$ be a subset consisting of elements which are δ-separated and ε-close to the unit circle, then*

$$S \le \exp(C_r \log M / \log\log M).$$

To complete the proof of Theorem 4.1, we apply Theorem 4.3 with $\varepsilon = O(\beta)$, $r = O_A(1)$, and $M = O(N^{O_A(1)})$ to conclude that the set $\mathcal{S}$ can be covered by $\exp(C_r \log N/\log\log N)$ disks of radius δ with $\delta = M\varepsilon^{1/C_r}$. Taking into account at most $4N^{\varepsilon_0}$ elements z_μ not included in $\mathcal{S}$, the set $\{\langle \mu, \gamma(t_1) - \gamma(t_2)\rangle\}_{\mu \in \mathcal{E}}$ can be covered by $4N^{\varepsilon_0} + \exp(C_r \log N/\log\log N) \le 5N^{\varepsilon_0}$ intervals of length $O(\delta)$. However, note that

$$\delta = M\varepsilon^{1/C_r} = O\left(N^{-C/C_r + O_A(1)}\right).$$

By choosing C sufficiently large, this would contradict with the equi-distribution assumption 1.9 on $\mathcal{E}$.

For the rest of this section we will justify Theorem 4.3. In this proof, C_r is a constant depending on r and may vary even within the same context. We denote the set of the coefficient vectors of S by

$$\mathcal{F} = \left\{ \bar{n} = (n_1, \dots, n_r) \in \mathbb{Z}^r : |n_i| < M, g_0 + \sum_{i=1}^r n_i g_i \in S \right\}.$$

Fix $\bar{m} \in \mathcal{F}$. Since $g_0 + \sum_{i=1}^r m_i g_i$ is ε-close to the unit circle, we have $|g_0 + \sum_{i=1}^r m_i g_i| \le 1 + \varepsilon$ and

$$\left| \sum_{i=1}^r (n_i - m_i) g_i \right| \le 2(1+\varepsilon) \quad \text{for all } \bar{n} \in \mathcal{F}. \tag{14}$$

Let $\langle \mathcal{F} - \bar{m} \rangle$ be the vector space generated by $\bar{n} - \bar{m}, \bar{n} \in \mathcal{F}$. We assume $\dim\langle \mathcal{F} - \bar{m}\rangle = r$, since otherwise we may reduce the rank of P without significantly changing the size of P (see [28, Chapter 3]).

Therefore, we can take r independent vectors $\bar{n}^{(1)}, \cdots, \bar{n}^{(r)} \in \mathcal{F}$ and use Cramer's rule to solve $g_1, \cdots, g_r$ in the following system of r equations.

$$\begin{aligned}
(n_1^{(1)} - m_1)g_1 + &\cdots + (n_r^{(1)} - m_r)g_r = c^{(1)} \\
&\cdots \\
&\cdots \\
&\cdots \\
(n_1^{(r)} - m_1)g_1 + &\cdots + (n_r^{(r)} - m_r)g_r = c^{(r)}
\end{aligned}$$

where $|c^{(1)}|, \cdots, |c^{(r)}| \le 2(1+\varepsilon) < 3$.

We obtain a bound

$$|g_1|, \ldots, |g_r| \le 3.2^r r! M^{r-1}, \tag{15}$$

and hence

$$|g_0| < \sum_i |n_i g_i| + 1 + \varepsilon < (3r)2^r r! M^r. \tag{16}$$

Next, assume that $|\mathcal{F}| \ge 2$. Then the separation assumption means that for any $\bar{m}, \bar{n} \in \mathcal{F}$ with $\bar{m} \ne \bar{n}$ we have $|\sum_{i=1}^r (m_i - n_i)g_i| > \delta$. Thus,

$$\max\{|g_1|, \ldots, |g_r|\} > \frac{\delta}{2rM}. \tag{17}$$

Without loss of generality, assume that the maximum above is attained by $|g_1|$.

Lemma 4.4. *There exist $z_0, z_1, \ldots, z_r, w_0, w_1, \ldots, w_r \in \mathbb{C}$ with $z_1 \ne 0$ such that for any $\bar{n} \in \mathcal{F}$*

$$\Big(z_0 + \sum_{i=1}^r n_i z_i\Big)\Big(w_0 + \sum_{i=1}^r n_i w_i\Big) = 1.$$

We next conclude Theorem 4.3 using this lemma. Let $\mathcal{A} = \{z_0 + \sum_{i=1}^r n_i z_i : \bar{n} \in \mathcal{F}\}$.

Applying Proposition 3 in [9] to the mixed progression

$$\{n_0 z_0 + n_0 w_0 + \sum_{i=1}^r n_i z_i + \sum_{i=1}^r n_i' w_i : |n_0|, |n_0'| < 2 \text{ and } |n_i|, |n_i'| < M\},$$

we have

$$|\mathcal{A}| \le \exp(D_r \log M / \log\log M),$$

for some positive constant D_r.

We next partition $\mathcal{F}$ as

$$\mathcal{F} = \bigcup_{a\in\mathcal{A}} \mathcal{F}_a, \text{ where } \mathcal{F}_a = \Big\{\bar{n} \in \mathcal{F} : z_0 + \sum_{i=1}^r n_i z_i = a\Big\}.$$

Let S be as in Theorem 4.3, we write

$$S = \Big\{g_0 + \sum_{i=1}^r n_i g_i : \bar{n} \in \mathcal{F}\Big\} = \bigcup_{a\in\mathcal{A}} S_a, \tag{18}$$

where

$$S_a := \{g_0 + \sum_{i=1}^r n_i g_i : \bar{n} \in \mathcal{F}_a\}.$$

Notice that $S_a \subset P_a := \{g_0 + \sum_{i=1}^r n_i g_i \in P : z_0 + \sum_{i=1}^r n_i z_i = a\}$. The gain here is that P_a is contained in a progression of rank at most $r-1$, that is,

$$g_0 + \sum_{i=1}^r n_i g_i = \left(g_0 + \frac{a - z_0}{z_1} g_1\right) + \sum_{i=2}^r n_i \left(g_i - \frac{z_i}{z_1} g_1\right)$$

so by induction

$$|S_a| \leq \exp(C_{r-1} \log M/\log\log M).$$

It thus follows from (18) that

$$|S| \leq \exp(C_r \log M/\log\log M),$$

for some appropriately chosen constant sequence C_r, completing the proof of Theorem 4.3.

We now prove Lemma 4.4. We will use the following effective form of Nullstellensatz [16].

Theorem 4.5. *Let $q, f_1, \dots, f_s \in \mathbb{Z}[x_1, \dots, x_n]$ with $\deg q, \deg f_i \leq d$ for all i such that q vanishes on the common zeros of $f_1, \cdots, f_s$ and $\mathbf{ht}(f_i) \leq H$. Then there exist $q_1, \dots, q_s \in \mathbb{Z}[x_1, \dots, x_n]$ and positive integers b, l such that*

$$b\, q^l = \sum_{i=1}^s q_i f_i \tag{19}$$

where

$$l \leq D = \max_{1\leq i\leq s}\{\deg q_i\} \leq 4nd^n$$

as well as

$$\max_{1\leq i\leq s}\{\log|b|, \mathbf{ht}(q_i)\} \leq 4n(n+1)d^n\big[H + \log s + (n+7)d\log(n+1)\big].$$

Here the height $\mathbf{ht}(f)$ of a polynomial $f \in \mathbf{Z}[x_1, \dots, x_n]$ is the logarithm of the maximum modulus of its coefficients.

Remark. Theorem 1 in [16] is stated for the case that $q = 1$ and that $f_1, \dots, f_s$ do not have common zeros. However, the standard proof of Nullstellensatz gives the above statement (see [2, Proposition 9] for instance.)

Now define the polynomial P over $\bar{n} \in \mathcal{F}$ as

$$P_{\bar{n}}(z_0, z_1, \dots, z_r, w_0, w_1, \dots, w_r) = \big(z_0 + \sum_{i=1}^{r} n_i z_i\big)\big(w_0 + \sum_{i=1}^{r} n_i w_i\big) - 1.$$

Assume that the claim of Lemma 4.4 does not hold, thus the polynomials $P_{\bar{n}}, \bar{n} \in \mathcal{F}$ have no common zeros with $z_1 \neq 0$.

By Theorem 4.5, with $n = 2r+2, s = |\mathcal{F}| \leq (2M)^r, d = 2, H \leq 2\log M$ we have

$$bz_1^l = \sum_{\bar{n} \in \mathcal{F}} P_{\bar{n}} Q_{\bar{n}}, \tag{20}$$

where $b \in \mathbb{Z}\backslash\{0\}$, $Q_{\bar{n}} \in \mathbb{Z}[z_0, \dots, z_r, w_0, \dots, w_r]$ such that

- $\deg(Q_{\bar{n}}), l \leq D \leq C_r'$
- the coefficients of $Q_{\bar{n}}$ are bounded by $M^{C_r'}$.

Now replacing $z_0, \dots, z_r$ and $w_0, \dots, w_r$ by $g_0, \dots, g_r$ and $\bar{g}_0, \dots, \bar{g}_r$ in (20), we have

$$|g_1|^l \leq \sum_{\bar{n} \in \mathcal{F}} |P_{\bar{n}}(g_0, \dots, g_r, \bar{g}_0, \dots, \bar{g}_d)|\, |Q_{\bar{n}}(g_0, \dots, g_r, \bar{g}_0, \dots, \bar{g}_d)|.$$

By (15), (16), (17) we then have

$$\left(\frac{\delta}{2rM}\right)^l \leq DM^{C_r'}(3.2^r r! rM^r)^D \sum_{\bar{n} \in \mathcal{F}} |P_{\bar{n}}(g_0, \dots, g_r, \bar{g}_0, \dots, \bar{g}_r)|.$$

On the other hand, by definition, $|P_{\bar{n}}(g_0, \dots, g_r, \bar{g}_0, \dots, \bar{g}_r)| \leq \varepsilon$ for any $\bar{n} \in \mathcal{F}$. It thus follows that

$$\left(\frac{\delta}{2rM}\right)^l \leq \left(\frac{\delta}{2rM}\right)^D \leq M^{C_r''}\varepsilon.$$

However, this is impossible with the choice of ε from Theorem 4.3.

5. Proof of Theorem 2.5: verification of Condition (4)

Let $\kappa = N^{-3}$. We will verify Condition (4) through the following deterministic lemma, which is of independent interest.

Theorem 5.1. *Suppose that $\gamma(t), t \in [0,1]$ is smooth and has non-vanishing curvature. Then there exist a constant c and a collection of at most N^2 intervals S_α each of length $O(\kappa)$ such that the following holds for almost all λ and for any eigenfunction $\Phi(x) = \sum_{\mu \in \mathcal{E}_\lambda} a_\mu e^{2\pi i \langle \mu, x \rangle}$ with $\sum_\mu |a_\mu|^2 = 1$.*

(1) The number of nodal intersections on $\cup S_\alpha$ is negligible

$$|N_\Phi \cap \cup\gamma(S_\alpha)| \ll \lambda N^{-1},$$

(2) Condition (4) on $[0,1] \setminus \cup S_\alpha$: for any $a \in [0,1] \setminus \cup_\alpha S_\alpha$, we have

$$|\{z \in B(a, N^7/\lambda) : \Phi(\gamma(z)) = 0\}| \ll N^7.$$

To prove Theorem 5.1 we first need a separation result (see also [4, Lemma 5]).

Lemma 5.2. *For almost all λ, we have*

$$\min_{\mu_1 \neq \mu_2 \in \mathcal{E}_\lambda} \|\mu_1 - \mu_2\| \gg \frac{\lambda}{\log^{3/2+\varepsilon} \lambda}. \tag{21}$$

Proof. (of Lemma 5.2) Let R be a parameter and $M = R(\log R)^{-3/2-\varepsilon}$. Then

$$\begin{aligned}
&\Big|\{(x,y) \in \mathbf{Z}^2 \times \mathbf{Z}^2 : \|x\| = \|y\| \leq R, 0 < \|x-y\| < M\}\Big| \\
&= \sum_{v \in \mathbf{Z}^2 \backslash \{0\}, \|v\| < M} |\{x \in \mathbf{Z}^2 : \|x\| = \|x+v\| \leq R\}| \\
&= \sum_{v \in \mathbf{Z}^2 \backslash \{0\}, \|v\| < M} \Big|\{\|x\| \leq R : 2\langle x, v\rangle + \|v\|^2 = 0\}\Big| \\
&\leq \sum_{v \in \mathbf{Z}^2 \backslash \{0\}, \|v\| < M} \Big|\{\|y\| \leq 3R : y_1 v_1 + y_2 v_2 = 0\}\Big|,
\end{aligned}$$

where $x = (x_1, x_2), v = (v_1, v_2)$ and $y = (y_1, y_2) = 2x + v$.

Now if $v_2 = 0$ then $y_1 = 0$. The contribution to the above sum is $O(MR)$. Similarly for $v_1 = 0$. For the other case that $v_1, v_2 \neq 0$, let $d = \gcd(v_1, v_2)$. Then $(v_1, v_2) = d(v_1', v_2')$ with $\gcd(v_1', v_2') = 1$. The equation $y_1 v_1' + y_2 v_2' = 0$ has $O(R/\|v'\|)$ solutions in y with $\|y\| < 3R$.

So by the Abel's summation formula, we have

$$\sum_{v \in \mathbf{Z}^2 \backslash \{0\}, \|v\| < M} |\{\|y\| \leq 3R : y_1 v_1 + y_2 v_2 = 0\}| \ll MR + \sum_{d<R} \sum_{v' \in \mathbf{Z}^2 \backslash \{0\}, \|v'\| < M/d} R/\|v'\|$$

$$= R \sum_{d<R} \sum_{n=1}^{M^2/d^2} \frac{r_2(n)}{\sqrt{n}} = R \sum_{d<R} \left[\frac{\sum_{n=1}^{M^2/d^2} r_2(n)}{M/d} + \sum_{N=1}^{M^2/d^2-1} (\sum_{n=1}^{N} r_2(n)) \left(\frac{1}{\sqrt{N}} - \frac{1}{\sqrt{N+1}} \right) \right].$$

By Gauss' formula

$$\sum_{n=0}^{x} r_2(n) = (\pi + o(1))x,$$

we have

$$\sum_{v \in \mathbf{Z}^2 \backslash \{0\}, \|v\| < M} |\{\|y\| \leq 3R : y_1 v_1 + y_2 v_2 = 0\}| \ll R \sum_{d<R} \frac{M}{d} \ll MR \log R.$$

Hence

$$|\{(x,y) \in \mathbf{Z}^2 \times \mathbf{Z}^2 : \|x\| = \|y\| \leq R, 0 < \|x-y\| < M\}| \ll MR \log R.$$

On the other hand,

$$|\{(x,y) \in \mathbf{Z}^2 \times \mathbf{Z}^2 : \|x\| = \|y\| \leq R, 0 < \|x-y\| < M\}| \geq \sideset{}{'}\sum_{E<R^2} \mathbf{1}_{\min\{\|x-y\|, \|x\|^2 = \|y\|^2 = E, x \neq y\} < M},$$

where $\sum'$ is the sum over E of sum of two squares. Note that by a classical result of Landau [18]

$$|\{E \in \mathbf{Z}, E < R^2, E = \text{sum of two squares}\}| \gg R^2/\sqrt{\log R}.$$

Recall that $M = R(\log R)^{-3/2-\varepsilon}$. Thus for almost all $E \leq R^2$ that are sum of two squares,

$$\min_{\|x\|^2=\|y\|^2=E, x\neq y} \|x-y\| \geq M \gg R(\log R)^{-3/2-\varepsilon} \gg \sqrt{E}(\log E)^{-3/2-\varepsilon}.$$

□

Recall that by Condition (1)(ii), the curve γ has an analytic continuation to $[0,1]+B(0,\varepsilon) \subset \mathbf{C}$. Arguing as in Sections 3.2 and 3.3, we get the following.

Lemma 5.3. *Let I be any interval with length $\delta = |I| < \varepsilon/2$. Then for any Φ as in Theorem 5.1*

$$|\{z \in I + B(0,\delta) : \Phi(\gamma(z)) = 0\}| \leq C\lambda\delta + \log N - \log \max_{t\in I} |\Phi(\gamma(t))|.$$

Proof. (of Lemma 5.3) For $z \in I + B(0, 2\delta), \exists t \in \mathbf{R}$ such that $|z-t| < 2\delta$,

$$|\gamma(z) - \gamma(t)| \leq c\delta.$$

Hence for $\mu \in \mathcal{E}_\lambda$,

$$\left|e^{i\langle\mu,\gamma(z)\rangle}\right| = \left|e^{i\langle\mu,\gamma(z)-\gamma(t)\rangle}\right| \leq e^{c\lambda\delta}.$$

Therefore

$$|\Phi(\gamma(z))| \leq (\sum_{\mu\in\mathcal{E}_\lambda} |a_\mu|)e^{c\lambda\delta} < \sqrt{N}e^{c\lambda\delta}.$$

Jensen's inequality then implies

$$\begin{aligned}|\{z \in I + B(0,\delta), \Phi(\gamma(z)) = 0\}| &\leq \log(\sqrt{N}e^{c\lambda\delta}) - \log \max_{t\in I} |\Phi(\gamma(t))| \\ &\leq c\lambda\delta + \log N - \log \max_{t\in I} |\Phi(\gamma(t))|.\end{aligned}$$

□

Now we want to bound $\max_{t\in I} |\Phi(\gamma(t))|$.

Lemma 5.4. *We have*

$$\frac{1}{|I|}\int_I |\Phi(\gamma(t))|^2 dt \geq 1/2,$$

provided that λ satisfied (21) of Lemma 5.2 and

$$|I| > \lambda^{-1/2}(\log\lambda)^{3/4+\varepsilon}N.$$

Proof. (of Lemma 5.4) We write

$$\int_I |\Phi(\gamma(t))|^2 dt = \int_I \left|\sum_\mu a_\mu e^{2\pi i \langle \mu, \gamma(t)\rangle}\right|^2 dt = |I| + \sum_{\mu\neq\mu'} a_\mu \bar{a}_{\mu'} \int_I e^{2\pi i\langle \mu-\mu',\gamma(t)\rangle}$$
$$\geq |I| - \sum_{\mu\neq\mu'} |a_\mu||a_{\mu'}|| \int_I e^{2\pi i\langle \mu-\mu',\gamma(t)\rangle}|.$$

By van der Corput's lemma on oscillatory integral (see for instance [6]),

$$\left|\int_I e^{2\pi i\langle \mu-\mu',\gamma(t)\rangle} dt\right| \leq \frac{1}{\|\mu-\mu'\|^{1/2}}.$$

Hence

$$\int_I |\Phi(\gamma(t))|^2 dt \geq |I| - \frac{\log^{3/4+\varepsilon}\lambda}{\lambda^{1/2}} N \gg |I|/2.$$

□

Recall the set of directions,

$$\mathcal{D} = \left\{\frac{\mu_1-\mu_2}{\|\mu_1-\mu_2\|}, \mu_1 \neq \mu_2, \mu_1,\mu_2 \in \mathcal{E}_\lambda\right\}.$$

We partition $[0,1]$ as follows: for every unit direction φ, let S_φ be the interval

$$S_\varphi := \{t \in [0,1], \angle(\gamma'(t),\varphi) < \kappa\}.$$

Claim 5.5. *Assume that the arc-length parametrized curve $\gamma(t)$ has curvature bounded from below by some $c > 0$ for all t. Then for each φ, S_φ is an interval and has size $O(\kappa)$, where the implied constant depends on c.*

Proof. Let $a(t)$ be the angle between $\gamma'(t)$ and φ. Then the curvature of γ at t is $|a'(t)|$ by definition. By continuity, the assumption that γ has curvature bounded from below by c implies that either $a'(t) \geq c$ for all t or $a'(t) \leq -c$ for all t. From either case, it is easy to deduce the claim.

□

Let $J = [0,1] \setminus \cup_{\varphi\in\mathcal{D}} S_\varphi$. We note that J depends on $\mathcal{E}_\lambda$ and γ but not on Φ. Now we prove Theorem 5.1. We first show that $|N_\Phi \cap \cup\gamma(S_\varphi)| \leq \lambda N^{-1}$.

Note that as $\kappa > \lambda^{-1/2}(\log\lambda)^{3/4+\varepsilon} N$, the condition of Lemma 5.4 holds. Thus

$$\max_{t\in S_\varphi} |\Phi(\gamma(t))| \geq \frac{1}{|S_\varphi|}\int_{S_\varphi} |\Phi(\gamma(t))|^2 dt \geq 1/2.$$

Lemma 5.3 implies that

$$|N_\Phi \cap \gamma(S_\varphi)| \ll \kappa\lambda + \log N - c \ll \kappa\lambda.$$

Hence

$$|N_\Phi \cap \cup_\varphi \gamma(S_\varphi)| \ll N^2\kappa\lambda \ll \lambda N^{-1}$$

proving the first part of Theorem 5.1.

Now for the second part, let $a \in J$. Let $\delta = N^7/\lambda$, $M = N^7$.

Denote $\tilde{I} = [a-\delta, a+\delta]$. Again, Lemma 5.3 implies that for $\delta = M/\lambda \le \lambda^{-1+\varepsilon}$
$$|\{z \in B(a,\delta) : \Phi(\gamma(z)) = 0\}| \le |\{z \in \tilde{I}+B(0,\delta) : \Phi(\gamma(z)) = 0\}| \le cM+\log N-\log \max_{t\in\tilde{I}} |\Phi(\gamma(t))|.$$

Since $a \in J, \angle(\gamma'(a), \varphi) \ge \kappa, \forall \varphi \in \mathcal{D}$. Thus for any $\mu \neq \mu'$,
$$|\langle \mu-\mu', \gamma'(a)\rangle| \ge \kappa\|\mu-\mu'\| \gg \delta\|\mu-\mu'\|.$$
On the other hand, with $\delta = M/\lambda \le \lambda^{-1+\varepsilon}$ and $t = a+\tau$, write
$$\langle \mu-\mu', \gamma(t)\rangle = \langle \mu-\mu', \gamma(a)\rangle + \langle \mu-\mu', \gamma'(a)\tau\rangle + O(\|\mu-\mu'\|\delta^2).$$

Because $|\langle \mu-\mu', \gamma'(a)\rangle| \ge \kappa\|\mu-\mu'\| \gg \|\mu-\mu'\|\delta$ and $\|\mu-\mu'\|\delta^2 \ll \lambda\lambda^{-2+\varepsilon} \ll \lambda^{-1+\varepsilon}$,
$$\left|\int_{-\delta}^{\delta} e^{i\langle \mu-\mu', \gamma'(a)\tau\rangle} d\tau\right| \le \frac{1}{|\langle \mu-\mu', \gamma'(a)\rangle|}.$$
We thus have
$$\frac{1}{|\tilde{I}|}\left|\int_{\tilde{I}} e^{i\langle(\mu-\mu'),\gamma'(a)\tau\rangle} d\tau\right| \le \frac{1}{\delta|\langle(\mu-\mu'),\gamma'(a)\rangle|} \le \frac{1}{\delta\kappa\|\mu-\mu'\|} \le \frac{\lambda}{M\kappa\|\mu-\mu'\|}.$$
Lemma 5.2 says that $\|\mu-\mu'\| \gg \frac{\lambda}{\log^{3/2+\varepsilon}\lambda}$. Hence
$$\frac{1}{|\tilde{I}|}\left|\int_{\tilde{I}} e^{i\langle(\mu-\mu'),\gamma(t)\rangle} dt\right| \le \frac{1}{|\tilde{I}|}\left|\int_{\tilde{I}} e^{i\langle(\mu-\mu'),\gamma'(a)\tau\rangle} d\tau\right| + O(\|\mu-\mu'\|\delta^2) \le \frac{N^3\log^{3/2+\varepsilon}\lambda}{M}.$$
Now we have
$$\frac{1}{|\tilde{I}|}\left|\int_{\tilde{I}} |\Phi(\gamma(t))dt\right|^2 \ge 1 - \sum_{\mu\neq\mu'} |a_\mu||a'_\mu|\frac{1}{|\tilde{I}|}|\int_{\tilde{I}} e^{i\langle(\mu-\mu'),\gamma(t)\rangle} dt| \gg 1.$$

So
$$\max_{t\in\tilde{I}} |\Phi(\gamma(t))| > 1/\sqrt{2}.$$
By Lemma 5.3, it follows that
$$|\{z \in B(a, N^7/\lambda) : \Phi(\gamma(z)) = 0\}| \le M + \log N + O(1) \ll N^7.$$

6. Checking Assumption 1.9 for $\mathcal{E}_\lambda$ for almost all λ

Assume otherwise that for some $r \in \mathbf{R}^2$ with $|r| = \frac{1}{2\pi\lambda}$, the set $\{\langle \mu, r\rangle, \mu \in \mathcal{E}_\lambda\}$ can be covered by $k = O(N^{\varepsilon_0})$ intervals $I_1, \ldots, I_k$ of length $\beta = N^{-1}$ each in $[0,1]$. Consider the disjoint intervals $J_j = (j/3k, (j+1)/3k), 0 \le j \le 3k-1$. Let $\varepsilon_0 < 1$, each interval $I_i, 1 \le i \le k$, intersects with at most two intervals J_{i_1}, J_{i_2}, and so there is one interval J_{j_0} which has no intersection with all $I_1, \ldots, I_k$. Thus there is no $\mu \in \mathcal{E}_\lambda$ such that
$$\langle \mu, r\rangle \in J_{j_0}. \tag{22}$$
In what follows we just use this simple consequence. Consider $\mathcal{E}_\lambda$ of $\mu = (\mu_1, \mu_2) \in \mathbf{Z}^2$ such that $\mu_1^2 + \mu_2^2 = m$.

Lemma 6.1. *For almost all number m up to x that can be written as a sum of two squares, the set $\mathcal{E}_\lambda$ satisfies Assumption 1.9.*

As Assumption 1.9 is on the angles α_μ of the vectors $(\mu_1, \mu_2) = \sqrt{m}e^{2\pi i \alpha_\mu}$ in $\mathcal{E}_\lambda$, it suffices to restrict to the set $G(x)$ of m of prime factors congruent with 1 modulo 4 (see [11]). Indeed, let D^2 denote the product of prime factors that are congruent with 3 modulo 4 of m, then in any representation of m as $a^2 + b^2$, we have $D|a$ and $D|b$, so that D does not affect the angles. Moreover, none of these angles is influenced by the power of 2 dividing m because if this power is even, the angles are unchanged and if it is odd there is a rotation by $\pi/4$. We define the discrepancy of the angles α_μ of the vectors (μ_1, μ_2) in $\mathcal{E}_\lambda$ as follows

$$\Delta_m = \max\Big\{\big|\#\{\alpha_\mu \in [\alpha_1, \alpha_2] \mod 1, \mu \in \mathcal{E}_\lambda\} - (\alpha_1 - \alpha_2) r_2(m)\big|, 0 \le \alpha_1 \le \alpha_2 \le 1\Big\}.$$

Denote also

$$R_0(x) = (A + o(1))\frac{x}{\sqrt{\log x}}, A = \frac{1}{2\sqrt{2}}\prod_p (1 - \frac{1}{p^2})^{1/2}.$$

Note that $R_0(x)$ is the number of $m \le x$ whose prime divisors are congruent with 1 mod 4 (see again [11]). Lemma 6.1 easily follows from the following result by Erdős and Hall.

Theorem 6.2. [11] *Let $\varepsilon > 0$ be fixed. Then for all but $o(R_0(x))$ integers $m \in G(x)$ we have*

$$\Delta_m < \frac{r_2(m)}{(\log x)^{\frac{1}{2}\log\frac{\pi}{2} - \varepsilon}}. \tag{23}$$

We can choose $\varepsilon = .001$ and apply this Theorem to a translation $[\alpha_1, \alpha_2]$ of J_{j_0} to get that the number of $\mu \in \mathcal{E}_\lambda$ with $\langle \mu, r\rangle \in J_{j_0}$ is at least

$$N|J_{j_0}| - \frac{r_2(m)}{(\log x)^{\frac{1}{2}\log\frac{\pi}{2} - \varepsilon}} = \frac{N}{3k} - \frac{N}{(\log x)^{\frac{1}{2}\log\frac{\pi}{2} - \varepsilon}}.$$

Since $\sum_{m \le x} r_2(m) = (\pi + o(1))x$, for almost all $m \in G(x)$ we have $N = r_2(m) \ll \log^{O(1)}(x)$. Thus in this case $k = o\left((\log x)^{\frac{1}{2}\log\frac{\pi}{2} - \varepsilon}\right)$, and so J_0 would contain at least one point of the set $\{\langle \mu, r\rangle, \mu \in \mathcal{E}_\lambda\}$, a contradiction.

7. Proof of Theorem 2.6

Under the assumption of Theorem 1.11, we deduce Theorem 2.6 from Theorem 2.4. The deduction of Theorem 2.6 from Theorem 2.5 under the setting of Theorem 1.8 is completely analogous.

The task is to pass from smooth test functions to indicator functions.

Let $l_j = |I_j| = O(1)$. Let c be the constant in Theorem 2.2, and let α be a sufficiently small constant depending on c and k. Let G_j be a smooth function that approximates the indicator function $\mathbf{1}_{[-l_j/2, l_j/2]}$; in particular, let G_j be supported on $[-l_j/2 - N^{-\alpha}, l_j/2 + N^{-\alpha}]$ such that $0 \le G_j \le 1$, $G_j = 1$ on $[-l_j/2, l_j/2]$, and $\|\nabla^a G_j\| \le CN^{C\alpha}$ for all $0 \le a \le 2k$.

Let x_j be the middle point of I_j. We will approximate $\mathcal{Z}_j$ by

$$\mathcal{T}_j := \sum G_j(\zeta - x_j)$$

where ζ runs over all roots of H.

By Theorem 2.4, we have

$$\mathbf{E}_{\varepsilon_\mu} \prod_{j=1}^{k} \mathcal{T}_j - \mathbf{E}_{\mathbf{g}} \prod_{j=1}^{k} \mathcal{T}_j = O\left(N^{-c+C\alpha}\right) = O\left(N^{-\alpha}\right) \tag{24}$$

by choosing α sufficiently small.

We will show that for each j,

$$\mathbf{E}_{\varepsilon_\mu} |\mathcal{T}_j - \mathcal{Z}_j|^k = O\left(N^{-\alpha}\right) \tag{25}$$

and for any constant α',

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{T}_j^k = O\left(N^{\alpha'}\right). \tag{26}$$

Set $\alpha' = \alpha/2k$. By Hölder's inequality and the triangle inequality, we have

$$\mathbf{E}_{\varepsilon_\mu} \prod_{j=1}^{k} \mathcal{Z}_j - \mathbf{E}_{\varepsilon_\mu} \prod_{j=1}^{k} \mathcal{T}_j = O\left(N^{-\alpha/k+\alpha'}\right) = O\left(N^{-\alpha/2k}\right).$$

Combining this with the same bound for the gaussian case and with (24), we obtain the desired result.

It remains to prove (25) and (26). The strategy is first to reduce to the Gaussian case using Theorem 2.4 and then work with the Gaussian case.

Let us prove (26). By Theorem 2.4, we have

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{T}_j^k - \mathbf{E}_{\mathbf{g}} \mathcal{T}_j^k = O\left(N^{-\alpha'}\right).$$

Therefore, it suffices to settle the Gaussian case. Note that $\mathcal{T}_j$ is bounded by X_j defined to be the number of roots of H in the interval $[x_j - l, x_j + l]$ for $l = l_j/2 + N^{-\alpha} = O(1)$.

By Jensen's inequality, we have

$$X_j = O(1) \log \frac{K}{|H(x)|}$$

where $K = \max_{z \in B(x_j, 2l)} |H(z)|$. Thus,

$$\mathbf{E}_{\mathbf{g}} X_j^k = O(1)\mathbf{E}|\log K|^k + O(1)\mathbf{E}|\log |H(x_j)||^k.$$

Since $H(x_j)$ is standard gaussian, $\mathbf{E}|\log |H(x_j)||^k = O(1)$.

Since $|H(x_j)| \le K = O\left(\frac{1}{\sqrt{N}} \sum_\mu |\varepsilon_{\mu,1}| + |\varepsilon_{\mu,2}|\right)$, we have

$$\mathbf{E}|\log |K||^k = O(\log^k N)$$

proving the desired bound.

Finally, we prove (25). Since $|\mathcal{T}_j - \mathcal{Z}_j|$ is less than the number of roots of H in a union of two intervals of length $N^{-\alpha}$. Approximating the indicator function of each of these intervals by a smooth test function supported on an interval of length $10N^{-\alpha}$ and applying Theorem 2.4 to this test function, it suffices to show that for any interval $J = [a, b]$ of length $b - a = O(N^{-\alpha})$, the number of roots of H in J, which is denoted by Y satisfies

$$\mathbf{E}_{\mathbf{g}} Y^k = O(N^{-\alpha}).$$

Assume that it holds for $k = 1$. That is $\mathbf{E}_{\mathbf{g}} Y = O(N^{-\alpha})$. We have

$$\mathbf{E}_{\mathbf{g}} Y^k \le \mathbf{E}_{\mathbf{g}} Y + \mathbf{E}_{\mathbf{g}} Y^k \mathbf{1}_{Y \ge 2}.$$

By Lemma 9.2, $\mathbf{P}_{\mathbf{g}}(Y \ge 2) = O\left(N^{-3\alpha/2}\right)$. Since Assumption (2.1) holds true, $Y \le N^{\alpha/k}$ with probability at least $1 - O\left(N^{-A}\right)$ for any constant A. Therefore, by condition (4) of Assumption (2.1),

$$\mathbf{E}_{\mathbf{g}}\left(Y^k \mathbf{1}_{Y \ge 2}\right) \le \mathbf{E}_{\mathbf{g}}\left(Y^k \mathbf{1}_{2 \le Y \le N^{\alpha/k}}\right) + \mathbf{E}_{\mathbf{g}}\left(Y^k \mathbf{1}_{Y \ge N^{\alpha/k}}\right) = O\left(N^{-\alpha/2}\right).$$

Thus, it remains to prove that $\mathbf{E}_{\mathbf{g}} Y = O(N^{-\alpha})$. By the Kac-Rice type formula (see, for instance, [14, Theorem 2.5]), one has for every $x \in \mathbf{R}$,

$$\mathbf{E}_{\mathbf{g}} Y \quad \le \quad \int_a^b \sqrt{\frac{\mathcal{S}(t)}{\mathcal{P}(t)^2}} dt,$$

where $\mathcal{P}(t) = \mathrm{Var}_{\mathbf{g}}(H(t)) = 1$, $\mathcal{Q}(t) = \mathrm{Var}_{\mathbf{g}}(H'(t)) = \frac{1}{N}\sum_\mu \left\langle \mu, \frac{1}{\lambda}\gamma'(t)\right\rangle^2 = O(1)$, $\mathcal{R}(t) = \mathrm{Cov}_{\mathbf{g}}(H(t), H'(t)) = 0$, and $\mathcal{S} = \mathcal{P}\mathcal{Q} - \mathcal{R}^2 = \mathcal{P}\mathcal{Q}$. And so, for every t,

$$\frac{\mathcal{S}(t)}{\mathcal{P}(t)^2} = \frac{\mathcal{Q}(t)}{\mathcal{P}(t)} = O(1)$$

and

$$\mathbf{E}_{\mathbf{g}} Y = O(1) \int_a^b 1 dt = O\left(N^{-\alpha}\right)$$

as desired.

8. Proof of Theorems 1.8 and 1.11

In this section, we deduce Theorems 1.8 and 1.11 from Theorem 2.6. To prove Theorem 1.11, we partition the interval $[0, \lambda]$ into λ intervals $I_1, \dots, I_\lambda$ of length 1 and apply Theorem 2.6 to every k-tuple of these intervals.

To prove Theorem 1.8, we partition the set $\mathcal{B}_2$ into $M = O(\lambda)$ intervals $I_1, \dots, I_M$ each of length $O(1)$. Applying Theorem 2.6 to every k-tuple of these intervals, we get

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}_{\mathcal{B}_2}^k = \mathbf{E}_{\mathbf{g}} \mathcal{Z}_{\mathcal{B}_2}^k + O(\lambda^k / N^c) \tag{27}$$

where $\mathcal{Z}_{\mathcal{B}_2}$ is the number of zeros of H in $\mathcal{B}_2$.

Let $\mathcal{Z}' = \mathcal{Z} - \mathcal{Z}_{\mathcal{B}_2}$ be the number of zeros of H in $[0, \lambda] \setminus \mathcal{B}_2$. By (26), the number of roots $\mathcal{Z}_j$ of H in each interval I_j satisfies

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}_j^h = O(N^\alpha)$$

for any small constant α and any $h \le k$. Thus, $\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}_{\mathcal{B}_2}^h = O(\lambda^h N^\alpha)$.

By Theorem 5.1, $\mathcal{Z}' \ll \lambda N^{-1}$ a.e. Hence,

$$\mathbf{E}_{\varepsilon_\mu} \mathcal{Z}^k - \mathbf{E}_{\varepsilon_\mu} \mathcal{Z}_{\mathcal{B}_2}^k \ll \lambda^k N^{-1+\alpha} \ll \lambda^k N^{-c}$$

by choosing $\alpha < 1 - c$. This together with (27) give the desired result.

9. Sketch of the proof of Theorem 2.2

To make the note self-consistent, we present here the main ideas of the proof; the reader is invited to conslute [22] for a complete treatment. We first show universality of the complex roots and then deduce Theorem 2.2 from it.

Theorem 9.1 (global universality, complex roots). *Let $H(z) = \sum_\mu f_\mu(z)$, with $H(z)$ be a random function with f_μ satisfying Assumption 2.1. Let k be an integer constant. For any complex numbers $z_1, \dots, z_k$ in $[0,T] \times [-c,c]$, and for every smooth function $G : \mathbb{C}^k \to \mathbb{C}$ supported on $B(0,c)^k$ with $|\nabla^a G(z)| \le 1$ for all $0 \le a \le 2k+4$ and $z \in \mathbf{C}^k$, we have*

$$\mathbf{E}_\xi \sum_{i_1,\dots,i_k} G(\zeta_{i_1}, \dots, \zeta_{i_k}) - \mathbf{E}_{\mathbf{g}} \sum_{i_1,\dots,i_k} G(\zeta_{i_1}, \dots, \zeta_{i_k}) = O(N^{-c}), \tag{28}$$

where the ζ_i are the roots of H, the sums run over all possible assignments of $i_1, \dots, i_k$ which are not necessarily distinct. The constant c here might be different from the constants in Assumption 2.1.

9.1. **Sketch of proof of Theorem 9.1.** By approximation arguments using Fourier expansion, we can reduce the problem to proving (28) for G of the form

$$G(w_1, \dots, w_m) = G_1(w_1) \dots G_k(w_k) \tag{29}$$

where for each $1 \le i \le k$, $G_i : \mathbb{C} \to \mathbb{C}$ is a smooth function supported in $B(0, 1/10)$ and $|\nabla^a G_i| \le 1$ for all $0 \le a \le 3$.

Let $X_j^H = \sum G_j(\zeta_i^H - z_j)$. By induction on k, it suffices to show that

$$\left| \mathbf{E} \prod_{j=1}^k X_j^H - \mathbf{E} \prod_{j=1}^k X_j^{\tilde{H}} \right| \le C\delta^c. \tag{30}$$

Let A be a large constant and c_1 be a small positive constant. By the Green's formula, one has

$$X_j^H = \sum_{i=1}^n G_j(\zeta_i^H - z_j) = -\frac{1}{2\pi} \int_{B(z_j,c)} \log |H(z)| \triangle G_j(z - z_j) dz. \tag{31}$$

In the next step, we show that the integral can be approximated by a finite sum with high probability. The technique is based on the Monte-Carlo Lemma, which is in fact a special case of Markov's inequality. In particular, let $w_{j,1}, \dots, w_{j,m_0}$ be drawn independently at random on the ball $B(z_j, c)$, and let S be the empirical average

$$S := \frac{1}{2c^2 m_0} \sum_{i=1}^{m_0} \log |H(w_{j,i})| \triangle G_j(w_{j,i} - z_j).$$

Then by Markov's inequality, we have

$$\mathbf{P}\left(\left|S-\frac{1}{2\pi}\int_{B(z_j,c)}\log|H(z)|\triangle G_j(z-z_j)\frac{dz}{\mathrm{Area}(B(z_j,c))}\right|\geq\lambda\right)$$
$$\leq\frac{1}{m\lambda^2}\int_{B(z_j,c)}|\log|H(z)|\triangle G_j(z-z_j)|^2\frac{dz}{\mathrm{Area}(B(z_j,c))}=\frac{O(1)}{m\lambda^2}\int_{B(z_j,c)}|\log|H(z)||^2\,dz.$$

Thus, to quantify the approximation of the integral by a finite sum, we need to control the 2-norm of $\log|H|$ on the balls $B(z_j,c)$. That is to bound the function $|H|$ from above and away from 0. These bounds are attained from conditions (2) and (3) of Assumption (2.1). Note that condition (2) only gives a lower bound of $|H|$ for a certain $x\in B(z_j,c)$. To pass from this to a bound that works for all $z\in B(z_j,c)$, one can make use of Harnack's inequality.

Note that on the tail event of conditions (2) and (3), the approximation is not valid. One has to instead show that the contribution of X_j^H on that event is negligible. That's when condition (4) becomes handy.

Going back to the good event when we can approximate the integral by a finite sum, we reduce the task of comparing X_j^H and $X_j^{\tilde{H}}$ to comparing $\sum_{i=1}^{m_0}\log|H(w_{j,i})|\triangle G_j(w_{j,i}-z_j)$ and $\sum_{i=1}^{m_0}\log|\tilde{H}(w_{j,i})|\triangle G_j(w_{j,i}-z_j)$. This is done by the Lindeberg swapping argument (see for instance [27] and the references therein). In particular, by smoothing the log function, we can further reduce the task to showing that for any deterministic $w_{j,i}$ with $1\leq j\leq k$, $1\leq i\leq m_0$, and for a smooth function $L:\mathbb{C}^{km_0}\to\mathbb{C}$,

$$\left|\mathbf{E}L\left(H(w_{j,i})\right)_{ji}-\mathbf{E}L\left(\tilde{H}(w_{j,i})\right)_{ji}\right|\leq CN^{-c}.$$

The swapping method uses the triangle inequality to bound the above difference by a sum of $2N$ differences each of which involves changing only one random variable to gaussian. For example, one of these differences is $\mathbf{E}L\left(H_0(w_{j,i})\right)_{ji}-\mathbf{E}L\left(H_1(w_{j,i})\right)$ where $H_0(z)=H(z)=\sum_\mu\xi_\mu f_\mu(z)$ and $H_1(z)=\tilde{\xi}_{\mu_1}f_{\mu_1}(z)+\sum_{\mu\neq\mu_1}\xi_\mu f_\mu(z)$. We then Taylor expand the function $L\left(H_0(w_{j,i})\right)_{ji}$ (and $L\left(H_1(w_{j,i})\right)_{ji}$) as a function of one variable ξ_μ (and $\tilde{\xi}_\mu$ respectively). Making use of the assumption that the first and second moments of ξ_μ and $\tilde{\xi}_\mu$ are the same, one can see that upon taking expectation, the first three terms in the Taylor expansions cancel out, leaving us with a small error term. Adding up these errors terms, one obtains N^{-c} as desired. The reader may notice that this is quite similar to a classical proof of the Central Limit Theorem using the swapping argument.

9.2. **Universality of real roots: sketch of proof of Theorem 2.2.** As in the proof of Theorem 9.1, we can reduce the problem to showing that

$$\left|\mathbf{E}\left(\prod_{j=1}^k X^H_{x_i,G_i,\mathbb{R}}\right)-\mathbf{E}\left(\prod_{j=1}^k X^{\tilde{H}}_{x_i,G_i,\mathbb{R}}\right)\right|\leq C'N^{-c},\tag{32}$$

where $X^H_{x_i,G_i,\mathbb{R}}=\sum_{\zeta_j^H\in\mathbb{R}}G_i(\zeta_j^H-x_i)$, ζ_j^H are the roots of H, and $H_i:\mathbb{R}\to\mathbb{C}$ are smooth functions supported on $[-c,c]$ and $B(0,c)$ respectively, such that $|\nabla^a G_i(x)|\leq 1$ for all $1\leq i\leq k$, $x\in\mathbf{R}$, and $0\leq a\leq 3$.

The idea is to reduce it to Theorem 9.1. This is done by showing that the number of complex zeros near the real axis is small with high probability.

Lemma 9.2. *We have*

$$\mathbf{P}\left(\mathcal{Z}_H B(x,\gamma) \geq 2\right) \leq C\gamma^{3/2}, \qquad \textit{for all } x \in [0,T]$$

where $\gamma = N^{-c}$ for any sufficiently small constant c.

Using Theorem 9.1, this lemma is reduced to the Gaussian case. Let $\tilde{H}(z) = \sum_\mu \tilde{\xi}_\mu f_\mu(z)$ where $\tilde{\xi}_\mu$ are standard gaussian. Let $g(z) = \tilde{H}(x) + \tilde{H}'(x)(z-x)$ and $p(z) = \tilde{H}(z) - g(z)$. By Rouché's theorem,

$$\mathbf{P}(\mathcal{Z}_{\tilde{H}} B(x,2\gamma) \geq 2) \leq \mathbf{P}\left(\min_{z\in\partial B(x,2\gamma)} |g(z)| \leq \max_{z\in\partial B(x,2\gamma)} |p(z)|\right).$$

Both $g(z)$ and $p(z)$ have zero mean. Condition (6) of Assumption (2.1) shows that for all $z \in B(x,2\gamma)$,

$$\mathrm{Var}(p(z)) = O\left(N^{-(4+\varepsilon)c}\mathrm{Var}(\tilde{H}(x))\right).$$

Thus with probability at least $1 - O\left(N^{-3c/2}\right)$,

$$\max_{z\in\partial B(x,2\gamma)} |p(z)| = O\left(N^{-(2+\varepsilon)c}\sqrt{\mathrm{Var}(\tilde{H}(x))}\right). \tag{33}$$

Now, for g, note that since g is a linear function with real coefficients, one has $\min_{z\in\partial B(x,2\gamma)} |g(z)| = \min |g(x \pm 2\gamma)|$. Condition 5 shows that $g(x \pm 2\gamma)$ is normally distributed with variance

$$\mathrm{Var}(g(x \pm 2\gamma)) \geq 1/2\mathrm{Var}(\tilde{H}(x)).$$

Therefore, with probability at least $1 - O\left(N^{-3c/2}\right)$,

$$|g(x \pm 2\gamma)| \geq N^{-3c/2}\sqrt{\mathrm{Var}(\tilde{H}(x))}$$

Combining this with (33), we obtain Lemma 9.2.

Acknowledgement. The authors are grateful to Z. Rudnick for helpful comments.

References

[1] Berry, M. V. Regular and irregular semiclassical wave functions. J. Phys. A 10 (1977), no. 12, 2083-2091.
[2] E. Bombieri, J. Bourgain and S. V. Konyagin, Roots of Polynomials in Subgroups of forumla and Applications to Congruences, Int Math Res Notices, 5, 802-834 (2009).
[3] J. Bourgain and Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces, C. R. Acad. Sci. Paris, Ser. I 347 (2009) 1249-1253.
[4] 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.
[5] J. Bourgain and Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces and nodal sets , Geometric and Functional Analysis: Volume 22, Issue 4 (2012), Page 878-937.
[6] J. Bourgain and Z. Rudnick, Nodal intersections and L_p restriction theorems on the torus. To appear in Israel J. Math., arXiv:1308.4247.

[7] J. Bourgain, Z. Rudnick and P. Sarnak, Spatial statistics for lattice points on the sphere I: Individual results, To appear in Bulletin of the Iranian Mathematical Society (BIMS) in honor of Freydoon Shahidi's 70th birthday, https://arxiv.org/abs/1606.05880.

[8] J. Cilleruelo and A. Granville, Lattice points on circles, squares in arithmetic progressions and sumsets of squares, in Additive Combinatorics, CRM Proceedings & Lecture Notes, Vol. 43, American Mathematical Society, Providence, RI, 2007, pp. 241-262.

[9] M.C. Chang, Factorization in generalized arithmetic progressions and application to the Erdős-Szemerédi sum-product problems, Geom. Funct. Anal. 13 (4), 720-736 (2003).

[10] Y. Do, O. Nguyen, and Van Vu, Roots of random polynomials with arbitrary coefficients, submitted, arXiv:1507.04994.

[11] P. Erdős and R. Hall, On the angular distribution of Gaussian integers with fixed norm, Discrete Mathematics 200 (1999), 87-94.

[12] P. Erdős and E. Szemerédi, On sums and products of integers, Studies in Pure Mathematics; To the memory of Paul Turán. P. Erdős, L. Alpár and G. Halász editors, Akadémiai Kiádó - Birkhauser Verlag, Budapest - Basel-Boston, Mass. 1983, 213-218.

[13] L. Fainsilber, P. Kurlberg and B. Wennberg, Lattice points on circles and discrete velocity models for the Boltzmann equation, SIAM J. Math. Anal. 37 no. 6 (2006), 1903-1922.

[14] Farahmand, Kambiz. Topics in random polynomials. Vol. 393. CRC Press, 1998.

[15] H. Iwaniec and E. Kowalski, Analytic number theory, American Mathematical Society, Colloquium publications, Volume 53.

[16] T. Krick, L. M. Pardo and M. Sombra, Sharp estimates for the arithmetic Nullstellensatz, Duke Math. J., 109, 521-598 (2001).

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

[18] E. Landau, Uber die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindestzahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate, Arch. der Math. u. Phys. (3). 13:305-312, 1908.

[19] V. Jarnik, Uber die Gitterpunkte auf konvexen Kurven, Math. Z. 24 (1) (1926) 500-518.

[20] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, Advances in Mathematics, Vol. 226 6 (2011), 5298-5319.

[21] H. Nguyen and V. Vu, Small probability, inverse theorems, and applications, Paul Erdős' 100th anniversary, Bolyai Society Mathematical Studies, Vol. 25 (2013).

[22] O. Nguyen and V. Vu, Local universality of zeros of random trigonometric polynomials, in preparation.

[23] Z. Rudnick and I. Wigman, Nodal Intersection for random eigenfunctions on the torus, Amer. J. of Mathematics, to appear.

[24] Z. Rudnick, I. Wigman and Nadav Yesha, Nodal intersections for random waves on the 3-dimensional torus, Annales de l'institut Fourier, to appear.

[25] Sally, J. Roots to research: a vertical development of mathematical problems. American Mathematical Soc., 2007.

[26] J. Toth and S. Zelditch, Counting nodal lines which touch the boundary of an analytic domain, Journal of Differential Geometry 81 (2009), 649-686.

[27] T. Tao and V. Vu, Local universality of zeroes of random polynomials, International Mathematics Research Notices, 0-84, 2014.

[28] T. Tao and V. Vu, Additive Combinatorics, Cambridge Univ. Press, 2006.

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210

E-mail address: nguyen.1261@math.osu.edu

Department of Mathematics, University of California, Riverside

E-mail address: mcc@math.ucr.edu