



HAL
open science

Nodal Replication of Planar Random Waves

Loïc Thomassey, Raphaël Lachièze-Rey

► **To cite this version:**

Loïc Thomassey, Raphaël Lachièze-Rey. Nodal Replication of Planar Random Waves. 2023. hal-04030237

HAL Id: hal-04030237

<https://hal.science/hal-04030237v1>

Preprint submitted on 15 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NODAL REPLICATION OF PLANAR RANDOM WAVES

LOIĆ THOMASSEY AND RAPHAËL LACHIEZE-REY

ABSTRACT. We study the *almost periods* of the eigenmodes of flat planar manifolds in the high energy limit. We prove in particular that the Gaussian Arithmetic Random Waves replicate almost identically at a scale at most $\ell_n := n^{-\frac{1}{2}} \exp(\mathcal{N}_n)$, where \mathcal{N}_n is the number of ways n can be written as a sum of two squares. It provides a qualitative interpretation of the *full correlation phenomenon* of the nodal length, which is known to happen at scales larger than $\ell'_n := n^{-1/2} \mathcal{N}_n^A$. We provide also a heuristic with a toy model pleading that the minimal scale of replication should be closer to ℓ'_n than ℓ_n .

CONTENTS

1. Introduction	1
2. Almost periodicity and replication	6
3. Dirichlet's theorem for almost periodic fields	13
4. Replication of the nodal lines	15
5. Optimality of Dirichlet's approximation theorem	19
6. Appendix	20
References	28

Keywords: Gaussian fields, Nodal sets, Arithmetic Random Waves, Almost periodic fields, Gauss circle problem

1. INTRODUCTION

This paper investigates random fields of the form

$$f(t) = \sum_{\lambda} a_{\lambda} e_{\lambda}(t)$$

where t belongs either to the Euclidean space \mathbb{R}^d or to the torus \mathbb{T}^d ,

$$e_{\lambda}(t) := \exp(2i\pi\langle t, \lambda \rangle), \quad t \in \mathbb{T}^d,$$

and the λ are finitely many deterministic wave vectors.

LABORATOIRE MAP5, UNIVERSITÉ PARIS-CITÉ

E-mail addresses: loic.thomassey@etu.u-paris.fr, raphael.lachieze-rey@parisdescartes.fr.

A primary motivation is the zero set of the random Laplace eigenfunctions on the two-dimensional flat torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. In this setting, the sequence of eigenvalues, or energy levels, are explicitly given by

$$E_n := 4\pi^2 n,$$

where n is expressible as a sum of two squares,

$$n \in \mathcal{S} := \{a^2 + b^2 : (a, b) \in \mathbb{Z}^2\}.$$

and the corresponding eigenspace \mathcal{E}_n is spanned by the \mathbb{L}^2 -orthonormal Fourier basis, with λ belonging to the set of frequencies

$$\Lambda_n := \{\lambda \in \mathbb{Z}^2 : \lambda_1^2 + \lambda_2^2 = n\}.$$

It is a finite-dimensional vector space whose dimension

$$\mathcal{N}_n := \text{card } \Lambda_n$$

is equal to the number of ways an integer can be written as a sum of two squares. The behaviour of \mathcal{N}_n is well-understood and it is known that $n \in \mathcal{S}$ if and only if every prime divisor of n congruent to 3 modulo 4 has an even valuation. In the latter case, \mathcal{N}_n has an explicit formula related to the prime decomposition of n , that is,

$$(1.1) \quad \mathcal{N}_n = 4 \prod_{i=1}^k (1 + \alpha_i),$$

where

$$n = 2^\alpha p_1^{\alpha_1} \dots p_k^{\alpha_k} q_1^{2\beta_1} \dots q_l^{2\beta_l}$$

with $p_i \equiv 1 \pmod{4}$ and $q_i \equiv 3 \pmod{4}$ prime numbers.

The preceding formula shows that the sequence \mathcal{N}_n is subject to large fluctuations. Yet, if we exclude some exceptional values, it is always possible to consider a subsequence of integers n for which the following assumptions hold.

There is a density 1 subset $\mathcal{S}' \subset \mathcal{S}$ of integers n such that

$$(1.2) \quad \mathcal{N}_n = \log(n)^{\log(2)/2+o(1)},$$

in the sense that for every $0 < \kappa < \frac{\log(2)}{2}$

$$(1.3) \quad \mathcal{N}_n \geq \log(n)^\kappa$$

for sufficiently large $n \in \mathcal{S}'$, and such that the angular distribution

$$\widetilde{\mu}_n := \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \delta_{\lambda/\sqrt{n}}$$

converges to the Haar measure σ on \mathbb{S}^1 for the weak-* topology as $n \in \mathcal{S}' \rightarrow \infty$. Moreover, the Kolmogorov distance between σ and $\widetilde{\mu}_n$ satisfies the bound

$$(1.4) \quad \text{Kol}(\widetilde{\mu}_n, \sigma) := \sup_{\Gamma \text{ arc on } \mathbb{S}^1} |\widetilde{\mu}_n(\Gamma) - \sigma(\Gamma)| \leq \frac{1}{\log(n)^\rho},$$

for any $\rho < \frac{1}{2} \log(\pi/2)$ and n large enough.

For the interested reader, discussions related to these results are available respectively in [BMW20] and [EH99, Theorem 1]. From now on, we only consider $n \in \mathcal{S}'$ so that equations (1.2)-(1.4) are satisfied throughout the article.

1.1. Planck scale Arithmetic Random Wave. To study the properties of a generic Laplace eigenfunctions on the torus, it is possible to embed the previous model in a probability space. Notice first that every real-valued eigenfunction of the Laplace-Beltrami operator can be uniquely written as

$$T_n(t) := \frac{1}{\sqrt{\mathcal{N}_n}} \sum_{\lambda \in \Lambda_n} a_\lambda e_\lambda(t), \quad t \in \mathbb{T}^2,$$

where (a_λ) is a sequence of complex numbers satisfying $a_{-\lambda} = \overline{a_\lambda}$. If the a_λ are chosen as a realisation of independent and identically distributed complex standard random normal variables, then the resulting random field is an infinitely differentiable, centered, stationary Gaussian process with covariance function

$$r_n(t) := \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \cos(2\pi\langle \lambda, t \rangle), \quad t \in \mathbb{T}^2.$$

As expected, it satisfies almost surely the partial differential equation

$$\Delta_{\mathbb{T}^2} T_n + E_n T_n = 0.$$

Alternatively, instead of describing the preceding field with its covariance function, it is sometimes preferable to work directly with its spectral measure

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

which, in that instance, is a purely atomic probability measure concentrated on the circle of radius \sqrt{n} .

It is natural to rescale the previous Gaussian process by a factor $n^{-1/2}$, also known as the *Planck scale*. The notations $\tilde{\cdot}$ will from now on refer to the rescaled variables.

$$\begin{aligned} \tilde{T}_n(t) &:= T_n\left(\frac{t}{\sqrt{n}}\right), \quad t \in \sqrt{n}\mathbb{T}^2, \\ \tilde{r}_n(t) &:= r_n\left(\frac{t}{\sqrt{n}}\right), \quad t \in \sqrt{n}\mathbb{T}^2, \\ \tilde{\mu}_n &:= \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \delta_{\frac{\lambda}{\sqrt{n}}}. \end{aligned}$$

As detailed later, the factor $n^{-1/2}$ can be viewed as the characteristic distance of the model. In domains whose diameter is below this threshold, T_n will behave almost like an eigenvalue of the Laplacian Δ on \mathbb{R}^2 in the sense that \tilde{T}_n will converge (in some sense) to a Berry Random Wave T , that is the unique Gaussian process on \mathbb{R}^2 which is stationary, rotation-invariant, ergodic and which satisfies almost-surely the equation

$$\Delta T + 4\pi^2 T = 0.$$

The covariance function associated with T is the so-called Bessel function of the first kind,

$$\tilde{r}(t) := J_0(2\pi|t|).$$

As a consequence of (1.4), \tilde{T}_n converges to T on a scale slightly larger than the Planck scale:

Lemma 1.1. [DNPR23, Proposition 5.1] *Let $\alpha \in \mathbb{N}^2$, then there is some $\rho > 0$ such that*

$$\sup_{|t| \leq \log(n)^\rho} |\partial^\alpha \tilde{r}_n(t) - \partial^\alpha \tilde{r}(t)| \longrightarrow 0.$$

1.2. Nodal set. An object intensively studied in the literature is the zero set $T_n^{-1}(0)$, known as the nodal set or nodal lines. Almost surely, it is a one-dimensional smooth manifold whose length on a subdomain $\Omega \subset \mathbb{T}^2$ is denoted by

$$\mathcal{L}_n(\Omega) := \text{len}(T_n^{-1}(0) \cap \Omega),$$

or simply

$$\mathcal{L}_{n,s} := \text{len}(T_n^{-1}(0) \cap B(s))$$

when Ω is $B(s)$, the centered ball of radius s .

Oravecz, Rudnick, Wigman [ORW08] were able to compute the expected nodal length

$$\mathbb{E}[\mathcal{L}_n(\Omega)] = \frac{|\Omega|}{2\sqrt{2}} \sqrt{E_n}.$$

Higher moments were harder to resolve as Kac-Rice formulas require in those cases a deep understanding of the behaviour of the correlation function r_n and its derivatives, turning the preceding integral computation into an arithmetical problem intrinsically related to the distribution of the spectral measure μ_n alongside the circle of radius \sqrt{n} . It was only in 2011 that Krishnapur, Kurlberg and Wigman [KKW13] were finally able to evaluate the asymptotic leading term of the variance of the nodal length on the full torus in the high energy limit :

$$\text{Var}(\mathcal{L}_n(\mathbb{T}^2)) \asymp \frac{1}{512} \frac{E_n}{\mathcal{N}_n^2}$$

The preceding variance is smaller than what was initially expected, but the result corroborates Berry's cancellation phenomenon, a principle related to the length of the nodal set of Gaussian Random Waves, which was first noticed by Berry in his seminal paper [Ber02]. The computation of the variance on subdomains is still possible but it requires a more careful analysis. It was done in [BMW20], an article in which the authors derived the following asymptotic

$$\text{Var}(\mathcal{L}_{n,s_n}) \asymp \frac{|B(s_n)|}{512} \frac{E_n}{\mathcal{N}_n^2},$$

valid as long as the radius s_n is above Planck-scale, that is $s_n > n^{-1/2+\varepsilon}$ where $\varepsilon > 0$. The preceding estimate immediately implies the full-correlation of the nodal length on the torus :

Theorem 1.2. *For every $\varepsilon > 0$,*

$$\sup_{s > n^{-1/2+\varepsilon}} \text{corr}(\mathcal{L}_n, \mathcal{L}_{n,s}) \longrightarrow 1.$$

This unexpected theorem has a nice consequence. In the high energy limit, one may retrieve the full nodal length on the torus \mathbb{T}^2 by simply observing the nodal length in a window slightly larger than Planck-scale. It is natural to ask whether the previous theorem still holds below Planck-scale and the answer is negative (see Section 2.1). As of today, the phase transition between no correlation and full-correlation is suspected to happen at some $s = n^{-1/2} \log(n)^A$ and some lower and upper bounds for A have been recently derived in [Sar22] and [DNPR23].

1.3. Full correlation and replication. The proof of theorem 1.2 relies heavily on a careful study of the so-called spectral quasi-correlations and it requires an extensive use of deep arithmetical results regarding the angular distribution $\widetilde{\mu}_n$. An explanation of this phenomenon, and of why full correlation does not occur on the sphere, can be found in the work of Todino [Tod20, Section 2.1].

The reader is encouraged to view this article as an alternative qualitative explanation of Theorem 1.2. We show that not only the length of the nodal line, but the whole field, replicates almost identically at a scale τ_n that we try to estimate, and we advocate that it should be close to the scale of full correlation. It makes sense that if the field replicates almost identically at scale τ_n , then by a continuity argument its nodal lines should replicate as well, and this is the purpose of Theorem 2.3 and Section 4.

A rough explanation of this result is that replication occurs on the torus because the dimension of the eigenspace on the torus is logarithmic, whereas it is polynomial on other manifolds, such as higher dimensional tori or the sphere (see Section 2.3). We stress that our results do not provide another proof of full correlation, for which quasi-correlations are the right tool [BMW20, Tod20]. Let us first explain what we mean by *almost replication*.

A sequence of almost periods (of the Arithmetic Random Wave) is a sequence $\tau_n \in \mathbb{T}^2$ such that a.s.

$$(1.5) \quad \sup_{t \in \mathbb{T}^2} |T_n(t + \tau_n) - T_n(t)| \xrightarrow{n \rightarrow +\infty} 0$$

with the complementary condition

$$(1.6) \quad \liminf \sqrt{n} |\tau_n| \geq 1.$$

This latter requirement ensures that the aforementioned almost periods are far from 0 with respect to the characteristic distance of the model. This is a necessary condition as nothing prevents otherwise a sequence of almost periods to be arbitrary close to 0. Hence, upon the existence of such τ_n , we shall say that the random waves *almost replicate* at scale τ_n , and we try in the rest of the paper to give an upper bound as small as possible on an almost period of τ_n .

1.3.1. Almost periodicity of eigenmodes. As an introduction, and instead of studying directly full-correlation on the torus \mathbb{T}^2 , one may start with a similar result on the 1-dimensional torus \mathbb{T}^1 . In this case, the set of eigenvalues consists of the energy levels $E_n = (2\pi)^2 n^2$, $n \in \mathbb{N}$, and an orthonormal basis of the associated eigenspace is given by the standard Fourier basis $\cos(2\pi nt)$ and $\sin(2\pi nt)$. In these settings, the Random Wave Model is a periodic centered Gaussian process whose covariance function is $\cos(2\pi nt)$ and an explicit formula is given by

$$X_n(t) := a \cos(2\pi nt) + b \sin(2\pi nt), \quad t \in \mathbb{T}^1,$$

where a, b are independent standard normal distribution.

The characteristic distance of this model is $1/n$, which is also the smallest period of T_n . In this instance, full-correlation has a straightforward explanation which boils down to one word, *periodicity*. In fact, if the zero set of T_n is known on an open ball slightly larger than Planck scale, then it is also known on a full period of X_n ,

and *a fortiori* on the whole torus. Last but not least, notice that the periodicity of X_n is encoded by the underlying periodicity of its correlation function.

It is very tempting to extend the previous reasoning to the torus \mathbb{T}^2 but unfortunately, the covariance function is no longer periodic at Planck scale. Yet, it is possible to circumvent in some ways this issue if the periodicity assumption is relaxed to almost periodicity. This idea is inspired by the study of almost periodic Gaussian processes in [LR22].

Say a function $f : \mathbb{R}^d \rightarrow \mathbb{C}$ is (Bohr-)almost periodic if and only if, for any $\varepsilon > 0$, there exists $T_0(\varepsilon) > 0$ such that every Euclidean ball of radius $T_0(\varepsilon)$ contains an ε -almost period τ in the sense that

$$\sup_{t \in \mathbb{R}^d} |f(t + \tau) - f(t)| \leq \varepsilon$$

The covariance function r_n of the Arithmetic Random Wave being a trigonometric polynomial, it is in particular an almost periodic function. Thus, if τ is a ε -almost period of r_n , then

$$r_n(\tau) = \text{corr}(T_n(t), T_n(t + \tau)) \geq 1 - \varepsilon.$$

$T_n(t)$ and $T_n(t + \tau)$ are hence heavily correlated and one might expect that if t is a zero of T_n , then there is probably another zero in some neighbourhood of $t + \tau$, both the size of the neighbourhood and the probability depending on the precision ε . Hence, if the zeros of T_n are known on a window Ω , then they are also known on the translated windows $\tau + \Omega$, $2\tau + \Omega$ and so forth, at least with high probability. If the analogy with the 1-dimensional torus \mathbb{T}^1 stills holds, the almost periods of T_n must be of order slightly larger than Planck-scale to get full-correlation.

2. ALMOST PERIODICITY AND REPLICATION

The sections introduces the main results proved in this article, and the discussion around the estimation of the pseudo-period. First, we recall that the covariance function of the Arithmetic Random Wave is a periodic function on the lattice \mathbb{Z}^2 . Our aim is to exhibit smaller almost periods of the order slightly above Planck scale, that is slightly above $n^{-1/2}$.

2.1. Existence of almost periods in the Arithmetic Random Wave. The first result of this article is the following theorem related to the existence of almost periods.

Theorem 2.1. *For $\alpha > 0$, there exists a sequence τ_n of almost periods such that asymptotically for $n \in \mathcal{S}'$*

$$(2.1) \quad |\tau_n| \leq \exp(\mathcal{N}_n^{1+\alpha}) n^{-1/2}$$

and the error term

$$\varepsilon_n := \sup_{t \in \mathbb{T}^2} |r_n(t + \tau_n) - r_n(t)|$$

is asymptotically bounded above by

$$(2.2) \quad \varepsilon_n \leq e^{-\log(n)^\kappa}$$

for $\kappa > 0$ depending on α .

The proof of Theorem 2.1 relies on a well-known deterministic principle in Diophantine approximation, known as Dirichlet's approximation Theorem, which in our framework can be stated in the following manner.

Theorem 2.2 (Dirichlet's approximation Theorem). *Let $\mathcal{N} \geq 1$ and consider an almost periodic function defined by*

$$f(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \cos(2\pi \langle \gamma_k, t \rangle), \quad t \in \mathbb{R}^d$$

with γ_k some vectors in \mathbb{R}^d .

If $\varepsilon > 0$, then there is an ε -almost period τ , that is a vector $\tau \in \mathbb{R}^d$ for which

$$\sup_{t \in \mathbb{R}^d} |f(t + \tau) - f(t)| \leq \varepsilon,$$

satisfying

$$1 \leq |\tau| \leq (2\pi)^{\mathcal{N}/d} \varepsilon^{-\mathcal{N}/d}.$$

We will now prove Theorem 2.1 assuming we have already proved Dirichlet's approximation Theorem.

Proof. Recall that

$$r_n(t) = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \cos(2\pi \langle \lambda, t \rangle).$$

Since $n \in \mathcal{S}'$,

$$\mathcal{N}_n = \log(n)^{\log(2)/2 + o(1)} \geq \log(n)^{\kappa/\alpha}$$

for n large enough and $\kappa < \alpha \log(2)/2$. Apply Theorem 2.2 to T_n and $\varepsilon_n = \exp(-\log(n)^\kappa)$. This ensures the existence of some almost period $\tau_n \in \mathbb{R}^2$ of r_n that satisfies asymptotically the inequality

$$\begin{aligned} |\tau_n| &\leq (2\pi)^{\mathcal{N}_n/2} \varepsilon_n^{-\mathcal{N}_n/2} \\ &= \exp\left(\frac{1}{2} \log(2\pi) \mathcal{N}_n + \frac{1}{2} \log(n)^\kappa \mathcal{N}_n\right) \\ &\leq \exp\left(\frac{1}{2} \log(2\pi) \mathcal{N}_n + \mathcal{N}_n^{1+\alpha/2}\right) \\ &\leq e^{\mathcal{N}_n^{1+\alpha}}. \end{aligned}$$

This finishes the proof. □

Notice that this latter result does not make use of any underlying arithmetical property satisfied by the wave vectors γ_k . This is in sharp contrast with the case of the Arithmetic Random Wave for which the wave vectors λ_k have a deep number theoretic flavour. This probably means that the bound obtained on the almost periods in Theorem 2.1 is not optimal and could be improved, this is the topic of the remainder of the section.

Let now τ_n be some sequence of almost periods as predicted by Theorem 2.1 and denote hereafter

$$T'_n(t) := T_n(\tau_n + t), \quad t \in \mathbb{T}^2,$$

the translated Gaussian process and

$$r'_n(t) := r_n(\tau_n + t), \quad t \in \mathbb{T}^2,$$

its corresponding covariance function.

The preceding theorem has a simple reformulation. In the high energy limit, $T_n(t)$ and $T'_n(t)$ are fully correlated and almost indistinguishable. The existence of almost periods for the correlation entails an approximate replication of the whole field at multiples of the so-called almost period. In fact, the Borell-TIS inequality [AT07, Theorem 2.1.1] yields that the supremum of the difference between the rescaled fields \tilde{T}_n and \tilde{T}'_n over the whole rescaled torus goes to 0, as well as for their derivatives, meaning that (1.5) is in order for some $\tau_n \leq \exp(\mathcal{N}_n^{1+\alpha})$. This has some interesting geometric consequences for nodal lines, which will be discussed now.

Theorem 2.3. *Let Ω be an open convex subset of \mathbb{R}^2 with compact closure and smooth boundary, say \mathcal{C}^1 . Denote by*

$$\begin{aligned}\tilde{\mathcal{Z}}_n(\Omega) &:= \{t \in \Omega : \tilde{T}_n(t) = 0\}, \\ \tilde{\mathcal{Z}}'_n(\Omega) &:= \{t \in \Omega : \tilde{T}'_n(t) = 0\},\end{aligned}$$

the respective zero sets of \tilde{T}_n and \tilde{T}'_n on Ω and let $H_1(dt)$ be the one-dimensional Hausdorff measure on \mathbb{R}^2 normalised so that any segment of length 1 has Hausdorff measure 1.

Then, for any continuous function $\varphi : \mathbb{R}^2 \mapsto \mathbb{R}$ with support in Ω ,

$$(2.3) \quad \int_{\tilde{\mathcal{Z}}_n(\Omega)} \varphi(t) H_1(dt) - \int_{\tilde{\mathcal{Z}}'_n(\Omega)} \varphi(t) H_1(dt) \longrightarrow 0,$$

the convergence holding in distribution.

In the high energy limit, the nodal sets of the Arithmetic Random Wave T_n and its translate T'_n are locally (that is at Planck scale) almost confounded and are geometrically very close. Therefore, the existence of almost periods for the covariance functions give rise to an asymptotic phenomenon of replication of the nodal lines.

Last but not least, instead of fixing an open set Ω with compact support, it is also possible to take a sequence of slowly increasing open sets Ω_n if φ has an unbounded support, under some decay assumptions. Unfortunately, it makes the proof less readable and enlightening and we decided to privilege clarity and concision.

Theorem 2.1 provides an upper bound on the existence of almost periods for the Arithmetic Random Waves. Similarly, there is a corresponding lower bound which is easily derived from (1.4).

Proposition 2.4. *Let τ_n be any sequence of almost periods of r_n . Then, there is some $A > 0$ such that*

$$|\tau_n| \geq \frac{\log(n)^A}{\sqrt{n}}.$$

for n large enough.

Contrary to the upper bound proof which depends only on the number of eigenvalues, the lower bound relies intrinsically on a careful understanding of the angular distribution μ_n . As such, it is only valid for the Arithmetic Random Wave model.

Proof. According to the conclusion of Lemma 1.1, there is some $\rho > 0$ such that

$$\sup_{|t| \leq \sqrt{n} \log(n)^\rho} |\tilde{r}_n(t) - \tilde{r}(t)| \longrightarrow 0.$$

Let $0 < A < \rho$ and assume that the conclusion of Proposition 2.4 does not hold. Then, we can find some sequence of almost periods τ_n and an increasing sequence n_k so that

$$|\tau_{n_k}| \leq \frac{\log(n_k)^A}{\sqrt{n_k}}.$$

In that case

$$\begin{aligned} |1 - r_{n_k}(\tau_{n_k})| &\geq |1 - J_0(2\pi\tilde{\tau}_{n_k})| - |\tilde{r}_{n_k}(\tilde{\tau}_{n_k}) - J_0(2\pi|\tilde{\tau}_{n_k}|)| \\ &\geq |1 - J_0(2\pi\tilde{\tau}_{n_k})| + o(1). \end{aligned}$$

But, as $|\tilde{\tau}_{n_k}| \geq \log(n_k)^A \rightarrow +\infty$ and J_0 converges to 0 at infinity,

$$\lim_{k \rightarrow +\infty} |1 - r_{n_k}(\tau_{n_k})| = 1.$$

This clearly contradicts the definition of τ_n as

$$1 - r_n(\tau_n) \rightarrow 0.$$

□

2.1.1. Discussion. If we combine Theorem 2.1 and Proposition 2.4, we see that the smallest almost period τ_n is located in an intermediate range, between $n^{-1/2} \log(n)^\rho$ and the subpolynomial range $n^{-1/2} \exp(\log(n)^{\kappa'})$ for $\kappa' \in (\log(2)/2, 1)$. This is very interesting as it is suspected that the Arithmetic Random Wave exhibits some kind of phase transition at a range $n^{-1/2} \log(n)^A$ for some $A > 0$. Below this threshold, T_n will behave similarly to the Berry Random Wave Model as it proved by Dierickx, Nourdin, Peccati and Rossi in [DNPR23]. Above this threshold, full correlation of the nodal length appears and our random eigenfunctions can no longer behave like a Berry Random Wave. This observation leads to a natural question, that is whether the phenomenon of nodal replication highlighted in this article is linked with the phenomenon of full correlation of the nodal lines. If that was the case, it would imply that one can find sequences of almost periods in the logarithmic range $n^{-1/2} \log(n)^A$. Such a result would mean that the upper bound in theorem 2.1 obtained via Dirichlet's approximation Theorem is far from optimal. That is not unlikely, yet it would be surprising as for most tuple of eigenvectors, this principle is indeed optimal. This is the topic of the forthcoming discussion.

Let $\mathcal{N} \geq 1$, $d \geq 1$ and $\gamma^\mathcal{N} = (\gamma_1, \dots, \gamma_\mathcal{N})$ be random elements in \mathbb{S}^{d-1} associated with the random covariance function

$$R_\mathcal{N}(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \cos(2\pi \langle \gamma_k, t \rangle), \quad t \in \mathbb{R}^d.$$

We make the following assumptions:

Assumption 2.5.

- γ is isotropic, i.e. for any rotation \mathcal{R} ,

$$\mathcal{R}\gamma^\mathcal{N} \stackrel{\mathcal{L}}{=} \gamma^\mathcal{N}.$$

- For a \mathcal{C}^∞ test function $h : \mathbb{R}^d \rightarrow [-1, 1]$, Hoeffding's inequality is satisfied

$$\mathbb{P} \left(\left| \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} h(\gamma_k) - \int_{\mathbb{S}^{d-1}} h(\gamma) \sigma(d\gamma) \right| > \varepsilon \right) \leq C \exp(-c\varepsilon^2 \mathcal{N}),$$

where c, C are constants that do not depend on h and σ is the uniform measure on \mathbb{S}^{d-1} .

The most natural model is that of i.i.d. random γ_i uniformly distributed on \mathbb{S}^1 . Now, if $\varepsilon_{\mathcal{N}}$ is a positive sequence converging to 0, then we let $\tilde{\tau}_{\mathcal{N}}$ be the smallest $\varepsilon_{\mathcal{N}}$ -almost period associated with $R_{\mathcal{N}}$, that is the smallest $|\tilde{\tau}_{\mathcal{N}}|$ such that

$$|\tilde{\tau}_{\mathcal{N}}| \geq 1 \quad \text{and} \quad R_{\mathcal{N}}(\tilde{\tau}_{\mathcal{N}}) > 1 - \varepsilon_{\mathcal{N}}.$$

According to Theorem 2.1, $\tilde{\tau}_{\mathcal{N}}$ satisfies almost surely, for any $\alpha > 0$, the inequality

$$(2.4) \quad |\tilde{\tau}_{\mathcal{N}}| \leq e^{\mathcal{N}^{1+\alpha}}$$

as long as \mathcal{N} is large enough.

The next proposition ensures that this latter inequality is indeed optimal for a generic sequence of wave vectors γ_k .

Proposition 2.6. *Under Assumption 2.5 and for any $a < c/d$ where c is the constant in Hoeffding's inequality,*

$$(2.5) \quad \tilde{\tau}_{\mathcal{N}} \geq e^{a\mathcal{N}}$$

with high probability in the sense that for $\varepsilon_{\mathcal{N}} \rightarrow 0$

$$\mathbb{P} \left(\sup_{1 \leq |t| \leq e^{a\mathcal{N}}} R_{\mathcal{N}}(t) \leq 1 - \varepsilon_{\mathcal{N}} \right) \rightarrow 1.$$

In particular, the preceding theorem is satisfied when $\gamma^{\mathcal{N}}$ is a sequence of independent and identically uniform random variables on \mathbb{S}^1 . Therefore, for a generic sequence of wave vectors, the bound obtained in Theorem 2.1 cannot be decently improved. The proof is at Section 5.

2.2. Degrees of freedom and Dirichlet's approximation Theorem optimality. Proposition 2.6 implies that the bound derived from Dirichlet's approximation Theorem is optimal and that one cannot do better for a generic sequence of wave vectors. Yet, modelling the frequencies of the Arithmetic Random Wave as the realisation of some independent and identically distributed uniform random variable is not an accurate representation of reality as all the arithmetic flavour of the model is lost in the process. Following the idea of Sartori [Sar22], it is possible to propose a more accurate model. Recall that n is expressible as the sum of two squares if

$$n = 2^\alpha \prod_{j=1}^k p_j^{\alpha_j} \prod_{j=1}^l q_j^{2\beta_j}$$

where p_i, q_k are prime numbers satisfying respectively $p \equiv 1 \pmod{4}$ and $q \equiv 3 \pmod{4}$. The preceding prime decomposition induces a prime decomposition in the ring of Gaussian integers $\mathbb{Z}[i]$, namely

$$n = 2^\alpha \prod_{j=1}^k \mathcal{P}_j^{\alpha_j} \overline{\mathcal{P}_j}^{\alpha_j} \prod_{j=1}^l q_j^{2\beta_j}$$

where \mathcal{P}_j is one the two squares roots of p_j in $\mathbb{Z}[i]$. Thus, if $n = \lambda_1^2 + \lambda_2^2$, it follows that

$$\lambda_1 + i\lambda_2 = \nu Z^\alpha \prod_{j=1}^k \mathcal{P}_j^{\gamma_j} \overline{\mathcal{P}_j}^{\alpha_j - \gamma_j} \prod_{j=1}^l q_j^{\beta_j}$$

for some $0 \leq \gamma_j \leq \alpha_j$ and $\nu \in \{1, -1, i, -i\}$. Here, Z refers to a square of 2 in $\mathbb{Z}[i]$.

For a generic integer n expressible as the sum of two squares, most of the valuations α_j are equal to 1 [Sar22]. We will make this assumption from now on. In that case, we denote by $\omega(n) = k$ the number of prime factors congruent to 3 (mod 4) in the decomposition n so that $2^{\omega(n)+2} = \mathcal{N}_n$.

If we write

$$\theta_j := \arg(\mathcal{P}_j), \quad 1 \leq j \leq k,$$

and

$$\theta := \arg \left((1 + \nu)^\alpha \prod_{j=1}^l q_j^{\beta_j} \right),$$

the correlation function of the Arithmetic Random Wave is then given by

$$(2.6) \quad \tilde{r}_n(t) = \frac{1}{\mathcal{N}_n} \sum_{\substack{\eta \in \{-1,1\}^{\omega(n)} \\ \nu \in \{0,1,2,3\}}} \cos \left(2\pi \left\langle \exp \left(i \frac{\pi}{2} \nu + i \sum_{j=1}^{\omega(n)} \eta_j \theta_j + i\theta \right), t \right\rangle \right).$$

In these settings, the wave vectors are dependent of the angles of the Gaussian primes. So, the system has only $\omega(n)$ degrees of freedom compared to \mathcal{N}_n before. This potentially allows to decrease drastically the bound on almost periods from Dirichlet's Arithmetic principle for a system with \mathcal{N} degrees of freedom. We were not able to prove rigorously this idea but it is possible to hint on why Dirichlet's Arithmetic Principle should not be optimal in these settings.

Consider the so-called *linearised* covariance function, that is

$$(2.7) \quad s_n(t) := \frac{4}{\mathcal{N}_n} \sum_{\eta \in \{-1,1\}^{\omega(n)}} \cos(2\pi \theta_\eta t), \quad t \in \mathbb{R}^2$$

where

$$(2.8) \quad \theta_\eta = \sum_{j=1}^{\omega(n)} \eta_j \theta_j \pmod{1}.$$

Let us stress that this covariance function is deterministic, on the contrary of Sartori which considers a version with $\omega(n)$ random arguments θ_j . In this model, the exponential terms in equation (2.6) were linearised around the origin, so that the number of degrees of freedom of the model, which can be interpreted as $\omega(n)$, the numbers of signs to choose, is the same as in the ARW. The latter model provides an interesting example of a deterministic almost periodic function where the bound given by Dirichlet's approximation (Theorem 2.2) is no longer optimal.

If (θ_η) was the realisation of some independent and identically random process, we would expect the smallest ε -almost period to be of order $\varepsilon^{-\mathcal{N}_n}$, by Proposition 2.6. But, the family (θ_η) is linearly dependent and using this observation, it is possible to improve drastically the upper bound.

Proposition 2.7. *There is $1 \leq \tilde{\tau}_n \leq c\mathcal{N}_n^{\log(\log(\log(n)))}$ such that*

$$|s_n(\tilde{\tau}_n) - 1| \rightarrow 0.$$

Hence this is a crude improvement over Dirichlet's Arithmetic Principle and closer in spirit to the scale \mathcal{N}_n^A where full correlation occurs, and Proposition 2.6 applied with $\mathcal{N} = \omega(n)$ yields that it is optimal if the angles θ_j can be considered as *generic* (i.e. behave like i.i.d. variables). Yet, it is not possible to extend this result to covariance functions of the form of (2.6) as the proof relies solely on linear considerations which are no longer satisfied outside the scope of this particular model. Yet, it shows that, for a model with $\omega(n)$ degrees of freedom, the bound from Dirichlet's approximation Theorem is no longer optimal. This leaves us with some unanswered but interesting questions.

Question 2.8. *What is the size of the smallest sequence of almost periods in model (2.7) ? Is it the same as for the actual ARW?*

2.3. Other manifolds and dimensions. The existence of sequences of almost periods for the correlation function r_n is directly related to the number of eigenvectors \mathcal{N}_n , and more precisely to the number of degrees of freedom of the system, this is a core idea of our bounds. The more distinct frequencies an almost periodic function has, the greater the almost period will likely be. In the case of the Arithmetic Random Wave, one has to remember that the norm of the almost period is constrained to live in the torus \mathbb{T}^2 , or in the dilated torus $\sqrt{n}\mathbb{T}^2$ after rescaling. Luckily, the number of frequencies \mathcal{N}_n grows logarithmically compared to \sqrt{n} . This is why it is possible to find almost periods slightly above Planck scale. Let us give other examples where the dimension of the eigenspace is polynomial and where there is no full correlation or almost replication.

2.3.1. Spherical harmonics. For the sphere \mathbb{S}^2 , there is no possibility of nodal replication. In fact, the correlation function r_n of the spherical harmonics of energy level $n(n+1)$ is related to the so-called Legendre Polynomials P_n and is given by

$$r_n(x, y) = P_n(\langle x, y \rangle), \quad x, y \in \mathbb{S}^1.$$

The characteristic distance in this model is $1/n$ and Hilb's formula [Sze75, Theorem 8.21.6]

$$P_n(\cos(\theta)) = \sqrt{\frac{\theta}{\sin(\theta)}} J_0\left(\left(n + \frac{1}{2}\right)\theta\right) + O\left(n^{-3/2}\right), \quad 0 \leq \theta \leq \pi - \varepsilon.$$

ensures that r_n converges uniformly to the Bessel function, except on small spherical caps located at the opposite pole. In particular, the latter formula prevents the existence of any almost period at a scale $1/n$ or slightly larger.

2.3.2. Higher dimensional tori. The situation is similar if one considers the Arithmetic Random Wave on the d -dimensional torus \mathbb{T}^d . In that case, the eigenvalues consist in the energy levels $E_n = -4\pi^2 n$ where n is a sum of d squares and the associated eigenspace has dimension $\mathcal{N}_n^{(d)}$, that is the number of ways an integer n can be written as sum of d squares. For $d \geq 5$, the precise asymptotic were derived first by Hardy and Littlewood and proved by Hua [Hua38] in 1938:

$$\frac{\Gamma(3/2)^d}{\Gamma(d/2)} n^{d/2-1} S(n),$$

where $S(n)$ is the so-called singular series, bounded above and below. The leading term of the asymptotics is $n^{d/2-1}$ whereas for $d = 2$, it was $\log(n)^{\log(2)/2}$. Hence,

for large d , the number of eigenvalues does no longer grows slowly compared to \sqrt{n} . Dirichlet's approximation Theorem will thus no longer ensure the existence of a sequence of almost periods slightly above Planck scale. The authors of this paper are not aware of any positive or negative result related to full-correlation in dimension $d \geq 5$, but, if the preceding heuristics is true, one might expect that nodal replication as well as full correlation no longer hold for $d \geq 5$ as there are too many eigenvalues.

3. DIRICHLET'S THEOREM FOR ALMOST PERIODIC FIELDS

The aim of the section is to prove the bounds on the smallest pseudo-period. To reach our goal, we will need to recall Dirichlet's approximation theorem. This is the content of the next paragraph.

3.1. Dirichlet's approximation theorem. Dirichlet's approximation theorem is a standard tool in Diophantine approximation and quantifies how well a generic vector $\mu \in \mathbb{R}^d$ can approximate a vector of integers. In this paper, one will need a generalization of this principle, which deals with simultaneous approximations. Before stating the aforementioned theorem, we introduce the following notation. For $\mu \in \mathbb{R}^d$, let $\text{dist}(\mu, \mathbb{Z})$ be the distance of μ to the nearest integer in \mathbb{Z} .

Theorem 3.1 (Dirichlet's approximation theorem). *Let $(\mu_k)_{1 \leq k \leq \mathcal{N}}$ a sequence of elements in \mathbb{R}^d . For any integer $m > 0$, there is $x \in \mathbb{Z}^d$ such that $1 \leq |x|_\infty \leq m^{\mathcal{N}/d}$ and*

$$\text{dist}(\langle \mu_k, x \rangle, \mathbb{Z}) \leq \frac{1}{m}, \quad 1 \leq k \leq \mathcal{N}.$$

$|\cdot|_\infty$ refers here to the supremum norm on \mathbb{R}^d .

The proof of this theorem relies on a clever use of the pigeonhole principle and is recalled briefly hereafter as it is quite interesting.

Proof. Let $N = m^{\mathcal{N}/d}$ and consider Q_N the set of lattice points $x \in \mathbb{Z}^d$ whose coordinates satisfy $1 \leq x_k \leq N$. Notice that Q_N contains $N^d = m^{\mathcal{N}}$ points.

There are now two possibilities. Either, there is some $x \in Q_N$ such that

$$\text{dist}(\langle \mu_k, x \rangle, \mathbb{Z}) \leq \frac{1}{m}, \quad 1 \leq k \leq n.$$

In that case, there is nothing to prove .

If not, split $[0, 1]^{\mathcal{N}}$ in m^n sub-squares of side length at most $\frac{1}{m}$ and associate with $x \in Q_N$ the vector $y_x \in \mathbb{R}^{\mathcal{N}}$ defined as

$$y_x = (\langle \mu_k, x \rangle \pmod{1})_{1 \leq k \leq \mathcal{N}}.$$

The set $\{y_x : x \in Q_N\}$ contains $N^d = m^{\mathcal{N}}$ elements and the above assumption ensures that none of these elements is contained in the lower sub-square $C = [0, \frac{1}{m}]^{\mathcal{N}}$. Hence, by the pigeonhole principle, they are two distinct elements y_{x_1} and y_{x_2} contained in the same square, that is

$$\text{dist}(\langle \mu_k, y_{x_1} - y_{x_2} \rangle, \mathbb{Z}) \leq \frac{1}{m}, \quad 1 \leq k \leq \mathcal{N}$$

In that case, $x = y_{x_2} - y_{x_1}$ works. □

As we have now proved Dirichlet's approximation theorem, it is time to concentrate on Theorem 2.2. In fact, we will prove a slight generalisation. Let $|\cdot|$ be euclidean norm. It will be preferable to work with this norm instead of the uniform norm. This is not a limitation as all norms are equivalent on \mathbb{R}^d and the conclusion of Dirichlet's approximation theorem remains the same, up to a constant we will neglect.

Corollary 3.2. *Let $\mathcal{N} \geq 1$ and define the almost periodic function*

$$f(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \cos(2\pi \langle \gamma_k, t \rangle), \quad t \in \mathbb{R}^d$$

with $\gamma_k \in \mathbb{R}^d$.

For every $\varepsilon > 0$, then there is an $\frac{1}{m}$ -almost period τ satisfying

$$1 \leq |\tau| \leq (2\pi)^{\mathcal{N}/d} m^{\mathcal{N}/d}$$

Moreover, if each γ_k belong to \mathbb{S}^{d-1} , then τ is also an almost period for the derivatives of f in the sense that for any $\alpha \in \mathbb{N}^d$, the following inequality is satisfied:

$$\sup_{t \in \mathbb{R}^d} |\partial^\alpha f(t + \tau) - \partial^\alpha f(t)| \leq \frac{(2\pi)^\alpha}{m}$$

where $x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}$.

Proof. Dirichlet approximation theorem ensures the existence of $1 \leq |\tau|_\infty \leq m^{n/d}$ such that the following inequalities are all satisfied:

$$\text{dist}(\langle \mu_k, \tau \rangle, \mathbb{Z}) \leq \frac{1}{m}, \quad 1 \leq k \leq \mathcal{N},$$

or equivalently

$$\text{dist}(2\pi \langle \mu_k, \tau \rangle, 2\pi\mathbb{Z}) \leq \frac{2\pi}{m}, \quad 1 \leq k \leq \mathcal{N}.$$

Now, for any $t \in \mathbb{R}^d$:

$$\begin{aligned} |f(t + \tau) - f(t)| &= \frac{1}{\mathcal{N}} \left| \sum_{k=1}^{\mathcal{N}} \Re e^{2i\pi \langle \mu_k, t + \tau \rangle} - \Re e^{2i\pi \langle \mu_k, t \rangle} \right| \\ &\leq \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \left| e^{2i\pi \langle \mu_k, \tau \rangle} - 1 \right| \\ (\text{mean value theorem}) &\leq \frac{1}{n} \sum_{k=1}^{\mathcal{N}} \text{dist}(2\pi \langle \mu_k, \tau \rangle, 2\pi\mathbb{Z}) \\ &\leq \frac{2\pi}{m}. \end{aligned}$$

If all the γ_k are located on \mathbb{S}^1 , then we can make use of the mean value inequality, noticing that

$$\begin{aligned} |\partial^\alpha f(t + \tau) - \partial^\alpha f(t)| &= \frac{(2\pi)^\alpha}{m} \left| \sum_{k=1}^{\mathcal{N}} \gamma_k^\alpha \left(\Re e^{2i\pi \langle \mu_k, t + \tau \rangle} - \Re e^{2i\pi \langle \mu_k, t \rangle} \right) \right| \\ &\leq \frac{(2\pi)^\alpha}{m} \sum_{k=1}^{\mathcal{N}} |\gamma_k^\alpha| \left| e^{2i\pi \langle \mu_k, \tau \rangle} - 1 \right| \end{aligned}$$

The conclusion follows from the inequality $|\gamma_k^\alpha| \leq 1$. \square

From now on, we will work with $\varepsilon > 0$ instead of $\frac{1}{m}$ as the latter notation is more readable. It is not problematic as we are mainly interested into applying the latter result to very small ε . In that case, we can always find an integer $m > 0$ such that $\varepsilon \approx \frac{1}{m}$.

Now, using Corollary 3.2, we can deduce a slight improvement over Theorem 2.1. Not only Corollary 3.2 implies the existence of sequence of almost periods τ_n associated with the covariance function of the Arithmetic Random Waves, but it also implies that τ_n is a sequence of almost periods for the derivatives of this function. This is a fundamental observation as derivatives of the covariance function control the derivatives of the underlying Gaussian field. This will be especially useful when dealing with the nodal lines. More precisely, we will need the content of the following proposition, which is an immediate consequence of the results above.

Corollary 3.3. *If τ_n is a sequence of almost periods satisfying the assumptions of Theorem 2.1, then for any $\alpha \in \mathbb{N}^2$,*

$$\sup_{t \in \sqrt{n}\mathbb{T}^2} |\partial^\alpha \tilde{r}_n(t + \tau_n) - \partial^\alpha \tilde{r}_n(t)| \rightarrow 0.$$

3.2. When Dirichlet's approximation theorem fails. This preceding theorem can be used to prove Proposition 2.7.

Proof. Let $0 < \varepsilon < 1$. According to Theorem 3.1 applied with $\mathcal{N} = \omega(n)$, $d = 1$, we can find some $|\tau| \leq \left(\frac{2\pi\omega(n)}{\varepsilon}\right)^{\omega(n)}$ such that

$$\text{dist}(\theta_i \tau, \mathbb{Z}) \leq \frac{\varepsilon}{2\pi\omega(n)}, \quad 1 \leq i \leq \omega(n).$$

For any $\eta \in \{-1, 1\}^{\omega(n)}$,

$$\text{dist}(\theta_\eta \tau, \mathbb{Z}) \leq \sum_{i=1}^{\omega(n)} \text{dist}(\theta_i \tau, \mathbb{Z}) \leq \frac{\varepsilon}{2\pi}.$$

Mirroring the proof of Corollary 3.2, we easily derive the inequality

$$|s_n(t + \tau) - s_n(t)| \leq \varepsilon, \quad t \in \mathbb{R}.$$

To finish the proof, one can choose for instance $\varepsilon_n = \omega(n)^{-1}$. □

4. REPLICATION OF THE NODAL LINES

In this section, the results related to the existence of almost periods are used to prove the phenomenon of the replication of the nodal lines as stated in Theorem 2.3. The proof is based on a nice application of the continuous mapping theorem. But, before dwelling into the proof, we give the main ideas.

At Planck scale, the Arithmetic Random Wave behaves almost exactly like a random eigenfunction of the Laplace operator on the plane \mathbb{R}^2 in the sense \tilde{T}_n that converges in distribution to a Berry Random Wave T for the topology of uniform convergence on Ω . On the other hand, we expect \tilde{T}'_n to behave very similarly to \tilde{T}_n in the high energy limit. It is thus natural to conjecture that $(\tilde{T}_n, \tilde{T}'_n)$ converges to (T, T) .

Unfortunately, the present formulation is too weak to derive any meaningful result related to the nodal lines of the Arithmetic Random Wave. What will be needed is a control on the behaviour of the derivatives of the underlying field and that control will be offered by Corollary 3.3. Therefore, instead of proving that $(\tilde{T}_n, \tilde{T}'_n)$ converges to (T, T) for the topology of uniform convergence, we will prove that the convergence holds also for the topology of uniform convergence of the first p derivatives.

Notation. In what follows, $\Omega \subset \mathbb{R}^2$ refers to a convex, bounded, open set with smooth boundary and $\mathcal{C}^p(\bar{\Omega})$ refers to the set of p times differentiable mappings on Ω whose partial derivatives extends continuously to the boundary. This space is endowed with the norm

$$\|f\|_p = \sum_{|\alpha| \leq p} \sup_{\Omega} |\partial^\alpha f|$$

where $|\alpha| = \alpha_1 + \alpha_2$. This norm turns $\mathcal{C}^p(\bar{\Omega})$ into a Polish space. Further details are recalled in the Appendix.

4.1. Proof of the replication phenomenon. Let

$$U_n(s, t) := (\tilde{T}_n(s), \tilde{T}'_n(t)), \quad s, t \in \bar{\Omega},$$

and

$$U(s, t) := (T(s), T(t)), \quad s, t \in \bar{\Omega}.$$

Both U_n and U define processes on the product space $\mathcal{C}^p(\bar{\Omega})^2$ endowed with the product topology. In particular, $\mathcal{C}^p(\bar{\Omega})^2$ is a Polish space.

Proposition 4.1. *Under the preceding assumptions, (U_n) converges in distribution to U .*

Proof. As $\mathcal{C}^p(\bar{\Omega})^2$ is Polish space, convergence in distribution is equivalent to tightness and unicity of the limit.

Step 1: tightness

The sequence (U_n) is tight if and only if its marginals (\tilde{T}_n) and (\tilde{T}'_n) are tight. As the Arithmetic Random Wave is a stationary process, \tilde{T}_n and \tilde{T}'_n have the same distribution. Hence, we only have to prove tightness for \tilde{T}_n in $\mathcal{C}^p(\bar{\Omega})$. Following the conclusion of Proposition 6.3, this amounts to prove that $(\partial^\alpha \tilde{T}_n)$ is tight in $\mathcal{C}^0(\bar{\Omega})$ for any $|\alpha| \leq p$. This problem is very tractable as tightness in $\mathcal{C}^0(\bar{\Omega})$ is a deeply studied topic.

We will use Kolmogorov tightness criterion [Kal02, Corollary 16.9]. $\partial^\alpha \tilde{T}_n$ is a stationary centered Gaussian process with correlation function

$$\tilde{r}_{\alpha, n}(t) := (-1)^{|\alpha|} \partial^{2\alpha} \tilde{r}_n(t), \quad t \in \mathbb{R}^2.$$

Fixed for a moment $t \in \Omega$. $\partial^\alpha \tilde{T}_n(t)$ is then a centered Gaussian random variable whose variance σ_n^2 is converging to

$$\sigma^2 := (2\pi)^{2|\alpha|} \int_{\mathbb{S}^1} \lambda^{2\alpha} \sigma(d\lambda) > 0.$$

where $|\alpha| = \alpha_1 + \alpha_2$ and σ is the uniform probability on \mathbb{S}^1 . Hence, $(\partial^\alpha \tilde{T}_n(t))$ converge in distribution to a centered normal random variable with variance σ^2 .

In the meantime, $\partial^\alpha \tilde{T}_n(t) - \partial^\alpha \tilde{T}_n(s)$ is also a centered Gaussian random variable with variance

$$\sigma^2(t-s) := 2(\tilde{r}_{\alpha,n}(0) - \tilde{r}_{\alpha,n}(t-s))$$

Combining the triangle inequality with

$$1 - \cos(t) \leq \frac{t^2}{2}, \quad t \in \mathbb{R}^2,$$

one has

$$\begin{aligned} |\tilde{r}_{\alpha,n}(0) - \tilde{r}_{\alpha,n}(t-s)| &\leq \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \left| 1 - \cos \left(2\pi \left\langle \frac{\lambda}{\sqrt{n}}, t-s \right\rangle \right) \right| \\ &\leq \frac{1}{2\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \left\langle \frac{\lambda}{\sqrt{n}}, t-s \right\rangle^2 \\ (\text{Cauchy Schwarz}) &\leq \frac{1}{2} |t-s|^2. \end{aligned}$$

In particular,

$$\begin{aligned} \mathbb{E} \left[\left(\partial^\alpha \tilde{T}_n(t) - \partial^\alpha \tilde{T}_n(s) \right)^4 \right] &= 3\sigma(t-s)^4 \\ &\leq 3|t-s|^4. \end{aligned}$$

The conditions of Kolmogorov criterion are met. Hence, the sequence $(\partial^\alpha \tilde{T}_n)$ is tight in $\mathcal{C}^0(\bar{\Omega})$.

Step 2: Unicity of the limit

The distribution of $\mathcal{C}^p(\bar{\Omega})^2$ is characterised by its finite dimensional distributions (see Lemma 6.4 and the associated discussion). It suffices thus to prove that

$$U_n^{(p)} = (U_n(t_1), \dots, U_n(t_p))$$

converges in distribution to

$$U^{(p)} = (U(t_1), \dots, U(t_p)),$$

where t_i is a finite sequence of points in Ω .

As both $U_n^{(p)}$ and $U^{(p)}$ are centered Gaussian processes, it is only needed to show that the covariance $U_n^{(p)}$ converges to the covariance of $U^{(p)}$. In terms of covariance function, this amounts to prove that

$$\begin{cases} \tilde{r}_n(t_j - t_i) \longrightarrow \tilde{r}(t_i - t_j), & 1 \leq i, j \leq n. \\ \tilde{r}'_n(t_j - t_i) \longrightarrow \tilde{r}'(t_i - t_j) \end{cases}$$

The latter statement is trivial as \tilde{r}_n converges uniformly to \tilde{r} (Lemma 1.1) and $\tilde{r}_n - \tilde{r}'_n$ converges uniformly to 0 (Theorem 2.1). This finishes the proof. \square

In order to prove nodal replication as introduced by Theorem 2.3, it is natural to consider the application

$$\Gamma : \begin{cases} \mathcal{C}^1(\bar{\Omega}, \mathbb{R}) & \longrightarrow & \mathbb{R} \cup \{\pm\infty\} \\ f & \longmapsto & \int_{\mathcal{Z}_f(\Omega)} \varphi(t) H_1(dt). \end{cases}$$

where $\mathcal{Z}_f(\Omega)$ is the zero set of f on Ω and $\varphi : \Omega \rightarrow \mathbb{R}$ is continuous. Instead of considering functions $f \in \mathcal{C}^p(\overline{\Omega})$, it will be preferable to work with a slightly larger domain U such that $\overline{\Omega} \subset U$ as this will allow some more leeway.

First, the co-area formula

$$\Gamma(f) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{\Omega} \mathbf{1}(|f(t)| \leq \varepsilon) \varphi(t) |\nabla f(t)| dt$$

ensures that Γ is a measurable mapping. This allows to define $\Gamma(\tilde{T}_n)$ and Theorem 2.3 amounts to show that $\Gamma(\tilde{T}_n) - \Gamma(\tilde{T}'_n)$ converges in distribution to 0. The latter difference could be ill-defined if both $\Gamma(\tilde{T}_n)$ and $\Gamma(\tilde{T}'_n)$ happens to be infinite at the same time, but this situation is excluded by Bulinskaya's lemma [AW09, Proposition 1.20] which ensures that almost surely \tilde{T}_n is regular.

Definition 4.2. $f \in \mathcal{C}^1(\overline{U})$ is regular if and only

$$\forall t \in \overline{U}, f(t) = 0 \implies \nabla f(t) \neq 0.$$

Regular functions have the property that their zero sets are 1 dimensional smooth manifold. This ensures in particular that $\Gamma(f)$ is not infinite and that the random variable $\Gamma(\tilde{T}_n) - \Gamma(\tilde{T}'_n)$ is well-defined. Theorem 2.3 will follow easily from the continuous mapping theorem [Kal02, Theorem 4.27] if we manage to prove that Γ is \mathbb{P}_T almost-surely continuous. This result is not immediate as Γ is not continuous everywhere.

Example 4.3. Let Ω be the open unit ball on \mathbb{R}^2 . let $f_\alpha(t) = \alpha - |t|^2$. It is easily checked that $\Gamma(f_1) = 0$. Yet, $\Gamma(f_\alpha) = 2\pi\alpha$ for $\alpha \rightarrow 1^-$ even though $f_\alpha \rightarrow f_1$ as α tends to 1^- .

In the preceding example, the zeros concentrate alongside the boundary and that is the root cause of the default of continuity. In our random settings, such situation is very unlikely to happen as it is testified by the following proposition.

Proposition 4.4.

$$H_1(\mathcal{Z}_T(\partial\Omega)) < +\infty, \quad a.s.$$

Proof. Denote by H_0 the counting measure. We will prove the stronger statement that the number of intersections between $\mathcal{Z}_f(\partial\Omega)$ is finite, that is $H_0(\mathcal{Z}_f(\partial\Omega)) < +\infty$ almost-surely.

As $\partial\Omega$ is compact, it suffices to show that the latter result holds locally. Fix $\omega \in \partial\Omega$ and parametrise $\partial\Omega$ in a neighbourhood of ω by a smooth map $\gamma : \overline{V} \rightarrow \partial\Omega$ where V is a bounded neighbourhood of 0. We have to show that the Gaussian process $T \circ \gamma$ has only a finite number of zeros in \overline{V} . But, this follows immediately from Kac-Rice formula [AW09, Theorem 6.2]. \square

Excluding the latter pathological edge cases, Γ is continuous. More precisely,

Proposition 4.5. *Let f be a regular function, φ continuous, such that*

$$(4.1) \quad H_1(\mathcal{Z}_f(\partial\Omega)) = 0,$$

then Γ is continuous in f .

The proof of this result is technical and not very enlightening. It is thus proved in the Appendix (see Proposition 6.6). Using the latter result, we deduce in particular that Γ is \mathbb{P}_T surely continuous and Theorem 2.3 follows easily for the continuous mapping theorem.

Before concluding this section, we can make a few comments. The requirement of taking a convex open set Ω with smooth boundary can be probably relaxed. Secondly, the result still holds, if instead of taking a fixed open set Ω , we take a sequence of slowly increasing open set Ω_n . The preceding proof no longer works in that case but it is possible to circumvent the issue with a careful use of Borel-Tis inequality or by some coupling arguments.

5. OPTIMALITY OF DIRICHLET'S APPROXIMATION THEOREM

The goal of this section is to investigate the optimality of Dirichlet's approximation theorem and prove Proposition 2.6.

In what follows, $\gamma^{\mathcal{N}}$ is a sequence of wave vectors satisfying Assumption 2.5. In particular, $\gamma^{\mathcal{N}}$ can be a sequence of independent uniformly distributed on \mathbb{S}^{d-1} . We start with a preliminary and immediate lemma.

Lemma 5.1. *Under the preceding assumptions, each component of $\gamma^{\mathcal{N}}$ follows a uniform distribution on the sphere \mathbb{S}^{d-1} and we let for $t \in \mathbb{R}^d$*

$$R(t) = \mathbb{E}[R_{\mathcal{N}}(t)] = \int_{\mathbb{S}^{d-1}} \cos(2\pi\langle\gamma, t\rangle) \sigma(d\gamma),$$

where σ is the uniform measure on the sphere.

Then, R defines a radial function which converges to 0 at infinity and which has a unique maximum at $t = 0$.

Proof. It is standard fact that the only rotation invariant probability on the sphere \mathbb{S}^{d-1} is the uniform measure Hausdorff measure on \mathbb{S}^{d-1} . Hence, each component of $\gamma^{\mathcal{N}}$ has to follow a uniform distribution.

We notice that, by symmetry,

$$R(t) = \int_{\mathbb{S}^{d-1}} e^{2i\pi\langle\gamma, t\rangle} \sigma(d\gamma)$$

so that $R(t) = \widehat{\sigma}(t)$ where $\widehat{\cdot}$ refers to the Fourier transform. In particular, this leads to an explicit formula for $R(t)$, that is

$$R(t) = \omega \frac{J_{\nu}(2\pi|t|)}{|t|^{\nu}}$$

where J_{ν} is the ν^{th} Bessel function of first kind, $\nu = d/2 - 1$ and ω a constant chosen so that $R(0) = 1$. The conclusion of Lemma 5.1 follows then from the standard properties of the Bessel functions. \square

We are now ready to prove Proposition 2.6. Let $a < c/d$ where c is the constant in Hoeffding's inequality and let $\varepsilon_{\mathcal{N}}$ be a positive sequence converging to 0. We define $\tau_{\mathcal{N}}$ as the smallest $\varepsilon_{\mathcal{N}}$ -almost period of $R_{\mathcal{N}}$. According to Hoeffding's inequality in Assumption 2.5,

$$(5.1) \quad \mathbb{P}(|R_{\mathcal{N}}(t) - R(t)| \geq 1 - \varepsilon_{\mathcal{N}}) \leq Ce^{-c(1-\varepsilon_{\mathcal{N}})^2\mathcal{N}}, \quad t \in \mathbb{R}^d,$$

To make use of the latter inequality, we cover the centered ball with radius $e^{a\mathcal{N}}$ with $d_{\mathcal{N}} = O(\mathcal{N}^d e^{da\mathcal{N}})$ balls of radius at most $\frac{1}{\mathcal{N}}$. Let $C_{i,\mathcal{N}}$ denote these latter balls and $t_{i,\mathcal{N}}$ be their respective centers. As both $R_{\mathcal{N}}$ and R are 2π -Lipschitz continuous, one has that

$$\sup_{t \in C_{i,n}} |R_{\mathcal{N}}(t) - R(t)| \leq |R_{\mathcal{N}}(t_{i,\mathcal{N}}) - R(t_{i,\mathcal{N}})| + \frac{2\pi}{\mathcal{N}}.$$

Hence, for any $\varepsilon > 0$,

$$\mathbb{P} \left(\limsup_{\mathcal{N} \rightarrow +\infty} \left\{ \sup_{|t| \leq e^{a\mathcal{N}}} |R_{\mathcal{N}}(t) - R(t)| \geq 1 - \varepsilon_{\mathcal{N}} \right\} \right) = \mathbb{P} \left(\limsup_{\mathcal{N} \rightarrow +\infty} \left\{ \sup_{1 \leq i \leq d_{\mathcal{N}}} |R_{\mathcal{N}}(t_{i,\mathcal{N}}) - R(t_{i,\mathcal{N}})| \geq 1 - \varepsilon_{\mathcal{N}} \right\} \right).$$

By the union bound and Equation (5.1),

$$\begin{aligned} \mathbb{P} \left(\sup_{1 \leq i \leq d_{\mathcal{N}}} |R_{\mathcal{N}}(t_{i,\mathcal{N}}) - R(t_{i,\mathcal{N}})| \geq 1 - \varepsilon_{\mathcal{N}} \right) &\leq d_{\mathcal{N}} C e^{-c(1-\varepsilon_{\mathcal{N}})^2 \mathcal{N}} \\ &= O \left(\mathcal{N}^d e^{-c(1-\varepsilon_{\mathcal{N}})^2 \mathcal{N} + ad\mathcal{N}} \right) \\ &= O \left(\mathcal{N}^d e^{-(c-ad)\mathcal{N}} \right) \end{aligned}$$

which is summable as $\varepsilon_{\mathcal{N}}$ converges to 0. Borel-Cantelli lemma implies that

$$\mathbb{P} \left(\limsup_{n \rightarrow +\infty} \left\{ \sup_{|t| \leq e^{a\mathcal{N}}} |R_n(t) - R(t)| \geq 1 - \varepsilon_{\mathcal{N}} \right\} \right) = 0.$$

Hence, $R_{\mathcal{N}}$ converges uniformly to R on balls of radius $e^{a\mathcal{N}}$. But, as $R(t)$ is bounded away from 1 everywhere except on a neighbourhood of 0, this implies that there is asymptotically no $\varepsilon_{\mathcal{N}}$ -almost periods smaller than $e^{a\mathcal{N}}$. In particular, Dirichlet's approximation theorem bound is the best possible in that case and Proposition 2.7 is proved.

6. APPENDIX

6.1. Topology of uniform convergence of derivatives. This first chapter recalls briefly the main results associated with the topology of uniform convergence of the first p derivatives on a compact set. In what follows, $\Omega \subset \mathbb{R}^2$ will designate a convex, bounded open set with locally \mathcal{C}^1 boundary. These latter assumptions can be relaxed, but in the framework of this article, there is no need to dwell into greater generality.

Notation. When $\alpha \in \mathbb{N}^2$, the notation $|\alpha| = \alpha_1 + \alpha_2$ will be used. In other context, it will refer to the standard euclidean norm.

Topology of uniform convergence of derivatives. Let p be a positive integer. We define the set $\mathcal{C}^p(\bar{\Omega})$ of p -times differentiable functions on $\bar{\Omega}$ as the set of functions $f : \bar{\Omega} \rightarrow \mathbb{R}$ satisfying

- (1) f is $\mathcal{C}^p(\Omega)$, that is f is p -times continuously differentiable on Ω .
- (2) $\partial^\alpha f$ extend continuously to $\bar{\Omega}$ for any $|\alpha| \leq p$.

It is natural to equip $\mathcal{C}^p(\bar{\Omega})$ with the norm

$$\|f\|_p := \sum_{|\alpha| \leq p} \|\partial^\alpha f\|_\infty$$

where $\|\cdot\|_\infty$ is the uniform norm on $\overline{\Omega}$.

In particular, the definition of $(\mathcal{C}^0(\overline{\Omega}), \|\cdot\|_0)$ coincides with the topology of uniform convergence on compact set for continuous function. We recall now some standard facts on $\mathcal{C}^p(\overline{\Omega})$.

Lemma 6.1. *$\mathcal{C}^p(\overline{\Omega})$ endowed with $\|\cdot\|_p$, is a Polish space, that is separable and complete.*

Proof. The family of polynomials $\mathbb{Q}[x, y]$ is a dense family in $\mathcal{C}^p(\overline{\Omega})$. This result is surely true, but it is not as obvious as it seems... Indeed, for pathological bounded open sets Ω , the density of polynomials may dramatically fail (see this discussion [Sau]).

In our setting, the regularity assumptions put on Ω forbids any pathological behaviour. As Ω is convex, it satisfies the assumptions of Whitney extension theorem [Whi34], which ensures that any $f \in \mathcal{C}^p(\Omega)$ can be extended to a function of class $\mathcal{C}^p(\mathbb{R}^2)$.

The separability follows then from a routine argument. Consider a mollifier, for instance the density of standard 2-dimensional random normal variable and let $\varphi_n(t) := 2^{-n}\varphi(2^nt)$. Then $\varphi_n \otimes f$ is a smooth function whose all partial derivatives of order $|\alpha| \leq p$ converge uniformly on compact sets to those of f . Weierstrass theorem ensures that φ_n can be well-approximated by a polynomial P_n on the euclidean unit ball B and we might choose the sequence P_n so that

$$\sup_B |P_n - \varphi_n| \rightarrow 0.$$

To conclude the proof, one has to check that

$$\mathbf{1}_B P_n \otimes f$$

is a polynomial which approximates arbitrarily closely f in the $\mathcal{C}^p(\overline{\Omega})$ topology.

We now prove the completeness of $\mathcal{C}^p(\overline{\Omega})$. Recall that $\mathcal{C}^0(\overline{\Omega})$ equipped with $\|\cdot\|_\infty$ is a Banach space.

Let (f_n) be a Cauchy sequence in $\mathcal{C}^p(\overline{\Omega})$. For every $|\alpha| \leq p$, $(\partial^\alpha f_n)$ is a Cauchy sequence in $\mathcal{C}^0(\overline{\Omega})$, hence there is a continuous function f_α on $\overline{\Omega}$ such that

$$\|\partial^\alpha f_n - f_\alpha\|_\infty \rightarrow 0.$$

We note $f := f_{(0,0)}$. To finish the proof, it suffices to show that f is $\mathcal{C}^p(\Omega)$ and $\partial^\alpha f = f_\alpha$. This follows immediately from the following well-known lemma.

Lemma 6.2. *Let (f_n) be a sequence of differentiable functions such that f_n converges pointwise to f and f'_n converges uniformly to g . Then, f is differentiable on $]a, b[$ and $f' = g$.*

□

Results related to the compactness of $\mathcal{C}^0(\overline{\Omega})$ such as Arzelà–Ascoli have a straightforward extension to $\mathcal{C}^p(\overline{\Omega})$. This derives from the following observation.

Lemma 6.3. *Let $A \subset \mathcal{C}^p(\overline{\Omega})$. The two following proposition are equivalent.*

- (1) *A is relatively compact in $\mathcal{C}^p(\overline{\Omega})$.*

(2) For any $|\alpha| \leq p$,

$$A(\alpha) := \{\partial^\alpha f : f \in A\}$$

is relatively compact in $\mathcal{C}^0(\overline{\Omega})$

Proof. The direct implication is immediate as

$$\|\partial^\alpha f\|_\infty \leq \|f\|_p.$$

For the converse, consider a subsequence f_n in A . Using the relative compactness of $A(\alpha)$ in $\mathcal{C}^0(\overline{\Omega})$, we can extract a subsequence $f_{\sigma(n)}$ such that

$$\|\partial^\alpha f_{\sigma(n)} - f_\alpha\|_\infty \rightarrow 0$$

where each f_α is some continuous function on $\overline{\Omega}$ and similarly to the proof of Lemma 6.1, one has $\partial^\alpha f = f_\alpha$ where $f := f_{(0,0)}$. \square

Borel sets and convergence in distribution in $\mathcal{C}^p(\overline{\Omega})$. In the previous paragraph, it was shown that $\mathcal{C}^p(\overline{\Omega})$ is Polish space. This is the natural framework to develop a well-behaved notion of convergence in distribution as Prokhorov's theorem is satisfied. The treatment of this question is inspired by Kallenberg [Kal02, Chapter 15].

There are two natural notions of measurable sets in $\mathcal{C}^p(\overline{\Omega})$, one is given by the Borel-algebra $\mathcal{B}(\mathcal{C}^p(\overline{\Omega}))$, the other is given by the cylindrical σ -algebra Cyl^p , that is the coarsest σ -algebra making the projection

$$\Pi_I : \begin{cases} \mathcal{C}^p(\overline{\Omega}) & \longrightarrow & \mathbb{R}^I \\ f & \longmapsto & \{f(t) : t \in I\} \end{cases}$$

measurable where I is any finite subset of Ω . Luckily, these two approaches can be unified into a unique one as they lead to the same measurable sets.

Lemma 6.4. $\mathcal{B}(\mathcal{C}^p(\overline{\Omega})) = \text{Cyl}^p$

Proof. Π_I is continuous, hence $\text{Cyl}^p \subset \mathcal{B}(\mathcal{C}^p(\overline{\Omega}))$.

For the converse, it suffices to show that any open set belongs to Cyl^p . As $\mathcal{C}^p(\overline{\Omega})$ is separable, it further reduces into showing that any closed ball B is Cyl^p -measurable. We denote by f the center of B and r its radius.

If $A := \{\alpha \in \mathbb{N}^2 : |\alpha| \leq p\}$ and D is a dense countable subset of Ω , then

$$B = \bigcap_{(t_\alpha) \in D^A} \left(g \in \mathcal{C}^p(\overline{\Omega}) : \sum_{|\alpha| \leq p} |\partial^\alpha g(t_\alpha) - \partial^\alpha f(t_\alpha)| \leq r \right).$$

Thus, proving the reverse inclusion amounts to prove that

$$\Pi_t^\alpha : \begin{cases} \mathcal{C}^p(\overline{\Omega}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \partial^\alpha f(t) \end{cases}$$

is Cyl^p -measurable.

Let (e_1, e_2) be the canonical basis of \mathbb{R}^2 and let

$$\Delta_h^i f(t) := \frac{f(t + he_i) - f(t)}{h}$$

be the discrete partial derivatives of order f . This definition extends naturally to higher derivatives Δ_h^α and Taylor formula ensures that

$$\lim_{h \rightarrow 0} \Delta_h^\alpha f(t) = \partial^\alpha f(t), \quad t \in \Omega.$$

As a pointwise limit of measurable functions, Π_t^α is thus Cyl^p -measurable. \square

In particular, the preceding lemma ensure that a distribution in $\mathcal{C}^p(\overline{\Omega})$ is characterised by its finite-dimensional distributions.

More generally, it ensures that a random variable defined on the product space $\mathcal{C}^p(\overline{\Omega})^2$ is also characterised by its finite dimensional. Indeed, as $\mathcal{C}^p(\overline{\Omega})$ is separable, one has

$$\begin{aligned} \mathcal{B}(\mathcal{C}^p(\overline{\Omega})^2) &= \mathcal{B}(\mathcal{C}^p(\overline{\Omega})) \otimes \mathcal{B}(\mathcal{C}^p(\overline{\Omega})) \\ &= \text{Cyl}^p \otimes \text{Cyl}^p \end{aligned}$$

This result was used in the proof of Proposition 4.1.

Alongside the cylindrical topology, we will need some results relative to the tightness of a sequence of random variables in $\mathcal{C}^p(\overline{\Omega})$.

Proposition 6.5. *Let (X_n) a sequence random variables on $\mathcal{C}^p(\overline{\Omega})$. Then, the two proposition are equivalent:*

- (1) (X_n) is tight.
- (2) For any $|\alpha| \leq p$, the sequence $(\partial^\alpha X_n)$ is tight for the topology of uniform convergence on $\overline{\Omega}$.

Proof. The proof relies on Prokhorov theorem [Kal02, Theorem 16.3]. In a Polish space, a sequence of random variable is tight if and only if it is relatively compact.

But according to Lemma 6.3, (X_n) is tight in $\mathcal{C}^p(\overline{\Omega})$ if and only if $(\partial^\alpha X_n)$ is tight in $\mathcal{C}^0(\overline{\Omega})$ for any $|\alpha| \leq p$. \square

6.2. Continuity along the nodal lines. The second part of this appendix is dedicated to the proof of Proposition 4.5. This work is a bit tedious and it is divided into smaller steps.

Notation. If g is continuous mapping, the notation $\mathcal{Z}_g(\Omega)$ will be used to denote the zero sets of g on Ω .

General framework. We recall first Proposition 4.5.

Let $U \subset \mathbb{R}^2$ be a bounded convex open set with smooth boundary and $f \in \mathcal{C}^1(\overline{U})$ be a regular function. If Ω is an open convex subset of U and φ a continuous function on Ω , then, provided that $f^{-1}(0)$ does not accumulate on $\partial\Omega$, the map

$$\Lambda(g) := \int_{\mathcal{Z}_g(\Omega)} f(t) H_1(dt)$$

is continuous at f . More precisely, the goal is to prove the following proposition.

Proposition 6.6. *Under the preceding assumptions and provided that*

$$(6.1) \quad H_1(\mathcal{Z}_f(\partial\Omega)) = 0,$$

the map Γ is continuous in f for the $\mathcal{C}^1(\overline{U})$ -topology.

The latter result can be extended to open sets Ω which are well-approximated above and below in the sense that there exists Ω_n and Ω^n two sequences of respectfully increasing and decreasing open sets such that

$$\bigcup_{n \geq 1} \Omega_n = \Omega$$

and

$$\bigcap_{n \geq 1} \Omega^n = \bar{\Omega},$$

provided some regularity on Ω . In the case of convex sets, these conditions are immediately satisfied.

Proof of continuity. By decomposing φ into its positive and negative parts, one has to show that Proposition 6.6 holds whenever φ is positive. Under this additional assumption, the latter proposition is induced by the following result.

Proposition 6.7. *Let $C \subset D$ be any convex open sets included in U such that $\bar{C} \subset D$. Then*

$$\limsup_{g \rightarrow f} \int_{\mathcal{Z}_g(C)} \varphi(t) H_1(dt) \leq \int_{\mathcal{Z}_f(D)} \varphi(t) H_1(dt)$$

and

$$\int_{\mathcal{Z}_f(C)} \varphi(t) H_1(dt) \leq \liminf_{g \rightarrow f} \int_{\mathcal{Z}_g(D)} \varphi(t) H_1(dt).$$

Before proving this result, we will show how Proposition 6.7 implies the continuity of Λ at f .

Proof of Proposition 6.6. Let Ω_n be a sequence of increasing convex sets such that $\bar{\Omega}_n \subset \Omega_{n+1}$ and

$$\bigcup_{n \geq 1} \Omega_n = \Omega.$$

From Proposition 6.7, one has

$$\limsup_{n \rightarrow +\infty} \int_{\mathcal{Z}_f(\Omega_n)} \varphi(t) H_1(dt) \leq \liminf_{g \rightarrow f} \int_{\mathcal{Z}_g(\Omega)} \varphi(t) H_1(dt).$$

The left hand part converges, in virtue of the dominated convergence theorem, to

$$\int_{\mathcal{Z}_f(\Omega)} \varphi(t) H_1(dt).$$

Hence,

$$\int_{\mathcal{Z}_f(\Omega)} \varphi(t) H_1(dt) \leq \liminf_{g \rightarrow f} \int_{\mathcal{Z}_g(\Omega)} \varphi(t) H_1(dt)$$

Note that this inequality is always true and is independent of condition (6.1). The latter condition will be needed for the converse inequality.

This time, we let Ω^n be a sequence of non-increasing convex open sets such that $\bar{\Omega}^{n+1} \subset \Omega^n$ and

$$\bigcap_{n \geq 1} \Omega^n = \bar{\Omega}.$$

Following the preceding reasoning, one has

$$\begin{aligned} \limsup_{g \rightarrow f} \int_{\mathcal{Z}_g(\Omega)} \varphi(t) H_1(dt) &\leq \liminf_{n \rightarrow +\infty} \int_{\mathcal{Z}(\Omega^n)} \varphi(t) H_1(dt) \\ &= \int_{\mathcal{Z}(\bar{\Omega})} \varphi(t) H_1(dt) \\ \text{(condition 6.1)} &= \int_{\mathcal{Z}(\Omega)} \varphi(t) H_1(dt). \end{aligned}$$

The two preceding inequality immediately implies that

$$\lim_{g \rightarrow f} \Gamma(g) = \Gamma(f)$$

which finishes the proof. \square

The latter part of this section is thus devoted to the proof of Proposition 6.7. The ideas involved into proving this result are summarised hereafter.

First, we establish for $g \in \mathcal{C}^1(\bar{\Omega})$ sufficiently close to f a one-to-one local correspondence between the connected components of $f^{-1}(0)$ and the ones of $g^{-1}(0)$. What is meant with local correspondence is the existence of a covering of $f^{-1}(0)$ by open sets V_i such that $\mathcal{Z}_g(V_i)$ consists in exactly one connected component γ_g . If we reformulate the latter statement, this means that the set V_i is crossed by exactly one nodal line of g . The rest of the proof consists into estimating the difference between γ_g and γ_f and showing that the latter is negligible as g tends to f . To translate this local result into a global one, a partition of unity argument is later invoked.

Technical lemmas. The first step of the proof consists of proving this local correspondence. We will need the following lemma, which establishes that the zeros of f and g cannot be too far away as long as g is sufficiently close to f .

First, we define some notations. As f is regular, there exists an open set $W \subset \Omega$ containing $\mathcal{Z}(\bar{C})$ such that

$$\inf_{t \in W} |\nabla f(t)| \geq m$$

with $m > 0$.

On the other hand, the uniform continuity on \bar{U} ensures that

$$|\nabla f(t) - \nabla f(s)| \leq \frac{m}{4}, \quad |s - t| \leq \delta.$$

for some $\delta > 0$.

Lemma 6.8. *Let $\varepsilon < \delta m/8$ and $g \in \mathcal{C}^1(\bar{U})$ such that $\|f - g\|_1 \leq \varepsilon$. If $f(t) = 0$ and $|\partial_1 f(t)| \geq m/2$, then there exists $|\eta| \leq 8\varepsilon/m$ such that*

$$g(t_1 + \eta, t_2) = 0.$$

In particular, if $f(t) = 0$, either $|\partial_1 f(t)|$ or $|\partial_2 f(t)|$ is no less than $m/2$ and the preceding lemma can be applied.

Corollary 6.9. *There is a neighbourhood V_f of $f \in \mathcal{C}^1(\bar{U})$ such that any $t \in \mathcal{Z}_f(\bar{C})$ can be associated with some $s \in \mathcal{Z}_g(D)$ such that $|s - t| \leq 8\varepsilon/m$.*

Proof. Without loss of generality, assume that $\partial_1 f(t) > 0$ and consider the function

$$F(\omega) := f(t_1 + \omega, t_2).$$

whose derivatives is given by

$$F'(\omega) = \partial_1 f(t_1 + \omega, t_2)$$

For $|\omega| \leq \delta$, the triangle inequality ensures that

$$F'(\omega) \geq \frac{m}{4},$$

so that $f(t_1 + 8\varepsilon/m, t_2)$ and $f(t_1 - 8\varepsilon/m, t_2)$ have opposite sign and their respective absolute value are at least 2ε . Thus, $g(t_1 + 8\varepsilon/m, t_2)$ and $g(t_1 - 8\varepsilon/m, t_2)$ have opposite sign and the conclusion follows from the Intermediate Value Theorem. \square

Lemma 6.10. *There is a finite open cover V_i of $\mathcal{Z}_f(\overline{C})$ and a neighbourhood V_f of f in $\mathcal{C}^1(\overline{U})$ such that for any $g \in V_f$, the following assertions are satisfied*

- (1) $\mathcal{Z}_g(V_i)$ has exactly one connected component.
- (2)

$$\mathcal{Z}_g(C) \subset \bigcup_i V_i \subset D.$$

In particular,

$$\mathcal{Z}_g(C) = \bigcup_i \mathcal{Z}_g(V_i) \subset \mathcal{Z}_g(D).$$

Proof. Let d_g be the minimal distance between two different connected components of $\mathcal{Z}_g(\overline{D})$. If there is no or only one connected component, d_g is arbitrarily set to 1.

As \overline{D} is compact, two different connected different components of f are always at a non zero distance and this bound holds uniformly in the sense there is a neighbourhood V_f of f such that

$$d := \inf_{g \in V_f} d_g > 0.$$

The proof of this result is omitted here but it requires the use of a quantitative version of the implicit theorem. Such statement yields an explicit neighbourhood, whose size depends only the first derivatives of the function, on which the conclusion of the implicit theorem holds. In particular, in the latter neighbourhood, the zero set of the related function is composed of exactly one connected component. Applying this result to a function g in a small neighbourhood of f , one sees that the size of the neighbourhood can be chosen so that it only depends on f . In particular, this prevents $\mathcal{Z}_g(\overline{D})$ from having two close connected components.

To construct the sets V_i , we choose a finite cover of $\mathcal{Z}_f(\overline{C})$ with open balls whose diameter is less than d and such that $\mathcal{Z}_f(V_i)$ is non empty. By definition, $\mathcal{Z}_g(V_i)$ has at most one connected component for $g \in V_f$. On the other hand, Corollary 6.9 ensures there is at least one. This finishes the proof. \square

In virtue of Corollary 6.10, the unique connected component of $\mathcal{Z}_g(V_i)$ can be parametrised by $(\gamma_{g,i}(t), t)$ or $(t, \gamma_{g,i}(t))$ depending on whether $|\partial_1 f| \geq m/2$ or $|\partial_2 f| \geq m/2$ on V_i . Note that it is always possible to choose V_i so that at least one of these two assumptions hold. Without loss of generality, we will only consider parametrisation of the second form.

Lemma 6.11. *As $g \rightarrow f$, $\gamma_{i,g}$ converges to $\gamma_{i,f}$ for the topology of uniform convergence of the first derivatives on compact sets.*

Proof. The uniform convergence of $\gamma_{i,g}$ to $\gamma_{i,f}$ is a straightforward consequence of Corollary 6.9 which bounds the distance between the cancellation points of f and g . As for the convergence of derivatives, it follows from the formula

$$\gamma'_{i,g}(t) = -\frac{\partial_1 g(t, \gamma_{i,g}(t))}{\partial_2 g(t, \gamma_{i,g}(t))} \rightarrow -\frac{\partial_1 f(t, \gamma_{i,f}(t))}{\partial_2 f(t, \gamma_{i,f}(t))} = \gamma'_{i,f}(t)$$

□

With Lemma 6.11 proved, all the necessary tools for the proof of Proposition 6.7 has been introduced.

Proof of Proposition 6.7. Consider a partition of unity associated with the open sets V_i , that is a family of continuous functions $0 \leq \phi_i \leq 1$ with compact support included in V_i satisfying

$$\sum_{i=1}^n \phi_i(t) \leq 1, \quad t \in D,$$

with equality when $t \in C$.

Let f_n be a sequence a sequence of functions converging to f in $C^1(\bar{U})$. The positivity of φ implies that

$$\begin{aligned} \int_{\mathcal{Z}_{f_n}(C)} \varphi(t) H_1(dt) &= \sum_i \int_{\mathcal{Z}_{f_n}(V_i \cap C)} \phi_i(t) \varphi(t) H_1(dt) \\ &\leq \sum_i \int_{\mathcal{Z}_{f_n}(V_i)} \phi_i(t) \varphi(t) H_1(dt) \\ &= \sum_i \int \phi(t, \gamma_{i,f_n}(t)) \varphi(t, \gamma_{i,f_n}(t)) \sqrt{1 + \gamma'_{i,f_n}(t, \gamma_{i,f_n}(t))^2} dt. \end{aligned}$$

Lemma 6.11 implies that the latter converges to

$$\begin{aligned} &\int \phi(t, \gamma_{i,f}(t)) \varphi(t, \gamma_{i,f}(t)) \sqrt{1 + \gamma'_{i,f}(t, \gamma_{i,f}(t))^2} dt \\ &= \sum_i \int_{\mathcal{Z}_f(V_i)} \phi_i(t) \varphi(t) H_1(dt) \\ &\leq \sum_i \int_{\mathcal{Z}_f(D)} \phi_i(t) \varphi(t) H_1(dt) \\ &\leq \int_{\mathcal{Z}_f(D)} \varphi(t) H_1(dt) \end{aligned}$$

Hence,

$$\limsup_{g \rightarrow f} \int_{\mathcal{Z}_g(C)} \varphi(t) H_1(dt) \leq \int_{\mathcal{Z}_f(D)} \varphi(t) H_1(dt),$$

which ultimately concludes the proof of Proposition 6.7. □

REFERENCES

- [AT07] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer, 2007.
- [AW09] Jean-Marc Azais and Mario Wschebor. *Level sets and extrema of random processes and fields*. 2009.
- [Ber02] M.V. Berry. Statistics of nodal lines and points in chaotic quantum billiards: Perimeter corrections, fluctuations, curvature. *Journal of Physics A: Mathematical and General*, 35, 2002.
- [BMW20] J. Benatar, D. Marinucci, and I. Wigman. Planck-scale distribution of nodal length of arithmetic random waves. *Journal d'Analyse Mathématique*, 141, 2020.
- [DNPR23] G. Dierickx, I. Nourdin, G. Peccati, and M. Rossi. Small scale clts for the nodal length of monochromatic waves. *Communications in Mathematical Physics*, 397, 2023.
- [EH99] P. Erdős and R. Hall. On the angular distribution of gaussian integers with fixed norm. *Discrete Mathematics*, 200, 1999.
- [Hua38] L.K. Hua. On Waring's problem. *The Quarterly Journal of Mathematics*, os-9, 01 1938.
- [Kal02] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [KKW13] M. Krishnapur, P. Kurlberg, and I. Wigman. Nodal length fluctuations for arithmetic random waves. *Annals of Mathematics*, 177, 2013.
- [LR22] R. Lachièze-Rey. Variance linearity for real gaussian zeros. *Annales de l'Institut Henri Poincaré*, 58, 2022.
- [ORW08] F. Oravecz, Z. Rudnick, and I. Wigman. The lera measure of nodal sets for random eigenfunctions on the torus. *Annales de l'Institut Fourier*, 151, 2008.
- [Sar22] A. Sartori. Spectral almost correlations and phase-transitions for the nodal length of arithmetic random waves. *International Mathematics Research Notices*, 2022, 2022.
- [Sau] M. Sauter. Density of polynomials in $c^k(\bar{\omega})$. MathOverflow.
- [Sze75] G. Szego. *Orthogonal Polynomials*. Colloquium Publications. American Mathematical Society, fourth edition, 1975.
- [Tod20] A. P. Todino. Nodal lengths in shrinking domains for random eigenfunctions on s^2 . *Bernoulli*, 26, 2020.
- [Whi34] H. Whitney. Functions differentiable on the boundaries of regions. *Annals of Mathematics*, 35, 1934.