



HAL
open science

On the geometry of random nodal sets

Louis Gass

► **To cite this version:**

Louis Gass. On the geometry of random nodal sets. Probability [math.PR]. Université de Rennes, 2022. English. NNT : 2022REN1S045 . tel-03906805

HAL Id: tel-03906805

<https://theses.hal.science/tel-03906805>

Submitted on 19 Dec 2022

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.

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Mathématiques et leurs interactions*

Par

Louis GASS

Sur la géométrie des ensembles nodaux aléatoires

Thèse présentée et soutenue à Rennes, le 6 Juillet 2022
Unité de recherche : IRMAR (UMR CNRS 6625)

Rapporteurs avant soutenance :

Domenico MARINUCCI Professeur, Université de Tor-Vergatta, Rome, Italie
Oanh NGUYEN Professeure, Brown University, Providence, USA

Composition du Jury :

Attention, en cas d'absence d'un des membres du Jury le jour de la soutenance, la composition du jury doit être revue pour s'assurer qu'elle est conforme et devra être répercutée sur la couverture de thèse

Examineurs :	François BOLLEY	Professeur, ENS Rennes, Rennes, France
	Thomas LETENDRE	Maître de conférences, Université Paris-Saclay, Orsay, France
	Giovanni PECCATI	Professeur, Université du Luxembourg, Luxembourg
	Maurizia ROSSI	Professeure, Université de Milan-Bicocca, Italie
Dir. de thèse :	Jürgen ANGST	Maître de conférences, HDR, Université de Rennes 1, France
Co-dir. de thèse :	Guillaume POLY	Maître de conférences, Université de Rennes 1, France

REMERCIEMENTS

L'achèvement de cette thèse clôt mes sept années de vie rennaise. C'est la fin d'un chapitre de ma vie qui m'aura profondément marqué, et je tiens du fond du cœur à remercier toutes les personnes qui ont contribué à l'aboutissement de ce papier, ainsi qu'à faire de moi la personne que je suis aujourd'hui.

Je remercie en premier lieu mes deux directeurs de thèse, Jürgen Angst et Guillaume Poly. Merci à vous deux pour la bienveillance infinie dont vous avez fait preuve avec moi durant ces trois dernières années. Je ne pouvais honnêtement pas rêver de meilleurs encadrants que vous deux. Cette thèse n'aurait pas eu la même saveur sans votre énergie, votre humour, nos discussions de montagnes enneigées et de paysages corses. Je suis extrêmement reconnaissant du petit bout de chemin que l'on a parcouru ensemble et je suis heureux de pouvoir vous considérer désormais comme mes collègues et mes amis.

Je remercie Domenico Marinucci et Oanh Nguyen pour avoir accepté de rapporter ma thèse. Vos relectures et vos remarques attentives ont permis d'améliorer le contenu de ce manuscrit. Je remercie l'ensemble des membres du jury pour leur participation : François Bolley, Thomas Letendre, Giovanni Peccati et Maurizia Rossi. Giovanni, je suis très heureux de poursuivre ma carrière mathématique à tes côtés durant les deux prochaines années, découvrir un nouveau laboratoire et un nouveau pays. Notre collaboration sera, je n'en doute pas, fructueuse sur tous les plans. Thomas, merci d'avoir été présent ces trois dernières années, humainement et mathématiquement. J'apprécie énormément les discussions avec toi et tes conseils avisés m'ont permis de prendre en maturité sur le monde de la recherche. Excuse-moi pour toutes les coquilles que tu as corrigées, j'espère pouvoir te rendre la pareille dans un futur proche !

Je remercie l'ensemble du personnel de l'IRMAR pour leur dévotion à créer une ambiance chaleureuse au sein du bâtiment, ainsi que l'ensemble des professeur.e.s de ma scolarité. Je remercie tous particulièrement quelques un.e.s de mes professeur.e.s de classe prépa : Marc Baron, Laurent Boré Annie Deléglise, Jean-Louis Lhuilier et Thierry Maillet. Vous avez contribué à développer mon amour pour les sciences et sans vous, jamais cette thèse n'aurait vu le jour.

Les maths, c'est aussi de jolies rencontres qui se transforment en amitiés au fil des années. Il vient un temps où chacun prend son propre chemin, mais les liens créés resteront gravés à jamais. Merci à Antoine, Camille, Fabrice, Grégoire, Juliette, Leo, Loulou, Lucile, Meriadec, Ronan et tous les autres camarades, avec qui j'ai partagé tant de moments, que ce soit au RU, à l'escalade ou autour d'un verre. Merci à mes trois cobureaux Emeline, Marc et Marie, avec qui les après-midis étaient tellement moins ennuyantes. Merci à Thibault, mon grand frère de thèse, partenaire de tennis hors pair et cuisinier de génie, qui est, au moment où j'écris ces lignes, entrain de préparer un marbré au chocolat juste à côté.

Je pense aussi à ces amitiés qui ont partagé mon quotidien durant ces dernières années. Merci à Armand et nos voyages à vélo, les sprints pancartes vachés au prolongateurs et les pâtés aux prunes. Merci à Antoine et Quentin, la coloc de l'ambiance. Je me suis vu grandir et m'épanouir à vos côtés, et trop souvent encore je ressens la mélancolie de nos moments passés tous les trois sous le même toit. Enfin, merci à Mégane, je n'ai pas les mots pour t'exprimer à quel point je suis heureux de toutes les épreuves que l'on a traversées ensemble durant ces deux dernières années.

Mens sana in corpore sano, jamais cette thèse n'aurait pu aboutir sans les innombrables séances défouloires à la salle d'escalade, et j'y ai rencontré parmi les plus belles personnes que je connaisse. Je remercie l'ensemble de la Team Chantal : Cabernet, Changlotte, Chanqui, Elen, Gadget, Léa, Mawa, le Moine, Marie, Pablito, Papy, Pierrot, et tous les autres pour les body trainings, les weekends chills et les vacances de folies. Ça va me manquer de ne plus vous voir lorsque j'irai grimper. Merci à vous deux, Marvin et Mika, pour les séances grimpe *a muerte* et les road-trips en van. Je n'imagine pas une vie sans continuer à vous revoir dès que possible, tellement ce que l'on vit ensemble est indescriptible. Enfin, merci à toi Norah, reine des crustacés. Je ne pourrais jamais te remercier assez d'être là et de m'accepter tel que je suis.

Parce qu'il seront toujours inconditionnellement là peu importe les épreuves, je remercie profondément toute ma famille. Même si l'on se voit peu, il n'y a pas un jour sans que je pense à vous. Merci à mes grands parents et mes petits adelpes que j'aime de tout mon cœur. Merci à vous deux Papa et Maman pour l'amour que vous me portez chaque jour.

Il reste encore tant de personnes à qui je voudrais exprimer ma gratitude et qui mériteraient bien plus qu'un simple paragraphe ici. Merci à vous tou.te.s, qui avez partagé un petit ou gros bout de ma vie.

Prologue	7
I Synthesis of the contributions	21
1 The nodal measure of random waves	22
1.1 Geometric and probabilistic settings	23
1.2 Literature on random waves and their nodal sets	34
1.3 Overview of the contributions	43
2 The zeros counting measure of real Gaussian processes	49
2.1 The combinatorics of cumulants	50
2.2 Moments and cumulants of the number of zeros	60
2.3 Literature on the zeros counting measure	68
2.4 Overview of the contributions	76
II Details of the contributions	82
3 Asymptotics for the nodal measure of Riemannian random waves	83
3.1 Introduction and main results	84
3.2 Salem–Zygmund CLT for Riemannian random waves	93
3.3 Almost sure asymptotics of nodal volume	101
3.4 Appendix	118
4 Variance for the number of zeros	132
4.1 Introduction	133
4.2 Kac–Rice formula for the second moment	136
4.3 Proof of the main theorems	143
5 Cumulants for the zeros counting measure	153
5.1 Introduction	154
5.2 Basics and notations	167
5.3 Kac–Rice formula for Gaussian processes	186
5.4 Asymptotics of the cumulants of the zeros counting measure	205

PROLOGUE

Dans cette thèse, nous nous intéressons aux ensembles nodaux aléatoires, c'est-à-dire aux lieux d'annulation de fonctions à valeurs réelles, dépendantes également d'un paramètre aléatoire. Notre principal modèle d'intérêt est celui des ondes aléatoires sur des variétés riemanniennes. L'étude de ce modèle, ainsi que de son ensemble nodal, est originellement motivé par de célèbres conjectures physiques et mathématiques, comme la conjecture de Berry. En dimension un, cette étude se réduit à celle des zéros des polynômes trigonométriques aléatoires. Elle s'inscrit dans la théorie plus générale des zéros de processus stochastiques unidimensionnels, elle aussi riche en applications : télécommunication, traitement du signal, etc.

Dans ce prologue, nous décrivons brièvement les problématiques auxquelles nous nous sommes intéressés, ainsi que les différentes contributions apportées à ce domaine durant cette thèse. L'architecture du manuscrit se décompose ensuite comme suit. La première partie résume l'ensemble des contributions apportées durant la thèse. Elle-même se scinde en deux chapitres qui correspondent à l'étude de l'ensemble nodal des ondes aléatoires riemanniennes, et à l'ensemble nodal de processus gaussiens réels. Ces deux chapitres s'articulent de manière similaires. Nous y présentons d'abord le contexte dans lequel nous nous plaçons ainsi que quelques outils mathématiques qui serviront à notre étude. Nous ferons ensuite un état de l'art de la littérature, ce qui nous permettra dans un dernier temps de détailler les différentes contributions apportées à chaque modèle durant cette thèse et de les mettre en relief par rapport au résultats préexistants.

Le modèle des ondes aléatoires riemanniennes. Nous commençons par décrire succinctement et de manière informelle le modèle des ondes aléatoires riemanniennes. Dans ce qui suit, \mathcal{M} désigne une variété riemannienne compacte sans bord de dimension d , et nous notons Δ son laplacien. Nous désignons par $(-\lambda_n^2)_{n \in \mathbb{N}}$ la suite ordonnée des valeurs propres du laplacien comptées avec multiplicité, ainsi que $(\varphi_n)_{n \in \mathbb{N}}$ une base de vecteurs propres unitaires associés, de sorte que pour $n \in \mathbb{N}$,

$$\Delta \varphi_n = -\lambda_n^2 \varphi_n.$$

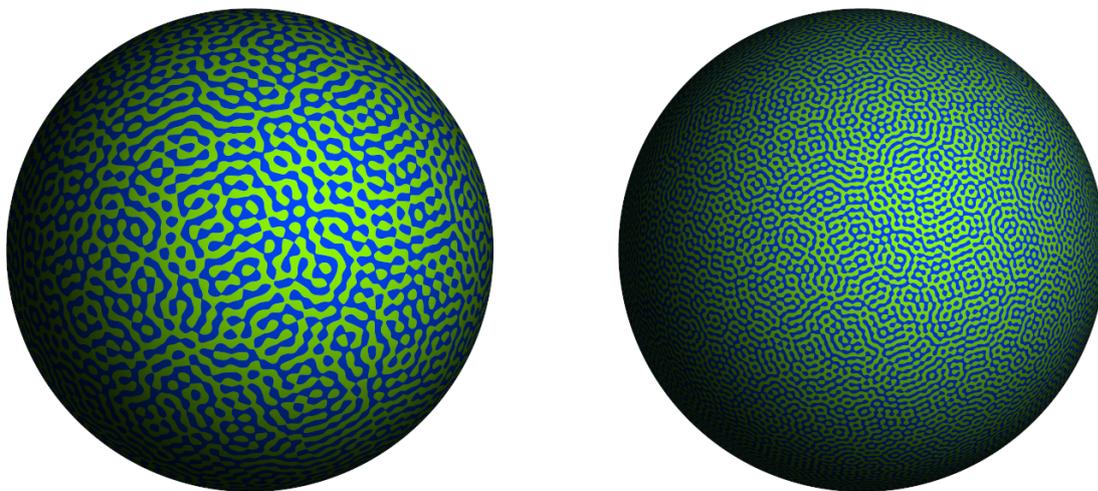


FIGURE 1 – Parties positives et négatives d’une onde aléatoire sphérique sur la sphère pour deux niveaux distincts d’énergies. La frontière entre les deux domaines correspond à l’ensemble nodal. Crédit : V. Beffara.

Soit $(a_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles gaussiennes indépendantes et identiquement distribuées (i.i.d.) définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Le modèle des ondes aléatoires riemanniennes est défini comme le processus stochastique sur \mathcal{M} suivant, pour $\lambda \in \mathbb{R}_+$ et $x \in \mathcal{M}$

$$f_\lambda(x) := \sum_{\lambda_n \leq \lambda} a_n \varphi_n(x).$$

L’ensemble nodal du processus f_λ est défini comme l’hypersurface aléatoire

$$Z_\lambda = \{x \in \mathcal{M} \mid f_\lambda(x) = 0\}.$$

Le modèle des ondes aléatoires riemanniennes et de son ensemble nodal est aujourd’hui un domaine de recherche mathématique florissant, qui s’inscrit dans la lignée plus générale des modèles d’ondes aléatoires. Leurs études sont originellement motivées par de nombreuses applications en physique et en mathématiques ainsi que par de célèbres conjectures. Citons par exemple une conjecture due au physicien M. Berry, qui suggère de manière informelle que les modes propres d’un billard chaotique se comportent, dans la limite de haute énergie, comme des ondes planes aléatoires.

Il est naturel de s’intéresser à diverses quantités topologiques ou métriques reliées à l’ensemble Z_λ , telles que son volume ou son nombre de composantes connexes. Une grande partie des résultats de la littérature concerne l’asymptotique de ces observables nodales en

moyenne. Par exemple P. Berard a montré dans [Bér85] que l'asymptotique en moyenne du volume $d - 1$ dimensionnel de l'ensemble des zéros, noté $\text{Vol}(Z_\lambda)$, est donnée par

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathbb{E}[\text{Vol}(Z_\lambda)]}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}. \quad (1)$$

Dans cette thèse nous nous sommes en premier lieu intéressés aux comportements presque sûrs de l'asymptotique de la mesure nodale. Nous montrons en particulier que l'asymptotique (1) peut être renforcée en une asymptotique presque sûre :

$$\lim_{\lambda \rightarrow +\infty} \frac{\text{Vol}(Z_\lambda)}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}. \quad (2)$$

Les résultats de convergences presque sûres sont généralement obtenus à partir de résultats de concentration et un argument de type Borel–Cantelli. Malheureusement, la question des fluctuations de la mesure nodale est délicate dans le cadre d'une variété riemannienne générale. Il n'existe actuellement pas de résultat suffisamment fin sur la variance qui permette de déduire des résultats de convergence presque sûrs concernant les ondes aléatoires riemanniennes et leurs mesures nodales. Il est donc nécessaire d'employer une nouvelle approche afin de prouver des résultats de convergence presque-sûre pour la mesure nodale. Notons que des résultats précis de fluctuations existent néanmoins pour des modèles d'ondes aléatoires particuliers, par exemple sur la sphère [Wig09 ; CMR18 ; MRW20], ou le tore [KKW13 ; Mar+16], où les calculs peuvent être menés explicitement.

L'approche Salem–Zygmund. Afin de prouver l'asymptotique presque sûre (2), il s'agit dans un premier temps de comprendre l'asymptotique du processus f_λ à la limite de haute énergie, c'est-à-dire lorsque le paramètre λ croît vers l'infini. Une réinterprétation probabiliste des travaux de Hörmander [Hör68] montrent que localement, et sous une renormalisation adéquate, le processus gaussien f_λ converge en loi (dans un bon espace fonctionnel) vers un processus limite universel f_∞ qui ne dépend que de la dimension ambiante. Cette approche, couplée avec la formule de Kac–Rice, permet de montrer à peu de frais la convergence en moyenne de la mesure nodale empirique vers la mesure riemannienne, et donc l'asymptotique (1).

Dans [SZ54], R. Salem et A. Zygmund montrent le résultat suivant. Soit $(a_n)_{n \geq 0}$ une suite de variables aléatoires gaussiennes indépendantes et X une variable aléatoire

uniforme sur $[0, 2\pi]$, indépendante de la suite de variables $(a_n)_{n \geq 0}$. Alors pour presque toutes réalisations des coefficients $(a_n)_{n \geq 0}$, la convergence en loi suivante vers une gaussienne est vérifiée

$$\sqrt{\frac{2}{n}} \sum_{k=0}^n a_k \cos(kX) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}_X} \mathcal{N}(0, 1).$$

Cette approche a été utilisée avec succès afin de montrer la convergence presque sûre de la mesure de comptage des zéros pour différents modèles de polynômes trigonométriques aléatoires dans la série d'articles [AP21 ; APP21]. Nous nous intéresserons à la généralisation de cette approche en dimension supérieure, où les polynômes trigonométriques seront remplacés par des combinaisons linéaires de fonctions propres du laplacien. Cette approche, comme nous le verrons au Chapitre 3, permet d'obtenir la convergence presque sûre de la mesure nodale en contournant la difficulté liée aux estimations de la fluctuation de la mesure nodale. Elle repose sur une technique de « dérandomisation » utilisée dans le cadre du tore par J. Bourgain [Bou14 ; BW16].

Le nombre de zéros d'un processus unidimensionnel. Sur le cercle, le modèle des ondes aléatoires riemanniennes coïncide avec celui des polynômes trigonométriques aléatoires et s'inscrit donc dans l'étude des zéros de processus stochastiques unidimensionnels, initiée par les travaux de M. Kac et S. O. Rice dans [Kac43 ; Ric45]. La formule de Kac (resp. de Kac–Rice) donne une expression explicite du nombre de zéros (resp. des moments du nombre de zéros) d'un processus stochastique suffisamment régulier.

Soit $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ des suites de variables aléatoires centrées réduites, et définissons le polynôme trigonométrique aléatoire pour $x \in \mathbb{R}$ par

$$f_n(x) = \frac{1}{\sqrt{n}} \sum_{k=0}^n a_k \cos(kx) + b_k \sin(kx).$$

Notons $Z_{f_n}[a, b]$ le nombre de zéros du processus f_n sur l'intervalle $[a, b]$. Une application directe de la formule de Kac–Rice montre que, dans le cas où $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ sont des variables gaussiennes i.i.d., le nombre moyen de zéros sur un intervalle $[a, b]$ satisfait l'asymptotique suivante

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[a, b]]}{n} = \frac{|b - a|}{\pi\sqrt{3}}.$$

Les travaux de Cuzick [Cuz76] ont permis d'obtenir l'asymptotique de la variance ainsi qu'un théorème central limite (TCL) pour le nombre de zéros d'un processus gaussien

stationnaire. Plus précisément, si f est un processus gaussien stationnaire, alors sous de bonnes hypothèses, il existe une constante $\sigma(f)$ dépendant du processus f telle que

$$\lim_{R \rightarrow +\infty} \frac{\text{Var}(Z_f[0, R])}{R} = \sigma(f).$$

Il a néanmoins fallu attendre les travaux récents de A. Granville et I. Wigman [GW11] afin de connaître l'asymptotique exacte de la variance et un TCL pour le nombre de zéros du modèle des polynômes trigonométriques aléatoires gaussiens i.i.d.. L'asymptotique de la variance est donnée par

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty),$$

où cette fois-ci f_∞ correspond à un processus stationnaire réel dont la fonction covariance est donnée par la fonction sinus cardinal (ou processus de Paley–Wiener). La preuve repose encore sur la formule de Kac–Rice pour la variance du nombre de zéros. Une autre approche, désormais standard, repose sur la décomposition en chaos de Wiener de la fonctionnelle donnant le nombre de zéros du processus f_n . Elle a permis de montrer des résultats similaires pour le nombre de zéros. On consultera à ce titre la série d'articles [AL13; ADL16].

La question de l'universalité. Dans notre cadre d'étude, *l'universalité* désigne la robustesse de l'asymptotique d'une quantité (par exemple le nombre moyen ou la variance du nombre de zéros d'un processus) lorsqu'on affaiblit certaines hypothèses du modèle. Dans [Fla17] est mis en évidence le phénomène d'universalité suivant. Pour toutes suites $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$ de variables i.i.d. centrées réduites (non nécessairement gaussiennes), le nombre de zéros sur un intervalle $[a, b]$ du polynôme trigonométrique f_n , défini plus haut, satisfait l'asymptotique suivante

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[a, b]]}{n} = \frac{|b - a|}{\pi\sqrt{3}}.$$

Autrement dit, le nombre moyen de zéros présente un caractère universel, dans le sens où l'asymptotique de cette quantité ne dépend pas de la loi sous-jacente des suites i.i.d. $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$. A contrario, la variance du nombre de zéros présente un caractère non-universel. En effet, les travaux de [BCP19; DNN20] montrent que pour toutes suites $(a_n)_{n \geq 0}$ et

$(b_n)_{n \geq 0}$ de variables i.i.d. centrées réduites, ayant un moment d'ordre suffisamment élevé,

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty) + \frac{2}{15}(\mathbb{E}[a_0^4] - 3).$$

Ainsi, la variance du nombre de zéros dépend asymptotiquement de la loi sous-jacente à travers son kurtosis, c'est-à-dire la quantité $\mathbb{E}[a_0^4] - 3$.

Dans une autre direction, nous pouvons étudier le caractère universel du nombre de zéros lorsqu'on relaxe l'hypothèse d'indépendance des suites de variables $(a_n)_{n \geq 0}$ et $(b_n)_{n \geq 0}$. Cette fois, on suppose que ces deux suites sont des processus stationnaires gaussiens indépendants, dont la corrélation est définie via une *mesure spectrale* μ . Pour k, l des entiers naturels,

$$\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \int_{-\pi}^{\pi} e^{i(k-l)x} d\mu(x).$$

Les auteurs dans [APP21] montrent que, dès lors que la mesure spectrale μ admet une densité *densité spectrale* ψ par rapport à la mesure de Lebesgue λ sur le tore qui est non-nulle et continue, le nombre de zéros moyen vérifie l'asymptotique

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[0, 2\pi]]}{n} = \frac{\lambda(\{\psi \neq 0\})}{\pi\sqrt{3}} + \frac{\lambda(\{\psi = 0\})}{\pi\sqrt{2}}.$$

En ce sens, le nombre moyen de zéros des polynômes trigonométriques aléatoires à coefficients dépendants est non-universel, l'asymptotique dépend de la mesure de Lebesgue du lieu d'annulation de la densité spectrale.

Dans le cas où la densité spectrale ψ est strictement positive, le dernier terme du membre de droite dans la formule ci-dessus s'annule, et on retrouve la même asymptotique que dans le cas de coefficients indépendants. Il est naturel de se demander si dans ce cadre, la variance du nombre de zéros suit la même asymptotique que dans le cas de coefficient gaussiens indépendants. Le résultat de non-universalité de la variance discutée en début de paragraphe nous indique que le caractère universel de la variance, dans ce contexte, n'est pas acquis. Sous l'hypothèse que la mesure spectrale admet une densité ψ par rapport à la mesure de Lebesgue qui soit continue et strictement positive, nous montrons au Chapitre 4 que

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty).$$

Autrement dit, de façon surprenante l'asymptotique de la variance est effectivement la même que dans le cas indépendant.

Les moments d'ordre supérieur La question des moments d'ordre supérieur du nombre de zéros d'un processus gaussien est plus délicate. Des conditions suffisantes assurant la finitude des moments ont été énoncées dans [Cuz75], mais leurs asymptotiques sont restées pendant longtemps des questions ouvertes. Récemment, les travaux de M. Ancona et T. Letendre [AL21a; AL21b] ont permis de montrer que pour un processus gaussien, dont la fonction de covariance est dans la classe de Schwartz des fonctions régulières à décroissance rapide, et tout entier $p \geq 1$,

$$\lim_{R \rightarrow +\infty} \mathbb{E} \left[\left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right)^p \right] = \mathbb{E}[N^p], \quad (3)$$

où N est une variable gaussienne centrée réduite. Cette approche donne donc une autre preuve du théorème central limite par la méthode des moments, dans ce cadre restreint de décorrélation rapide. Malheureusement, le processus de Paley–Wiener et le modèle des polynômes trigonométriques aléatoires échappent largement à ce cadre. Dans cette thèse, nous nous sommes donc naturellement intéressés à l'asymptotique des moments d'ordre supérieur pour les moments des zéros de polynômes trigonométriques à coefficients gaussiens, qui échappaient jusqu'à présent à la littérature. Nous montrons au Chapitre 5 que le p -ème cumulants du nombre de zéros d'un polynôme trigonométrique aléatoire satisfait l'asymptotique suivante

$$\lim_{n \rightarrow +\infty} \frac{\kappa_p(Z_{f_n}[a, b])}{n} = \gamma_p,$$

où γ_p est une constante explicite. Le lien entre moments et cumulants permet de retrouver l'asymptotique des moments centrés (3) et donc un TCL. L'approche par les moments a l'avantage, comparé à la décomposition en chaos, de prouver la concentration polynomiale à tout ordre, ainsi que la convergence presque sûre du nombre de zéros vers sa moyenne renormalisée. Notre approche est suffisamment générale pour permettre de prouver l'asymptotique des moments du nombre de zéros d'une large classe de processus gaussiens. Citons par exemple le modèle des polynômes trigonométriques aléatoires avec coefficients dépendants, ou encore celui des polynômes orthogonaux aléatoires.

PROLOGUE

In this thesis, we are interested in the nodal sets of random functions, that is in the vanishing locus of real functions, also depending on a random parameter. Our main model of interest is the so-called Riemannian random waves model. It is initially motivated by celebrated conjectures in both physics and mathematics, such as the Berry conjecture. In dimension one, this study reduces to the study of the zeros set of random trigonometric polynomials. It is a part of the more general theory of zeros of one dimensional stochastic processes, also rich in applications: telecommunication, signal processing, etc.

In this prologue, we briefly describes our questions of interest, and the different contributions to the domain obtained during this thesis. The manuscript is structured as follow. The first part of this thesis gives a general overview of our contributions to the domain. It is itself split into two chapters, which corresponds to the study of Riemannian random waves, and to the study of the nodal set associated to real Gaussian processes. These two chapters are organized similarly. We first present our framework and a few useful tools for our study. Then, we make a quick survey of the literature, which allows us, at last, to present the different contributions to each model during this thesis, and to compare them to existing results.

The Riemannian random waves model. We first briefly describes the Riemannian random waves model. In what follows, \mathcal{M} is a smooth compact Riemannian manifold of dimension d , without boundary. We denote the Laplace operator on \mathcal{M} by Δ . Let $(-\lambda_n^2)_{n \in \mathbb{N}}$ be the sorted sequence of Laplace eigenvalues counted with multiplicity, and $(\varphi_n)_{n \in \mathbb{N}}$ a basis of unitary eigenfunctions, so that for $n \in \mathbb{N}$,

$$\Delta \varphi_n = -\lambda_n^2 \varphi_n.$$

Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of real Gaussian random variables, independent and identically distributed (i.i.d.), defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The Riemannian random

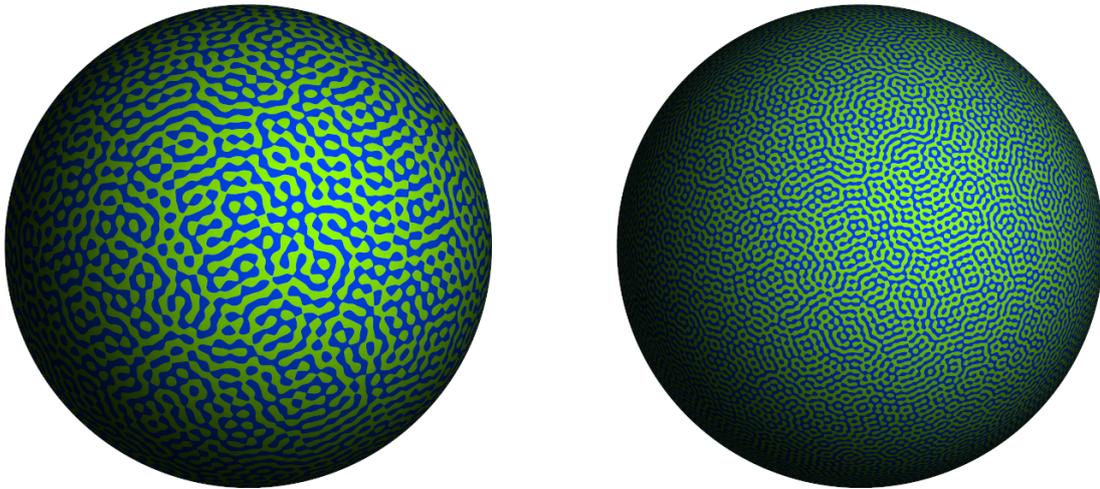


Figure 2 – Positive and negative parts of a spherical random wave for two distinct level of energies. The boundary between the two domains corresponds to the nodal set. Credit : V. Beffara.

waves model is defined as the following stochastic process on \mathcal{M} , for $\lambda \in \mathbb{R}_+$ and $x \in \mathcal{M}$

$$f_\lambda(x) := \sum_{\lambda_n \leq \lambda} a_n \varphi_n(x).$$

The nodal set of the process f_λ is defined as the random hypersurface

$$Z_\lambda = \{x \in \mathcal{M} \mid f_\lambda(x) = 0\}.$$

The Riemannian random waves model, and its associated nodal set, is a flourishing research domain in mathematics. It is part of the more general framework of random waves models. Their studies are originally motivated by numerous applications in physics and mathematics, as well as celebrated conjectures. Let us cite for instance a conjecture stated by the physicist M. Berry, which informally suggests that the eigenmodes of a chaotic billiard asymptotically behaves, in the high limit energy, as random plane waves.

One can study several topological quantities related to the nodal set Z_λ , such as its volume or its number of connected components. A large part of the results in the literature concerns the mean asymptotics of this quantities. For instance, P. Berard proved in [Bér85] that the mean asymptotics of the $d - 1$ dimensional volume of the zeros set, denoted

$\text{Vol}(Z_\lambda)$, is given by

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathbb{E}[\text{Vol}(Z_\lambda)]}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}. \quad (4)$$

During this thesis, we were primarily interested in the almost sure behavior of the nodal measure in the high energy limit. We in particular shows that the above convergence (4) can be reinforced in an almost sure convergence :

$$\lim_{\lambda \rightarrow +\infty} \frac{\text{Vol}(Z_\lambda)}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}. \quad (5)$$

Almost sure results are usually recovered from concentration estimates and a Borel–Cantelli like argument. Unfortunately, studying the fluctuations of the nodal measure is a delicate task. Currently, there isn’t any sufficiently fine results on the variance, which would imply almost-sure results concerning the Riemannian random waves model and its associated nodal set. One must then take a new approach in order to prove almost sure results for the nodal measure. Note that if not available in general, precise results about the fluctuations of the nodal measure for particular models of random waves, for instance on the sphere [Wig09; CMR18; MRW20], or on the torus [KKW13; Mar+16], where one can make explicit computations.

The Salem–Zygmund approach. A first step towards the almost sure asymptotics (5) is to explore the asymptotics of the process f_λ in the high energy limit, that is, when the parameter λ grows to infinity. A probabilistic reinterpretation of the works of Hörmander [Hör68] shows that locally, with a proper rescaling, the Gaussian process f_λ converges in distribution (in a good function space) towards a process f_∞ . This last process is universal in the sense that it depends only on the ambient dimension. This approach, combined with Kac–Rice formula, allows us to show with little effort the convergence in expectation of the empirical nodal measure towards the Riemannian measure, and thus the asymptotics (4).

In [SZ54], R. Salem and A. Zygmund shows the following result. Let $(a_n)_{n \geq 0}$ be a sequence of independent random Gaussian variables, and X be a uniform random variable on $[0, 2\pi]$, independent of the sequence $(a_n)_{n \geq 0}$. Then, for almost all realization of the coefficients $(a_n)_{n \geq 0}$, the following convergence in distribution holds

$$\sqrt{\frac{2}{n}} \sum_{k=0}^n a_k \cos(kX) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}_X} \mathcal{N}(0, 1).$$

This approach has been successfully extended in order to show the almost sure convergence of the zeros counting measure for different models of random trigonometric polynomials, in the series of article [AP21; APP21]. We are interested in the generalization of this approach to the higher dimension, where trigonometric polynomials are naturally replaced by linear combinations of Laplace eigenfunctions. It allows us, as we will see in Chapter 3, to obtain the almost sure convergence of the nodal measure. This method bypasses variances estimates for the nodal measure, and is rather related to a “derandomization” technique, already used in the torus case by J. Bourgain [Bou14; BW16].

The number of zeros of a one dimensional real process. On the circle, the model of Riemannian random waves coincides with the model of random trigonometric polynomials and is thus part of the study of the zeros of one-dimensional stochastic processes, initiated by the work of M. Kac and S. O. Rice in [Kac43; Ric45]. The formula of Kac (resp. Kac–Rice) gives an explicit expression for the number of zeros (resp. moments of the number of zeros) of a sufficiently regular stochastic process.

Let $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$ be sequences of centered random variables with unit variance, and let us define the random trigonometric polynomial for $x \in \mathbb{R}$ by

$$f_n(x) = \frac{1}{\sqrt{n}} \sum_{k=0}^n a_k \cos(kx) + b_k \sin(kx).$$

Let $Z_{f_n}[a, b]$ be the number of zeros of the process f_n on the interval $[a, b]$. A direct application of the Kac–Rice formula shows that, in the case where $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$ are i.i.d. Gaussian variables, the average number of zeros over an interval $[a, b]$ satisfies the following asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[a, b]]}{n} = \frac{|b - a|}{\pi\sqrt{3}}.$$

The work of Cuzick in [Cuz76] has provided the asymptotics of the variance and a central limit theorem (CLT) for the number of zeros of a stationary Gaussian process. More precisely, if f is a stationary Gaussian process, then under suitable assumptions, there exists a constant $\sigma(f)$ depending on the process f such that

$$\lim_{R \rightarrow +\infty} \frac{\text{Var}(Z_f[0, R])}{R} = \sigma(f).$$

Nevertheless, we had to wait for the recent work of A. Granville and I. Wigman [GW11]

in order to recover the exact asymptotics of the variance and a CLT for the number of zeros of the model of i.i.d. Gaussian random trigonometric polynomials. The variance asymptotics is given by

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty),$$

where this time f_∞ corresponds to a real stationary process whose covariance function is given by the cardinal sine function (the so-called Paley–Wiener process). The proof again relies on the Kac–Rice formula for the variance of the number of zeros. Another approach, which has now become standard, is based on the Wiener chaos decomposition of the functional giving the number of zeros of the process f_n . It allows to show similar results for the number of zeros. One can consult the series of articles [AL13; ADL16].

The question of universality. In our framework, the term *universality* refers to the robustness of the asymptotics of a quantity (e.g. the mean number or the variance of the number of zeros of a process) when one relaxes some assumptions on the model. In [Fla17] the following universality phenomenon is proved. For all sequences $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$ of i.i.d. variables (not necessarily Gaussian), the number of zeros on an interval $[a, b]$ of the trigonometric polynomial f_n , defined above, satisfies the following asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[a, b]]}{n} = \frac{|b - a|}{\pi\sqrt{3}}.$$

In other words, the average number of zeros has a universal behavior, in the sense that the asymptotics of this quantity does not depend on the underlying distribution of the sequences $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$. On the other hand, the variance of the number of zeros has a non-universal behavior. Indeed, the work of [BCP19; DNN20] shows that, for all sequences $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$ of centered i.i.d. variables, having a sufficiently high moment,

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty) + \frac{2}{15}(\mathbb{E}[a_0^4] - 3).$$

Thus, the variance of the number of zeros depends asymptotically on the underlying law through its kurtosis, i.e. the quantity $\mathbb{E}[a_0^4] - 3$.

In another direction, we can study the universal behavior of the number of zeros when we relax the independence hypothesis on the sequences of variables $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$. This time, we assume that these two sequences are independent Gaussian stationary

processes, whose correlation is defined via a spectral measure μ . For k, l natural numbers,

$$\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \int_{-\pi}^{\pi} e^{i(k-l)x} d\mu(x).$$

The authors in [APP21] show that, as soon as the spectral measure μ admits a non-zero absolutely continuous part with spectral density ψ with respect to the Lebesgue measure λ on the torus, the number of average zeros satisfies the asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[0, 2\pi]]}{n} = \frac{\lambda(\{\psi \neq 0\})}{\pi\sqrt{3}} + \frac{\lambda(\{\psi = 0\})}{\pi\sqrt{2}}.$$

In this sense, the average number of zeros of random trigonometric polynomials with dependent coefficients is non-universal. The asymptotics this time depends on the Lebesgue measure of the vanishing locus of the spectral density.

In the case where the spectral density ψ is positive, the last term on the right hand side in the above formula cancels, and we recover the same asymptotics as in the independent case. It is natural to ask whether, in this framework, the variance of the number of zeros follows the same asymptotics as in the case of independent Gaussian coefficients. The result of non-universality of the variance discussed at the beginning of this paragraph tells us that the universality of the variance, in this context, is not acquired. Under the assumption that the spectral measure admits a density ψ with respect to the Lebesgue measure, which is continuous and positive, we show in Chapter 4 that

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty).$$

In other words, and quite surprisingly, the asymptotics of the variance is in fact the same as in the independent case.

Moment of higher order The question of higher order moments of the number of zeros of a Gaussian process is more challenging. Sufficient conditions ensuring the finiteness of higher moments have been stated in [Cuz75], but their asymptotics have remained for a long time open questions. Recently, the work of M. Ancona and T. Letendre [AL21a; AL21b] have shown that for a Gaussian process, whose covariance function is in the

Schwartz class of smooth and rapidly decreasing functions, and any integer $p \geq 1$,

$$\lim_{R \rightarrow +\infty} \mathbb{E} \left[\left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right)^p \right] = \mathbb{E}[N^p], \quad (6)$$

where N is a standard Gaussian variable. This approach thus gives another proof of the central limit theorem by the method of moments, in this restricted framework of fast decorrelation. Unfortunately, the Paley–Wiener process and the random trigonometric polynomial model largely escape this framework. In this thesis, we are therefore naturally interested in the asymptotics of higher order moments for the moments of the zeros of trigonometric polynomials with Gaussian coefficients, which escaped the literature until now. We show in Chapter 5 that the p -th cumulant of the number of zeros of a random trigonometric polynomial satisfies the following asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\kappa_p(Z_{f_n}[a, b])}{n} = \gamma_p,$$

where γ_p is an explicit constant. The link between moments and cumulants allows to recover the asymptotics of centered moments (6) and thus a CLT. The moments-based approach has the advantage, compared to the Wiener chaos decomposition, of proving the polynomial concentration at any order, as well as the almost sure convergence of the number of zeros to its renormalized expectation. Our approach is sufficiently general to allow to prove the asymptotics of the moments of the number of zeros of a large class of Gaussian processes. As representative examples, one can cite the model of random trigonometric polynomials with dependent coefficients, or the model of random orthogonal polynomials.

PART I

Synthesis of the contributions

THE NODAL MEASURE OF RANDOM WAVES

In this chapter, we summarize the different contributions of this thesis to the study of the nodal volume associated with Riemannian random waves. These contributions will be detailed in Chapter 3, which consists in the first published article [Gas21a]. We first define the geometric and probabilistic frameworks we place ourselves into, and we present a few tools in order to study the nodal sets of random waves. We then make a quick survey of the literature on random waves and their associated nodal sets. At last, we discuss our contributions to the field.

1.1 Geometric and probabilistic settings	23
1.1.1 The local Weyl law on a Riemannian manifold	23
1.1.2 The probabilistic model of random waves	27
1.1.3 Co-area formula and Kac–Rice formula for random fields	29
1.2 Literature on random waves and their nodal sets	34
1.2.1 The remainder in the local Weyl law	35
1.2.2 About some conjectures	36
1.2.3 Asymptotics related to the nodal set of random waves	38
1.3 Overview of the contributions	43
1.3.1 The Salem–Zygmund approach	43
1.3.2 Almost sure convergence of the nodal volume	47

1.1 Geometric and probabilistic settings

In this section, we detail the model of random waves on a general Riemannian manifold, informally defined in introduction. At first, we discuss the Riemannian geometry framework and in particular the local Weyl law, which gives the asymptotics of the spectral projector on the first eigenspaces in the high energy limit.

In the following (\mathcal{M}, g) is a smooth compact Riemannian manifold of dimension d , without boundary. The underlying Riemannian distance is denoted by dist , and the Riemannian measure is denoted by μ . The manifold is naturally equipped with the Laplace–Beltrami operator Δ , which generalizes the Laplace operator defined on Euclidean spaces. The second order differential operator Δ is self-adjoint and has compact resolvent. Standard results in spectral theory ensures the existence of an L^2 -orthonormal basis $(\varphi_n)_{n \in \mathbb{N}}$ of eigenfunctions of Δ associated with the ordered eigenvalues with multiplicities $(-\lambda_n^2)_{n \in \mathbb{N}}$, so that for $n \geq 0$,

$$\Delta \varphi_n = -\lambda_n^2 \varphi_n.$$

Such an orthonormal basis is not unique due to sign changes and possible multiplicity of the eigenvalues, but we choose one for the rest of the chapter. For all $n \in \mathbb{N}$, we can assume that λ_n is non-negative and $\|\varphi_n\|_2 = 1$.

1.1.1 The local Weyl law on a Riemannian manifold

A fundamental tool in spectral analysis is the *local Weyl law*, first proved by Hörmander in [Hör68]. It describes the precise asymptotics of the spectral projector on the eigenspaces associated with energy levels up to λ , in the high energy limit. Given $x, y \in \mathcal{M}$ and $\lambda \in \mathbb{R}_+$, we define

$$K_\lambda(x, y) = \sum_{\lambda_n \leq \lambda} \varphi_n(x) \varphi_n(y) \quad \text{and} \quad K_\lambda(x) := \sum_{\lambda_n \leq \lambda} \varphi_n^2(x),$$

the *two-points spectral kernel projector* on the eigenspace generated by the eigenfunctions up to order λ , also known as the *large band kernel*. Integrating the function $x \mapsto K_\lambda(x)$ on \mathcal{M} we obtain the *eigenvalues counting function*

$$K(\lambda) := \text{Card} \{n \in \mathbb{N} \mid \lambda_n \leq \lambda\} = \int_{\mathcal{M}} K_\lambda(x) d\mu(x). \quad (1.1)$$

Let $B_d(0, 1)$ be the unit ball in \mathbb{R}^d and σ_d its volume, explicitly given by

$$\text{Vol}(B_d(0, 1)) = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2} + 1\right)}.$$

We introduce the function

$$\begin{aligned} \mathcal{B}_d : \mathbb{R} &\longrightarrow \mathbb{R} \\ r &\longrightarrow \frac{1}{\sigma_d} \int_{B_d(0,1)} e^{ir\xi_1} d\xi. \end{aligned}$$

The function \mathcal{B}_d can be seen as the Fourier transform of the indicator of the unit ball, which is invariant by rotation and hence depends only on the radius of its argument. It is related to the Bessel function of the first kind $\mathcal{J}_{\frac{d}{2}}$ by the formula

$$\mathcal{B}_d(r) = \frac{1}{\sigma_d} \left(\frac{2\pi}{r}\right)^{\frac{d}{2}} \mathcal{J}_{\frac{d}{2}}(r).$$

The following theorem, known as the local Weyl law, see [Hör68], describes the asymptotics of the two-points spectral kernel projector in the high energy limit.

Theorem 1.1.1 (Local Weyl law). *Uniformly on $x, y \in \mathcal{M}$, as λ goes to infinity,*

$$K_\lambda(x, y) = \frac{\sigma_d \text{Vol}(\mathcal{M})}{(2\pi)^d} \lambda^d \mathcal{B}_d(\lambda \text{dist}(x, y)) + O(\lambda^{d-1}).$$

For an arbitrary number of derivatives in x and y , one has

$$\partial_{\alpha,\beta} K_\lambda(x, y) = \frac{\sigma_d \text{Vol}(\mathcal{M})}{(2\pi)^d} \lambda^d \partial_{\alpha,\beta} [\mathcal{B}_d(\lambda \text{dist}(x, y))] + O(\lambda^{d+\alpha+\beta-1}).$$

This result is remarkable in the sense that the first order asymptotics of the spectral projector K_λ depends on the manifold \mathcal{M} only through its dimension d and its volume $\text{Vol}(\mathcal{M})$. By the relation (1.1), we obtain the following asymptotics for the eigenvalues counting function $K(\lambda)$

$$K(\lambda) = \frac{\sigma_d \text{Vol}(\mathcal{M})}{(2\pi)^d} \lambda^d + O(\lambda^{d-1}).$$

It implies following first-order asymptotics for the n -th eigenvalue as n grows to infinity, known as the classical Weyl law, see [Wey11].

Theorem 1.1.2 (Classical Weyl law). *One has the equivalent as n grows to infinity*

$$\lambda_n \simeq 2\pi \left(\frac{n}{\sigma_d \text{Vol}(\mathcal{M})} \right)^{1/d}.$$

Up to now we have considered the large band kernel K_λ associated to the energy windows $[0, \lambda]$. Let $\tau(\lambda)$ be a positive and non-decreasing function, such that $\tau(\lambda) = O(\sqrt{\lambda})$. One can also consider the *short band kernel* associated with energy windows $]\lambda - \tau(\lambda), \lambda]$, defined by

$$k_\lambda(x, y) = \sum_{\lambda_n \in]\lambda - \tau(\lambda), \lambda]} \varphi_n(x) \varphi_n(y) \quad \text{and} \quad k(\lambda) := \text{Card} \{n \in \mathbb{N} \mid \lambda - \tau(\lambda) \leq \lambda_n \leq \lambda\}.$$

This function is also called the *monochromatic kernel*, though some authors employ this denomination only for the case $\tau(\lambda) = 1$. We define the function

$$\mathcal{S}_d = \mathcal{B}_{d-2}.$$

The function \mathcal{S}_d can be seen as the Fourier transform of the spherical measure on the sphere \mathbb{S}^{d-1} embedded in \mathbb{R}^d . Then the local Weyl law directly implies the asymptotics

$$k_\lambda(x, y) = \frac{d\sigma_d}{(2\pi)^d} \lambda^{d-1} \tau(\lambda) \mathcal{S}_d(\lambda \text{ dist}(x, y)) + O(\lambda^{d-1}).$$

Particular cases : the torus and the sphere. The spectral decomposition $(\lambda_n, \varphi_n)_{n \geq 0}$ can be explicitly solved in a few specific manifolds, notably the d -dimensional torus and sphere, which we detail now.

We denote the scalar product in \mathbb{R}^d by $\langle \cdot, \cdot \rangle$ and the associated norm by $\|\cdot\|_2$. Let \mathbb{T}^d be the flat torus of dimension d , seen as the quotient $\mathbb{R}^d / (2\pi\mathbb{Z})^d$. For a vector $n \in \mathbb{Z}^d$ with integer coordinates, the functions

$$x \mapsto \cos\langle n, x \rangle \quad \text{and} \quad x \mapsto \sin\langle n, x \rangle$$

are solutions to the Laplace eigenvalue problem

$$\Delta\varphi = -\|n\|_2^2 \varphi.$$

By Fourier analysis, this collection of functions forms a Hilbert basis of $L^2(\mathbb{T}^d)$ and every solution to the Laplace problem on \mathbb{T}^d is a linear combination of these solutions. In particular, the dimension $k(\lambda)$ of the eigenspace $\ker(\Delta + \lambda^2 \text{Id})$ is exactly

$$k(\lambda) := \text{Card} \left\{ n \in \mathbb{Z}^d \mid \|n\|_2 = \lambda \right\}.$$

It corresponds to the number of integer points on the sphere of radius λ . The classical Weyl law on the torus has the following interpretation. The number of integer points in a ball of radius λ is asymptotically equivalent, for large λ , to the volume of such a ball.

Let \mathbb{S}^d be the Euclidean sphere of dimension d in \mathbb{R}^{d+1} . A vector $x \in \mathbb{R}^d \setminus \{0\}$ can be written in polar coordinates $x = (r, \theta)$, where $r \in \mathbb{R}_+^*$ and $\theta \in \mathbb{S}^d$. The Laplace-Beltrami operators on \mathbb{R}^{d+1} and \mathbb{S}^d are related by the following formula in polar coordinates

$$\Delta_{\mathbb{R}^{d+1}} = \frac{\partial^2}{\partial r^2} + \frac{d}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{\mathbb{S}^d}.$$

Now, let H_n be a homogeneous polynomial of degree n . In polar coordinates, one can then writes

$$H_n(r, \theta) = r^n H_n(\theta).$$

Now assume that the homogeneous polynomial H_n is harmonic, that is $\Delta_{\mathbb{R}^{d+1}} H_n = 0$. Applying the Laplace operator on H_n one gets

$$\Delta_{\mathbb{S}^d} H_n = -n(n + d - 1) H_n.$$

A homogeneous harmonic polynomial of degree n is thus an eigenfunction of the Laplace operator on \mathbb{S}^d , associated to the eigenvalue $-n(n + d - 1)$. One can show by Stone–Weierstrass Theorem that the space of homogeneous harmonic polynomials restricted to the sphere \mathbb{S}^d forms a dense subspace of $L^2(\mathbb{S}^d)$.

Choose $\lambda = \sqrt{n(n + d - 1)}$. The dimension $k(\lambda)$ of the eigenspace generated by the solutions to the eigenvalue problem $\Delta\varphi = -\lambda^2\varphi$ on \mathbb{S}^d is then exactly the dimension of

the space of homogeneous harmonic polynomial of degree n . One can show that

$$k(\lambda) = \frac{2n + d - 1}{n} \binom{n + d - 2}{n - 1}.$$

One can write an explicit basis of eigenfunctions by introducing the family of Gegenbauer polynomials $(P_{n,d})_{n \geq 0}$, which reduces to the family Legendre polynomials when $d = 2$. Then the two-points spectral kernel projector on the eigenspace associated to the eigenvalue $\lambda = \sqrt{n(n + d - 1)}$ is given by

$$k_\lambda : (x, y) \mapsto P_{n,d}(\cos(\text{dist}(x, y))).$$

For a complete exposition of the spherical harmonics, we refer the reader to [Sze75; Mor98].

1.1.2 The probabilistic model of random waves

Let us now describe our two main probabilistic models, classically known as the Riemannian random waves and monochromatic random waves model. We use the framework and notations of the previous subsection. Let us consider $(a_n)_{n \geq 0}$ a sequence of independent and identically distributed standard Gaussian random variables on a probability space $(\Omega_a, \mathcal{F}_a, \mathbb{P}_a)$. We will denote by \mathbb{E}_a the associated expectation. Recall from the previous section that $K(\lambda)$ (resp. $k(\lambda)$) is the number of Laplace eigenvalues lower than λ (resp. in some energy windows $]\lambda - \tau(\lambda), \lambda]$). The model of Riemannian random waves is defined as the following Gaussian combination of eigenfunctions

$$f_\lambda : x \mapsto \frac{1}{\sqrt{K(\lambda)}} \sum_{\lambda_n \leq \lambda} a_n \varphi_n(x).$$

Similarly, the model of monochromatic random wave is defined as

$$\tilde{f}_\lambda : x \mapsto \frac{1}{\sqrt{k(\lambda)}} \sum_{\lambda_n \in]\lambda - \tau(\lambda), \lambda]} a_n \varphi_n(x).$$

These two models give a probabilistic interpretation of the two-points projector kernels introduced in the previous section. Indeed, for all $x, y \in \mathcal{M}$, one has

$$\mathbb{E}_a[f_\lambda(x)f_\lambda(y)] = \frac{K_\lambda(x, y)}{K(\lambda)} \quad \text{and} \quad \mathbb{E}_a[\tilde{f}_\lambda(x)\tilde{f}_\lambda(y)] = \frac{k_\lambda(x, y)}{k(\lambda)}.$$

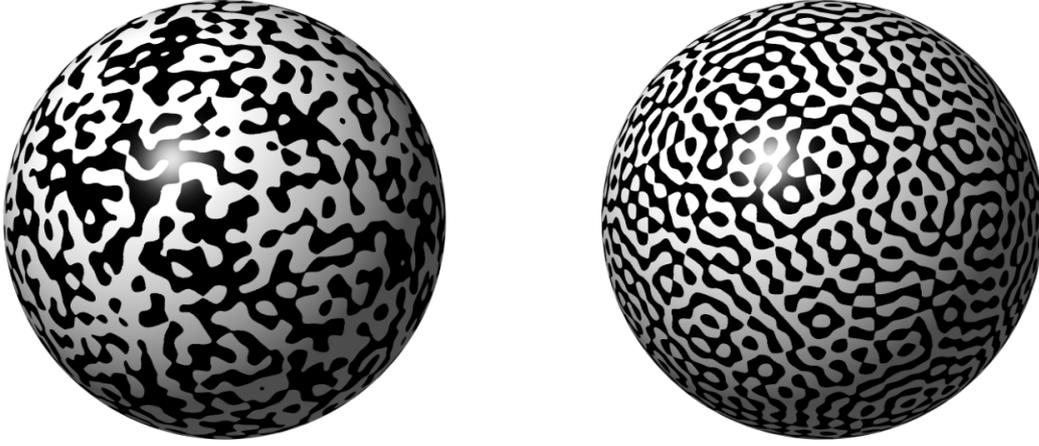


Figure 1.1 – Positive and negative part of Riemannian (on the left) and monochromatic (on the right) random waves on the sphere. Credit : A. Barnett.

As indicated by the local Weyl law of Theorem 1.1.1, a non trivial scaling limit for this process is attained at scale $1/\lambda$. The naive way to proceed is to consider at a point x a flattened version of the process f_λ (resp. \tilde{f}_λ) on the tangent space $T_x\mathcal{M}$, given by the pullback by the Riemannian exponential. It will be convenient to define the flattened version of our stochastic process around each point x , on a fixed Euclidean space that does not depend on x . To this end, we choose

$$I_x : \mathbb{R}^d \longrightarrow T_x\mathcal{M},$$

an isometry between the canonical Euclidean space \mathbb{R}^d and the tangent space at x . We only require the mapping $x \mapsto I_x$ to be measurable. Regarding the torus \mathbb{T}^d , one can choose I_x to be the canonical isometry given by the projection. However, there is no canonical choice (nor even a continuous choice) of a family $(I_x)_{x \in \mathcal{M}}$ in a general manifold. All the following mentioned results are independent of this choice of isometry. Denoting the Riemannian exponential based at $x \in \mathcal{M}$ by \exp_x , we define

$$\Phi_x := \exp_x \circ I_x.$$

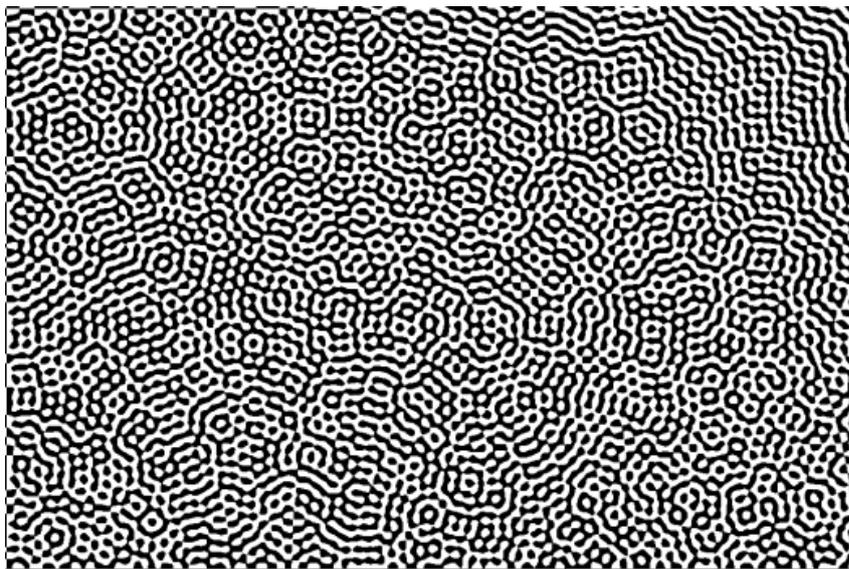


Figure 1.2 – Positive and negative part of a planar random wave. Notice the resemblance with monochromatic random waves on the sphere in Figure 1.1. Credit : E. Bogomolny and C. Schmit.

This map allows us to define a rescaled and flattened version of f_λ and \tilde{f}_λ (or any function on \mathcal{M}) around some point $x \in \mathcal{M}$ by setting

$$\begin{aligned}
 g_\lambda^x : \mathbb{R}^d &\longrightarrow \mathbb{R} & \tilde{g}_\lambda^x : \mathbb{R}^d &\longrightarrow \mathbb{R} \\
 v &\longrightarrow f_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right] & v &\longrightarrow \tilde{f}_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right].
 \end{aligned}$$

We also define the isotropic Gaussian processes on \mathbb{R}^d g_∞ and \tilde{g}_∞ with covariance functions \mathcal{B}_d and \mathcal{S}_d respectively, introduced in the previous section. A probabilistic reinterpretation of the local Weyl law then shows that these two processes converge in distribution (at a fixed point x) towards the Gaussian processes g_∞ and \tilde{g}_∞ , respectively. In particular the limiting processes only depend on the topological dimension d and are independent of the base manifold \mathcal{M} . The process \tilde{g}_∞ is sometimes called the *random plane wave process*.

1.1.3 Co-area formula and Kac–Rice formula for random fields

In this section, we detail a few tools for studying the zeros set of Riemannian random waves. The following material is standard and can be found in [Fed69; AW09]. We start from the classical co-area formula and we decline it for stochastic processes, leading to the celebrated Kac–Rice formula for random fields. We recall first the definition of Hausdorff

measures. Let (X, ρ) be a metric space, and U a subset of X . We denote the *diameter* of U by $\text{diam}(U)$, defined by

$$\text{diam}(U) := \sup_{x, y \in U} \rho(x, y),$$

The *d-dimensional Hausdorff measure* of a subset S of X is the well-defined quantity

$$\mathcal{H}^d(S) := \sup_{\delta > 0} \inf \left\{ \sum_{i=1}^{+\infty} \text{diam}(U_i)^d \mid S \subset \bigcup_{i=1}^{+\infty} U_i \quad \text{and} \quad \text{diam}(U_i) < \delta \right\}.$$

The definition of the *d-dimensional Hausdorff measure* coincides with the Lebesgue measure for subsets of a Riemannian manifold of dimension d .

Co-area formula. In the following, \mathcal{M} is a Riemannian manifold of dimension d , and f is a regular function from \mathcal{M} to $\mathbb{R}^{d'}$, with $d \geq d'$. Sard Theorem implies that for almost all $u \in \mathbb{R}^{d'}$, the level set $f^{-1}(u)$ is a submanifold of \mathcal{M} of dimension $d - d'$. Then the $d - d'$ dimensional Hausdorff measure of almost all level sets of f is locally finite. From a measure-theoretic point of view, this statement is expressed with the co-area formula. Let $\nabla f(x)$ be the gradient of f at a point $x \in \mathcal{M}$. It is a linear application from the tangent space at x to $\mathbb{R}^{d'}$. We then define

$$\|\nabla f(x)\| := \sqrt{\det[(\nabla_x f)^T (\nabla_x f)]}.$$

Theorem 1.1.3 (Co-area formula). *Let $f : \mathcal{M} \rightarrow \mathbb{R}^{d'}$ be a function of class \mathcal{C}^1 on \mathcal{M} and $g : \mathcal{M} \rightarrow \mathbb{R}$ a non-negative measurable function. Then*

$$\int_{\mathcal{M}} g(x) \|\nabla f(x)\| dx = \int_{\mathbb{R}^{d'}} \left(\int_{f^{-1}(u)} g(y) d\mathcal{H}^{d-d'}(y) \right) du.$$

In particular, if A is a measurable subset of \mathcal{M} and φ is a non-negative measurable function from $\mathbb{R}^{d'}$ to \mathbb{R} then

$$\int_A \varphi(f(x)) \|\nabla f(x)\| dx = \int_{\mathbb{R}^{d'}} \varphi(u) \mathcal{H}^{d-d'}(f^{-1}(u) \cap A) du.$$

A rigorous proof of this theorem as well as some generalizations can be found in [Fed69],

but we sketch it here for completeness. Note that the second statement follows from the first one by taking $g = (\varphi \circ f)\mathbb{1}_A$.

Proof. We first assume that $\mathcal{M} = \mathbb{R}^d$, and $f(x) = Ax$ for some surjective matrix A of size $d \times d'$. The matrix $A^T A$ is then invertible, and for $x \in \mathbb{R}$,

$$\|\nabla f(x)\| = \sqrt{\det A^T A} \quad \text{and} \quad f^{-1}(x) = \ker(A) + {}^T A(A^T A)^{-1}x.$$

Then we compute, by change of variable and basic linear algebra

$$\begin{aligned} \int_{\mathbb{R}^{d'}} \left(\int_{f^{-1}(u)} g(y) d\mathcal{H}^{d-1}(y) \right) du &= \int_{\mathbb{R}^{d'}} \left(\int_{\ker(A)} g(y + {}^T A(A^T A)^{-1}u) dy \right) du \\ &= \int_{\text{im}({}^T A)} \left(\int_{\ker(A)} \sqrt{\det A^T A} g(y + v) dy \right) dv \\ &= \int_{\mathbb{R}^d} g(x) \sqrt{\det A^T A} dx. \end{aligned}$$

Thus, the co-area formula holds true for surjective linear applications. For a general function $f : \mathcal{M} \rightarrow \mathbb{R}^{d'}$ of class \mathcal{C}^1 , we use the fact that locally, f is well-approximated by its gradient. The conclusion roughly follows by the previous linear case and a standard approximation with a partition of unity on \mathcal{M} . \square

Kac formula. The Kac formula gives an explicit formula for the nodal volume of f , that is the quantity $\mathcal{H}^{d-d'}(f^{-1}(0))$. It consists in taking in the co-area formula of Theorem 1.1.3, an approximation of the Dirac at 0. We say that 0 is a regular value for a function f if the Jacobian of f at each point of the nodal set of f is surjective. The result is contained in the following theorem.

Theorem 1.1.4 (Kac formula). *The mapping*

$$\begin{aligned} \mathcal{C}^1(\mathcal{M}, \mathbb{R}^{d'}) &\longrightarrow \mathbb{R} \\ f &\longrightarrow \mathcal{H}^{d-d'}(f^{-1}(0)) \end{aligned}$$

is continuous for the \mathcal{C}^1 topology on the set of functions such that 0 is a regular value. For such a function f ,

$$\mathcal{H}^{d-d'}(f^{-1}(0)) = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^{d'}} \int_{\mathcal{M}} \mathbb{1}_{] -\varepsilon, \varepsilon[}^{d'}(f(x)) \|\nabla f(x)\| dx.$$

A full proof of this lemma in the case where \mathcal{M} is a torus can be found in [APP18], but for the sake of completeness we sketch it here.

Proof. Let f be a function such that 0 is a regular value for f . Having 0 as regular value is an open condition, thus one can consider a family of functions $(f_t)_{t \in] -\varepsilon, \varepsilon[}$ such that 0 is a regular value, and for which the application $t \rightarrow f_t$ is continuous for the \mathcal{C}^1 topology of uniform convergence. Let G be the mapping

$$\begin{aligned} G :] -\varepsilon, \varepsilon[\times \mathcal{M} &\longrightarrow \mathbb{R}^{d'} \\ (t, x) &\longrightarrow f_t(x) \end{aligned}$$

By the implicit function Theorem, $G^{-1}(0)$ is locally the graph of a function H of class \mathcal{C}^1 defined on $] -\varepsilon, \varepsilon[\times \mathbb{R}^{d-d'}$. Up to taking ε small enough, one has, for $t \in] -\varepsilon, \varepsilon[$, the local equality

$$f_t^{-1}(0) = \text{graph } H(t, \cdot). \tag{1.2}$$

Now the $d - d'$ -dimensional Hausdorff measure of the graph of a \mathcal{C}^1 function defined on an open set U of $\mathbb{R}^{d-d'}$ is given by the following formula, generalizing the arc-length in dimension one

$$\mathcal{H}^{d-d'}(\text{graph } H(t, \cdot) \cap (U \times \mathbb{R}^{d'})) = \int_U \sqrt{1 + \|\nabla H(t, y)\|^2} dy.$$

The first statement of Theorem 1.1.4 then follows from (1.2) and the continuity of the

application $t \rightarrow \nabla H(t, \cdot)$. It follows that the function

$$u \mapsto \mathcal{H}^{d-1}(f^{-1}(u))$$

is continuous in a neighborhood of 0. The second statement then directly follows from the co-area formula, by taking $\varphi = \frac{1}{(2\varepsilon)^{d'}} \mathbb{1}_{]u-\varepsilon, u+\varepsilon[}^{d'}$ and letting ε go to zero. \square

Kac–Rice formula. Let f be a centered stochastic Gaussian process with \mathcal{C}^1 paths defined on \mathcal{M} and with values in \mathbb{R} , such that $\text{Var}(f(x)) > 0$ for all $x \in \mathcal{M}$. We say that the process f is isotropic when its covariance function depends only on the distance between its arguments. The Kac–Rice formula gives an explicit formula for the mean nodal volume of the process f , which reduces to a very simple expression when f is an isotropic Gaussian process.

Theorem 1.1.5 (Kac–Rice formula). *Let f be a stochastic Gaussian process with \mathcal{C}^2 paths defined on \mathcal{M} , such that for all $x \in \mathcal{M}$, the Gaussian vector $(f(x), f'(x))$ is non degenerate. Then for $u \in \mathbb{R}$ and A a Borel subset of \mathcal{M} ,*

$$\mathbb{E} \left[\mathcal{H}^{d-1}(f^{-1}(u) \cap A) \right] = \int_A \mathbb{E} \left[\|\nabla f(x)\| \mid f(x) = u \right] \frac{1}{\sqrt{2\pi \text{Var } f(x)}} e^{-\frac{u^2}{2 \text{Var}(f(x))}} dx.$$

If f is an isotropic Gaussian process, then

$$\mathbb{E} \left[\mathcal{H}^{d-1}(f^{-1}(u) \cap A) \right] = \text{Vol}(A) \mathbb{E} \left[\|\nabla f\| \right] \frac{1}{\sqrt{2\pi \text{Var } f}} \exp \left(-\frac{u^2}{2 \text{Var } f} \right).$$

The technical step of the proof consists in exchanging the expectation and the limit in the Kac formula, see Theorem 1.1.4. We will sketch the proof and omit this point. A rigorous proof can be found in [AW09].

Proof. The hypotheses on the Gaussian process f imply that the any real number u is almost surely a regular value for f . Then the Kac formula implies that

$$\mathcal{H}^{d-1}(f^{-1}(u)) = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^{d'}} \int_{\mathcal{M}} \mathbb{1}_{]u-\varepsilon, u+\varepsilon[}^{d'}(f(x)) \|\nabla f(x)\| dx.$$

We then pass to expectation and exchange the limit to get

$$\mathbb{E} \left[\mathcal{H}^{d-1}(f^{-1}(u) \cap A) \right] = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^{d'}} \int_{\mathcal{M}} \mathbb{E} \left[\mathbb{1}_{|u-\varepsilon, u+\varepsilon|^{d'}}(f(x)) \|\nabla f(x)\| \right] dx.$$

By Gaussian conditioning, one gets

$$\mathbb{E} \left[\mathcal{H}^{d-1}(f^{-1}(u)) \right] = \int_{\mathcal{M}} \mathbb{E} \left[\|\nabla f(x)\| \mid f(x) = u \right] \frac{1}{\sqrt{2\pi \operatorname{Var} f(x)}} e^{-\frac{u^2}{2\operatorname{Var} f(x)}} dx,$$

hence the first statement. If the process f is isotropic, then for all $x \in \mathcal{M}$, the random variables $f(x)$ and $\nabla f(x)$ are independent, and there exists constants λ_0 and λ_2 such that

$$\operatorname{Var} f = \lambda_0 \quad \text{and} \quad \operatorname{Cov}(\nabla f) = \lambda_2 \operatorname{Id}.$$

In that case, we deduce that

$$\begin{aligned} \mathbb{E} \left[\mathcal{H}^{d-1}(f^{-1}(u)) \right] &= \operatorname{Vol}(\mathcal{M}) \mathbb{E} \left[\|\nabla f\| \right] \frac{1}{\sqrt{2\pi \operatorname{Var} f}} e^{-\frac{u^2}{2\operatorname{Var} f}} \\ &= \operatorname{Vol}(\mathcal{M}) \sqrt{\frac{\lambda_2}{\pi \lambda_0}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \exp\left(-\frac{u}{2\lambda_0}\right). \end{aligned}$$

□

For instance, Kac–Rice formula applied to the isotropic Gaussian processes g_∞ and \tilde{g}_∞ defined in Section 1.1.2 on a ball B of unit volume leads to

$$\mathbb{E} \left[\mathcal{H}^{d-1}(g_\infty^{-1}(0) \cap B) \right] = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}, \quad (1.3)$$

and

$$\mathbb{E} \left[\mathcal{H}^{d-1}(\tilde{g}_\infty^{-1}(0) \cap B) \right] = \frac{1}{\sqrt{\pi d}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}.$$

1.2 Literature on random waves and their nodal sets

In this section, we make a brief survey of the literature on the model of random waves and its associated nodal set. Heuristically, the more accurate the estimates of the remainder in the local Weyl law are, the more precise are the results one can obtain in the study of

random waves. We first discuss possible improvements for the remainder in the local Weyl law stated in Theorem 1.1.3. We then make a quick digression on famous conjectures in the field of random waves, which were early motivations in the study of random waves and their associated nodal sets. At last, we discuss recent advanced in the study of random waves on general Riemannian manifolds and on the particular cases of the sphere and the torus.

1.2.1 The remainder in the local Weyl law

The study of the remainder in the local Weyl law is a very active field of research and has been investigated in many recent papers (see [Ivr16] for a nice survey). Notice first that the remainder in the Weyl law is sharp on the d -sphere, since the eigenvalues are concentrated on the points $(\sqrt{n(n+d-1)})_{n \geq 0}$. This last property is common to manifolds such that all their geodesics are closed (they are called Zoll manifolds, see [Zel97] and the references therein). In the spherical case, a precise remainder in the local Weyl law is given by Darboux asymptotic formula, see [Sze75, Thm. 8.21.8].

In the torus case, giving a better estimate on the classical Weyl law is equivalent to the Gauss circle problem, which consists in giving a sharp estimate of the number of integer points inside a ball of large radius. To this end, we define

$$R_d(\lambda) := \text{Card} \{n \in \mathbb{Z}^d \mid \lambda = \|n\|_2\} - \sigma_d \lambda^d.$$

The classical Weyl law asserts that in the high energy limit, $R_d(\lambda) = O(\lambda^{d-1})$. This remainder geometrically corresponds to the size of the “uncertainty zone” near the boundary of a ball of radius λ . Heuristically, integer points near the boundary of the ball tend to be equidistributed inside and outside the ball, hence reducing the order of the remainder. Estimates of the remainder rely on arithmetic properties of the sum of d squares and is well-understood in large dimensions, but much less in low dimensions. It has been proved for $d \geq 5$ that $R_d(\lambda) \simeq \lambda^{d-2}$. A similar result holds for $d = 4$, up to a logarithmic term. In dimensions $d = 2$ and $d = 3$, it has only been shown that $R_d(\lambda) = O(\lambda^c)$ for some non-optimal constant $c < d - 1$. The Gauss circle conjecture, still open to this day, states that for any $\varepsilon > 0$

$$R_2(\lambda) = o(\lambda^{\frac{1}{2}+\varepsilon}).$$

We refer to [Ivi+06] for a complete survey of the Gauss circle problem and recent develop-

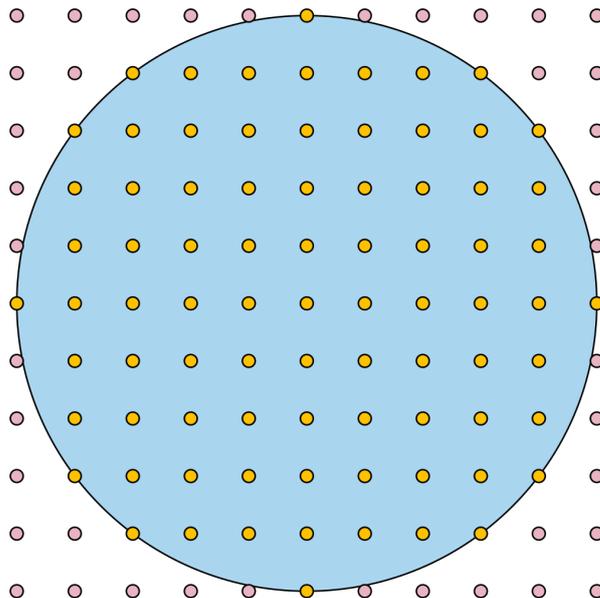


Figure 1.3 – Visualization of the Gauss circle problem. A circle of radius 5 contains 81 integers points. The area of the disk is approximately 78.54, which gives a relative error of 3%. Credit : D. Eppstein.

ment on this topic.

A slightly improved estimate of remainder in the local Weyl law, by a factor $\log(\lambda)$, has been proved in [Bér77; Kee19] for manifolds such that the set of closed geodesic has zero measure in the unit tangent bundle of \mathcal{M} . For such manifold, one has

$$K_\lambda(x, y) = \frac{\sigma_d \text{Vol}(\mathcal{M})}{(2\pi)^d} \lambda^d \mathcal{B}_d(\lambda \text{dist}(x, y)) + O\left(\frac{\lambda^{d-1}}{\log(\lambda)}\right).$$

Examples of such manifolds comprise for instance of manifolds with negative curvatures, such as quotient of hyperbolic spaces. In that case, one can take in the definition of the monochromatic random wave the energy windows $]\lambda, \lambda + 1]$, that is $\tau(\lambda) = 1$. The local Weyl law then holds with a remainder of logarithmic order. For “generic” manifolds, one should expect, see for instance [Non08], the remainder in the (local) Weyl law to be much smaller than the crude bound $O(\lambda^{d-1})$.

1.2.2 About some conjectures

In this section, we discuss some famous conjectures in the field of random waves and the nodal sets of eigenfunctions, that have motivated intense research in these fields. We

will mention the Berry conjecture on the chaotic behavior of an eigenfunction in the high energy limit, the Yau conjecture on the nodal volume of eigenfunctions, and the Bogomolny–Schmit conjecture on percolation for nodal set of random waves.

The Berry conjecture. One of the earlier motivation for the introduction of random waves is a conjecture, or more accurately a heuristic, given by M. Berry in his seminal paper [Ber77]. It states that for a generic manifold, an eigenfunction φ_n behaves like a random plane wave in the high energy limit. To formalize the chaotic behavior of a (deterministic) eigenfunction in the high energy limit, we introduce a uniform random variable X on \mathcal{M} . Then one should expect the following convergence in distribution

$$\varphi_n(X) \xrightarrow[n \rightarrow +\infty]{\Longrightarrow} \mathcal{N}(0, 1),$$

and more precisely, recalling the definition of the function Φ_x of previous section, one should expect the following convergence (in distribution) of stochastic processes

$$\varphi_n \left[\Phi_X \left(\frac{\cdot}{\lambda_n} \right) \right] \xrightarrow[n \rightarrow +\infty]{\Longrightarrow} \tilde{g}_\infty. \quad (1.4)$$

A precise formulation of Berry heuristics is still subject to discussion. We refer the reader to [Ing21] and the references therein for more insight on this topic.

The Yau conjecture. The study of nodal sets associated with Laplace eigenfunctions is the object of a vast literature in particular thanks to Yau’s conjecture. In his paper [Yau82], S. T. Yau made the following famous conjecture. Given a compact smooth manifold \mathcal{M} , there are positive constants c and C such that for all $n \geq 0$, the following bound holds for the nodal volume associated to the Laplace eigenfunction of energy λ_n

$$c \lambda_n \leq \text{Vol} \{ \varphi_n = 0 \} \leq C \lambda_n.$$

The Yau conjecture has been proved in the framework of real-analytic manifolds by H. Donnelly and C. Fefferman [DF88]. In the smooth setting, the best result so far were recently obtained by A. Logunov and E. Mallinikova, where they proved the lower bound in Yau’s conjecture [Log18b] and a polynomial upper bound for the nodal volume [Log18a]. We refer the reader to the nice survey [LM20] of these authors about old and recent developments on Yau’s conjecture.

The Bogomolny–Schmit conjecture. In their series of papers [BS02; BS07], E. Bogomolny and C. Schmit introduced a percolation-like model, see Figure 1.4. They heuristically argued that it should be close to the model of planar random waves (and more generally of monochromatic random waves on a general Riemannian manifolds). In particular they made several conjectures about the topology of nodal sets in connection with critical bond percolation. By this analogy, the number of nodal domains of a random plane wave on the ball of volume R , denoted $N(R)$, should satisfy the asymptotic behavior

$$\lim_{R \rightarrow +\infty} \frac{N(R)}{R^2} = \frac{3\sqrt{3} - 5}{4}, \quad (1.5)$$

where the right constant is computed from percolation theory (see [BS02] and the references therein). Some of these conjectures have been numerically checked in [BK13] and the references therein. It appears that the Bogomolny–Schmidt prediction (1.5) for the number of nodal domains is close to the real value, but slightly “off” according to experimental simulations. The study of percolation on nodal sets of planar random waves is difficult due to the long range correlations of planar random waves and the absence of FKG inequality due to the oscillatory behavior of the covariance function, see [BG17]. One shouldn’t push the analogy between the Bogomolny–Schmit model and planar random waves too far, since this percolation model completely ignore the long range correlation of planar random waves.

1.2.3 Asymptotics related to the nodal set of random waves

The study of nodal set associated to random waves has been a well-established field of research for the last two decades, in particular motivated by celebrated conjectures presented in the last section. We discuss some of the old and recent advances in the field.

Nodal set of random waves on general manifolds. The model of Riemannian random waves and the study of its nodal sets were initiated by S. Zelditch in [Zel09], partially inspired by the Berry conjecture and its link with quantum chaos. We have seen in Section 1.1.2 that the model of random waves locally converges in distribution towards the universal isotropic Gaussian process g_∞ . The Kac–Rice formula stated in Theorem

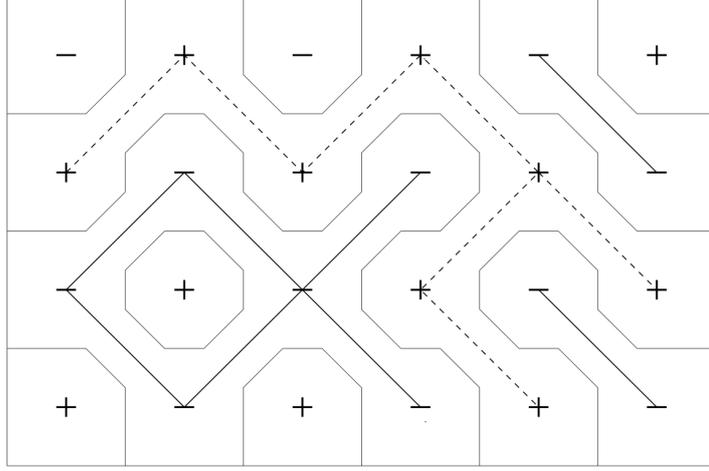


Figure 1.4 – The percolation-like model introduced by E. Bogomolny and C. Schmit. We start with a square grid on the plane, filled alternatively with + and -. Each vertex of the plane connects with probability 1/2 the two opposites +, and with probability 1/2 the two opposites -. Credit : E. Bogomolny and C. Schmit.

1.1.5 above asserts that

$$\mathbb{E} \left[\mathcal{H}^{d-1}(f_\lambda^{-1}(0)) \right] = \int_{\mathcal{M}} \mathbb{E} \left[\|\nabla f_\lambda(x)\| \mid f_\lambda(x) = 0 \right] \frac{1}{\sqrt{2\pi \operatorname{Var} f_\lambda(x)}} dx.$$

The local Weyl law implies that the random vector $(f_\lambda(x), \frac{1}{\lambda} \nabla f_\lambda(x))$ converges in distribution towards the Gaussian vector $(g_\infty(0), \nabla g_\infty(0))$. As a consequence of formula (1.3), we deduce the following asymptotics, for a Borel subset A of the manifold \mathcal{M}

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathbb{E} \left[\mathcal{H}^{d-1}(f_\lambda^{-1}(0) \cap A) \right]}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \operatorname{Vol}(A). \quad (1.6)$$

A rigorous proof can be found in [Zel09]. A similar asymptotics also holds in the monochromatic model. It implies in particular that the nodal set, in mean, tends to equidistribute on the manifold, in the high energy limit. The author also conjectures that this equidistribution holds almost surely, but is only able to prove it upon averaging on the set of eigenvalues, see [Zel09, Cor. 2].

More recently, the authors in [KSW21] have shown the universality of the expected nodal volume in the real-analytic framework. They were able to prove the exact same asymptotics (1.6) when considering a non-Gaussian combination of eigenfunctions, only

assuming that the iid random variables are centered with unit variance. It generalizes the universality result of [APP18] on the torus, to the larger class of real-analytic manifolds. The question of universality in the smooth setting is still open.

The question of the variance and the central limit theorem is more delicate. The remainder in the Weyl law is not sufficiently accurate in order to deduce the exact asymptotic of the variance. One can still show with the Kac–Rice for the variance formula, the existence of a positive constant c , depending only on the ambient dimension, such that

$$\text{Var} \left(\frac{\mathcal{H}^{d-1}(f_\lambda^{-1}(0))}{\lambda} \right) = O(\lambda^{-c}).$$

Nevertheless, there exists “small-scale results” (see [Die+20; NPR19] and the references therein) where the exact asymptotics of the variance and a CLT for the nodal length of monochromatic random waves on geodesic balls of size $\frac{\log(\lambda)}{\lambda}$ have been obtained, in concordance with Berry’s predictions for planar random waves [Ber02]. Passing from small-scale results to global scale seems out of reach in a general Riemannian framework, due to the lack of accurate estimates concerning the remainder in the local Weyl law.

Random spherical harmonics. On the d -dimensional sphere \mathbb{S}^d , we consider a Gaussian linear combination of eigenfunctions associated to energy level λ , with $\lambda = \sqrt{n(n+d-1)}$ and n a positive integer. It defines the model of *random spherical harmonics* \tilde{f}_λ . This model has brought a lot of attention in the past decades since it is fairly explicit, see [Ros19] for a survey of recent developments. As explained in Section 1.1.2, one can describe exactly the Laplace spectrum and the two-points spectral kernel projector, using Legendre polynomials in dimension 2, and Gegenbauer polynomials in higher dimension. Using Darboux asymptotics formula [Sze75, Theorem 8.21.8], one has the following asymptotics for the n -th Gegenbauer polynomial in dimension d , and $\theta \in]0, \pi[$

$$P_{n,d}(\cos(\theta)) = \frac{1}{\sqrt{\pi n}} \left(\frac{2}{\sin(\theta)} \right)^{\frac{d-1}{2}} \cos \left[\left(n + \pi \frac{d-1}{2} \right) \theta - \pi \frac{d-1}{4} \right] + O(n^{-3/2}).$$

This asymptotics provides a very accurate remainder in the local Weyl law, which allows to prove a variety of results concerning the asymptotics of the nodal measure of random spherical harmonics. Notice first that the model of random spherical harmonics \tilde{f}_λ , with $\lambda = \sqrt{n(n+d-1)}$, is an isotropic Gaussian process, and the Kac–Rice formula of Theorem

1.1.5 yields the following explicit expression for the expectation of the nodal volume

$$\mathbb{E}[\mathcal{H}^1(\tilde{f}_\lambda^{-1}(0))] = \frac{2\pi^{d/2}}{\sqrt{d} \Gamma\left(\frac{d}{2}\right)} \lambda.$$

The variance of the nodal volume can also be explicitly computed in dimension 2. One could conjecture that the variance is also of order λ , as for the Bargman–Fock random field on the plane. In fact, one has the following variance asymptotics in the high energy limit

$$\text{Var}(\mathcal{H}^1(\tilde{f}_\lambda)) \simeq \frac{1}{32} \log(\lambda).$$

This result has been proved in [Wig09] using the Kac–Rice formula, and have been strengthened to a quantitative central limit theorem in [MRW20] by the Wiener chaos method. It is conjectured in [Wig09] that for the d -sphere with $d \geq 3$, one has the following asymptotics for the variance of the nodal volume

$$\text{Var}(\mathcal{H}^{d-1}(\tilde{f}_\lambda^{-1}(0))) \simeq \frac{c}{\lambda^{d-2}},$$

for some positive constant c . The slow growth of the variance is characteristic of the so-called *Berry cancellation phenomenon*, observed by M. Berry in [Ber02] for the nodal fluctuation of planar random waves. From the Kac–Rice viewpoint, it is a consequence of the cancellation of two terms in the Kac–Rice formula, producing a lower order for the variance of the nodal length than other planar models with short-range correlations. The Wiener chaos decomposition gives another interpretation of this phenomenon. It is due to the domination of the fourth order Wiener chaos. We refer the reader to [NPR19] and the references therein for more details on the Berry cancellation phenomenon.

In agreement with the Bogomolny-Schmidt conjecture on the expected number of nodal domains, see (1.5), we close the paragraph with the beautiful result of S. Nazarov and M. Sodin in [NS09], on the number nodal domains of a random spherical harmonic in dimension 2, denoted $N(\tilde{f}_\lambda)$. They showed the existence of a positive constant a (the so-called “Nazarov–Sodin constant”) such that for every positive ε , there are constant $c(\varepsilon)$ and $C(\varepsilon)$ such that

$$\mathbb{P}\left(\left|\frac{N(\tilde{f}_\lambda)}{\lambda^2} - a\right| \geq \varepsilon\right) \leq C(\varepsilon)e^{-c(\varepsilon)n}.$$

Arithmetic random waves. On the torus \mathbb{T}^d , one usually considers a Gaussian linear combination of eigenfunctions associated to energy level λ , with $\lambda^2 \in \mathbb{N}$. It defines the model of *arithmetic random waves* \tilde{f}_λ . As explained in section 1.1.2, the local Weyl law on the torus is deeply intertwined with arithmetic and number theoretic considerations, leading to results of different nature compared to its spherical counterpart, described in the previous paragraph.

As in the spherical case, the covariance function is stationary and Kac–Rice formula yields the following explicit formula for the expectation of nodal volume

$$\mathbb{E}[\mathcal{H}^{d-1}(\tilde{f}_\lambda^{-1}(0))] = c(d)\lambda,$$

where $c(d)$ is an explicit constant that depends on the ambient dimension. The question of the variance is more subtle than in the spherical case. We denote \mathcal{N}_n cardinal of the eigenspace associated to the eigenvalue $\lambda = \sqrt{n}$, where n is a sum of d squares. The two dimensional torus has the particular property that lattice points on the circle do not equidistribute when \mathcal{N}_n is large, contrary to the higher dimensional torus (see [DS90] and the references therein). The authors in [KKW13] have shown on the two-dimensional torus, the following asymptotics for the variance of the nodal length as the quantity \mathcal{N}_n grows to infinity

$$\text{Var}(\mathcal{H}^1(\tilde{f}_\lambda^{-1}(0))) = c_n \frac{n}{\mathcal{N}_n^2} (1 + o(1)).$$

Here, c_n is a positively bounded quantity that depends on the angular distribution of lattice points on the circle of radius n . In that sense, the asymptotics of the variance is non-universal since it also depends on the eigenspace through the constant c_n . This non-universality phenomenon does not occur for the three-dimensional torus \mathbb{T}^3 , due to the aforementioned equidistribution of lattice points on the sphere, as the quantity \mathcal{N}_n grows to infinity. The authors in [BM19] have shown on the three-dimensional torus, the following asymptotics for the variance of the nodal area as the quantity \mathcal{N}_n grows to infinity

$$\text{Var}(\mathcal{H}^2(\tilde{f}_\lambda^{-1}(0))) = c \frac{n}{\mathcal{N}_n^2} (1 + o(1)),$$

for some explicit constant c that does not depends on n . In both cases, the variance is of lower order than expected and is again a manifestation of the Berry cancellation phenomenon (see [Ber02] and the discussion in the previous paragraph). Note that both variance asymptotics has been strengthened into non-CLT in [Mar+16; Cam19], which

quite differs from the CLT known for random spherical harmonics. We again refer to [Ros19] for a deeper insight on this topic.

1.3 Overview of the contributions

In this section, we describe our contributions to asymptotics related to Riemannian random waves, which led to the first paper [Gas21a] written during this thesis. In this paper, we consider the Riemannian random wave model of Gaussian linear combinations of Laplace eigenfunctions on a general compact Riemannian manifold. With probability one with respect to the Gaussian coefficients, we establish that, both for large and small band models, the process properly rescaled and evaluated at an independently and uniformly chosen point X on the manifold, converges in distribution under the sole randomness of X towards an universal Gaussian field as the frequency tends to infinity. This result extends the celebrated central limit Theorem of Salem–Zygmund for trigonometric polynomials, to the more general framework of compact Riemannian manifolds.

We then deduce from the above convergence the almost-sure asymptotics of the nodal volume associated with the random wave. To the best of our knowledge, in the real Riemannian case, these asymptotics were only known in expectation and not in the almost sure sense due to the lack of sufficiently accurate variance estimates. This addresses in particular a question of S. Zelditch regarding the almost sure equidistribution of nodal volume.

1.3.1 The Salem–Zygmund approach

As mentioned in the prologue above, R. Salem and A. Zygmund show in [SZ54] the following result. If $(a_n)_{n \geq 0}$ is a sequence of independent Gaussian random variables and X is a uniform random variable on $[0, 2\pi]$, then almost surely with respect to the coefficients $(a_n)_{n \geq 0}$ the following convergence in distribution holds

$$\sqrt{\frac{2}{n}} \sum_{k=0}^n a_k \cos(kX) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}_X} \mathcal{N}(0, 1).$$

This approach has been revisited in the series of articles [AP21; APP21], where the authors established both a quantitative version and a functional version of Salem–Zygmund Theorem and then used these results to deduce the almost sure asymptotics of the number

of zeros of random trigonometric polynomials with symmetric coefficients. More specifically, we define the process g_n on \mathbb{R} by

$$g_n(v) = \sqrt{\frac{2}{n}} \sum_{k=0}^n a_k \cos \left(k \left(X + \frac{v}{n} \right) \right).$$

Then almost surely with respect to the coefficients $(a_n)_{n \geq 0}$, the process g_n converges in distribution towards the so-called Paley–Wiener process, that is the process g_∞ with covariance function

$$\mathbb{E}[g_\infty(u)g_\infty(v)] = \frac{\sin(u-v)}{u-v}.$$

We show that this theorem can be extended to the framework of Riemannian random waves defined in the previous section. Let \mathcal{M} be a compact Riemannian manifold of dimension d , and $(\lambda_n, \varphi_n)_{n \geq 0}$ be the spectral decomposition of the Laplace operator. Recall the definition of the Riemannian and monochromatic random waves

$$f_\lambda : x \mapsto \frac{1}{\sqrt{K(\lambda)}} \sum_{\lambda_n \leq \lambda} a_n \varphi_n(x) \quad \text{and} \quad \tilde{f}_\lambda : x \mapsto \frac{1}{\sqrt{k(\lambda)}} \sum_{\lambda_n \in]\lambda-\tau(\lambda), \lambda]} a_n \varphi_n(x).$$

as well as their flattened and rescaled version at a point $x \in \mathcal{M}$,

$$g_\lambda^x : v \mapsto f_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right] \quad \text{and} \quad \tilde{g}_\lambda^x : v \mapsto \tilde{f}_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right].$$

Let X be a uniformly distributed random variable on \mathcal{M} . The following generalization of Salem–Zygmund Theorem then holds.

Theorem 1.3.1. *Almost surely with respect to the probability \mathbb{P}_a , the two processes $(g_\lambda^X(v))_{v \in \mathbb{R}^d}$ and $(\tilde{g}_\lambda^X(v))_{v \in \mathbb{R}^d}$ converge in distribution under \mathbb{P}_X with respect to the \mathcal{C}^∞ topology, towards isotropic Gaussian processes $(g_\infty(v))_{v \in \mathbb{R}^d}$ and $(\tilde{g}_\infty(v))_{v \in \mathbb{R}^d}$ with respective covariance functions*

$$\mathbb{E}_X[g_\infty(u)g_\infty(v)] = \mathcal{B}_d(\|u-v\|) \quad \text{and} \quad \mathbb{E}_X[\tilde{g}_\infty(u)\tilde{g}_\infty(v)] = \mathcal{S}_d(\|u-v\|).$$

This theorem could be seen as a weak justification of Berry heuristics 1.4. Instead of considering the behavior of a single eigenfunction in the high energy limit, we consider the behavior of a random sum of eigenfunctions in the high energy limit. The added randomness allows us to recover the asymptotics predicted by Berry for a single eigenfunction. The

proof is of general nature and could certainly be applied to other frameworks, as it is solely based on the local Weyl law and the associated polynomial remainder.

As usual, a proof of convergence in distribution for stochastic processes can be split into two parts, namely the convergence of finite-dimensional distributions, and a tightness criterion. Here, the tightness criterion is relatively immediate since our estimates are valid in the \mathcal{C}^∞ topology. The convergence of finite dimensional distributions is based upon the convergence of characteristic functions. For a positive integer p , let $t = (t_j)_{1 \leq j \leq p}$ be a collection of real numbers and $v = (v_j)_{1 \leq j \leq p}$ be a collection of vectors in \mathbb{R}^d . We define

$$N_\lambda^x(v, t) = \sum_{j=1}^p t_j g_\lambda^x(v_j) \quad \text{and} \quad N_\infty(v, t) = \sum_{j=1}^p t_j g_\infty(v_j).$$

We will omit the dependence in v and t when appropriate. It suffices to prove that

$$\lim_{n \rightarrow +\infty} \mathbb{E}_X \left[e^{iN_\lambda^X} - \mathbb{E}_a \left[e^{iN_\lambda^X} \right] \right] = 0. \quad (1.7)$$

Indeed, the local Weyl law asserts that the right term inside the expectation converges towards the characteristic function of N_∞ , uniformly on the random parameter X . The proof of (1.7) is based on the following quantitative convergence of moments, for any positive integer q

$$\lim_{n \rightarrow +\infty} \mathbb{E}_a \left[\left| \mathbb{E}_X \left[e^{iN_\lambda^X} - \mathbb{E}_a \left[e^{iN_\lambda^X} \right] \right] \right|^{2q} \right] = O \left(\frac{1}{\lambda^q} \right).$$

The conclusion follows from a Borel–Cantelli argument in order to remove the Gaussian expectation. This last convergence relies on an explicit -but technical- expansion of the q -th power. Let us sketch the proof for $q = 1$, since it contains the main ideas of the proof. We define

$$\Delta_X = e^{iN_\lambda^X} - \mathbb{E}_a \left[e^{iN_\lambda^X} \right].$$

Let Y another uniformly distributed random variable on \mathcal{M} , independent of X . Then

$$\begin{aligned}
 \mathbb{E}_a \left[|\mathbb{E}_X[\Delta_X]|^2 \right] &= \mathbb{E}_a \left[\mathbb{E}_{X,Y}[\Delta_X \overline{\Delta_Y}] \right] \\
 &= \mathbb{E}_{X,Y} \left[\mathbb{E}_a[\Delta_X \overline{\Delta_Y}] \right] \\
 &= \mathbb{E}_{X,Y} \left[\mathbb{E}_a \left[\left(e^{iN_\lambda^X} - \mathbb{E}_a \left[e^{iN_\lambda^X} \right] \right) \left(e^{-iN_\lambda^Y} - \mathbb{E}_a \left[e^{-iN_\lambda^Y} \right] \right) \right] \right] \\
 &= \mathbb{E}_{X,Y} \left[e^{-\frac{1}{2}(\mathbb{E}_a[N_\lambda^X]^2 + \mathbb{E}_a[N_\lambda^Y]^2)} \left(e^{-\mathbb{E}[N_\lambda^X N_\lambda^Y]} - 1 \right) \right] \\
 &\leq \mathbb{E}_{X,Y} \left[|\mathbb{E}_a[N_\lambda^X N_\lambda^Y]| \right],
 \end{aligned}$$

where the last inequality is a direct consequence of the mean-value theorem for the exponential function. This last term is easily controlled with the remainder in the local Weyl law and the decay of the limit covariance kernel. It suffices to observe the bound

$$\begin{aligned}
 \mathbb{E}_{X,Y} [|\mathbb{E}_a[f_\lambda(X)f_\lambda(Y)]|] &= \iint_{\mathcal{M} \times \mathcal{M}} |K_\lambda(x, y)| dx dy \\
 &= \iint_{\mathcal{M} \times \mathcal{M}} |\mathcal{B}_d(\lambda \operatorname{dist}(x, y))| dx dy + O\left(\frac{1}{\lambda}\right) \\
 &= O\left(\frac{1}{\lambda}\right).
 \end{aligned}$$

A similar computation holds in the monochromatic framework. Note that the Borel–Cantelli argument does not allow in fact to recover the dependence in t and v in the convergence given by Theorem 1.3.1. We prove the more precise result.

Theorem 1.3.2. *Let K a compact subset of \mathbb{R}^d and $\varepsilon > 0$. Then there is a constant $C(\omega)$ depending only on K , ε and the Gaussian sequence $(a_k)_{k \geq 0}$ such that*

$$\sup_{v \in K} \left| \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] - e^{-\frac{1}{2}\mathbb{E}_X[N_\infty(v,t)^2]} \right| \leq C(\omega) \frac{1 + \|t\|^{2+\varepsilon}}{\lambda^{1/2-\varepsilon}}.$$

This quantitative result reinforces the convergence in distribution given by Theorem 1.3.1 in three directions. Firstly, it gives a rate of convergence of order $\sqrt{\lambda}$ for the stochastic convergence, which is reminiscent of the classical rate of the CLT given by Berry–Essen type bounds. Secondly, we have a control of the convergence in frequency, given by the polynomial growth in the parameter t . Using results from [Arr+17], it allows us to reinforce the convergence in distribution to the convergence in 4-Wassertein metric, and subsequently

in Kolmogorov distance. At last, we recover the uniformity with respect to the parameter v on compact subsets of \mathbb{R}^d . The proof of Theorem 1.3.2 is mainly based on Sobolev inequalities in order to control the supremum norms by more manageable L^1 norms of derivatives.

1.3.2 Almost sure convergence of the nodal volume

The stochastic convergence given by the previous Theorem 1.3.1 allows us to deduce the almost-sure convergence of the nodal measure towards an explicit constant, which constitutes the main results of the article [Gas21a]. It reinforces the previously known convergence in expectation for the nodal volume of random waves obtained in [Zel09]. To this end, we define $(\mu_\lambda)_{\lambda>0}$ (resp. $(\tilde{\mu}_\lambda)_{\lambda>0}$) be the random nodal measures associated with the random waves model $(f_\lambda)_{\lambda>0}$ (resp. $(\tilde{f}_\lambda)_{\lambda>0}$). For a continuous function $h : \mathcal{M} \rightarrow \mathbb{R}$, one has

$$\int_{\mathcal{M}} h(x) d\mu_\lambda(x) := \int_{\{f_\lambda=0\}} h(x) d\mathcal{H}^{d-1}(x).$$

Theorem 1.3.3. *Almost surely with respect to the probability \mathbb{P}_a , the sequence $(\frac{1}{\lambda}\mu_\lambda)_{\lambda>0}$ converges weakly to the Riemannian volume measure μ , up to an explicit multiplicative factor. That is to say, for every continuous function $h : \mathcal{M} \rightarrow \mathbb{R}$,*

$$\lim_{\lambda \rightarrow +\infty} \frac{1}{\lambda} \int_{\mathcal{M}} h(x) d\mu_\lambda(x) = \left(\int_{\mathcal{M}} h(x) d\mu(x) \right) \mathbb{E}[\mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B)],$$

and

$$\lim_{\lambda \rightarrow +\infty} \frac{1}{\lambda} \int_{\mathcal{M}} h(x) d\tilde{\mu}_\lambda(x) = \left(\int_{\mathcal{M}} h(x) d\mu(x) \right) \mathbb{E}[\mathcal{H}^{d-1}(\{\tilde{g}_\infty = 0\} \cap B)].$$

The right-hand side can be explicitly computed by Kac–Rice formula for random fields, see Theorem 1.1.5. In particular, choosing $h = \mathbf{1}_{\mathcal{M}}$ we get

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} = \frac{1}{\sqrt{\pi(d+2)}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \text{Vol}(\mathcal{M}),$$

and

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{\tilde{f}_\lambda = 0\})}{\lambda} = \frac{1}{\sqrt{\pi d}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \text{Vol}(\mathcal{M}).$$

This result thus improves the results in [Zel09] or [Let16] about the asymptotics of the expected nodal volume. In particular, this result positively answers a question raised by S. Zelditch in [Zel09, Cor. 2], about the almost sure equidistribution of the random nodal measure.

The proof relies on the following stochastic representation formula between the nodal volumes of the processes f_λ and g_λ^X , in the spirit of Bourgain derandomization, see [Bou14; BW16]. For large λ , we have

$$\frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} \simeq \mathbb{E}_X[Z_\lambda], \quad \text{with} \quad Z_\lambda = \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B). \quad (1.8)$$

By Theorem 1.3.1 and the continuity of the random nodal volume for the \mathcal{C}^1 -topology, the continuous mapping Theorem asserts that the nodal measure of g_λ^X , denoted Z_λ , converges in distribution (under \mathbb{P}_X) towards the nodal measure of the Gaussian process g_∞ . A rather involved anti-concentration lemma, that relies on the quantitative CLT given by Theorem 1.3.2, allows us to prove the uniform integrability of the family $(Z_\lambda)_{\lambda>0}$. It directly implies the convergence of expectations under \mathbb{P}_X in (1.8) and Theorem 1.3.3. In particular the proof bypasses variance estimates for the nodal volume that lacks in the literature, and relies on more geometric considerations. The details of the proof will be found in Chapter 3 of this thesis, which is based on the article [Gas21a].

The question of universality in this almost-sure context seem to be a delicate topic. Recently, the authors in [KSW21] have shown the universality of the expected nodal volume in the real-analytic framework. On the sphere, some informal numerical evidence (communicated to me by D. Marinucci) suggests that such universality results do not hold in the almost-sure sense. Numerically, the nodal measure does not seem to converge towards the Riemannian measure on the sphere, when one replaces Gaussian distributions by Rademacher or uniform distributions.

THE ZEROS COUNTING MEASURE OF REAL GAUSSIAN PROCESSES

In this chapter, we summarize the different contributions of this thesis to the study of the zeros counting measure of real Gaussian processes, and in particular of random trigonometric polynomials. These contributions will be detailed in Chapters 4 and 5, which consists in the second and third submitted articles of this thesis [Gas21b; Gas21c]. We first introduce a few tools in order to analyze the zeros counting measure of a stochastic process. We then make a quick survey of the literature concerning the zeros counting measure of random processes before giving an overview of our contributions to the field.

2.1	The combinatorics of cumulants	50
2.1.1	Möebius inversion on a lattice	51
2.1.2	The lattice of partitions of a finite set	52
2.1.3	Cumulants of a random variable	56
2.2	Moments and cumulants of the number of zeros	60
2.2.1	Kac–Rice formula for the factorial moments	61
2.2.2	Gaussian conditioning and Kac density	64
2.3	Literature on the zeros counting measure	68
2.3.1	Expectation	68
2.3.2	Variance and CLT	70
2.3.3	Higher moments and cumulants	73
2.4	Overview of the contributions	76
2.4.1	Variance asymptotics	76
2.4.2	Cumulants asymptotics	79

2.1 The combinatorics of cumulants

Let $(X_n)_{n \geq 0}$ be a sequence of centered random variables with unit variance, and N be a standard Gaussian random variable. There are several available methods in order to prove a central limit theorem (CLT) for the sequence $(X_n)_{n \geq 0}$, that is the convergence in distribution of the form

$$X_n \underset{n \rightarrow +\infty}{\Longrightarrow} N.$$

We focus on one of these methods, namely the method of moments. Assume that $(X_n)_{n \geq 0}$ has finite moments of all orders. Then, the CLT is equivalent to the convergence of each moments of X_n towards the corresponding moments of N , see [Bil95, Thm. 30.2]. For each integer $p \geq 0$, it suffices to show that

$$\lim_{n \rightarrow +\infty} \mathbb{E}[X_n^p] = \mathbb{E}[N^p].$$

The method of moments is well-advised when one has an explicit expression for the p -th moment, which in our framework is given by the Kac–Rice formula, see Theorem 2.2.1. The moments of a standard Gaussian variables are given by

$$\mathbb{E}[N^p] = \begin{cases} \frac{p!}{2^{p/2} (\frac{p}{2})!} & \text{if } p \text{ is even,} \\ 0 & \text{if } p \text{ is odd.} \end{cases}$$

This expression has a nice combinatoric interpretation. The quantity $\mathbb{E}[N^p]$ is the number of partition into pairs of a set with p elements. For instance, $\mathbb{E}[N^4] = 3$, there are 3 partitions into pair of the set $\{1, 2, 3, 4\}$:

$$\{\{1, 2\}, \{3, 4\}\}, \quad \{\{1, 3\}, \{2, 4\}\} \quad \text{and} \quad \{\{1, 4\}, \{2, 3\}\}.$$

Proving the asymptotics of every moments often uses this combinatoric interpretation of Gaussian moments. One has to find a way to make this combinatoric explicit in the expression of $\mathbb{E}[X_n^p]$, this is exactly the role of the cumulants of a random variable.

The cumulants of a random variable are classically defined from the power series expansion of the logarithm of the moment generating function. In the following subsections, we give a combinatoric introduction of the cumulants based on the Möebius inversion formula on the lattice of partitions. This approach is largely inspired by the paper of T. P. Speed [Spe83]. It has the advantage of proving usual properties of cumulants with minimal

computation. Though not obviously related to the study of zeros of random functions, the use of these combinatoric tools will appear clearer in Section 2.2 below. For further applications of this combinatoric-based approach of cumulants, one can also refer to the book [PT11].

2.1.1 Möbius inversion on a lattice

Let (P, \preceq) be a finite lattice, that is a partially ordered finite set such that two elements have a greatest lower bound – a *meet* – and a least upper bound – a *join* –. Given two elements x, y in P with $x \preceq y$, we define the closed interval $[x, y]$ as

$$[x, y] = \{z \in P \mid x \preceq z \preceq y\}.$$

We define recursively the Möbius function μ of an interval $[x, y]$ as the quantity

$$\mu([x, y]) = \begin{cases} 1 & \text{if } x = y, \\ - \sum_{x \preceq z \prec y} \mu([x, z]) & \text{else.} \end{cases}$$

By induction, the Möbius function satisfies the two following identities

$$\sum_{x \preceq z \preceq y} \mu([x, z]) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{else,} \end{cases} \quad \text{and} \quad \sum_{x \preceq z \preceq y} \mu([z, y]) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{else.} \end{cases} \quad (2.1)$$

From these two formulas, we deduce the following principle, known as *Möbius inversion*.

Theorem 2.1.1 (Möbius inversion). *Let f and g be two functions from P to \mathbb{R} . We have the following equivalence.*

$$\left(\forall x \in P, \quad g(x) = \sum_{y \succeq x} f(y) \right) \iff \left(\forall x \in P, \quad f(x) = \sum_{y \succeq x} \mu([x, y])g(y) \right).$$

Similarly, we have

$$\left(\forall x \in P, \quad g(x) = \sum_{y \preceq x} f(y) \right) \iff \left(\forall x \in P, \quad f(x) = \sum_{y \preceq x} \mu([y, x])g(y) \right).$$

Proof. Assume that for all $x \in P$, $g(x) = \sum_{y \succeq x} f(y)$. Then

$$\sum_{y \succeq x} \mu([x, y])g(y) = \sum_{y \succeq x} \mu([x, y]) \left(\sum_{z \succeq y} f(z) \right) = \sum_{z \succeq x} f(z) \left(\sum_{x \preceq y \preceq z} \mu([x, y]) \right).$$

Relations (2.1) imply that the right hand term equals $f(x)$. The converse sense as well as the second statement can be proved in a similar fashion. \square

A well-known application of Möebius inversion formula is on the ring \mathbb{Z} of integers equipped the divisibility as partial order, which is a fundamental tool in arithmetic and number theory. One can also cite the lattice of subsets of some finite set, with inclusion as partial order. In that case, Möebius inversion yields the *principle of inclusion-exclusion*. We will concentrate from now on the lattice of partitions of a finite set, which we define in the following subsection.

2.1.2 The lattice of partitions of a finite set

In the following, A is a non-empty finite set. We say that \mathcal{I} is a *partition* of A if it is a collection of disjoint and non-empty subsets of A such that their union equals to A . Elements of a partition are called *cells*. We denote the set of partitions of A by \mathcal{P}_A . For instance,

$$\mathcal{P}_{\{a,b,c\}} = \left\{ \left\{ \{a\}, \{b\}, \{c\} \right\}, \left\{ \{a\}, \{b, c\} \right\}, \left\{ \{b\}, \{a, c\} \right\}, \left\{ \{c\}, \{a, b\} \right\}, \left\{ \{a, b, c\} \right\} \right\}.$$

The set of partitions of A is naturally equipped with a partial order \preceq . Given \mathcal{I} and \mathcal{J} two partitions of A , we say that \mathcal{I} is *finer* than \mathcal{J} (or that \mathcal{J} is *coarser* than \mathcal{I}) and we denote it by $\mathcal{I} \preceq \mathcal{J}$ (or $\mathcal{J} \succeq \mathcal{I}$), if

$$\forall I \in \mathcal{I}, \exists J \in \mathcal{J} \text{ such that } I \subset J.$$

In other words, the partition \mathcal{I} is finer than the partition \mathcal{J} if the partition \mathcal{I} is a refinement of the partition \mathcal{J} . If J is a cell of the partition \mathcal{J} , we can then define

$$\mathcal{I}_J = \{I \in \mathcal{I} \mid I \subset J\}. \tag{2.2}$$

The set \mathcal{I}_J is then a partition of the set J . Note that there is a one-to-one correspondence between the set of partitions of A coarser than a partition \mathcal{I} , and the set of partitions of

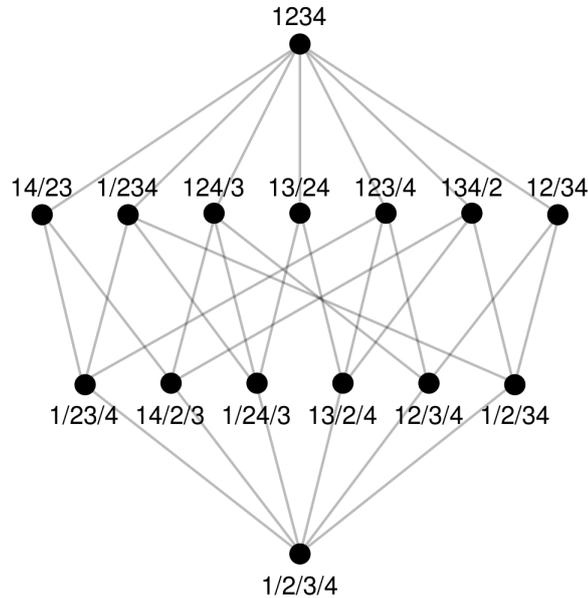


Figure 2.1 – Lattice of partitions of the set $\{1, 2, 3, 4\}$. Credit : Ed g2s.

the set \mathcal{I} , given by the application

$$\mathcal{J} \mapsto \{\mathcal{I}_J \mid J \in \mathcal{J}\}. \quad (2.3)$$

Given two partitions \mathcal{I} and \mathcal{J} , one can define its *meet* $\mathcal{I} \wedge \mathcal{J}$ as the coarsest partition that refines both \mathcal{I} and \mathcal{J} , and its *join* $\mathcal{I} \vee \mathcal{J}$ as the finest partition that is refined both by \mathcal{I} and \mathcal{J} . Explicitly, the cells of $\mathcal{I} \wedge \mathcal{J}$ are all the non-empty intersections of a block in \mathcal{I} and a block in \mathcal{J} . The cells of $\mathcal{I} \vee \mathcal{J}$ are exactly the smallest subsets of A that are both union of blocks of \mathcal{I} and union of blocks of \mathcal{J} . These two properties turn the partially ordered set (\mathcal{P}_A, \preceq) into a finite lattice. The following proposition makes explicit the Möbius function on the lattice of partitions \mathcal{P}_A .

Proposition 2.1.2. *Let \mathcal{I}, \mathcal{J} be two partitions of the set A with $\mathcal{I} \preceq \mathcal{J}$. Then*

$$\mu([\mathcal{I}, \mathcal{J}]) = (-1)^{|\mathcal{I}| - |\mathcal{J}|} \prod_{J \in \mathcal{J}} (|\mathcal{I}_J| - 1)! .$$

Proof. Given two integers n and p , we define the factorial power

$$[n]^p = n(n-1)\dots(n-p+1).$$

We apply Möebius inversion of Theorem 2.1.1 to the following identity, for $n \in \mathbb{N}$ and $\mathcal{I} \in \mathcal{P}_A$,

$$n^{|\mathcal{I}|} = \sum_{\mathcal{J} \succeq \mathcal{I}} n^{|\mathcal{J}|}. \quad (2.4)$$

This identity can be deduced by considering the number of n -colorings of the cells of the partitions \mathcal{I} , with $n \in \mathbb{N}$. There are naively $n^{|\mathcal{I}|}$ such colorings. We can also count them by gathering cells of same color, leading to the formula on the right-hand side of (2.4). Möebius inversion then implies, for all $n \in \mathbb{N}$,

$$n^{[|\mathcal{I}|]} = \sum_{\mathcal{J} \succeq \mathcal{I}} \mu([\mathcal{I}, \mathcal{J}]) n^{|\mathcal{J}|}.$$

Identifying the linear term in n in both sides leads to

$$\mu([\mathcal{I}, \{A\}]) = (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \quad .$$

More generally, let \mathcal{I} and \mathcal{J} be two partitions with $\mathcal{I} \preceq \mathcal{J}$. The one-to-one correspondence (2.3) implies the factorization

$$[\mathcal{I}, \mathcal{J}] = \prod_{J \in \mathcal{J}} [\mathcal{I}_J, \{J\}],$$

which implies in turn

$$\mu([\mathcal{I}, \mathcal{J}]) = \prod_{J \in \mathcal{J}} \mu([\mathcal{I}_J, \{J\}]) = (-1)^{|\mathcal{I}|-|\mathcal{J}|} \prod_{J \in \mathcal{J}} (|\mathcal{I}_J| - 1)! \quad .$$

□

Let $(m_B)_{B \subset A}$ and $(\kappa_B)_{B \subset A}$ be two families of numbers indexed by the collection of subsets of A . For a collection \mathcal{I} of subsets of A we define

$$m_{\mathcal{I}} = \prod_{I \in \mathcal{I}} m_I \quad \text{and} \quad \kappa_{\mathcal{I}} = \prod_{I \in \mathcal{I}} \kappa_I.$$

In that context, Möebius inversion takes the following form.

Proposition 2.1.3. *We have*

$$\forall B \subset A, \quad m_B = \sum_{\mathcal{I} \in \mathcal{P}_B} \prod_{I \in \mathcal{I}} \kappa_I$$

if and only if

$$\forall B \subset A, \quad \kappa_B = \sum_{\mathcal{I} \in \mathcal{P}_B} (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \prod_{I \in \mathcal{I}} m_I.$$

Proof. We prove the direct implication. The converse sense is proved similarly. Let \mathcal{J} be a partition of a subset B of A . Then by the correspondence (2.3) one has

$$\begin{aligned} m_{\mathcal{J}} &= \prod_{J \in \mathcal{J}} \left(\sum_{\mathcal{I} \in \mathcal{P}_J} \kappa_{\mathcal{I}} \right) \\ &= \sum_{\mathcal{I} \preceq \mathcal{J}} \kappa_{\mathcal{I}} \end{aligned}$$

By Möebius inversion on the lattice \mathcal{P}_B , we get

$$\kappa_{\mathcal{J}} = \sum_{\mathcal{I} \preceq \mathcal{J}} (-1)^{|\mathcal{I}|-|\mathcal{J}|} \left(\prod_{J \in \mathcal{J}} (|\mathcal{I}_J| - 1)! m_J \right)$$

We get the conclusion by choosing $\mathcal{J} = \{B\}$. □

Proposition 2.1.4. *Let $(m_B)_{B \subset A}$ and $(\kappa_B)_{B \subset A}$ be two families of numbers related by one of the equivalent formulas in Proposition 2.1.3. Assume the existence of a partition $\mathcal{J} \neq \{A\}$ such that*

$$\forall B \subset A, \quad m_B = \prod_{J \in \mathcal{J}} m_{J \cap B}.$$

Then

$$\kappa_A = 0.$$

Proof. One has

$$\begin{aligned}
 \kappa_A &= \sum_{\mathcal{I} \in \mathcal{P}_A} \mu([\mathcal{I}, \{A\}]) m_{\mathcal{I}} \\
 &= \sum_{\mathcal{I} \in \mathcal{P}_A} \mu([\mathcal{I}, \{A\}]) \prod_{J \in \mathcal{J}} \left(\prod_{I \in \mathcal{I}} m_{I \cap J} \right) \\
 &= \sum_{\mathcal{I} \in \mathcal{P}_A} \mu([\mathcal{I}, \{A\}]) m_{\mathcal{I} \wedge \mathcal{J}} \\
 &= \sum_{\mathcal{I} \in \mathcal{P}_A} \mu([\mathcal{I}, \{A\}]) \left(\sum_{\mathcal{K} \preceq \mathcal{I} \wedge \mathcal{J}} \kappa_{\mathcal{K}} \right) \\
 &= \sum_{\mathcal{K} \preceq \mathcal{J}} \kappa_{\mathcal{K}} \left(\sum_{\mathcal{I} \succeq \mathcal{K}} \mu([\mathcal{I}, \{A\}]) \right).
 \end{aligned}$$

Since $\mathcal{J} \neq \{A\}$, then we also have $\mathcal{K} \neq \{A\}$. Identity (2.1) then implies that

$$\left(\sum_{\mathcal{I} \succeq \mathcal{K}} \mu([\mathcal{I}, \{A\}]) \right) = 0,$$

and the conclusion follows. \square

2.1.3 Cumulants of a random variable

We now apply the previous results on the Möebius inversion on the lattice of partitions to define the cumulants of a random variable and prove some useful properties. To this end, we consider A a finite set and $X = (X_a)_{a \in A}$ be a collection of real random variables indexed by A . Assume that these random variables have finite moments up to order $|A|$. For a subset B of A , we define $X_B = (X_b)_{b \in B}$,

$$m(X_B) = \mathbb{E} \left[\prod_{b \in B} X_b \right] \quad \text{and} \quad \kappa(X_B) = \sum_{\mathcal{I} \in \mathcal{P}_B} (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \prod_{I \in \mathcal{I}} m(X_I). \quad (2.5)$$

The quantity $m(X_B)$ (resp. $\kappa(X_B)$) is called the *joint moment* (resp. *joint cumulant*) of the collection of random variables $(X_b)_{b \in B}$. One has by Möebius inversion of Proposition 2.1.3, that

$$m(X_B) = \sum_{\mathcal{I} \in \mathcal{P}_B} \prod_{I \in \mathcal{I}} \kappa(X_I).$$

Proposition 2.1.5 (Cancellation of cumulants). *Let \mathcal{J} be a partition of A different from the trivial partition $\{A\}$. Assume that the collection of random vectors $((X_j)_{j \in J})_{J \in \mathcal{J}}$, indexed by the cells of the partition \mathcal{J} , are mutually independent. Then*

$$\kappa(X) = 0.$$

Proof. This is a direct application of Proposition 2.1.4, since the independence assumption implies that

$$m(X_B) = \prod_{J \in \mathcal{J}} m(X_{J \cap B}).$$

□

In order to study the distribution of the collection of random variables $X = (X_a)_{a \in A}$, we will consider the joint cumulants of repetitions of the variables $(X_a)_{a \in A}$. To this end, we use the following multindex notations. For a vector $p = (p_a)_{a \in A}$ of \mathbb{N}^A and $t \in \mathbb{R}^A$, we define

$$p! := \prod_{a \in A} p_a!, \quad |p| := \sum_{a \in A} p_a, \quad \text{and} \quad t^p := \prod_{a \in A} t_a^{p_a}.$$

We also define

$$A^p := \{(a, i) \mid a \in A, 1 \leq i \leq p_a\}.$$

We introduce, when defined, the quantities

$$m_p(X) := \mathbb{E} \left[\prod_{a \in A} X_a^{p_a} \right] \quad \text{and} \quad \kappa_p(X) := \kappa \left(\bigcup_{a \in A} \underbrace{(X_a, \dots, X_a)}_{p_a \text{ times}} \right),$$

with the convention $m_0(X) = 1$ and $\kappa_0(X) = 0$. For a single variable X and $p \in \mathbb{N}$, we have

$$m_p(X) = \mathbb{E}[X^p] \quad \text{and} \quad \kappa_p(X) = \kappa(\underbrace{X, \dots, X}_{p \text{ times}}).$$

Given the definition of the set A^p , we have the following equality

$$m_p(X) = m(X^{(p)}) \quad \text{and} \quad \kappa_p(X) = \kappa(X^{(p)}),$$

where $X^{(p)}$ is the random vector indexed by A^p such that $X_{a,i}^{(p)} = X_a$. Assume that the

moments of the random vector X are finite. For $t \in \mathbb{R}^A$, we define the *moment generating function* as the formal series

$$M_X(t) := \sum_{p \in \mathbb{N}^A} m_p(X) \frac{t^p}{p!}.$$

Note that this series does not necessarily converges. Since $m_0(X) = 1$ we can define the *cumulant generating function* $H_X(t)$ as the formal series

$$K_X(t) := \log M_X(t).$$

Proposition 2.1.6. *The formal Taylor series of the function K_X is*

$$K_X(t) = \sum_{p \in \mathbb{N}^A} \kappa_p(X) \frac{t^p}{p!}.$$

Proof. It suffices to show that

$$\exp \left(\sum_{p \in \mathbb{N}^A} \kappa_p(X) \frac{t^p}{p!} \right) = M_X(t).$$

We compute the left hand term to get

$$\begin{aligned} \exp \left(\sum_{p \in \mathbb{N}^A} \kappa_p(X) \frac{t^p}{p!} \right) &= \sum_{n \geq 0} \frac{1}{n!} \left(\sum_{p \in \mathbb{N}^A} \kappa_p(X) \frac{t^p}{p!} \right)^n \\ &= \sum_{n \geq 0} \frac{1}{n!} \sum_{p^{(1)}, \dots, p^{(n)} \in \mathbb{N}^A} \left(\prod_{i=1}^n \kappa_{p^{(i)}}(X) \right) \frac{t^{p^{(1)} + \dots + p^{(n)}}}{p^{(1)}! \dots p^{(n)}!}. \end{aligned}$$

Denoting the term of order p by $c_p(X)$, we get the identity

$$p! c_p(X) = \sum_{n \geq 0} \sum_{p^{(1)} + \dots + p^{(n)} = p} \left(\prod_{i=1}^n \kappa_{p^{(i)}}(X) \right) \frac{1}{n!} \frac{p!}{p^{(1)}! \dots p^{(n)}!}.$$

By enumerating partitions of the set A^p by the number of cells, a standard combinatoric

argument shows that this sum is exactly

$$p! c_p(X) = \sum_{\mathcal{I} \in \mathcal{P}_{AP}} \left(\prod_{I \in \mathcal{I}} \kappa_I(X^{(p)}) \right).$$

Möbius inversion then implies the identity

$$c_p(X) = m_p(X).$$

□

Note that the logarithm and the exponential respect the graduation. As a consequence, the above proof remains valid if one replace the logarithm and exponential functions by their truncated definition up to some order $p \in \mathbb{N}^A$. That is, this formal series relation between moments and cumulants works even though moments (and thus cumulants) are defined up to some order $p \in \mathbb{N}^A$. We then have the following proposition, that characterizes Gaussian moments.

Proposition 2.1.7. *A random vector $X = (X_a)_{a \in A}$ is Gaussian if and only if*

$$\forall p \in \mathbb{N}^A \text{ with } |p| \geq 3, \quad \kappa_p(X) = 0.$$

Proof. If X is a Gaussian vector $\mathcal{N}(m, \Sigma)$, then its moment generating function converges and is given by

$$M_X(t) = \exp \left(\langle m, t \rangle + \frac{1}{2} \langle t, \Sigma t \rangle \right).$$

Taking the logarithm of this expression, one obtains by Proposition 2.1.6 the next expression for the cumulant generating function

$$K_X(t) = \langle m, t \rangle + \frac{1}{2} \langle t, \Sigma t \rangle,$$

and the conclusion follows. The converse sens is similar, since the cumulant generating function necessarily has the above form. □

The joint cumulant of a collection of random variables $(X_a)_{a \in A}$ must be seen as the “pure order $|A|$ ” dependence between the random variables $(X_a)_{a \in A}$. If the variables $(X_a)_{a \in A}$ can be split into two independent collections of random variables, then there is no “pure

order $|A|$ dependence”. This explains why their joint cumulant is zero and the conclusion of Proposition 2.1.5. The joint moment, by the formula (2.5), is obtained by adding up all possible mutual dependences between random variables. In that sense, cumulants are more natural than moments in order to explore the mutual dependence between random variables. If X is a random variable, then

$$\kappa_1(X) = \mathbb{E}[X] \quad \text{and} \quad \kappa_2(X) = \text{Var}(X).$$

The cumulant $\kappa_3(X)$ and $\kappa_4(X)$ are respectively called the *skewness* and the *kurtosis* of the random variable X . They measure the asymmetry and the flatness of a given distribution, respectively.

Proposition 2.1.7 implies that Gaussian distributions are the only distributions such that all the interactions (i.e. cumulants) are of order at most 2 (expectation and covariance). It is a combinatoric interpretation of Gaussian distribution. In that sense, the collection of cumulants higher than 2 can be seen as a measure of non Gaussianity of a given distribution.

2.2 Moments and cumulants of the number of zeros

In this section, we make explicit the cumulants of the random variable that counts the number of zeros of a Gaussian process. In the following, f is a real one-dimensional Gaussian stochastic process on an probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with \mathcal{C}^1 path, such that for all $x \in \mathbb{R}$, $\text{Var}(f(x)) > 0$. For indexes of derivation $(u, v) \in \mathbb{N}^2$, we define the correlation function and its derivatives (when it makes sense) by

$$r^{(u,v)}(x, y) := \mathbb{E}[f^{(u)}(x)f^{(v)}(y)].$$

We denote the number of zeros of the process f on the interval $[a, b]$ by

$$Z_f[a, b] = \{x \in [a, b] \mid f(x) = 0\}.$$

In the previous Chapter 1, we introduced the Kac–Rice formula in Theorem 1.1.5, which gives an expression for the mean number of zeros of a stochastic process. It states that for

an interval $[a, b]$ of the real line, one has

$$\mathbb{E}[Z_f[a, b]] = \int_a^b \mathbb{E}[|f'(x)| \mid f(x) = 0] \frac{1}{\sqrt{2\pi \operatorname{Var}(f(x))}} dx.$$

There is also an expression for the higher moments of the random variable $Z_f[a, b]$, called Kac–Rice formula for higher moments. This formula expresses the factorial moment of order p of the number of zeros, as the integral of the Kac density of order p , denoted by ρ_p . The function ρ_p is defined in term of conditional expectation, and thanks to the explicit formula for Gaussian conditioning, one can make explicit the function ρ_p in terms of the covariance function r .

2.2.1 Kac–Rice formula for the factorial moments

For integers n, p we denote $n^{[p]}$ the p -th factorial power of n defined by

$$n^{[p]} := n(n-1)\dots(n-p+1).$$

Let $\rho_p : \mathbb{R}^p \rightarrow \mathbb{R}$ be the p -th order Kac density, defined for a collection $x = (x_i)_{1 \leq i \leq p}$ of p distinct points in \mathbb{R} by

$$\rho_p(x) := \mathbb{E} \left[\prod_{i=1}^p |f'(x_i)| \mid f(x_1) = \dots = f(x_p) = 0 \right] \frac{1}{\sqrt{(2\pi)^p \det \operatorname{Cov}(f(x_1), \dots, f(x_p))}}. \quad (2.6)$$

Theorem 2.2.1 (Kac–Rice formula for factorial moments). *Let f be a real Gaussian stochastic process with \mathcal{C}^1 paths, and p a positive integer. Assume that the Gaussian vector $(f(x_1), \dots, f(x_p))$ does not degenerate for any distinct numbers x_1, \dots, x_p in an interval $[a, b]$. Then*

$$\mathbb{E}[Z_f[a, b]^{[p]}] = \int_{[a, b]^p} \rho_p(x) dx.$$

A full proof can be found in [AW09], but we sketch it here for completeness.

Proof. The proof relies on Kac formula applied to the stochastic process $G : \mathbb{R}^p \rightarrow \mathbb{R}^p$

defined by

$$G(x_1, \dots, x_p) = \prod_{i=1}^p f(x_i).$$

We have the immediate equality

$$\text{Card} \{x \in [a, b]^p \mid G(x) = 0\} = \text{Card} \{x \in [a, b] \mid f(x) = 0\}^p.$$

One could hope to get the p -th moment of the number of zeros on $[a, b]$ thanks to this last formula, and Kac formula of Theorem 1.1.4 applied to the function G . Unfortunately, the process G is degenerated on the large diagonal Δ of \mathbb{R}^p defined by

$$\Delta = \{(x_1, \dots, x_p) \in \mathbb{R}^p \mid \exists i \neq j \text{ such that } x_i = x_j\},$$

and this approach does not immediately work. The strategy consists then in removing the diagonal Δ in order to apply Kac formula to a non-degenerate Gaussian process. Let us observe first the equality

$$\text{Card} \{x \in [a, b]^p \setminus \Delta \mid G(x) = 0\} = \text{Card} \{x \in [a, b] \mid f(x) = 0\}^{[p]}.$$

This can be seen directly by noting that choosing a zero of G outside of Δ is equivalent to choosing p distinct zeros of f . Let J be a compact subset of $[a, b]^p \setminus \Delta$. For any $(x_1, \dots, x_p) \in J$, the vector $(f(x_1), \dots, f(x_p))$ does not degenerate. By Kac formula of Theorem 1.1.4 applied to the function G , one gets

$$\text{Card} \{x \in J \mid G(x) = 0\} = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^p} \int_J \mathbb{1}_{|-\varepsilon, \varepsilon|^p}(G(x)) |\det \nabla G(x)| dx.$$

One can show that one can take the expectation and exchange the limits, to get by Gaussian conditioning

$$\begin{aligned} \mathbb{E} [\text{Card} \{x \in J \mid G(x) = 0\}] &= \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^p} \int_J \mathbb{E} [\mathbb{1}_{|-\varepsilon, \varepsilon|^p}(G(x)) |\det \nabla G(x)|] dx \\ &= \int_J \mathbb{E} [|\det \nabla G(x)| \mid G(x) = 0] \frac{1}{\sqrt{(2\pi)^p \det \text{Cov } G(x)}} dx. \end{aligned} \tag{2.7}$$

Note that

$$\det \nabla G(x) = \prod_{i=1}^p f(x_i) \quad \text{and} \quad \text{Cov}(G(x)) = \text{Cov}(f(x_1), \dots, f(x_p)).$$

Since (2.7) is valid for all compact subsets J of $[a, b]^p \setminus \Delta$,

$$\text{Card} \{x \in [a, b]^p \setminus \Delta \mid G(x) = 0\} = \int_{[a, b]^p \setminus \Delta} \rho_p(x) dx.$$

The set Δ has Lebesgue measure zero in \mathbb{R}^p and the conclusion follows. \square

We are now in position to compute the p -th cumulant of the random variable $Z_f[a, b]$ in terms of its factorial moments. We define the *cumulative Kac density* of order p on distinct points x_1, \dots, x_p as the quantity

$$F_p(x_1, \dots, x_p) := \sum_{\mathcal{I} \in \mathcal{P}_{\{1, \dots, p\}}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}|-1} \prod_{I \in \mathcal{I}} \rho_{|I|}((x_i)_{i \in I}).$$

Proposition 2.2.2. *One has*

$$\kappa_p(Z_f[a, b]) = \sum_{\mathcal{J} \in \mathcal{P}_{\{1, \dots, p\}}} \int_{[a, b]^{|\mathcal{J}|}} F_{|\mathcal{J}|}(x) dx.$$

Proof. By identity (2.4), one has for non negative integers k, p

$$k^p = \sum_{\mathcal{J} \in \mathcal{P}_{\{1, \dots, p\}}} k^{|\mathcal{J}|}.$$

We then compute

$$\begin{aligned} \mathbb{E}[Z_f[a, b]^p] &= \sum_{\mathcal{J} \in \mathcal{P}_{\{1, \dots, p\}}} \mathbb{E}[Z_f[a, b]^{|\mathcal{J}|}] \\ &= \sum_{\mathcal{J} \in \mathcal{P}_{\{1, \dots, p\}}} \int_{[a, b]^{|\mathcal{J}|}} \rho_{|\mathcal{J}|}(x) dx. \end{aligned}$$

The formula for cumulants and the equivalence (2.2) then yields

$$\begin{aligned}
 \kappa_p(Z_f[a, b]) &= \sum_{\mathcal{I} \in \{1, \dots, p\}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}|-1} \sum_{\mathcal{J} \preceq \mathcal{I}} \prod_{I \in \mathcal{I}} \int_{[a, b]^{|\mathcal{J}_I|}} \rho_{|\mathcal{J}_I|}(x) dx \\
 &= \sum_{\mathcal{J} \in \{1, \dots, p\}} \sum_{\mathcal{I} \succeq \mathcal{J}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}|-1} \prod_{I \in \mathcal{I}} \int_{[a, b]^{|\mathcal{J}_I|}} \rho_{|\mathcal{J}_I|}(x) dx \\
 &= \sum_{\mathcal{J} \in \{1, \dots, p\}} \int_{[a, b]^{|\mathcal{J}|}} \sum_{\mathcal{I} \succeq \mathcal{J}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}|-1} \prod_{I \in \mathcal{I}} \rho_{|\mathcal{J}_I|}((x_j)_{j \in \mathcal{J}_I}) dx \\
 &= \sum_{\mathcal{J} \in \{1, \dots, p\}} \int_{[a, b]^{|\mathcal{J}|}} F_{|\mathcal{J}|}(x) dx.
 \end{aligned}$$

□

For instance, applying the above Proposition 2.2.2, with $p = 2$, yields the following formula for the variance of the number of zeros

$$\text{Var}(Z_f[a, b]) = \left(\int_a^b \rho_1(x) dx \right) + \left(\int_a^b \int_a^b \rho_2(x, y) - \rho_1(x)\rho_1(y) dx dy \right). \quad (2.8)$$

2.2.2 Gaussian conditioning and Kac density

The expression of the Kac density ρ_p , defined in (2.6), makes use of conditional expectation. In the specific Gaussian case, one can make explicit the conditional distribution of a Gaussian vector. We consider two jointly centered Gaussian vectors X and Y , and we assume that the Gaussian vector X is non-degenerate. We define

$$\Sigma^{11} := \text{Cov}(X), \quad \Sigma^{22} := \text{Cov}(Y), \quad \Sigma^{12} := \text{Cov}(X, Y),$$

and

$$\Sigma := \text{Cov}[(X, Y)] = \begin{pmatrix} \Sigma^{11} & \Sigma^{12} \\ {}^T \Sigma^{12} & \Sigma^{22} \end{pmatrix}.$$

Since the Gaussian vector X is non-degenerate, one can define the Schur complement of Σ^{11} in Σ as the matrix

$$\Sigma^c := \Sigma^{22} - {}^T \Sigma^{12} (\Sigma^{11})^{-1} \Sigma^{12}. \quad (2.9)$$

Lemma 2.2.3. *One has*

$$\text{Law}(Y|X = 0) \sim \mathcal{N}(0, \Sigma^c).$$

Proof. We define the Gaussian vector

$$Y^c = Y - {}^T\Sigma^{12}(\Sigma^{11})^{-1}X.$$

A direct computation shows that

$$\text{Cov}(X, Y^c) = 0 \quad \text{and} \quad \text{Cov}(Y^c) = \Sigma^c.$$

Since for Gaussian vectors, decorrelation implies independence, we have the following equality of conditional distributions

$$\text{Law}(Y|X) \sim \mathcal{N}({}^T\Sigma^{12}(\Sigma^{11})^{-1}X, \Sigma^c).$$

In particular,

$$\text{Law}(Y|X = 0) \sim \mathcal{N}(0, \Sigma^c).$$

□

We now apply Lemma 2.2.3 to get an explicit expression for the function ρ_p . Let $x = (x_i)_{1 \leq i \leq p}$ be a collection of distinct points in \mathbb{R} . We choose

$$X = ((f(x_1), \dots, f(x_p))) \quad \text{and} \quad Y = (f'(x_1), \dots, f'(x_p)),$$

so that

$$\Sigma(x) = \text{Cov}(f(x_1), \dots, f(x_p), f'(x_1), \dots, f'(x_p)) = \begin{pmatrix} \Sigma^{11}(x) & \Sigma^{12}(x) \\ \Sigma^{21}(x) & \Sigma^{22}(x) \end{pmatrix}.$$

We assume that the vector $(f(x_1), \dots, f(x_p))$ is non-degenerate, so that one can consider the Schur complement $\Sigma^c(x)$ as in (2.9). Then Lemma 2.2.3 yields the following expression for the Kac density of order p .

Proposition 2.2.4. *We have*

$$\text{Law}(f'(x_1), \dots, f'(x_p) | f(x_1) = \dots = f(x_p) = 0) \sim \mathcal{N}(0, \Sigma^c(x)),$$

and

$$\rho_p(x) = \frac{1}{(2\pi)^p \sqrt{\det \Sigma(x)}} \int_{\mathbb{R}^p} |u_1| \dots |u_p| \exp\left(-\frac{1}{2} {}^T u (\Sigma^c(x))^{-1} u\right) du.$$

Proof. The first assertion is a direct application of Lemma 2.2.3. As for the second assertion, one has

$$\begin{aligned} \rho_p(x) &= \mathbb{E} \left[\prod_{i=1}^p |f'(x_i)| \mid f(x_1) = \dots = f(x_p) = 0 \right] \frac{1}{\sqrt{(2\pi)^p \det \text{Cov}(f(x_1), \dots, f(x_p))}} \\ &= \frac{1}{(2\pi)^p \sqrt{(\det \Sigma^c(x))(\det \Sigma^{11}(x))}} \int_{\mathbb{R}^p} |u_1| \dots |u_p| \exp\left(-\frac{1}{2} {}^T u (\Sigma^c(x))^{-1} u\right) du. \end{aligned}$$

The conclusion follows from the identity

$$(\det \Sigma^c(x))(\det \Sigma^{11}(x)) = \det \Sigma(x), \quad (2.10)$$

which can be deduced from a row reduction of the matrix Σ . \square

For instance, if $x \in \mathbb{R}$ then the distribution of $f'(x)$ given that $f(x) = 0$ is Gaussian, with covariance

$$\frac{r(x, x)r^{(1,1)}(x, x) - (r^{(1,0)}(x, x))^2}{r(x, x)}. \quad (2.11)$$

For p in $\{1, 2\}$, the function ρ_p has a closed formula. For distinct real numbers x, y we make explicit the conditional variance defined in (2.9) for $p = 2$ in term of the matrix

$$\Sigma^c(x, y) = \begin{pmatrix} a(x, y) & b(x, y) \\ b(x, y) & c(x, y) \end{pmatrix}.$$

Proposition 2.2.5. *One has*

$$\rho_1(x) = \frac{1}{\pi} \frac{\sqrt{(r(x, x)r^{(1,1)}(x, x) - r^{(1,0)}(x, x))^2}}{r(x, x)},$$

and

$$\rho_2(x, y) = \frac{1}{\pi^2 \det \Sigma^{11}(x, y)} \left(\sqrt{\det \Sigma^c(x, y)} + b(x, y) \arcsin \left(\frac{b(x, y)}{\sqrt{a(x, y)c(x, y)}} \right) \right).$$

Proof. The first assertion follows from (2.11), the formula

$$\frac{1}{2} \int_{\mathbb{R}} |u| \exp \left(-\frac{u^2}{2} \right) du = 1,$$

and a change of variable. As for the second formula, we can reduce by a change of variable the computation of $\rho_2(x)$ to the integral

$$I = \iint_{\mathbb{R}^2} |u||v| \exp \left(\frac{1}{2}(u^2 + v^2 + 2\delta uv) \right) dudv,$$

for some $\delta \in]-1, 1[$. We set

$$w = \frac{1}{\sqrt{2(1+\delta)}}(u+v) \quad \text{and} \quad z = \frac{1}{\sqrt{2(1-\delta)}}(u-v),$$

to get

$$I = \frac{1}{2(1-\delta^2)^{3/2}} \iint_{\mathbb{R}^2} |(1-\delta)w^2 - (1+\delta)z^2| \exp \left(-\frac{w^2 + z^2}{2} \right) dw dz.$$

Passing to polar coordinates, we obtain

$$\begin{aligned}
 I &= \frac{1}{2(1-\delta^2)^{3/2}} \left(\int_{\mathbb{R}^+} r^3 \exp\left(-\frac{r^2}{2}\right) dr \right) \left(\int_0^{2\pi} |(1-\delta)\cos^2\theta - (1+\delta)\sin^2\theta| d\theta \right) \\
 &= \frac{1}{(1-\delta^2)^{3/2}} \int_0^{2\pi} |\delta - \cos(2\theta)| d\theta \\
 &= \frac{2}{(1-\delta^2)^{3/2}} \left(\int_0^{\arccos\delta} (\cos(\theta) - \delta) d\theta + \int_{\arccos\delta}^{\pi} (\delta - \cos(\theta)) d\theta \right) \\
 &= \frac{2}{(1-\delta^2)^{3/2}} (2\sin(\arccos\delta) + \delta(\pi - 2\arccos\delta)) \\
 &= \frac{4}{(1-\delta^2)^{3/2}} (\sqrt{1-\delta^2} + \delta \arcsin\delta).
 \end{aligned}$$

The second statement follows, using the identity (2.10). □

For $p \geq 3$, there is no known closed expression for the Kac density ρ_p , which *a priori* makes the analysis of the higher moments harder than the expectation and variance. Indeed, we cannot rely anymore on a direct analysis of the expressions given by Proposition 2.2.5.

2.3 Literature on the zeros counting measure

We discuss here existing results concerning the zeros counting measure of random processes, and in particular of the model of random trigonometric polynomials, defined for a positive integer n by

$$f_n(x) = \frac{1}{\sqrt{n}} \sum_{k=0}^n a_k \cos(kx) + b_k \sin(kx), \quad (2.12)$$

where $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are sequences of random variables. We are interested in the asymptotics of the zeros counting measure as n grows to infinity. The results can be classified into three categories: asymptotics of the expectation and in the almost-sure sense, asymptotics of the variance and the associated CLT, and at last asymptotics of higher moments and cumulants.

2.3.1 Expectation

The expectation of the number of zeros of a Gaussian random process is easily computed thanks to Kac–Rice formula. For a Gaussian process f with \mathcal{C}^1 paths, the expected number

of zeros for this process is given by the first statement of Proposition 2.2.5. For a stationary Gaussian process f , the function ρ_1 is constant, and

$$\mathbb{E}[Z_f[a, b]] = \frac{|b - a|}{\pi} \sqrt{-\frac{r''(0)}{r(0)}}.$$

We apply this formula to the model of random trigonometric polynomials (2.12) with i.i.d. Gaussian coefficients, to get

$$\mathbb{E}[Z_{f_n}[a, b]] = \frac{|b - a|}{\pi} \sqrt{\frac{1}{n} \sum_{k=0}^n k^2} = \frac{|b - a|}{\pi} \sqrt{\frac{(n+1)(2n+1)}{6}}.$$

We then deduce the following asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[a, b]]}{n} = \frac{|b - a|}{\pi\sqrt{3}}.$$

This result admits extensions in several directions. First, the exact same asymptotics has been proved in [Fla17], in the case where the sequences of random variables $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are centered with unit variance. It implies a universality result for the expected number of zeros of random trigonometric polynomials. This result has also been strengthened to the following almost sure convergence in [AP21]

$$\lim_{n \rightarrow +\infty} \frac{Z_{f_n}[a, b]}{n} = \frac{|b - a|}{\pi\sqrt{3}},$$

It entails the equidistribution of the real zeros counting measure towards the rescaled Lebesgue measure on the torus, up to a factor $2/\sqrt{3}$.

In another direction, the series of papers [ADP19; APP21] have explored universality questions when one removes the independence assumption on the Gaussian sequences $(a_k)_{k \geq 0}$, $(b_k)_{k \geq 0}$. Assume that these two sequences are independent sequences of stationary Gaussian random variables with correlation given by the spectral measure μ :

$$\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \int_0^{2\pi} e^{it(k-l)} d\mu(t).$$

We assume that the measure μ admits an absolutely continuous non-zero component ψ with respect to the Lebesgue measure. Then the following asymptotics in expectation

holds

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}[Z_{f_n}[0, 2\pi]]}{n} = \frac{\lambda(\{\psi \neq 0\})}{\pi\sqrt{3}} + \frac{\lambda(\{\psi = 0\})}{\pi\sqrt{2}}.$$

The expected number of zeros for this model is non-universal and the asymptotics depends on the Lebesgue measure of the zeros set of the spectral density ψ . Nevertheless, when the spectral density is positive everywhere one recovers the same asymptotics as in the independent framework. The almost sure convergence has also been explored, as well as some particular cases of singular distributions, see [Pau20].

2.3.2 Variance and CLT

As for the expectation, the variance of the number of zeros of a real Gaussian process f also has an explicit expression in terms of the covariance function given by the second statement of Proposition 2.2.5.

The first problem which arises is the finiteness of the variance. We see in the definition of Kac density (2.6), and more explicitly in the second statement of Proposition 2.2.5, that the limit as $|x - y| \rightarrow 0$ of the quantity $\rho_2(x, y)$ is an undefined ratio. Assuming that the covariance function $r : \mathbb{R}^2 \rightarrow \mathbb{R}$ is four times differentiable, one can see by a Taylor expansion of the function ρ_2 near $x = y$ that

$$\rho_2(x, y) = O(|x - y|). \quad (2.13)$$

This ensures the local integrability of the function ρ_2 near the diagonal, and thus the finiteness of the variance. Note that (2.13) is an evidence for the local repulsion of the zeros of a regular Gaussian process. To see this, let us compute the probability that the Gaussian process f has at least two zeros in the interval $[x, x + \varepsilon]$.

$$\begin{aligned} \mathbb{P}(Z_f[x, x + \varepsilon] \geq 2) &\leq \sum_{n=0}^{+\infty} n(n-1)\mathbb{P}(Z_f[x, x + \varepsilon] = n) \\ &\leq \mathbb{E}[Z_f[x, x + \varepsilon]^{[2]}] \\ &\leq \iint_{[x, x+\varepsilon]^2} \rho_2(x, y) dx dy. \\ &= O(\varepsilon^3). \end{aligned}$$

Whereas for two independent random variables X and Y with a positive density, the probability that there are both in an interval $[x, x + \varepsilon]$ is of order ε^2 . This means that

when one looks at the zeros of f , they seem more repulsed than a sample of independent random points on the line.

In the case of a stationary process, D. Geman showed in [Gem72] that the finiteness of the variance of the number of zeros on a bounded interval is equivalent to the finiteness of the following integral, for a small parameter ε ,

$$\int_0^\varepsilon \frac{r''(t) - r''(0)}{t} < +\infty. \quad (2.14)$$

One can numerically observe the asymptotic linear growth of the variance of the number of zeros with respect to the size of the interval, for a large class of regular Gaussian processes. Given the expression (2.8), the linear growth of the variance is equivalent to the integrability of the function $F_2 : x \mapsto \rho_2(x) - \rho_1^2$ near infinity. Let x and y be two points such that the Gaussian vectors $(f(x), f'(x))$ and $(f(y), f'(y))$ are independent. Then the explicit expressions of the functions ρ_1 and ρ_2 given by Lemma 2.2.5 imply that $F_2(x - y) = 0$. This computational explanation is not satisfactory, and here is a more probabilistic heuristics of this fact. The definition of the functions ρ_1 and ρ_2 , and conditional independence imply

$$\begin{aligned} \rho_2(x - y) &= \frac{\mathbb{E}[|f'(x)||f'(y)| \mid \mid f(x) = f(y) = 0]}{2\pi\sqrt{\det \text{Cov}(f(x), f(y))}} \\ &= \frac{\mathbb{E}[|f'(x)| \mid \mid f(x) = 0]}{\sqrt{2\pi\mathbb{E}[f(x)^2]}} \frac{\mathbb{E}[|f'(y)| \mid \mid f(y) = 0]}{\sqrt{2\pi\mathbb{E}[f(y)^2]}} \\ &= \rho_1^2, \end{aligned} \quad (2.15)$$

thus $F_2(x - y) = 0$. For Gaussian processes, decorrelation implies independence and one can expect the function F_2 to be controlled by the magnitude of the functions r , r' and r'' . Indeed, a direct analysis of the function F_2 shows that for some constant C and a point x away from the origin, one has

$$|F_2(x)| \leq C \sup_{i \in \{0,2\}} |r^{(i)}(x)|^2. \quad (2.16)$$

Thus, integrability of the function F_2 is implied by the square integrability of the covariance function and its second derivative. Combined with Geman condition (2.14), it implies the

following asymptotics

$$\lim_{R \rightarrow +\infty} \frac{\text{Var}(Z_f[0, R])}{R} = \sigma(f),$$

where

$$\sigma(f) = \rho_1 + \int_{\mathbb{R}} (\rho_2(x) - \rho_1^2) dx. \quad (2.17)$$

This fact was first proved by J. Cuzick in [Cuz76]. He showed under the additional assumption that $\sigma(f) > 0$ (this assumption is satisfied for large class of Gaussian processes f , see [Lac20]) the following CLT for the number of zeros on a growing interval

$$\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\text{Var}(Z_f[0, R])} \xrightarrow{R \rightarrow +\infty} N,$$

where N is a standard Gaussian random variable. Note that the violation of the square integrability condition, for instance for processes with covariance function decreasing like $1/x^\varepsilon$, can lead to superlinear growth of the variance and non-central limit theorems, see [Slu94].

Later on, it has been established in [GW11], for the model of trigonometric polynomials with i.i.d. Gaussian coefficients, the following asymptotics of the variance of the number of zeros

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_f[a, b])}{n} = |b - a| \sigma(f_\infty),$$

where $\sigma(f_\infty)$ is given by (2.17), with f_∞ being the Paley–Wiener process, that is a stationary Gaussian process whose correlation function is given by the cardinal sine function. Heuristically, the process f_n locally converges towards the Paley–Wiener process: uniformly for $x \in \mathbb{R}$ and s, t in compact subsets of \mathbb{R} , one has

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[f_n \left(x + \frac{s}{n} \right) f_n \left(x + \frac{t}{n} \right) \right] = \frac{\sin(t - s)}{t - s} = \mathbb{E}[f_\infty(s) f_\infty(t)].$$

Hence one should expect that the number of zeros of these two processes are asymptotically related, given that the number of zeros is an additive quantity. The authors in [GW11] also show a CLT, using a method of approximation by m -dependent processes.

The fluctuations of the number of zeros of various models of Gaussian processes has been studied since, either by the Kac–Rice formula that we presented in the previous section, or by the method of decomposition into Wiener chaos. This last method relies on a L^2 decomposition on a Gaussian space of the Kac counting formula, giving the number

of zeros of the Gaussian process f . It allows to recover in a robust way a CLT for the number of zeros thanks to the fourth moment theorem of I. Nourdin, D. Nualart and G. Peccati, see [NP12]. As representative examples of applications of the methods of Gaussian chaos, we can cite the CLT for the model of random trigonometric polynomials (2.12) in [AL13], the half sum

$$\frac{1}{\sqrt{n}} \sum_{k=0}^n a_k \cos(kx),$$

see [ADL16], and the model of random orthogonal polynomials,

$$\frac{1}{\sqrt{n}} \sum_{k=0}^n a_k P_k(x),$$

where $(P_k)_{k \geq 0}$ is a sequence of orthogonal polynomials with respect to some measure on the real line, see [LP21; Do+21]. In higher dimension, it has been successfully used to obtain results on fluctuations of the nodal volume in models of arithmetic (resp. monochromatic) random waves on the torus (resp. the sphere). We refer the reader to the first part of this thesis for more details and references on these models of random waves.

In another direction, one can explore the universality of the fluctuation of the number of zeros of random trigonometric polynomials. The series of papers [BCP19; DNN20] show that when the random coefficients $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are two i.i.d. sequences of centered random variables with unit variance –non necessarily Gaussian– admitting a sufficiently high moment, one has the following variance asymptotics

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[0, 2\pi])}{n} = \sigma(f_\infty) + \frac{2}{15}(\mathbb{E}[a_0^4] - 3). \quad (2.18)$$

The variance's asymptotics is thus non-universal: it depends on the kurtosis (i.e. the fourth cumulant) $\mathbb{E}[a_0^4] - 3$ of the distribution of the random variable a_0 . Note that this term cancels in the Gaussian framework.

2.3.3 Higher moments and cumulants

As introduced in the beginning of this Section, the method of moments consists in proving the convergence towards a Gaussian random variable by proving the convergence of every moments towards the moments of a Gaussian random variable. From the equivalent point of view of cumulants, one must prove the convergence of all cumulants greater than two towards zero. In our context, the method of Wiener chaos does not easily allow to

recover the convergence of higher moments of the number of zeros. One has to prove the CLT directly by using the Kac formula for higher cumulants, given by Theorem 2.2.1 and Proposition 2.2.2.

The question of the finiteness of higher moments is more involved than the variance and the condition (2.14). A sufficient and necessary condition has been given in [Bel66; Cuz75], but it is not as easy to check as the Geman condition (2.14) for the variance. One can show by Rolle Lemma that a regularity assumption on the process f , up to a sufficiently high order, implies the finiteness of the p -th of the number of zeros, see [AW09, Thm. 3.6]. But regularity of the process is not a necessary condition. Consider a stationary process whose covariance function has the following Taylor expansion near the origin

$$r(t) = 1 - \frac{t^2}{2} + |t^3| + o(t^3).$$

In [Cuz75], J. Cuzick shows that the random variable $Z_f[a, b]$ has finite moments of every order, despite not being three times differentiable.

For moments of order $p \geq 3$, there is no closed formula for the Kac density ρ_p as for the first and second order Kac density, see Proposition 2.2.5. The analysis of the Kac density of greater order cannot rely upon such a formula anymore. The approach, taken by the authors for instance in the articles [GW11; LP21] to prove variance asymptotics, is not adaptable to higher moments. Nevertheless, the heuristics explained in the previous Subsection 2.3.2, leading to the control (2.16), can be adapted to higher moments. On some regularity condition on the covariance function, one can show, in a similar fashion, the control for $i \in \{1, \dots, p\}$ and a point $x = (x_i)_{1 \leq i \leq p}$ away from the large diagonal Δ

$$F_p(x_1, \dots, x_p) \leq C \max_{u, v \in \{0, 1\}} \max_{j \neq i} |r^{(u, v)}(x_i - x_j)|^2. \quad (2.19)$$

As for the variance, the linear growth of the p -th cumulant is directly linked to integrability properties of the cumulative Kac densities F_j for $j \leq p$, according to Proposition 2.2.2. If the covariance function r and its first partial derivatives decrease like $1/x^p$, then one expects the following asymptotics for $k \leq p$ (assuming that the process is stationary)

$$\lim_{R \rightarrow +\infty} \frac{1}{R} \int_{[0, R]^k} F_k(x_1, \dots, x_k) dx_1 \dots dx_k = \int_{\mathbb{R}^{k-1}} F_k(x_1, \dots, x_{k-1}, 0) dx_1 \dots dx_{k-1}. \quad (2.20)$$

This implies the linear growth of the cumulant of order p . Note that the linear growth of

every cumulant implies that for every $p \geq 3$,

$$\kappa_p \left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right) = O \left(\frac{1}{\sqrt{R}} \right), \quad (2.21)$$

and thus the CLT by the method of cumulants, see Proposition 2.1.7. This is more or less the path adopted in the series of papers of Ancona and Letendre [AL21a; AL21b], even though they chose the approach of moments instead of cumulants (which implies more involved combinatorics than the cumulants-based approach). The proof is heuristically simple, but the near-diagonal analysis is not straightforward and requires the introduction of divided differences, originally used in the papers [Bel66; Cuz75], in order to give alternative non-singular formulas near the large diagonal Δ .

Considering the discussion of the last two paragraphs, we deduce that the method of moments in order to prove the CLT for the number of zeros can be applied to a stationary Gaussian process whose covariance function r belongs to the Schwartz class of infinitely differentiable and fast decreasing functions. This is the main theorem of the paper [AL21b], which proves that in that framework

$$\lim_{R \rightarrow +\infty} \mathbb{E} \left[\left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right)^p \right] = \mathbb{E}[N^p],$$

where N is a standard Gaussian random variable. As a representative process in the Schwartz class, one can cite the celebrated Bargman–Fock stationary process, whose covariance function is given by

$$r(t) = \exp \left(-\frac{t^2}{2} \right).$$

This process naturally appears, in the context of random algebraic geometry, as the local scaling limit of Kostlan polynomials.

In the context of random trigonometric polynomials, the local scaling limit is the Paley–Wiener process, with sinc covariance function. The function sinc is slowly decreasing and is far out of reach of the previous framework. The estimate (2.19) is thus not sufficient to prove the asymptotics of moments and cumulants for such processes, and thus largely falls outside the scope of [AL21b].

2.4 Overview of the contributions

In this section, we describe our contributions to the asymptotics of the numbers of zeros of Gaussian processes on the real line, which lead to the series of papers [Gas21b; Gas21c].

In the first paper [Gas21b], we prove the asymptotics of the variance of the number of zeros for a large class of Gaussian processes, unifying results contained in the series of articles [GW11; AL13; ADL16; LP21], which treat the variance’s asymptotics of specific models of Gaussian processes. As representative examples, we can cite the class of stationary Gaussian processes on the real line, the model of random trigonometric polynomials with independent coefficients or the model of random orthogonal polynomials. Moreover, we explore the model of random trigonometric polynomials with dependent coefficients and we show that under a mild condition on the dependency, the asymptotics of the variance of its number of zeros is the same as in the independent framework. It can thus be seen as a continuation of the series of papers [ADP19; APP21], where the asymptotics of the expected number of zeros for this model has been established.

The second paper [Gas21c] is the natural continuation of the first one. We prove the exact asymptotics of the cumulants to any order for the number of zeros of a Gaussian process on the real line, and more generally of linear statistics associated with the zeros counting measure. Representative examples include the ones previously cited, and in particular the sinc process and the model of Gaussian trigonometric polynomials, for which the asymptotics of higher moments were only conjectured so far. This paper substantially refines the article of M. Ancona and T. Letendre [AL21b], where asymptotics of higher moments of the zeros counting measure are proved in the more restrictive framework of stationary Gaussian processes with correlation function in the Schwartz space.

2.4.1 Variance asymptotics

In the second paper [Gas21b] of this thesis, we were originally concerned with the variance’s asymptotics of the number of zeros of the model of random trigonometric polynomials with dependent coefficients defined by

$$f_n(x) = \sum_{k=0}^n a_k \cos(kt) + b_k \sin(kt),$$

where the sequences $(a_k)_{k \geq 0}$, $(b_k)_{k \geq 0}$ are independent sequences of stationary Gaussian random variables, whose correlation is given by a spectral measure μ . For natural integers k, l ,

$$\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \int_0^{2\pi} e^{it(k-l)} d\mu(t).$$

As explained in Section 2.3.1, the expected number of zeros for this model is universal as soon as the spectral measure μ has a positive density with respect to the Lebesgue measure. But the universality of the fluctuations of the number of zeros for this model is not granted, given the non-universality result (2.18) when one relaxes the Gaussian assumption. The question remained open whether the fluctuations of the number of zeros were universal with respect to the underlying dependency of the sequence of Gaussian random variables $(a_k)_{k \geq 0}$.

The first theorem of the second paper ensures the universality of the variance of the number of zeros for the model of Gaussian trigonometric polynomials with dependent coefficients.

Theorem 2.4.1. *Suppose that the spectral measure μ has a positive continuous density ψ with respect to the Lebesgue measure on the torus \mathbb{T} . Then there for any subinterval $[a, b]$ of the torus,*

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{f_n}[a, b])}{n} = |b - a| \sigma(f_\infty).$$

The variance's asymptotics, as for the expectation, is thus the same as in the Gaussian independent framework, and does not depend on the spectral measure μ , as long as it admits a continuous positive density. The heuristics of proof is very similar to the one given in Section 2.3.2 for stationary Gaussian processes, except that one must check that estimates are uniform with respect to the parameter n , and one must also deal with the fact that the process f_n is not stationary.

Firstly, one can control the near-diagonal behavior of the second order Kac density by the regularity of the process f_n (which implies the Geman condition, see (2.12) and the above discussion). Secondly, one has for a point x of the torus, and $(s, t) \in \mathbb{R}^2$ the convergence

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[f_n \left(x + \frac{t}{n} \right) f_n \left(x + \frac{s}{n} \right) \right] = \psi(x) \frac{\text{sinc}(t - s)}{t - s}.$$

In other words, the process f_n locally converges (around a point x) towards the Paley–

Wiener process f_∞ , with a factor $\psi(x)$, the (positive) spectral density at x . The covariance function of this limiting stationary process is square integrable, and the variance for its number of zeros is thus asymptotically linear with the size of the domain, see the discussion in Section 2.3.2. One can also show a uniform L^2 control on the rescaled covariance function $f_n(\cdot/n)$ and its partial derivatives, which implies, in a similar fashion, the linear growth of the variance. Note that multiplying a function by a positive quantity does not change its number of zeros. It explains why the limiting asymptotics for the variance only depends on the limit factor $\sigma(f_\infty)$, but not on the spectral density ψ .

In fact, the proof of Theorem 2.4.1 goes beyond the framework of trigonometric polynomials, and can be adapted to the previously cited models for which the asymptotics of the variance of the number of zeros is known, such as the model of random orthogonal polynomials. It leads to a more general theorem, and we refer the reader to Chapter 4 for a rigorous statement. The needed ingredients to obtain the linear asymptotics of the variance are the following one. It comprises of a regularity hypothesis on the sequence of processes in order to ensure the finiteness of the variance, the uniform square integrability of the covariance functions and its first partial derivatives, and the local convergence towards a stationary Gaussian process.

It seems likely that, under the same ingredients listed above, one could prove a CLT by the usual method of Wiener chaos expansion. Even if we would have enhanced Theorem 2.4.1 by giving the exact nature of the fluctuations, we did not choose this path for two reasons.

Firstly, the proof shows that one can obtain an exact asymptotics for the variance with only two simple natural hypothesis that are comparable (though slightly stronger) to the hypotheses of Cuzick in [Cuz76] for the stationary case. We believe that the proof is short, self-contained and more importantly does not rely on anything else than Gaussian conditioning and a little bit of not-so-involved analysis.

Secondly, we avoid using the explicit formulas of Proposition 2.2.5 and the tedious calculations that necessarily goes with, which does not bring any insight to the heuristics that hides behind the asymptotics of the variance. It also leads the author to believe that a similar proof could be given in order to prove the asymptotics of higher moments for a model of random processes satisfying good regularities assumptions and a square-integrability condition. In that sense, the variance case can be seen as a first step towards the more challenging problem of tackling the exact asymptotics of higher moments and

cumulants for the number of zeros of trigonometric polynomials, which is described below.

2.4.2 Cumulants asymptotics

The third and last paper [Gas21c] of this thesis is precisely concerned with the cumulant asymptotics of the number of zeros (and more generally the zeros counting measure) of Gaussian processes with slowly decreasing covariance function. As discussed in Section 2.3.3, the joint asymptotics of all the p -th order moment for the number of zeros of a Gaussian process is only known in the very restrictive case where the covariance function is in the Schwartz class, thus largely ruling out processes with slowly decreasing covariance functions, such as the Paley–Wiener process and the models of Gaussian trigonometric/orthogonal polynomials. In the following, N denotes a standard Gaussian random variable.

The following theorem proves the asymptotics of the higher moments and cumulants for a regular stationary Gaussian process with a covariance function satisfying a mild integrability condition.

Theorem 2.4.2. *Let f be a stationary Gaussian process with C^∞ paths and $p \geq 2$. If the covariance function r and its derivatives are square-integrable, then*

$$\lim_{R \rightarrow +\infty} \mathbb{E} \left[\left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right)^p \right] = \mathbb{E}[N^p].$$

Under the additional assumption that the covariance function r and its derivatives are in $L^q(\mathbb{R})$ for all $q > 1$, then there is an explicit constant $\gamma_p(f)$ such that

$$\lim_{R \rightarrow +\infty} \frac{\kappa_p(Z_f[0, R])}{R} = \gamma_p(f).$$

The sinc function is in $L^q(\mathbb{R})$ for all $q > 1$, and this theorem can be in particular applied to prove the asymptotics of higher moments for the Paley–Wiener process f_∞ with sinc covariance function. In that case, the asymptotics (2.21) is valid and the error term in the convergence of the moments is of order $1/\sqrt{R}$. It not only proves the polynomial concentration at any order for the number of zeros around its mean, but it also implies a rate of convergence for the moment metric, which is reminiscent of the Berry–Essen bound for more classical CLT.

In fact, our proof is sufficiently robust to be extended to models of Gaussian processes converging in some sense towards a limiting stationary process satisfying the hypotheses of Theorem 2.4.2. It leads to a more general formulation and we refer the reader to Chapter 5. It implies, for instance, the linear growth of the cumulants of the number of zeros for the model of Gaussian trigonometric polynomial with independent coefficients defined in (2.12).

Theorem 2.4.3. *For any subinterval $[a, b]$ of the torus and any positive integer p , one has*

$$\lim_{n \rightarrow +\infty} \frac{\kappa_p(Z_{f_n}[a, b])}{n} = |b - a| \gamma_p(f_\infty).$$

Thus, the asymptotics of cumulants for this model are the same as for the Paley–Wiener process f_∞ . This last theorem can also be adapted to cover the models of Gaussian trigonometric polynomials with dependent coefficients and the model of Gaussian orthogonal polynomials. In particular it proves a CLT for these models, which has only very recently been proved in the second model in [Do+21].

The proof of Theorem 2.4.2 roughly follows the heuristics given in Section 2.3.3, except that we improve the bound (2.19) the following way. Let p be a positive integer and \mathcal{I} a partition of $\{1, \dots, p\}$. Let x_1, \dots, x_p be distinct real numbers such that the collections of Gaussian vectors $((f(x_i), f'(x_i))_{i \in I})_{I \in \mathcal{I}}$, indexed by the cells of the partitions \mathcal{I} , are mutually independent. Then, a similar computation as in (2.15) shows that

$$\rho_p(x_1, \dots, x_p) = \prod_{I \in \mathcal{I}} \rho_{|I|}((x_i)_{i \in I}).$$

The cumulant cancellation property given by Proposition 2.1.5 then implies that the cumulative Kac density F_p at point x is zero, and the order of cancellation can be made precise. For a partition \mathcal{I} , we define $[i]_{\mathcal{I}}$ the cell of the partition to which the integer i belongs. One can prove the following control, away from the large diagonal Δ ,

$$F_p(x_1, \dots, x_p) \leq C \max_{u, v \in \{0, 1\}} \max_{\substack{1 \leq i, j \leq p \\ [i]_{\mathcal{I}} \neq [j]_{\mathcal{I}}}} |r^{(u, v)}(x_i - x_j)|^2,$$

valid for every partition \mathcal{I} of the set $\{1, \dots, p\}$. Note that this last bound coincides with the bound (2.19) when the partition \mathcal{I} is of the form $\{\{i\}, \{1, \dots, p\} \setminus \{i\}\}$. Together

with a fine interpolation inequality, these sets of bounds indexed by all the partitions of the set $\{1, \dots, p\}$ imply the following control, for $p \geq 3$, in the case where the covariance function r and its derivatives are square integrable,

$$\int_{[0,R]^p} F_p(x_1, \dots, x_p) dx_1 \dots dx_p = o(R^{p/2}).$$

The homogeneity of cumulants then implies that

$$\lim_{R \rightarrow +\infty} \kappa_p \left(\frac{Z_f[0, R] - \mathbb{E}[Z_f[0, R]]}{\sqrt{\text{Var}(Z_f[0, R])}} \right) = 0.$$

Under the additional hypothesis that r and its derivatives are in $L^{\frac{p}{p-1}}$, one even recovers the linear growth of the p -th cumulant (2.20).

PART II

Details of the contributions

ASYMPTOTICS FOR THE NODAL MEASURE OF RIEMANNIAN RANDOM WAVES

This chapter consists in the first published paper of this thesis [Gas21a], concerning the almost-sure asymptotics for the nodal measure of Riemannian random waves.

3.1 Introduction and main results	84
3.1.1 Introduction	84
3.1.2 Geometric and probabilistic settings	85
3.1.3 Statement of the results and outline of the proofs	90
3.2 Salem–Zygmund CLT for Riemannian random waves	93
3.2.1 Finite dimensional convergence and decorrelation estimates	94
3.2.2 Tightness estimates	100
3.3 Almost sure asymptotics of nodal volume	101
3.3.1 A Stochastic representation formula	102
3.3.2 Application of the Central Limit Theorem	103
3.3.3 Negative moment estimates for the random field	106
3.3.4 Uniform moment estimates for the nodal volume	113
3.4 Appendix	118

3.1 Introduction and main results

3.1.1 Introduction

The central limit theorem by R. Salem and A. Zygmund [SZ54] asserts that, when properly rescaled and evaluated at a uniform random point on the circle, a generic real trigonometric polynomial converges in distribution towards a Gaussian random variable. This classical result was recently revisited in [AP21] where the authors established both a quantitative version and a functional version of Salem–Zygmund theorem and then use these results to deduce the almost sure asymptotics of the number of zeros of random trigonometric polynomials with symmetric coefficients. The goal of the present article is to extend the latter results to the general Riemannian framework and more particularly to the so-called Riemannian random wave model, where random trigonometric polynomials are naturally replaced by random, Gaussian, linear combinations of Laplace eigenfunctions.

The study of nodal sets associated with Laplace eigenfunctions is the object of a vast literature, in particular thanks to Yau’s conjecture, see [Yau82; DF88] and [LM16; Log18a; Log18b] for recent breakthroughs. The introduction of probabilistic models in this context has numerous motivations in mathematical physics among which quantum chaos heuristics [Zel10], Berry’s conjecture [Ber77], which suggests that quantum chaotic eigenfunctions asymptotically behave like Euclidean random waves, and in percolation, as attests the Bogomolny-Schmit conjecture [BS02]. The most common probabilistic model then consists in considering random linear combinations of Laplace eigenfunctions, whose coefficients are independent and identically distributed standard Gaussian variables, see for instance [RW08], [Wig09] in the case of toral and spherical harmonics or [Zel09] for the case of a general Riemannian manifold.

There is a vast literature on the asymptotic behavior of nodal observables in the complex setting, see for instance [Zel09] and the references therein. In the complex case, the fast decay of the limit covariance kernels appearing in the models provides strong concentration estimates that naturally lead to almost sure asymptotics. There exist a significant technical difference between the complex case and the real case, explaining why the Riemannian analogue of long known results in the complex domain have evaded proof. The covariance kernels in the Riemannian setting are oscillatory and of rather slow decay, while in the complex domain they have exponential decay away from the diagonal. This is the main reason why, in the case of real Riemannian manifolds, most results in

the literature concern the asymptotics of nodal observables in expectation [CH20; Let16]. Coupling Borel–Cantelli Lemma with available concentration/variance estimates, one can then provide almost sure asymptotics, but only along sufficiently decreasing subsequences, see [CH20], or [MRW20] in the spherical case. In this paper, we establish the almost sure asymptotics of the nodal volume associated with Riemannian random waves, without considering a subsequence of eigenvalues.

Indeed, we consider here a generic Gaussian combination of Laplace eigenfunctions and this combination being fixed, we evaluate it at a uniform and independent random point on the manifold. Under the sole randomness of this evaluation point, we then prove that when properly normalized and localized in the neighborhood of the point, the random field statistically converges towards an explicit universal Euclidean random wave, see Section 3.2 below for precise statements. This result generalizes Salem–Zygmund’s central limit theorem to the Riemannian framework. Our method is inspired by [AP21] and makes a crucial use of Weyl type estimates and some decorrelation estimates of the limit field.

Starting from a stochastic representation formula of the nodal volume, in the spirit of Bourgain derandomization technique [Bou14; BW16], we then deduce from the above convergence, the almost-sure asymptotics of the nodal volume of a Riemannian random wave towards an explicit universal limit. This last result answers a question raised by S. Zelditch in [Zel09] about the almost sure convergence of random nodal measure. Moreover, it allows to recover and reinforce the asymptotics in expectation obtained so far in the literature, see e.g. [Let16; CH20]. Note that our approach is only based upon the almost sure convergence in distribution of the random field and some uniform moment bounds, and it does not require any variance nor concentration estimates.

3.1.2 Geometric and probabilistic settings

In order to state our main results, let us describe the geometric and probabilistic contexts and fix our notations.

Conventions

Let \mathbb{N} be the set of non-negative integers, \mathbb{R} be the set of real numbers and \mathbb{R}_+ be the set of non-negative real numbers. Let f, g, h three functions defined on an unbounded subset of \mathbb{R}_+ . We use the conventional notations o , O and \simeq the following way. We say

that $f(\lambda) = g(\lambda) + O(h(\lambda))$ if there is a constant C such that

$$\forall \lambda \in \mathbb{R}_+, \quad |f(\lambda) - g(\lambda)| \leq Ch(\lambda).$$

We say that $f(\lambda) = g(\lambda) + o(h(\lambda))$ if there is a function $\varepsilon(\lambda)$ such that

$$\forall \lambda \in \mathbb{R}_+, \quad |f(\lambda) - g(\lambda)| \leq \varepsilon(\lambda)h(\lambda) \quad \text{and} \quad \lim_{\lambda \rightarrow +\infty} \varepsilon(\lambda) = 0.$$

At last, we say that $f(\lambda) \simeq g(\lambda)$ if $\lim_{\lambda \rightarrow +\infty} |f(\lambda) - g(\lambda)| = 0$.

Geometric setting

Let (\mathcal{M}, g) be a smooth compact manifold without boundary of dimension $d \geq 2$. Without loss of generality we will assume that the associated volume measure μ is normalized i.e. $\mu(\mathcal{M}) = 1$. It is naturally equipped with the Laplace–Beltrami operator denoted Δ . The second order differential operator Δ is autoadjoint and has compact resolvent. Spectral theory asserts the existence of an orthonormal basis $(\varphi_n)_{n \in \mathbb{N}}$ of eigenfunctions of Δ associated with the ordered eigenvalues with multiplicities $(-\lambda_n^2)_{n \in \mathbb{N}}$. For all $n \in \mathbb{N}$, we can assume that λ_n is non-negative and

$$\Delta \varphi_n = -\lambda_n^2 \varphi_n \quad \text{and} \quad \int_{\mathcal{M}} \varphi_n^2 d\mu = 1.$$

Given $x, y \in \mathcal{M}$ and $\lambda \in \mathbb{R}_+$, we define

$$K_\lambda(x, y) = \sum_{\lambda_n \leq \lambda} \varphi_n(x)\varphi_n(y) \quad \text{and} \quad K_\lambda(x) := \sum_{\lambda_n \leq \lambda} \varphi_n^2(x), \quad (3.1)$$

the two-point spectral kernel projector on the eigenspace generated by the eigenfunctions up to order λ . Integrating the function $x \mapsto K_\lambda(x)$ on \mathcal{M} we obtain

$$K(\lambda) := \text{Card} \{n \in \mathbb{N} \mid \lambda_n \leq \lambda\} = \int_{\mathcal{M}} K_\lambda(x) d\mu(x),$$

the eigenvalue counting function. A fundamental tool in spectral analysis is the local Weyl law, first proved by Hörmander in [Hör68]. It describes the precise asymptotics of the two-point spectral projector. Let σ_d be the volume of the unit ball in \mathbb{R}^d :

$$\sigma_d = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2} + 1\right)},$$

and define for $x \in \mathbb{R}^d$ the function $\mathcal{B}_d : \mathbb{R} \rightarrow \mathbb{R}$ by

$$\mathcal{B}_d(\|x\|) := \frac{1}{\sigma_d} \int_{|\xi| \leq 1} e^{i\langle x, \xi \rangle} d\xi.$$

It is well-defined since the right-hand side is invariant by rotation. The local Weyl law asserts that uniformly on $x, y \in \mathcal{M}$, as λ goes to infinity,

$$K_\lambda(x, y) = \frac{\sigma_d}{(2\pi)^d} \lambda^d \mathcal{B}_d(\lambda \operatorname{dist}(x, y)) + O(\lambda^{d-1}). \quad (3.2)$$

The limit kernel \mathcal{B}_d only depends on the dimension d . It is related to the Bessel function of the first kind \mathcal{J} by the formula

$$\mathcal{B}_d(\|x\|) = \frac{1}{\sigma_d} \left(\frac{2\pi}{\|x\|} \right)^{d/2} \mathcal{J}_{\frac{d}{2}}(\|x\|).$$

The result of Hörmander goes beyond since the Weyl asymptotics is also true in the \mathcal{C}^∞ topology. For an arbitrary number of derivatives in x and y , one has

$$\partial_{\alpha, \beta} K_\lambda(x, y) = \frac{\sigma_d}{(2\pi)^d} \lambda^d \partial_{\alpha, \beta} [\mathcal{B}_d(\lambda \operatorname{dist}(x, y))] + O(\lambda^{d+|\alpha|+|\beta|-1}), \quad (3.3)$$

and the remainder is also uniform on x and y . Taking $x = y$ in the local Weyl law, one gets the following classical Weyl law on the number of eigenvalues of magnitude lower than λ :

$$K_\lambda(x) = \frac{\sigma_d}{(2\pi)^d} \lambda^d + O(\lambda^{d-1}) \quad \text{and} \quad K(\lambda) = \frac{\sigma_d}{(2\pi)^d} \lambda^d + O(\lambda^{d-1}), \quad (3.4)$$

from which one can deduce a first-order asymptotics for the n -th eigenvalue given by

$$\lambda_n \simeq 2\pi \left(\frac{n}{\sigma_d} \right)^{1/d}. \quad (3.5)$$

The function K_λ is also known as the *large band kernel* since the sum in (3.1) is over the sets of eigenvalues in $[0, \lambda]$. Now let $0 < \tau \leq 1/2$ be an exponent. One can also consider a small band setting by defining

$$k_\lambda(x, y) = \sum_{\lambda_n \in]\lambda - \lambda^\tau, \lambda]} \varphi_n(x) \varphi_n(y) \quad \text{and} \quad k(\lambda) := \operatorname{Card} \{n \in \mathbb{N} \mid \lambda - \lambda^\tau \leq \lambda_n \leq \lambda\}.$$

We can translate the local Weyl Law to this setting, at the cost of a worst rest. Precisely,

let

$$\mathcal{S}_d : \|x\| \mapsto \frac{1}{d\sigma_d} \int_{|\xi|=1} e^{i\langle x, \xi \rangle} d\xi = \mathcal{B}_{d-2}(\|x\|).$$

One has the following asymptotics, valid in the \mathcal{C}^∞ topology:

$$k_\lambda(x, y) = \frac{d\sigma_d}{(2\pi)^d} \lambda^{d+\tau-1} \mathcal{S}_d(\lambda \operatorname{dist}(x, y)) + O(\lambda^{d-1}). \quad (3.6)$$

Note that the model of monochromatic random waves, corresponding to the energy windows $[\lambda, \lambda + 1]$ is more challenging, since the rest in the Weyl law is of the same order as the dominant term. On manifolds with no conjugate points, it has been proved in [CH20; Kee19] a refined version of Weyl law that gives a rest of the form $O(\lambda^{d-1}/\log(\lambda))$ in (3.4). It is plausible that a logarithmic decay suffices to prove Theorem 3.1.1 and Theorem 3.1.2, at least in an analytic setting, but the proof would be of very different nature. The model of monochromatic random waves has been thoroughly investigated on the torus and the sphere, see for instance [Wig09; RW08; Die+20], where explicit calculations can be carried out, and refined versions of the Weyl law are available. We believe that in these particular models, some ideas of this paper can also be adapted in order to prove almost sure results concerning the convergence of nodal measure.

Probabilistic models

Let us now describe our main probabilistic models, classically known as the large band and small band Riemannian random waves models. Let us consider $(a_n)_{n \geq 0}$ a sequence of independent and identically distributed standard Gaussian random variables on a probability space $(\Omega_a, \mathcal{F}_a, \mathbb{P}_a)$. We will denote by \mathbb{E}_a the associated expectation. The two models are defined as the following Gaussian combination of eigenfunctions:

$$f_\lambda : x \mapsto \frac{1}{\sqrt{K(\lambda)}} \sum_{\lambda_n \leq \lambda} a_n \varphi_n(x) \quad \text{and} \quad \tilde{f}_\lambda : x \mapsto \frac{1}{\sqrt{k(\lambda)}} \sum_{\lambda_n \in]\lambda - \lambda^\tau, \lambda]} a_n \varphi_n(x).$$

We could also have introduced an intermediate band regime, see for instance [BW18], for which $\lambda_n \in]\alpha\lambda, \lambda]$ and $\alpha \in]0, 1[$, but for the simplicity of the statements, we choose to focus on the two "extreme" cases defined above. These processes give a probabilistic interpretation of the projector kernels introduced above since they coincide with the

covariance kernel of these processes. For all $x, y \in \mathcal{M}$ we have indeed

$$\mathbb{E}_a[f_\lambda(x)f_\lambda(y)] = \frac{K_\lambda(x, y)}{K(\lambda)} \quad \text{and} \quad \mathbb{E}_a[\tilde{f}_\lambda(x)\tilde{f}_\lambda(y)] = \frac{k_\lambda(x, y)}{k(\lambda)}.$$

Consider the canonical Euclidean space \mathbb{R}^d . For all $x \in \mathcal{M}$ we define

$$I_x : \mathbb{R}^d \longrightarrow T_x\mathcal{M},$$

an isometry between \mathbb{R}^d and the tangent space at x . We only require the mapping $x \mapsto I_x$ to be measurable. For the torus \mathbb{T}^d we can choose for I_x the canonical isometry, but in all generality there is no canonical choice (nor even a continuous choice) of a family $(I_x)_{x \in \mathcal{M}}$. Denoting \exp_x the Riemannian exponential based at $x \in \mathcal{M}$ we define

$$\Phi_x := \exp_x \circ I_x.$$

This map allows us to define a rescaled and flattened version of f_λ and \tilde{f}_λ (or any function on \mathcal{M}) around some point $x \in \mathcal{M}$ by setting

$$\begin{aligned} g_\lambda^x : \mathbb{R}^d &\longrightarrow \mathbb{R} & \tilde{g}_\lambda^x : \mathbb{R}^d &\longrightarrow \mathbb{R} \\ v &\longrightarrow f_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right] & v &\longrightarrow \tilde{f}_\lambda \left[\Phi_x \left(\frac{v}{\lambda} \right) \right]. \end{aligned}$$

In the literature the processes g_λ^x and \tilde{g}_λ^x have already been studied, see for instance [BW18; CH20; Zel09]. Thanks to the Weyl law, they converge in distribution (at a fixed point x) towards an isotropic Gaussian process whose covariance function is given by the function \mathcal{B}_d and \mathcal{S}_d respectively. In particular the limit process only depends on the topological dimension d and is independent of the base manifold \mathcal{M} .

Let us now consider a random variable X on another probability space $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$. We denote by \mathbb{P}_X its distribution and \mathbb{E}_X the associated expectation. We will assume \mathbb{P}_X to be the uniform distribution, i.e. $\mathbb{P}_X = \mu$, unless otherwise specified. The random variables $(X, (a_n)_{n \geq 0})$ lives in the product probability space $(\Omega_a \times \Omega_X, \mathcal{F}_a \wedge \mathcal{F}_X, \mathbb{P}_a \otimes \mathbb{P}_X)$. More informally, X is a random variable independent from the sequence $(a_n)_{n \geq 0}$. Randomizing on the spatial parameter x we define the following processes on \mathbb{R}^d :

$$g_\lambda^X : v \mapsto f_\lambda \left(\Phi_X \left(\frac{v}{\lambda} \right) \right) \quad \text{and} \quad \tilde{g}_\lambda^X : v \mapsto \tilde{f}_\lambda \left(\Phi_X \left(\frac{v}{\lambda} \right) \right), \quad v \in \mathbb{R}^d. \quad (3.7)$$

3.1.3 Statement of the results and outline of the proofs

The first main result of the article is the following functional central limit theorem which generalizes [AP21, Thm. 3] to the case of a general compact Riemannian manifold of dimension $d \geq 2$.

Theorem 3.1.1. *Let \mathcal{M} be a smooth compact manifold of dimension $d \geq 2$. Almost surely with respect to the probability \mathbb{P}_a , the two processes $(g_\lambda^X(v))_{v \in \mathbb{R}^d}$ and $(\tilde{g}_\lambda^X(v))_{v \in \mathbb{R}^d}$ converge in distribution under \mathbb{P}_X with respect to the \mathcal{C}^∞ topology, towards isotropic Gaussian processes $(g_\infty(v))_{v \in \mathbb{R}^d}$ and $(\tilde{g}_\infty(v))_{v \in \mathbb{R}^d}$ with respective covariance functions*

$$\mathbb{E}_X[g_\infty(u)g_\infty(v)] = \mathcal{B}_d(\|u - v\|) \quad \text{and} \quad \mathbb{E}_X[\tilde{g}_\infty(u)\tilde{g}_\infty(v)] = \mathcal{S}_d(\|u - v\|).$$

The result still holds if X is a random variable on \mathcal{M} with a bounded density with respect to the volume measure μ .

Let us emphasize that in the literature these kind of results are known only under Gaussian expectation. The concentration result obtained in [CH20] allows the authors to prove a similar result up to a subsequence of polynomial growth. Our result is new in the sense that the sole randomization on the uniform random variable X suffices to recover the asymptotic behavior of f_λ (without extracting a subsequence), and open the door to almost-sure results concerning functionals of f_λ , as demonstrates the next Theorem 3.1.2 concerning almost-sure asymptotics of the nodal volume.

The proof of Theorem 3.1.1 is the object of the next Section 3.2 and it is based upon convergence of characteristic functions. Taking the expectation under \mathbb{P}_a , the Gaussian framework allows us to make – technical but explicit – computations of characteristic functions. By a Borel–Cantelli argument we recover an almost sure convergence under \mathbb{P}_a . The proof could certainly be applied to more general settings as it uses mostly the following two main ingredients :

- The local Weyl law, which gives the limit distribution of g_λ^x (as a Gaussian process) towards the Gaussian process g_∞ .
- The statistical decorrelation of Lemma 3.2.4, which roughly states that if X and Y are independent uniform random variables on \mathcal{M} , then the associated Gaussian processes g_λ^X and g_λ^Y statistically decorrelate as λ goes to $+\infty$. It is a consequence of the decaying rate of the limit kernel \mathcal{B}_d .

As usual, a proof of convergence for stochastic processes splits into two parts. The convergence of finite dimensional distributions given by Theorem 3.2.1, and a tightness property given by Theorem 3.2.7.

The second main result of the article is the following almost-sure asymptotics of the nodal volume associated with the random fields f_λ and \tilde{f}_λ . Almost surely, for λ large enough, the nodal sets $\{f_\lambda = 0\}$ and $\{\tilde{f}_\lambda = 0\}$ are random smooth submanifolds of codimension one, thanks to Bulinskaya Lemma, see [AW09, p. 34]. We denote by \mathcal{H}^{d-1} the $(d - 1)$ -dimensional Hausdorff measure, and we set

$$\delta := (\sigma_d)^{-1/d} \quad \text{and} \quad B := B(0, \delta), \quad (3.8)$$

the Euclidean ball centered at zero with radius δ . Recall that the quantity σ_d is the volume of the unit ball in \mathbb{R}^d . The parameter δ is chosen such that the ball B has unit volume.

Theorem 3.1.2. *Almost surely with respect to the probability \mathbb{P}_a ,*

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} = \mathbb{E}_X[\mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B)],$$

and

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{\tilde{f}_\lambda = 0\})}{\lambda} = \mathbb{E}_X[\mathcal{H}^{d-1}(\{\tilde{g}_\infty = 0\} \cap B)].$$

This result improves the result [Zel09, Thm. 1] or [Let16, Thm. 1.1] about the convergence of nodal volume in expectation under \mathbb{P}_a . Passing from an almost-sure convergence to a convergence in expectation is a short corollary of our proof (see Corollary 3.3.12). This result can even be strengthened to the weak convergence of nodal measure. Let $(\mu_\lambda)_{\lambda>0}$ (resp. $(\tilde{\mu}_\lambda)_{\lambda>0}$) be the (random) nodal measure associated with $(f_\lambda)_{\lambda>0}$ (resp. $(\tilde{f}_\lambda)_{\lambda>0}$). For a continuous function $h : \mathcal{M} \rightarrow \mathbb{R}$,

$$\int_{\mathcal{M}} h(x) d\mu_\lambda(x) := \frac{1}{\lambda} \int_{\{f_\lambda=0\}} h(x) d\mathcal{H}^{d-1}(x),$$

and

$$\int_{\mathcal{M}} h(x) d\tilde{\mu}_\lambda(x) := \frac{1}{\lambda} \int_{\{\tilde{f}_\lambda=0\}} h(x) d\mathcal{H}^{d-1}(x).$$

Theorem 3.1.3. *Almost surely with respect to the probability \mathbb{P}_a , the sequence $(\mu_\lambda)_{\lambda>0}$ converges weakly to the Riemannian volume measure μ , up to an explicit factor. That is, for every continuous function $h : \mathcal{M} \rightarrow \mathbb{R}$,*

$$\lim_{\lambda \rightarrow +\infty} \int_{\mathcal{M}} h(x) d\mu_\lambda(x) = \left(\int_{\mathcal{M}} h(x) d\mu(x) \right) \mathbb{E}_X[\mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B)],$$

and

$$\lim_{\lambda \rightarrow +\infty} \int_{\mathcal{M}} h(x) d\tilde{\mu}_\lambda(x) = \left(\int_{\mathcal{M}} h(x) d\mu(x) \right) \mathbb{E}_X[\mathcal{H}^{d-1}(\{\tilde{g}_\infty = 0\} \cap B)].$$

This result positively answers the question raised by S. Zelditch in [Zel09, Cor. 2], about the asymptotics of random nodal measure. In [Let16] is considered the more general framework of random submanifolds (that is, intersection of independent Riemannian random waves). It is shown that an analogous statement of Theorem 3.1.3 for random submanifolds holds in expectation under \mathbb{P}_a . The author believes that the strategy of proof in the present paper could be translated to the framework of [Let16] to show the almost-sure asymptotics of the nodal volume of random submanifolds.

The right-hand side in Theorem 3.1.2 can be explicitly computed by the Kac–Rice formula for random fields (see the Remark 3.3.2), which yields

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d+2}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)},$$

and

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathcal{H}^{d-1}(\{\tilde{f}_\lambda = 0\})}{\lambda} = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}.$$

The proof relies on the connection between the nodal volumes of the processes f_λ and g_λ^X given by Lemma 3.3.1 which states that for large λ ,

$$\frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} \simeq \mathbb{E}_X[Z_\lambda], \quad \text{with} \quad Z_\lambda = \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B).$$

By Theorem 3.1.2 and the continuity of the random nodal volume for the \mathcal{C}^1 -topology, the continuous mapping theorem asserts that the nodal volume of g_λ^X on the ball B , denoted Z_λ , converges in distribution towards the nodal volume of g_∞ on the ball B .

To recover convergence of expectations and thus Theorem 3.1.2, it is then sufficient to prove the uniform integrability of the family $(Z_\lambda)_{\lambda>0}$. Unfortunately the process g_λ^X is not Gaussian under \mathbb{P}_X , and sufficient conditions for the boundedness of power moments in the literature (as the ones given in [Arm+19]) are too restrictive for our purpose. The approach we use to prove finiteness of all positive moments in Theorem 3.3.10 do not rely on the Kac–Rice formula, ill-devised for non Gaussian processes, but on more geometric considerations.

Thanks to a variant of the Crofton formula given by Lemma 3.3.8, we can relate the nodal volume of g_λ^X to the anti-concentration of g_λ^X around zero on deterministic points. The anti-concentration bound is given in Lemma 3.3.3 by the finiteness of a small negative moment of g_λ^X . The proof of the existence of a negative moment uses the explicit rate of convergence of characteristic function given in Lemma 3.2.8. It allows us to rewrite the convergence in term of the so-called smooth Wasserstein distance in Lemma 3.3.4 (following the approach in [Arr+17]), which is a stronger notion of convergence than the convergence in distribution.

Throughout the different proofs, C will denote a generic constant which does not depend on λ nor the sequence $(a_n)_{n>0}$, and $C(\omega)$ will denote a constant which does not depend on λ but may depend on the sequence $(a_n)_{n>0}$ (generally, a constant that comes from a Borel–Cantelli argument).

At last, we will prove the above theorems in the long band regime, that is for the process g_λ^X , but the proofs apply almost verbatim in the small band regime. The minor differences arising between the two cases will be detailed in the proofs.

3.2 Salem–Zygmund CLT for Riemannian random waves

In this section we give the proof of Theorem 3.1.1, and a few corollary results which will be of use in the study of the almost sure asymptotics of nodal volume in next Section 3.3. As usual, the proof of the functional convergence splits into the convergence of finite dimensional marginals and some tightness estimates.

3.2.1 Finite dimensional convergence and decorrelation estimates

We first establish a quantitative version of the convergence of the finite dimensional marginals of g_λ^X (resp. \tilde{g}_λ^X) towards those of g_∞ (resp. \tilde{g}_∞). We set

$$\eta(\lambda) = \begin{cases} 1 & \text{in the large band regime,} \\ \lambda^{1-\tau} & \text{in the small band regime.} \end{cases} \quad (3.9)$$

Fix an integer $p \geq 1$, $v = (v_1, \dots, v_p) \in (\mathbb{R}^d)^p$ and $t = (t_1, \dots, t_p) \in \mathbb{R}^p$, and define in the large band regime

$$N_\lambda(v, t) := \sum_{i=1}^p t_i g_\lambda^X(v_i) \quad \text{and} \quad N_\infty(v, t) := \sum_{i=1}^p t_i g_\infty(v_i),$$

and respectively in the small band regime

$$N_\lambda(v, t) := \sum_{i=1}^p t_i \tilde{g}_\lambda^X(v_i) \quad \text{and} \quad N_\infty(v, t) := \sum_{i=1}^p t_i \tilde{g}_\infty(v_i).$$

We will omit the dependence in (v, t) of $N_\lambda(v, t)$ and $N_\infty(v, t)$ when appropriate. Note that these linear combinations are Gaussian random variables under \mathbb{P}_a . We prove that the characteristic function of N_λ under \mathbb{P}_X converges to the one of N_∞ as λ goes to infinity.

Theorem 3.2.1. *Almost surely with respect to the probability \mathbb{P}_a , one has for all $t \in \mathbb{R}^p$ and $v \in (\mathbb{R}^d)^p$, in the large band regime,*

$$\lim_{\lambda \rightarrow +\infty} \mathbb{E}_X \left[e^{iN_\lambda(v, t)} \right] = \mathbb{E}_X \left[e^{iN_\infty(v, t)} \right] = \exp \left(-\frac{1}{2} \sum_{i, j=1}^p t_i t_j \mathcal{B}_a(\|v_i - v_j\|) \right),$$

and in the small band regime,

$$\lim_{\lambda \rightarrow +\infty} \mathbb{E}_X \left[e^{iN_\lambda(v, t)} \right] = \mathbb{E}_X \left[e^{iN_\infty(v, t)} \right] = \exp \left(-\frac{1}{2} \sum_{i, j=1}^p t_i t_j \mathcal{S}_a(\|v_i - v_j\|) \right).$$

Since $N_\lambda(v, t)$ is a Gaussian random variable under \mathbb{P}_a , the explicit formula for the characteristic function of a Gaussian variable gives

$$\mathbb{E}_a \left[e^{iN_\lambda(v, t)} \right] = e^{-\frac{1}{2} \mathbb{E}[N_\lambda(v, t)^2]}, \quad \text{and similarly,} \quad \mathbb{E}_X \left[e^{iN_\infty(v, t)} \right] = e^{-\frac{1}{2} \mathbb{E}[N_\infty(v, t)^2]}.$$

In order to quantify the convergence rate, for any integer $q > 0$, we set

$$\Delta_\lambda^{(q)} := \mathbb{E}_a \left[\left| \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] - \mathbb{E}_X \left[e^{iN_\infty(v,t)} \right] \right|^{2q} \right]. \quad (3.10)$$

Let K be a compact subset of \mathbb{R}^d . In the following we will assume that the vectors v_1, \dots, v_p belong to K . Recall the definition of $\eta(\lambda)$ in (3.9).

Theorem 3.2.2. *There is a constant C depending only on \mathcal{M} , K and q , such that*

$$\Delta_\lambda^{(q)} \leq C(1 + \|t\|)^{4q} \left(\frac{\eta(\lambda)}{\lambda} \right)^q. \quad (3.11)$$

The proof of Theorem 3.2.1 is a direct consequence of the second assertion in Theorem 3.2.6 at the end of this section, but observe that Theorem 3.2.2 implies a weak version of Theorem 3.2.1 and gives the core idea of the proof. Indeed, let us recall from Equation (3.5) that the sequence $(\lambda_n)_{n \geq 0}$ of eigenvalues grows as $Cn^{1/d}$. Fix some $t \in \mathbb{R}^p$ and let $\varepsilon > 0$ be a small parameter. Markov inequality implies that

$$\mathbb{P}_a \left(\left| \mathbb{E}_X \left[e^{iN_{\lambda_n}(v,t)} \right] - \mathbb{E}_X \left[e^{iN_\infty(v,t)} \right] \right| > \lambda_n^\varepsilon \sqrt{\frac{\eta(\lambda_n)}{\lambda_n}} \right) \leq \frac{\Delta_{\lambda_n}^{(q)}}{\lambda_n^{2q\varepsilon}} \left(\frac{\lambda_n}{\eta(\lambda_n)} \right)^q = O \left(n^{-2q\varepsilon/d} \right).$$

For $q > d/(2\varepsilon)$, the left-hand term is summable and Borel–Cantelli Lemma implies the existence a constant $C(\omega, v, t)$ such that

$$\left| \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] - \mathbb{E}_X \left[e^{iN_\infty(v,t)} \right] \right| \leq C(\omega, v, t) \frac{\sqrt{\eta(\lambda)}}{\lambda^{\frac{1}{2}-\varepsilon}}. \quad (3.12)$$

In particular, for a fixed $t \in \mathbb{R}^p$ and $v \in \mathbb{R}^p$, this proves the convergence in distribution of $N_\lambda(v, t)$ towards $N_\infty(v, t)$, almost surely with respect to the probability \mathbb{P}_a . Note that Theorem 3.2.1 states that the convergence holds almost surely under \mathbb{P}_a , simultaneously for all $t \in \mathbb{R}^d$ and $v \in K^p$, and thus requires the inversion of quantifiers. We deal with this issue in Theorem 3.2.6 at the end of Section 3.2, which makes explicit the dependence of $C(\omega, v, t)$ in Equation (3.12) with respect to v and t .

Proof of Theorem 3.2.2. Define

$$\tilde{\Delta}_\lambda^{(q)} := \mathbb{E}_a \left[\left| \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] - \mathbb{E}_a \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] \right|^{2q} \right]. \quad (3.13)$$

By triangular inequality, we have

$$\Delta_\lambda^{(q)} \leq 4^q \left(\tilde{\Delta}_\lambda^{(q)} + \left| e^{-\frac{1}{2}\mathbb{E}_X[N_\infty(v,t)^2]} - \mathbb{E}_X \left[e^{-\frac{1}{2}\mathbb{E}_a[N_\lambda(v,t)^2]} \right] \right|^{2q} \right).$$

Using the 1–Lipschitz regularity of $x \mapsto e^{-x}$, we then get

$$\Delta_\lambda^{(q)} \leq 4^q \tilde{\Delta}_\lambda^{(q)} + 4^{q-1} \left| \mathbb{E}_X \left[N_\infty(v,t)^2 \right] - \mathbb{E}_X \mathbb{E}_a \left[N_\lambda(v,t)^2 \right] \right|^{2q}. \quad (3.14)$$

The last term in Equation (3.14) can be evaluated as follows. The following direct computation is done in the large band regime with limit kernel \mathcal{B}_d , but it remains true in the small band regime with limit kernel \mathcal{S}_d . We have first

$$\begin{aligned} \mathbb{E}_a[N_\lambda^2] &= \mathbb{E}_a \left[\left(\sum_{i=1}^p t_i g_\lambda^X(v_i) \right)^2 \right] \\ &= \sum_{i,j=1}^p t_i t_j \frac{1}{K(\lambda)} \sum_{\lambda_n \leq \lambda} \varphi_n \left[\Phi_X \left(\frac{v_i}{\lambda} \right) \right] \varphi_n \left[\Phi_X \left(\frac{v_j}{\lambda} \right) \right] \\ &= \sum_{i,j=1}^p t_i t_j \frac{K_\lambda \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_X \left(\frac{v_j}{\lambda} \right) \right)}{K(\lambda)}. \end{aligned}$$

Using Weyl law and the fact that v lives in a compact set, we obtain

$$\begin{aligned} \left| \mathbb{E}_a \left[N_\lambda^2 \right] - \mathbb{E}_X \left[N_\infty^2 \right] \right| &\leq \sum_{i,j=1}^p |t_i| |t_j| \left| \frac{K_\lambda \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_X \left(\frac{v_j}{\lambda} \right) \right)}{K(\lambda)} - \mathcal{B}_d(\|v_i - v_j\|) \right| \\ &\leq \sum_{i,j=1}^p |t_i| |t_j| \left| \mathcal{B}_d \left[\lambda \operatorname{dist} \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_X \left(\frac{v_j}{\lambda} \right) \right) \right] - \mathcal{B}_d(\|v_i - v_j\|) \right| + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right) \\ &\leq \sum_{i,j=1}^p |t_i| |t_j| \left| \lambda \operatorname{dist} \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_X \left(\frac{v_j}{\lambda} \right) \right) - \|v_i - v_j\| \right| + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right). \end{aligned}$$

The last line is justified by the fact that the function \mathcal{B}_d is Lipschitz continuous. The differential of the exponential map at 0 is the identity, which implies the following

asymptotic, uniformly on v in a compact subset:

$$\left| \lambda \operatorname{dist} \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_X \left(\frac{v_j}{\lambda} \right) \right) - \|v_i - v_j\| \right| = O \left(\frac{1}{\lambda} \right),$$

and we deduce

$$\left| \mathbb{E}_a [N_\lambda^2] - \mathbb{E}_X [N_\infty^2] \right| = \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right).$$

Injecting this estimate in Equation (3.14), we get

$$\Delta_\lambda^{(q)} \leq 4^q \tilde{\Delta}_\lambda^{(q)} + \|t\|^{4q} O \left(\left(\frac{\eta(\lambda)}{\lambda} \right)^{2q} \right).$$

The conclusion of Theorem 3.2.2 then follows from the following lemma. □

Lemma 3.2.3. *There is a constant C depending only on \mathcal{M} , K and q , such that*

$$\tilde{\Delta}_\lambda^{(q)} \leq C(1 + \|t\|^{4q}) \left(\frac{\eta(\lambda)}{\lambda} \right)^q. \quad (3.15)$$

The proof of Lemma 3.2.3 is rather technical and for the sake of readability, it is postponed until Section 3.4 of the Appendix. To give the reader a taste of the arguments involved, the proof is essentially based on explicit computations of characteristic functions and the key argument is the following decorrelation Lemma 3.2.4. With the same notations as above, let Y be a uniform random variable in \mathcal{M} , independent of X and of the Gaussian coefficients (a_k) . Let us set

$$N_\lambda^X := \sum_{j=1}^p t_j g_\lambda^X(v_j), \quad N_\lambda^Y := \sum_{j=1}^p t_j g_\lambda^Y(v_j),$$

in the large band regime and respectively in the small band regime

$$N_\lambda^X := \sum_{j=1}^p t_j \tilde{g}_\lambda^X(v_j), \quad N_\lambda^Y := \sum_{j=1}^p t_j \tilde{g}_\lambda^Y(v_j).$$

Lemma 3.2.4. *There is a constant C depending only on \mathcal{M} and K , such that*

$$\mathbb{E}_X \left[\left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| \right] \leq C \|t\|^2 \frac{\eta(\lambda)}{\lambda}.$$

Proof of Lemma 3.2.4. An explicit computation and the Weyl law give

$$\begin{aligned} \left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| &= \left| \sum_{i,j=1}^p t_i t_j \frac{1}{K_\lambda} \sum_{\lambda_n \leq \lambda} \varphi_n \left[\Phi_X \left(\frac{v_i}{\lambda} \right) \right] \varphi_n \left[\Phi_Y \left(\frac{v_j}{\lambda} \right) \right] \right| \\ &\leq \sum_{i,j=1}^p |t_i| |t_j| \left| \mathcal{B}_d \left(\lambda \operatorname{dist} \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_Y \left(\frac{v_j}{\lambda} \right) \right) \right) \right| + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right), \end{aligned}$$

and the remainder is uniform on X, Y . Again, the above computation is done in the large band regime with limit kernel \mathcal{B}_d , but it holds in the small band regime with limit kernel \mathcal{S}_d . Define

$$c_\lambda := \lambda \operatorname{dist} \left(\Phi_X \left(\frac{v_i}{\lambda} \right), \Phi_Y \left(\frac{v_j}{\lambda} \right) \right) - \lambda \operatorname{dist}(X, Y).$$

By triangle inequality, $|c_\lambda|$ is bounded by $2|K|$, where $|K|$ is the diameter of the compact subset K in which lives v_1, \dots, v_p . It follows that

$$\left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| \leq \sum_{i,j=1}^p |t_i| |t_j| \left| \mathcal{B}_d \left(\lambda \operatorname{dist}(X, Y) + c_\lambda \right) \right| + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right).$$

Taking the expectation with respect to X we obtain

$$\begin{aligned} \mathbb{E}_X \left[\left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| \right] &\leq \int_M \sum_{i,j=1}^p |t_i| |t_j| \left| \mathcal{B}_d \left(\lambda \operatorname{dist}(x, Y) + c_\lambda \right) \right| d\mu(x) + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right) \\ &\leq \sum_{i,j=1}^p |t_i| |t_j| \left(\int_{\operatorname{dist}(x,Y) \leq \varepsilon} \left| \mathcal{B}_d \left(\lambda \operatorname{dist}(x, Y) + c_\lambda \right) \right| d\mu(x) \right. \\ &\quad \left. + \int_{\operatorname{dist}(x,Y) > \varepsilon} \left| \mathcal{B}_d \left(\lambda \operatorname{dist}(x, Y) + c_\lambda \right) \right| d\mu(x) \right) + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right) \\ &\leq \sum_{i,j=1}^p |t_i| |t_j| (I_1 + I_2) + \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right), \end{aligned} \tag{3.16}$$

where I_1 and I_2 are the two integrals appearing in the last expression. For ε small enough

we can pass in local polar coordinates into the first integral I_1 . We obtain

$$\begin{aligned} I_1 &\leq d \sigma_d \int_0^\varepsilon \sup_{c \in [-2|K|, 2|K|]} |\mathcal{B}_d(\lambda r + c)| (1 + O(r^2)) r^{d-1} dr \\ &\leq \frac{C}{\lambda^d} \int_0^{\lambda \varepsilon} \sup_{c \in [-2|K|, 2|K|]} |\mathcal{B}_d(u + c)| u^{d-1} du. \end{aligned} \quad (3.17)$$

We use the following asymptotics for \mathcal{B}_d and \mathcal{S}_d at infinity:

$$\begin{aligned} \mathcal{B}_d(u) &= C u^{-\frac{d+1}{2}} \sin\left(u - \frac{d-1}{4}\pi\right) + O\left(u^{-\frac{d+3}{2}}\right), \\ \mathcal{S}_d(u) &= C u^{-\frac{d-1}{2}} \sin\left(u - \frac{d-3}{4}\pi\right) + O\left(u^{-\frac{d+1}{2}}\right). \end{aligned}$$

Injecting these asymptotics into expression (3.17) we obtain

$$I_1 = O\left(\frac{\eta(\lambda)}{\lambda}\right).$$

For the second integral and λ large enough, we use the fact that $|c_\lambda| \leq 2|K|$ and the asymptotic formula for \mathcal{B}_d (resp. \mathcal{S}_d) to obtain

$$I_2 \leq \sup_{t \geq \varepsilon} \sup_{c \in [-2|K|, 2|K|]} |\mathcal{B}_d(\lambda t + c)|,$$

from which we deduce

$$I_2 = O\left(\frac{\eta(\lambda)}{\lambda}\right).$$

Finally we recover from inequality (3.16) and the definition (3.9) that

$$\mathbb{E}_X \left[\left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| \right] \leq C \|t\|^2 \frac{\eta(\lambda)}{\lambda}.$$

□

Remark 3.2.5. if X is not a uniform random variable on M , but has a bounded density h with respect to the volume measure μ , then Equation (3.16) becomes

$$\mathbb{E}_X \left[\left| \mathbb{E}_a \left[N_\lambda^X N_\lambda^Y \right] \right| \right] \leq \|h\|_\infty \sum_{i,j=1}^p |t_i| |t_j| (I_1 + I_2) + \|t\|^2 O\left(\frac{\eta(\lambda)}{\lambda}\right),$$

and the end of the proof of Lemma 3.2.4 remains unchanged. Throughout this section, this is the only difference that arises when X is not a uniform distribution on \mathcal{M} .

In the following application of Theorem 3.1.1 to nodal volume, we will need finer estimates on the constant $C(\omega, v, t)$ in Equation (3.12). The Borel–Cantelli Lemma does not allow to track the dependence of $C(\omega, v, t)$ with respect to the parameters v and t . It is the content of the following theorem, proved in Appendix 3.4. The proof relies of Sobolev injections in order to control the supremum norm by some $W^{k,1}$ norm, which is more convenient to work with when taking the expectation under \mathbb{P}_a .

Theorem 3.2.6. *Fix $\varepsilon > 0$. There is a constant $C(\omega)$ depending only K and ε , such that*

$$\sup_{v \in K} \left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right| \leq C(\omega)(1 + |t|^{2+\varepsilon}) \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2-\varepsilon}.$$

And more generally,

$$\sup_{v \in K} \left| \mathbb{E}_X \left[e^{iN_\lambda(v,t)} \right] - e^{-\frac{1}{2}\mathbb{E}_X[N_\infty(v,t)^2]} \right| \leq C(\omega)(1 + \|t\|^{2+\varepsilon}) \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2-\varepsilon}.$$

3.2.2 Tightness estimates

We now turn to the proof of the tightness for the family $(g_\lambda^X)_{\lambda>0}$. Recall the definition of the ball B in (3.8).

Theorem 3.2.7. *Almost surely with respect to the probability \mathbb{P}_a , the family of stochastic processes $(g_\lambda^X)_{\lambda>0}$ is tight with respect to the Frechet topology on $\mathcal{C}^\infty(B)$.*

The tightness in \mathcal{C}^1 topology is sufficient for the rest of the article but the proof of \mathcal{C}^∞ tightness does not cost any more calculations. The proof is short once we proved the following lemma.

Lemma 3.2.8. *Let p be a positive integer, and α a d -dimensional multi-index. There is a constant $C(\omega)$ depending only p and α such that*

$$\mathbb{E}_X \left[\int_B |\partial_\alpha g_\lambda^X(v)|^{2p} dv \right] \leq C(\omega).$$

The proof of Lemma 3.2.8 is given in the Appendix 3.4 and relies on hypercontractivity and a Borel–Cantelli argument.

Proof of Theorem 3.2.7. By Kolmogorov tightness criterion for stochastic processes (see [Kun97, p. 39]) in dimension d with C^∞ topology, it suffices to show that for every multi-index of differentiation β , for some $p > d/2$, and for all $u, v \in B$,

$$\mathbb{E}_X \left[\left| \partial_\beta g_\lambda^X(v) - \partial_\beta g_\lambda^X(u) \right|^{2p} \right] \leq C(\omega) \|v - u\|^{2p}.$$

We use the mean-value Theorem and Sobolev injection to get

$$\begin{aligned} \mathbb{E}_X \left[\left(\frac{\partial_\beta g_\lambda^X(v) - \partial_\beta g_\lambda^X(u)}{\|v - u\|} \right)^{2p} \right] &\leq C \sum_{k=1}^d \mathbb{E}_X \left[\left(\sup_{u \in B} |\partial_k \partial_\beta g_\lambda^X| \right)^{2p} \right] \\ &\leq C \sum_{k=1}^d \mathbb{E}_X \left[\left(\|\partial_k \partial_\beta g_\lambda^X\|_{W^{d+1,1}} \right)^{2p} \right] \\ &\leq C \sum_{|\alpha| \leq |\beta| + d + 2} \mathbb{E}_X \left[\left(\int_B |\partial_\alpha g_\lambda^X(u)| du \right)^{2p} \right] \\ &\leq C \sum_{|\alpha| \leq |\beta| + d + 2} \mathbb{E}_X \left[\int_B |\partial_\alpha g_\lambda^X(u)|^{2p} du \right]. \end{aligned}$$

From Lemma 3.2.8, we have then

$$\mathbb{E}_X \left[\int_B |\partial_\alpha g_\lambda^X(u)|^{2p} du \right] \leq C(\omega),$$

hence the result. \square

3.3 Almost sure asymptotics of nodal volume

As already mentioned above, almost surely in the random coefficients, the nodal sets $\{f_\lambda = 0\}$ and $\{\tilde{f}_\lambda = 0\}$ associated with the random wave models are random smooth

submanifolds of codimension one. The object of this section is to give the proof of Theorem 3.1.2 on the almost sure asymptotics of the associated nodal volume.

3.3.1 A Stochastic representation formula

The first step in the proof of Theorem 3.1.2 consists in connecting the zeros of f_λ (resp. \tilde{f}_λ) to the zeros of g_λ^X (resp. \tilde{g}_λ^X). This is the object of the following lemma. In the following, $\delta = (\sigma_d)^{-1/d}$ (recall its definition (3.8)).

Lemma 3.3.1. *Let $f : \mathcal{M} \rightarrow \mathbb{R}$ a smooth function such that 0 is a regular value of f . Then*

$$\mathbb{E}_X \left[\mathcal{H}^{d-1} \left(\{f = 0\} \cap B \left(X, \frac{\delta}{\lambda} \right) \right) \right] = \left(1 + O \left(\frac{1}{\lambda^2} \right) \right) \frac{\mathcal{H}^{d-1}(\{f = 0\})}{\lambda^d}.$$

The remainder term is uniform on the function f . More generally, if h is a continuous function on \mathcal{M} , then

$$\begin{aligned} \mathbb{E}_X \left[h(X) \mathcal{H}^{d-1} \left(\{f = 0\} \cap B \left(X, \frac{\delta}{\lambda} \right) \right) \right] &= \left(1 + O \left(\frac{1}{\lambda^2} \right) \right) \frac{1}{\lambda^d} \int_{\{f=0\}} h(x) \mathcal{H}^{d-1}(x) \\ &\quad + O \left(\omega_h \left(\frac{\delta}{\lambda} \right) \right) \frac{\mathcal{H}^{d-1}(\{f = 0\})}{\lambda^d}, \end{aligned}$$

where ω_h denotes the modulus of continuity of h .

Proof. Since 0 is a regular value of f , then $\mathcal{H}^{d-1}(\{f = 0\}) < +\infty$. We have

$$\begin{aligned} \mathbb{E}_X \left[h(X) \mathcal{H}^{d-1} \left(\{f = 0\} \cap B \left(X, \frac{\delta}{\lambda} \right) \right) \right] &= \int_{\mathcal{M}} h(x) \mathcal{H}^{d-1} \left(\{f = 0\} \cap B \left(x, \frac{\delta}{\lambda} \right) \right) d\mu(x) \\ &= \int_{\mathcal{M}} \int_{\{f=0\}} h(x) \mathbf{1}_{B(x, \frac{\delta}{\lambda})}(y) d\mathcal{H}^{d-1}(y) d\mu(x) \\ &= \int_{\{f=0\}} \left(\int_{B(y, \frac{\delta}{\lambda})} h(x) d\mu(x) \right) d\mathcal{H}^{d-1}(y) \end{aligned}$$

The uniform continuity of h implies the following equality

$$\int_{B(y, \frac{\delta}{\lambda})} h(x) d\mu(x) = \left(h(y) + O \left(\omega_h \left(\frac{\delta}{\lambda} \right) \right) \right) \text{Vol}_{\mathcal{M}} \left(B \left(y, \frac{\delta}{\lambda} \right) \right).$$

Standard comparison theorem for geodesic ball asserts that uniformly on x ,

$$\begin{aligned} \text{Vol}_{\mathcal{M}} \left(B \left(x, \frac{\delta}{\lambda} \right) \right) &= \text{Vol}_{\mathbb{R}^d} \left(B \left(0, \frac{\delta}{\lambda} \right) \right) \left(1 + O \left(\frac{1}{\lambda^2} \right) \right) \\ &= \frac{1}{\lambda^d} \left(1 + O \left(\frac{1}{\lambda^2} \right) \right), \end{aligned}$$

from which we conclude. \square

Note that, alternatively, we could have proved the asymptotic representation formula given by Lemma 3.3.1 using the closed Kac–Rice formula for manifolds in [Jub19].

3.3.2 Application of the Central Limit Theorem

The next step in the proof of Theorem 3.1.2 then consists in using the central limit theorem as established in Section 3.2. We define the mapping

$$\begin{aligned} \Phi_x^{(\lambda)} : B &\longrightarrow \mathcal{M} \\ v &\longrightarrow \Phi_x \left(\frac{v}{\lambda} \right). \end{aligned}$$

Choosing $f = f_\lambda$ in Lemma 3.3.1 and recalling the relation (3.7) between g_λ and f_λ , we obtain

$$\begin{aligned} \frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda^d} &= \mathbb{E}_X \left[\mathcal{H}^{d-1} \left(\{f_\lambda = 0\} \cap B \left(X, \frac{\delta}{\lambda} \right) \right) \right] \left(1 + O \left(\frac{1}{\lambda^2} \right) \right) \\ &= \mathbb{E}_X \left[\mathcal{H}^{d-1} \left[\Phi_X^{(\lambda)} \left(\{g_\lambda^X = 0\} \cap B \right) \right] \right] \left(1 + O \left(\frac{1}{\lambda^2} \right) \right). \end{aligned} \quad (3.18)$$

The mapping $\Phi_x^{(\lambda)}$ is a diffeomorphism onto its image for λ small enough and uniformly on $x \in \mathcal{M}$. The exponential map is a local diffeomorphism and its differential at zero is the identity. We deduce that the mapping $\Phi_x^{(\lambda)}$ is bi-Lipschitz, and uniformly on $x \in \mathcal{M}$,

$$\text{Lip} \left(\Phi_x^{(\lambda)} \right) = \frac{1}{\lambda} \left(1 + O \left(\frac{1}{\lambda} \right) \right) \quad \text{and} \quad \text{Lip} \left((\Phi_x^{(\lambda)})^{-1} \right) = \lambda \left(1 + O \left(\frac{1}{\lambda} \right) \right).$$

Using scaling properties of Hausdorff measures under bi-Lipschitz mappings we obtain

$$\mathcal{H}^{d-1} \left[\Phi_X^{(\lambda)} \left(\{g_\lambda^X = 0\} \cap B \right) \right] = \frac{1}{\lambda^{d-1}} \mathcal{H}^{d-1} \left[\{g_\lambda^X = 0\} \cap B \right] \left(1 + O \left(\frac{1}{\lambda} \right) \right),$$

and from expression (3.18) it follows that

$$\frac{\mathcal{H}^{d-1}(\{f_\lambda = 0\})}{\lambda} = \mathbb{E}_X \left[\mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B) \right] \left(1 + O\left(\frac{1}{\lambda}\right) \right), \quad (3.19)$$

and more generally for a continuous function h on \mathcal{M} ,

$$\begin{aligned} \frac{1}{\lambda} \int_{\{f=0\}} h(x) \mathcal{H}^{d-1}(x) &= \mathbb{E}_X \left[h(X) \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B) \right] \left(1 + O\left(\frac{1}{\lambda}\right) \right) \\ &\quad + O\left(\omega_h\left(\frac{\delta}{\lambda}\right)\right) \frac{\mathcal{H}^{d-1}(\{f = 0\})}{\lambda}. \end{aligned} \quad (3.20)$$

The function $g \mapsto \mathcal{H}^{d-1}(\{g = 0\} \cap B)$ is continuous on the set of functions such that 0 is a regular value on B , endowed with the \mathcal{C}^1 topology (see [APP18, Thm. 3]). The limit process g_∞ is non-degenerate since the limit kernels \mathcal{B}_d and \mathcal{S}_d are positive definite covariance functions, and Bulinskaya Lemma (see [AW09, p. 34]) asserts that \mathbb{P}_a -almost surely, the point 0 is a regular value for the process g_∞ . Define

$$Z_\lambda := \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B) \quad \text{and} \quad Z_\infty := \mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B).$$

The continuous mapping theorem and the convergence in distribution of Theorem 3.1.1 imply the following convergence in distribution under \mathbb{P}_X :

$$\mathbb{P}_a - \text{a.s.}, \quad Z_\lambda \xrightarrow{\mathbb{P}_X} Z_\infty. \quad (3.21)$$

Theorem 3.1.2 about convergence of nodal volume is proved if we can pass to the convergence of expectations under \mathbb{P}_X in (3.21), according to the stochastic representation formulas (3.19). Passing to the expectation follows from the uniform integrability (with respect to \mathbb{P}_X) of the family of random variables $(Z_\lambda)_{\lambda>0}$. This last point is the object of the next Sections 3.3.3 and 3.3.4.

Let \tilde{X} be another random variable on \mathcal{M} with a density h with respect to the volume measure μ . Then

$$\mathbb{E}_X \left[h(X) \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap B) \right] = \mathbb{E}_{\tilde{X}} \left[\mathcal{H}^{d-1}(\{g_\lambda^{\tilde{X}} = 0\} \cap B) \right].$$

We deduce that Theorem 3.1.3 about convergence of nodal measure is proved if we can pass to the convergence of expectations under \mathbb{P}_X in (3.21), but this time with X a random

variable on \mathcal{M} with a continuous density h with respect to the volume measure μ . Since the functional central limit Theorem (3.1.1) and all the theorems of Section 3.2 remain valid when X is a random variable with bounded density h (see Remark 3.2.5), the proof of Theorem 3.1.3 is the same as the proof of Theorem 3.1.2.

Remark 3.3.2. The quantity $\mathbb{E}_X[\mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B)]$ in Theorem 3.1.2 has an explicit value, thanks to the Kac–Rice formula. We roughly sketch the proof here (see [AW09, p. 177] for more details). Taking the expectation in the co-area formula gives

$$\int_{\mathbb{R}} \varphi(y) \mathbb{E}_X [\mathcal{H}^{d-1}(\{g_\infty = y\} \cap B)] dy = \int_B \mathbb{E}_X [\varphi(g_\infty(x)) \|\nabla_x g_\infty\|] d\mu(x).$$

The Gaussian process g_∞ is stationary, hence its law does not depend on the point x . The Gaussian variables $g_\infty(x), \partial_1 g_\infty(x), \dots, \partial_d g_\infty(x)$ are independents. Hence,

$$\begin{aligned} \int_{\mathbb{R}} \varphi(y) \mathbb{E}_X [\mathcal{H}^{d-1}(\{g_\infty = y\} \cap B)] dy &= \text{Vol}(B) \mathbb{E}_X[\varphi(g_\infty)] \mathbb{E}_X[\|\nabla g_\infty\|] \\ &= \mathbb{E}_X[\|\nabla g_\infty\|] \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(y) e^{-\frac{y^2}{2}} dy, \end{aligned}$$

and we deduce that for almost all $y \in \mathbb{R}$,

$$\mathbb{E}_X [\mathcal{H}^{d-1}(\{g_\infty = y\} \cap B)] = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \mathbb{E}_X[\|\nabla g_\infty\|].$$

It is actually true for all $y \in \mathbb{R}$, and this is the difficult part of the proof which we do not detail. An direct computation gives

$$\mathbb{E}_X [(\partial_1 g_\infty)^2] = \dots = \mathbb{E}_X [(\partial_d g_\infty)^2] = \frac{1}{d+2},$$

and

$$\mathbb{E}_X[\|\nabla g_\infty\|] = \sqrt{\frac{2}{d+2}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}.$$

Taking $y = 0$ we deduce

$$\mathbb{E}_X [\mathcal{H}^{d-1}(\{g_\infty = 0\} \cap B)] = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d+2}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)},$$

When $d = 1$ we recover the classical asymptotics $\frac{1}{\pi\sqrt{3}}$ for the number of real roots of a

random trigonometric polynomial. For the process \tilde{g}_∞ , we have

$$\mathbb{E}_X [(\partial_1 g_\infty)^2] = \dots = \mathbb{E}_X [(\partial_d g_\infty)^2] = \frac{1}{d},$$

which gives

$$\mathbb{E}_X [\mathcal{H}^{d-1}(\{\tilde{g}_\infty = 0\} \cap B)] = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d}} \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})}.$$

3.3.3 Negative moment estimates for the random field

The uniform integrability of the volume of the nodal set can be deduced from anti-concentration of the stochastic process g_λ^X around zero. If the manifold were real-analytic, it would be sufficient to have the finiteness of a logarithmic moment, which is the approach taken in [AP21], see Remark 3.3.11 below. Since we consider here \mathcal{C}^∞ manifolds, we need a stronger control, given by the following lemma.

Lemma 3.3.3. *Let $\nu < \frac{1}{20\tau d}$ a small exponent. There is a constant $C(v, \omega)$ such that*

$$\sup_{\lambda > 0} \mathbb{E}_X [|g_\lambda^X(v)|^{-\nu}] < C(v, \omega).$$

Let $\alpha > 0$, $\varepsilon > 0$ and $(v_i)_{i \in \mathbb{N}}$ be any sequence in B . There is a constant $C(\omega)$ (also depending on α , ε and the sequence $(v_i)_{i \in \mathbb{N}}$) such that

$$\sup_{\lambda > 0} \int_1^{+\infty} \frac{1}{t^{1+\alpha+\varepsilon}} \sum_{i=0}^{\lfloor t^\alpha \rfloor} \mathbb{E}_X [|g_\lambda^X(v_i)|^{-\nu}] dt < C(\omega).$$

The second technical assertion is a refinement of the first one and will be used in the final step of the proof of uniform integrability. It compensates the fact that the constant $C(v, \omega)$ may depend on v , see also Remark 3.3.7 below.

The proof of Lemma 3.3.3 relies on the two following lemmas, which relate the speed of convergence of characteristic functions given in Theorem 3.2.6 to more classical distances on the space of measures. The first lemma compares the Kolmogorov distance and the so-called smooth Wasserstein distance.

Lemma 3.3.4. *Given two random variables X, Y , and $\alpha \in \mathbb{N}$, we set*

$$\text{Wass}_{(\alpha)}(X, Y) := \sup \left\{ |\mathbb{E}[\phi(X)] - \mathbb{E}[\phi(Y)]| \mid \phi \in \mathcal{C}^\alpha(\mathbb{R}), \|\phi\|_\infty \leq 1, \dots, \|\phi^{(\alpha)}\|_\infty \leq 1 \right\},$$

and

$$\text{Kol}(X, Y) := \sup_{t \in \mathbb{R}} |\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)|.$$

If Y has a density bounded by M , there is a constant C depending only on M and α such that :

$$\text{Kol}(X, Y) \leq \min \left(1, C \text{Wass}_{(\alpha)}(X, Y)^{\frac{1}{\alpha+1}} \right).$$

Proof. Fix some $t \in \mathbb{R}$. Let $0 < \varepsilon < 1$, and consider $\varphi \in \mathcal{C}^\alpha(\mathbb{R})$ a nonincreasing function such that

$$\varphi(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 0 & \text{if } x \geq 1 \end{cases}.$$

Define $\varphi_\varepsilon : x \mapsto \varphi((x - t)/\varepsilon)$, which is an upper \mathcal{C}^α approximation of $\mathbf{1}_{]-\infty, t]}$. Then

$$\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t) \leq (\mathbb{E}[\varphi_\varepsilon(X)] - \mathbb{E}[\varphi_\varepsilon(Y)]) + (\mathbb{E}[\varphi_\varepsilon(Y)] - \mathbb{P}(Y \leq t)).$$

For the first term, observe that $\|\varphi_\varepsilon^{(k)}\|_\infty = \varepsilon^{-k} \|\varphi^{(k)}\|_\infty$, and thus there is a constant C such that

$$\mathbb{E}[\varphi_\varepsilon(X)] - \mathbb{E}[\varphi_\varepsilon(Y)] \leq \frac{C}{\varepsilon^\alpha} \text{Wass}_{(\alpha)}(X, Y).$$

For the second term,

$$\mathbb{E}[\varphi_\varepsilon(Y)] - \mathbb{P}(Y \leq t) \leq M\varepsilon.$$

We can make the same computations with a lower \mathcal{C}^α approximation of $\mathbf{1}_{]-\infty, t]}$, which gives a similar lower bound on the quantity $\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)$. Optimizing in ε we obtain the desired bound. \square

The second lemma relates the smooth Wasserstein distance and the rate of convergence of characteristic functions.

Lemma 3.3.5. *Let $(X_n)_{n \geq 0}$ be a sequence of random variables converging in distribution towards a random variable X . Assume that for some exponents $m \in \mathbb{N}$ and $\alpha \in \mathbb{R}_+$ there is a constant C such that*

$$\left| \mathbb{E} \left[e^{itX_n} \right] - \mathbb{E} \left[e^{itX} \right] \right| \leq C \frac{1 + |t|^m}{n^\alpha},$$

and for some exponent $\beta > 0$:

$$\sup_{n \in \mathbb{N}} \mathbb{E}[|X_n|^\beta] < +\infty.$$

Then there is a constant C depending on m, α, β such that :

$$\text{Wass}_{(m+1)}(X_n, X) \leq C n^{-\frac{2\alpha\beta}{2\beta+1}}.$$

A general form of the theorem can be found in [Arr+17], but we prove it in Appendix 3.4 for completeness.

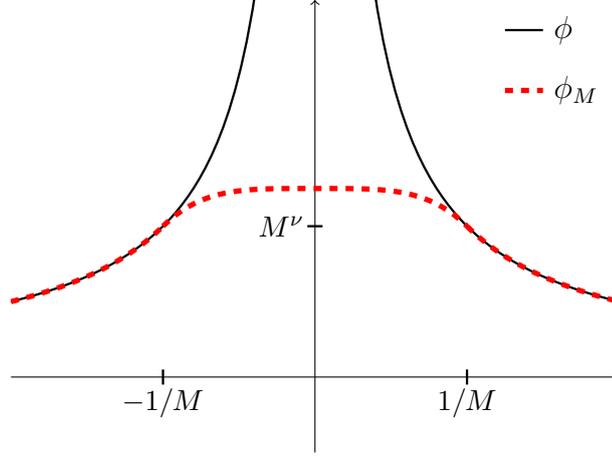
Remark 3.3.6. Denote $\text{Wass}_{(\alpha)}^X$ is the smooth Wasserstein distance under \mathbb{P}_X , and let N be a standard Gaussian random variable. Lemma 3.3.5 and the rate of convergence given by Theorem 3.2.6 imply that for some $\varepsilon > 0$ the existence of a constant $C(\omega)$ independent of $v \in B$, such that

$$\text{Wass}_{(4)}^X(g_\lambda^X(v), N) \leq C(\omega) \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2-\varepsilon}.$$

The moment condition is satisfied for every $\beta > 0$ and uniformly in $v \in B$, by Sobolev injection and Lemma 3.2.8.

We are now in position to give the proof of Lemma 3.3.3 on the negative moment of the random field g_λ .

Proof of Lemma 3.3.3 . We define $\phi : x \mapsto |x|^{-\nu}$. Let ϕ_M be a $\mathcal{C}^\infty(\mathbb{R})$ approximation of ϕ , which coincide on $\mathbb{R} \setminus [-\frac{1}{M}, \frac{1}{M}]$. We can choose the function ϕ_M such that for all $p \in \mathbb{N}$, $\|\phi_M^{(p)}\|_\infty \leq C_p M^{\nu+p}$ (see Figure 3.1).


 Figure 3.1 – The functions ϕ and ϕ_M .

Let N be a standard Gaussian random variable. We write

$$\begin{aligned} \mathbb{E}_X \left[|g_\lambda^X(v)|^{-\nu} \right] - \mathbb{E}_X \left[|N|^{-\nu} \right] &= \underbrace{\mathbb{E}_X [\phi(g_\lambda^X(v)) - \phi_M(g_\lambda^X(v))]}_{\Delta_1} + \underbrace{\mathbb{E}_X [\phi_M(g_\lambda^X(v)) - \phi_M(N)]}_{\Delta_2} \\ &\quad + \underbrace{\mathbb{E}_X [\phi(N) - \phi_M(N)]}_{\Delta_3}. \end{aligned}$$

For the term Δ_3 , we use Cauchy–Schwarz inequality to obtain

$$|\mathbb{E}_X[\phi(N)] - \mathbb{E}_X[\phi_M(N)]| \leq \mathbb{E}_X \left[(\phi - \phi_M)(N) \mathbb{1}_{|N| \leq \frac{1}{M}} \right] \leq \sqrt{\frac{\mathbb{E}_X[|N|^{-2\nu}]}{M}} = \frac{C}{\sqrt{M}}. \quad (3.22)$$

For the term Δ_2 , we use the smooth Wasserstein estimate in Lemma 3.3.5 and Remark 3.3.6. We have

$$|\mathbb{E}_X[\phi_M(g_\lambda^X(v))] - \mathbb{E}[\phi_M(N)]| \leq \max_{p \leq 4} \|\phi_M^{(p)}\|_\infty \text{Wass}_{(4)}(g_\lambda^X, N) \leq CM^{\nu+4} \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2-\varepsilon}. \quad (3.23)$$

For the more difficult term Δ_1 , we use Cauchy–Schwarz inequality to obtain

$$|\mathbb{E}_X[\phi(g_\lambda^X(v))] - \mathbb{E}[\phi_M(g_\lambda^X(v))]| \leq \sqrt{\mathbb{E}_X[|g_\lambda^X(v)|^{-2\nu]} \cdot \sqrt{\mathbb{P}_X \left(|g_\lambda^X(v)| < \frac{1}{M} \right)}. \quad (3.24)$$

For the right-hand term, using Kolmogorov distance and Lemma 3.3.4 we have

$$\begin{aligned} \mathbb{P}_X \left(|g_\lambda^X(v)| < \frac{1}{M} \right) &\leq \mathbb{P} \left(|N| < \frac{1}{M} \right) + 2 \operatorname{Kol}(g_\lambda^X(v), N) \\ &\leq \frac{C}{M} + C \operatorname{Wass}_{(4)}(g_\lambda^X(v), N)^{\frac{1}{5}} \\ &\leq \frac{C}{M} + C \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/10-\varepsilon}. \end{aligned}$$

For the left-hand term we fix $\theta = \nu d + \varepsilon$ with $\varepsilon > 0$, and

$$p = \frac{1}{2\nu + \frac{\varepsilon}{d}}.$$

The exponent p satisfies

$$2\nu p < 1 \quad \text{and} \quad 2\theta p > d.$$

We compute

$$\begin{aligned} \mathbb{P}_a \left(\mathbb{E}_X[|g_\lambda^X(v)|^{-2\nu}] > \lambda^{2\theta} \right) &\leq \frac{\mathbb{E}_a \left[\mathbb{E}_X[|g_\lambda^X(v)|^{-2\nu}]^p \right]}{\lambda^{2p\theta}} \\ &\leq \frac{\mathbb{E}_X \mathbb{E}_a[|g_\lambda^X(v)|^{-2\nu p}]}{\lambda^{2p\theta}}. \end{aligned}$$

Recall that g_λ is a Gaussian variable under \mathbb{P}_a , whose variance approaches 1 uniformly in X and v . Since $2\nu p < 1$ we obtain

$$\mathbb{P}_a \left(\mathbb{E}_X[|g_{\lambda_n}^X(v)|^{-2\nu}] > \lambda_n^{2\theta} \right) \leq C \frac{\mathbb{E}_X \left[\mathbb{E}_a \left[(g_{\lambda_n}^X(v))^2 \right]^{-\nu p} \right]}{\lambda_n^{2\theta p}} \leq \frac{C}{\lambda_n^{2\theta p}}.$$

Since $\lambda_n \simeq Cn^{1/d}$ the left-hand side is summable and Borel–Cantelli lemma asserts the existence of a constant $C(v, \omega)$ such that

$$\mathbb{E}_X[|g_{\lambda_n}^X(v)|^{-2\nu}] \leq C(v, \omega) \lambda_n^{2\theta}.$$

Finally, bounding the terms in (3.24) we obtain

$$\left| \mathbb{E}_X[\phi(g_\lambda^X(v))] - \mathbb{E}[\phi_M(g_\lambda^X(v))] \right| \leq C(v, \omega) \lambda^{d\nu+\varepsilon} \sqrt{\frac{1}{M} + \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/10-\varepsilon}}. \quad (3.25)$$

Adding the bounds on Δ_1 , Δ_2 and Δ_3 given by the expressions (3.22), (3.23) and (3.25), we obtain the following bound:

$$\begin{aligned} \mathbb{E}_X \left[|g_\lambda^X(v)|^{-\nu} \right] &\leq \mathbb{E}_X[|N|^{-\nu}] + \frac{C}{\sqrt{M}} + CM^{\nu+4} \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2-\varepsilon} \\ &\quad + C(v, \omega) \lambda^{d\nu+\varepsilon} \sqrt{\frac{1}{M} + \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/10-\varepsilon}}. \end{aligned}$$

Recall that

$$\eta(\lambda) = \begin{cases} 1 & \text{in the large band regime} \\ \lambda^{1-\tau} & \text{in the small band regime.} \end{cases}$$

We then choose

$$\nu < \frac{\tau}{20d} \quad \text{and} \quad M = \left(\frac{\lambda}{\eta(\lambda)} \right)^{1/10}.$$

Since a Gaussian random variable has bounded negative moments for exponents $\nu > -1$, we deduce

$$\sup_{\lambda > 0} \mathbb{E}_X \left[|g_\lambda^X(v)|^{-\nu} \right] \leq C(v, \omega). \quad (3.26)$$

It remains to prove the second technical part of Lemma 3.3.3. We cannot directly apply the first bound since the constant obtained in (3.26) may depend on v . Mimicking the previous computation, we write

$$\int_1^{+\infty} \frac{1}{t^{\alpha+2}} \sum_{i=0}^{\lfloor t^\alpha \rfloor} \mathbb{E}_X[|g_\lambda^X(v_i)|^{-\nu}] dt = \Delta_1 + \Delta_2 + \Delta_3.$$

Estimates (3.22) and (3.23) for Δ_1 and Δ_2 remain unchanged. For the quantity Δ_3 , we keep the previous notations. We define the event

$$A(\lambda) = \left\{ \int_1^{+\infty} \frac{1}{t^{1+\alpha+\varepsilon}} \sum_{i=0}^{\lfloor t^\alpha \rfloor} \mathbb{E}_X[|g_\lambda^X(v_i)|^{-2\nu}] dt > \lambda^{2\theta} \right\}.$$

We have, using Markov inequality in the first line, and Hölder inequality in the second line,

$$\begin{aligned}
 \mathbb{P}_a(A(\lambda)) &\leq \frac{1}{\lambda^{2p\theta}} \mathbb{E}_a \left[\left(\int_1^{+\infty} \frac{1}{t^\alpha} \sum_{i=0}^{\lfloor t^\alpha \rfloor} \mathbb{E}_X [|g_\lambda^X(v_i)|^{-2\nu}] \frac{dt}{t^{1+\varepsilon}} \right)^p \right] \\
 &\leq \frac{C}{\lambda^{2p\theta}} \int_1^{+\infty} \frac{\lfloor t^\alpha \rfloor^{p-1}}{t^{p\alpha+1+\varepsilon}} \sum_{i=0}^{\lfloor t^\alpha \rfloor} \mathbb{E}_X [\mathbb{E}_a [|g_\lambda^X(v_i)|^{-2\nu p}]] dt \\
 &\leq \frac{C}{\lambda^{2p\theta}} \int_1^{+\infty} \frac{\lfloor t^\alpha \rfloor^p}{t^{p\alpha+1+\varepsilon}} dt \\
 &\leq \frac{C}{\lambda^{2p\theta}}.
 \end{aligned}$$

The end of the proof remains unchanged. \square

Remark 3.3.7. The dependence in v of the constant $C(v, \omega)$ given in Equation (3.26) is not entirely satisfactory, and is a consequence of Borel–Cantelli lemma in Equation (3.25). We were not able to give a bound on the quantity

$$\mathbb{E}_a \left[\sup_{v \in B} \mathbb{E}_X [|g_\lambda^X(v)|^{-\nu}] \right].$$

It does not impact the rest of the article since the second part of Lemma 3.3.5 suffices to carry out our computations, but let us give a little more insight about what happens from a measure-theoretic point of view.

The Sobolev trick we used before to obtain the uniformity on v does not apply here due to the lack of regularity of the function $x \mapsto |x|^{-\nu}$. Nevertheless it may happen in particular cases that we can recover uniformity. If we are on a torus \mathbb{T}^d endowed with any flat metric, we can choose for the isometry I_x the canonical embedding into \mathbb{R}^d and the mapping Φ_x is the usual sum. If X is a uniform random variable on \mathbb{T}^d , then so is $X + v$ for any $v \in \mathbb{T}^d$. It follows that under \mathbb{P}_X and for all $v, v' \in \mathbb{R}^d$,

$$g_\lambda^X(v) \stackrel{\mathcal{L}}{=} g_\lambda^X(v'),$$

and quantities such as $\mathbb{E}_X [|g_\lambda^X(v)|^{-\nu p}]$ do not depend on v , which gives the uniformity in v . Denote by μ_v the pushforward of the measure μ under the mapping $x \mapsto \Phi_x(v)$. For all $f \in \mathcal{C}^0(\mathcal{M})$,

$$\int_{\mathcal{M}} f(\Phi_x(v)) d\mu(x) = \int_{\mathcal{M}} f(x) d\mu_v(x).$$

In the torus case, μ_v is the canonical measure and does not depend on the parameter

v . In all generality, few can be said about μ_v . It does not always admit a density with respect to the Riemannian measure since the function $x \mapsto \Phi_x(v)$ may have support on a 1-dimensional subspace for an ill-chosen choice of isometries $(I_x)_{x \in \mathcal{M}}$. Nevertheless, if the measure μ_v has a density h_v belonging to $L^p(\mathcal{M})$ space for some $p > 1$ and uniformly on $v \in B$, then

$$\begin{aligned} \int_{\mathcal{M}} |g_\lambda^X(v)|^{-\nu} d\mu(x) &= \int_{\mathcal{M}} |f_\lambda(x)|^{-\nu} h_{\frac{v}{\lambda}}(x) d\mu(x) \\ &\leq \left(\int_{\mathcal{M}} |f_\lambda(x)|^{-\nu q} \right)^{1/q} \sup_{v \in B} \|h_v\|_p \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

Taking the expectation under \mathbb{P}_a , and choosing $\nu < \frac{1}{q}$, we obtain

$$\mathbb{E}_a \left[\sup_{v \in B} \mathbb{E}_X \left[|g_\lambda^X(v)|^{-\nu} \right] \right] \leq \sup_{v \in B} \|h_v\|_p \int_{\mathcal{M}} \mathbb{E}_a \left[|f_\lambda(x)|^{-\nu q} \right]^{1/q} dx < +\infty.$$

At last, given a smooth compact Riemannian manifold \mathcal{M} , it is always possible to construct a family of isometries $(I_x)_{x \in \mathcal{M}}$ such that the family $(\mu_v)_{v \in B}$ has a density uniformly bounded in L^∞ .

3.3.4 Uniform moment estimates for the nodal volume

In order to complete the proof of the uniform integrability of nodal volume, we now introduce a geometric lemma which relates the nodal volume of a function to the number of zero of this function on a straight line passing through predefined points. It is a variant of the Crofton formula (see [ÁF07] for a general presentation of the various Crofton formulæ), and a d -dimensional extension of [APP18, Thm. 6].

Lemma 3.3.8. *Let E be a \mathcal{C}^2 -hypersurface in \mathbb{R}^d , intersecting the cube $D = [0, a]^d$. Assume that E has bounded curvature on D . Then there exists a segment S passing through one of the vertices of the cube D and such that*

$$\text{Card}(E \cap S \cap D) \geq c \frac{\mathcal{H}^{d-1}(E \cap D)}{a^{d-1}}, \quad \text{with } c = \frac{1}{2^{d+1}d}.$$

The proof of Lemma 3.8 is postponed to Appendix 3.4, and relies on a probabilistic method to shows the existence of a segment S satisfying the conclusion of Lemma 3.3.8.

Remark 3.3.9. If $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is a smooth function and 0 is a regular value of g , then $g^{-1}(\{0\})$ is a smooth manifold and we can apply Lemma 3.3.8 to deduce the existence of a segment S passing through one of the vertices of the cube $D = [0, a]^d$ and such that

$$\text{Card}(\{g = 0\} \cap S \cap D) \geq c \frac{\mathcal{H}^{d-1}(\{g = 0\} \cap D)}{a^{d-1}}. \quad (3.27)$$

Denote g_S its restriction on S , and suppose that g_S cancels at least p times at points $\mathbf{w}_1, \dots, \mathbf{w}_p$. By the generalized Rolle lemma, for all $v \in S$, there exists a point c_v in S such that

$$|g(v)| = \frac{\prod_{j=1}^p \|v - \mathbf{w}_j\|}{p!} |g_S^{(p)}(c_v)|.$$

Hence if the segment S passes through the vertex v_j on the cube D then

$$\begin{aligned} |g(v_j)| &\leq \frac{\mathcal{H}^1(S)^p}{p!} \sum_{|\alpha|=p} \|\partial_\alpha g\|_\infty \\ &\leq C a^p \sum_{|\alpha|=p} \|\partial_\alpha g\|_\infty. \end{aligned}$$

To sum up, if we have

$$\mathcal{H}^{d-1}(\{g = 0\} \cap D) \geq \frac{a^{d-1}}{c} p,$$

then for at least one of the vertices v_j of the cube,

$$|g(v_j)| \leq C a^p \sum_{|\alpha|=p} \|\partial_\alpha g\|_\infty.$$

In [Arm+19, Thm. 5.2] the authors proved the finiteness of moments of nodal volume under the requirement of joint bounded density of first k derivatives. This hypothesis is too strong for our purpose, since our process under \mathbb{P}_X depends only on the randomness of X and we cannot expect a joint bounded density of the first derivatives.

Theorem 3.3.10. \mathbb{P}_a -almost surely, the family of random variables $(Z_\lambda)_{\lambda>0}$ is uniformly integrable. More precisely, for all $\gamma > 0$,

$$\sup_{\lambda>0} \mathbb{E}_X[Z_\lambda^{1+\gamma}] \leq C_\gamma(\omega).$$

Conjointly with the convergence in distribution of the nodal volume, it implies the convergence of all moments of Z_λ to those of Z_∞ .

Proof of Theorem 3.3.10. For all $A > 0$ (to be fixed later),

$$\begin{aligned} \mathbb{E}_X[Z_\lambda^{1+\gamma}] &= (1 + \gamma) \int_0^{+\infty} t^\gamma \mathbb{P}_X(Z_\lambda > t) dt \\ &\leq C_A + (1 + \gamma) \int_A^{+\infty} t^\gamma \mathbb{P}_X(Z_\lambda > t) dt, \end{aligned} \quad (3.28)$$

hence we need to estimate the quantity $\mathbb{P}_X(Z_\lambda > t)$ for all t greater than some constant A . Up to embedding the ball B in a cube we will consider that the vector v lives in a hypercube (of size 1 for simplicity).

Consider a rectangular grid on $[0, 1]^d$ of size $\frac{1}{\lfloor t^\theta \rfloor}$ with $\theta = 1 - \varepsilon$. The hypercube $[0, 1]^d$ is split into $\lfloor t^\theta \rfloor^d$ smaller cubes, which we number by $(D_i)_{1 \leq i \leq \lfloor t^\theta \rfloor^d}$ (see Figure 3.2).

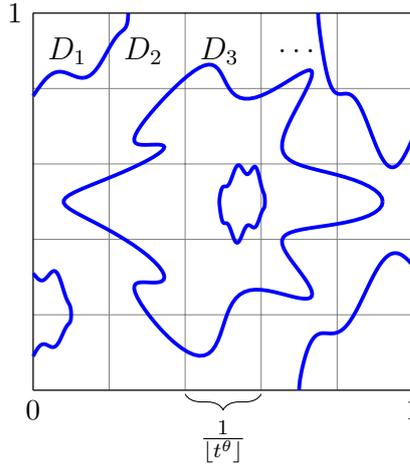


Figure 3.2 – The grid defined on $[0, 1]^d$.

Let $(v_j)_{1 \leq j \leq \lfloor t^\theta \rfloor^d}$ be the vertices of the grid. For all $i \in \{1, \dots, \lfloor t^\theta \rfloor^d\}$, let $(v_{ij})_{1 \leq j \leq 2^d}$ be the vertices of the i -th cube, and

$$Z_\lambda^{(i)} := \mathcal{H}^{d-1}(\{g_\lambda^X = 0\} \cap D_i),$$

be the volume of zeros contained in the i -th cube. By the pigeonhole principle,

$$\{Z_\lambda > t\} \subset \bigcup_{i=1}^{\lfloor t^\theta \rfloor^d} \left\{ Z_\lambda^{(i)} > \frac{t}{\lfloor t^\theta \rfloor^d} \right\}. \quad (3.29)$$

Let p be an integer to be fixed later. We use Lemma 3.3.8 and then Remark 3.3.9, keeping the same notations, to deduce that if $t \geq A$, with

$$A := \left(\frac{p}{c}\right)^{1/(1-\theta)}.$$

and $a = \lfloor t^\theta \rfloor^{-1}$, then

$$\left\{ Z_\lambda^{(i)} > \frac{t}{\lfloor t^\theta \rfloor^d} \right\} \subset \left\{ Z_\lambda^{(i)} > \frac{p}{c \lfloor t^\theta \rfloor^{d-1}} \right\} \subset \bigcup_{j=1}^{2^d} \left\{ |g_\lambda^X(v_{ij})| \leq \frac{C}{\lfloor t^\theta \rfloor^p} \sum_{|\alpha|=p} \|\partial_\alpha g_\lambda^X\|_\infty \right\}.$$

Fix $k \geq 1$. Taking the expectation with respect to \mathbb{P}_X we obtain

$$\begin{aligned} \mathbb{P}_X \left(Z_\lambda^{(i)} > \frac{t}{\lfloor t^\theta \rfloor^d} \right) &\leq \sum_{j=1}^{2^d} \mathbb{P}_X \left(|g_\lambda^X(v_{ij})| \leq \frac{C}{\lfloor t^\theta \rfloor^p} \sum_{|\alpha|=p} \|\partial_\alpha g_\lambda^X\|_\infty \right) \\ &\leq \sum_{j=1}^{2^d} \mathbb{P}_X \left(|g_\lambda^X(v_{ij})| \leq \frac{t^\varepsilon}{\lfloor t^\theta \rfloor^p} \right) + 2^d \mathbb{P}_X \left(C \sum_{|\alpha|=p} \|\partial_\alpha g_\lambda^X\|_\infty > t^\varepsilon \right) \\ &\leq \left(\frac{t^\varepsilon}{\lfloor t^\theta \rfloor^p} \right)^\nu \sum_{j=1}^{2^d} \mathbb{E}_X [|g_\lambda^X(v_{ij})|^{-\nu}] + C \frac{\mathbb{E}_X \left[\sum_{|\alpha|=p} \|\partial_\alpha g_\lambda^X\|_\infty^k \right]}{t^{k\varepsilon}} \quad (3.30) \\ &\leq \left(\frac{t^\varepsilon}{\lfloor t^\theta \rfloor^p} \right)^\nu \sum_{j=1}^{2^d} \mathbb{E}_X [|g_\lambda^X(v_{ij})|^{-\nu}] + \frac{C_k(\omega)}{t^{k\varepsilon}}. \end{aligned}$$

In the last line we used the Sobolev injection and estimate of Lemma 3.2.8 to bound the right hand side. Taking the expectation in expression (3.29) and using the union bound we obtain

$$\mathbb{P}_X(Z_\lambda > t) \leq C_k(\omega) \frac{\lfloor t^\theta \rfloor^d}{t^{k\varepsilon}} + \left(\frac{t^\varepsilon}{\lfloor t^\theta \rfloor^p} \right)^\nu 2^d \sum_{i=1}^{\lfloor t^\theta \rfloor^d} \mathbb{E}_X [|g_\lambda^X(v_i)|^{-\nu}].$$

Recalling expression (3.28) we obtain

$$\mathbb{E}_X[Z_\lambda^{1+\gamma}] \leq C + C_k(\omega) \int_A^{+\infty} t^\gamma \frac{\lfloor t^\theta \rfloor^d}{t^{k\varepsilon}} dt + C \int_A^{+\infty} t^\gamma \left(\frac{t^\varepsilon}{\lfloor t^\theta \rfloor^p} \right)^\nu \sum_{j=1}^{\lfloor t^\theta \rfloor^d} \mathbb{E}_X [|g_\lambda^X(v_i)|^{-\nu}] dt.$$

Choosing k and p such that

$$k > \frac{\gamma + \theta d + 1}{\varepsilon} \quad \text{and} \quad \nu(p\theta - \varepsilon) - (d\theta + \gamma) > 1,$$

we can apply the second part of Lemma 3.3.3 to deduce the existence of a constant $C(\omega)$ such that

$$\mathbb{E}_X[Z_\lambda^{1+\gamma}] \leq C(\omega).$$

□

Remark 3.3.11. If we were in an analytic setting, we could use the same argument as the one in [AP21, Thm. 9], which roughly relies on the convergence of Taylor expansion of eigenfunctions. In the C^∞ setting we can only apply the generalized Rolle lemma with a fixed p , and it explains why we used the partitioning of the cube $[0, 1]^d$. A careful analysis of the proof shows that it requires a manifold of finite regularity C^k for k large enough.

Corollary 3.3.12. For all $p \geq 1$:

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathbb{E}_a \left[\left(\mathcal{H}^{d-1}(\{f_\lambda = 0\}) \right)^p \right]}{\lambda^p} = \left(\frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d+2}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \right)^p,$$

and

$$\lim_{\lambda \rightarrow +\infty} \frac{\mathbb{E}_a \left[\left(\mathcal{H}^{d-1}(\{\tilde{f}_\lambda = 0\}) \right)^p \right]}{\lambda^p} = \left(\frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{d}} \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \right)^p.$$

Proof. Passing from almost-sure convergence to convergence in expectation is a consequence of the dominated convergence. It suffices to show that

$$\sup_{\lambda > 0} \mathbb{E}_a \left[(\mathbb{E}_X[Z_\lambda])^p \right] < +\infty,$$

which can be seen by raising to the power p and taking the expectation under \mathbb{P}_a in Equation (3.30). In more direct way, all the almost-sure estimates are deduced from Borel-Cantelli lemma and Markov inequality applied to the power function with arbitrary large exponent, and thus remain true under expectation. A similar argument holds for higher moments. □

3.4 Appendix

Proof of decorrelation estimates

In this first part of the Appendix, we give the proof of Lemma 3.2.3 and Theorem 3.2.6 stated in Section 3.2.

Proof of Lemma 3.2.3

Let X_1, \dots, X_{2q} be independent copies of X . The expectation with respect to the random variables X_1, \dots, X_{2q} will be noted $\mathbb{E}_{\mathbf{X}}$. To enhance the dependence with respect to X_k , we set for all $k \in \{1, \dots, q\}$,

$$N_{\lambda}^{(k)} = \sum_{j=1}^p t_j g_{\lambda}^{X_k}(v_j),$$

and for all $k \in \{q+1, \dots, 2q\}$,

$$N_{\lambda}^{(k)} = - \sum_{j=1}^p t_j g_{\lambda}^{X_k}(v_j).$$

Then, for $k \neq l$, applying Lemma 3.2.4 with $X = X_k$ and $Y = X_l$, we have uniformly in X_l

$$\mathbb{E}_{X_k} \left[\left| \mathbb{E}_a \left[N_{\lambda}^{(k)} N_{\lambda}^{(l)} \right] \right| \right] = \|t\|^2 O \left(\frac{\eta(\lambda)}{\lambda} \right). \quad (3.31)$$

The following lemma, based on an explicit computation of Gaussian characteristic functions and integral Taylor formula, gives an explicit expression of $\tilde{\Delta}_{\lambda}^{(q)}$, which is the object of Lemma 3.2.3. For $s \in [0, 1]^{2q}$, let

$$\begin{aligned} f(s) &:= \sum_{k=1}^{2q} \mathbb{E}_a \left[\left(N_{\lambda}^{(k)} \right)^2 \right] + \sum_{\substack{k,l=1 \\ k \neq l}}^{2q} s_k s_l \mathbb{E}_a \left[N_{\lambda}^{(k)} N_{\lambda}^{(l)} \right] \\ &= \mathbb{E}_a \left[\left(\sum_{k=1}^{2q} s_k N_{\lambda}^{(k)} \right)^2 \right] + \sum_{k=1}^{2q} (1 - s_k^2) \mathbb{E}_a \left[\left(N_{\lambda}^{(k)} \right)^2 \right]. \end{aligned}$$

Lemma 3.4.1. *We have*

$$\tilde{\Delta}_\lambda^{(q)} = \mathbb{E}_\mathbf{X} \left[\int_{[0,1]^{2q}} \partial_1 \dots \partial_{2q} \left(\exp \left(-\frac{1}{2} f \right) \right) (s) ds \right].$$

Proof of Lemma 3.4.1. From mutual independence of the family (X_1, \dots, X_{2q}) ,

$$\tilde{\Delta}_\lambda^{(q)} = \mathbb{E}_\mathbf{X} \mathbb{E}_a \left[\prod_{k=1}^{2q} \left(e^{iN_\lambda^{(k)}} - \mathbb{E}_a \left[e^{iN_\lambda^{(k)}} \right] \right) \right].$$

We define

$$\tilde{\Delta}_\lambda^{(q)}(s) := \mathbb{E}_\mathbf{X} \mathbb{E}_a \left[\left(\prod_{k=1}^{2q} \mathbb{E}_a \left[e^{i\sqrt{1-s_k^2} N_\lambda^{(k)}} \right] \right) \left(\prod_{k=1}^{2q} \left(e^{is_k N_\lambda^{(k)}} - \mathbb{E}_a \left[e^{is_k N_\lambda^{(k)}} \right] \right) \right) \right]. \quad (3.32)$$

Developing the product, and using the characteristic function of a Gaussian random variable, a direct computation shows that

$$\begin{aligned} \tilde{\Delta}_\lambda^{(q)}(s) &:= \mathbb{E}_\mathbf{X} \left[\exp \left(-\frac{1}{2} \sum_{k=1}^{2q} \mathbb{E}_a \left[(N_\lambda^{(k)})^2 \right] \right) \right. \\ &\quad \left. \times \sum_{A \subset \{1, \dots, 2q\}} (-1)^{|A|} \exp \left(-\frac{1}{2} \sum_{\substack{k, l \in A \\ k \neq l}} s_k s_l \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right) \right]. \end{aligned} \quad (3.33)$$

If $s_k = 0$ for some $k \in \{1, \dots, 2q\}$, then from expression (3.32),

$$\tilde{\Delta}_\lambda^{(q)}(s) = 0.$$

In other words, the function $s \mapsto \tilde{\Delta}_\lambda^{(q)}(s)$ cancels if one of its coordinates is zero. By integral Taylor formula,

$$\tilde{\Delta}_\lambda^{(q)} = \tilde{\Delta}_\lambda^{(q)}(1, \dots, 1) = \int_{[0,1]^{2q}} \partial_1 \dots \partial_{2q} \tilde{\Delta}_\lambda^{(q)}(s) ds.$$

But from expression (3.33), the only term that depends on all coordinates (and thus won't

be canceled after differentiation) is the term corresponding to $A = \{1, \dots, 2q\}$, which is

$$\mathbb{E}_{\mathbf{X}} \left[\exp \left(-\frac{1}{2} \sum_{k=1}^{2q} \mathbb{E}_a \left[(N_\lambda^{(k)})^2 \right] \right) \exp \left(-\frac{1}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^{2q} s_k s_l \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right) \right] = \mathbb{E}_{\mathbf{X}} \left[\exp \left(-\frac{1}{2} f(s) \right) \right].$$

□

Now, for a set A , denote $\Pi(A)$ the collection of partitions of A into groups of two elements. A direct computation shows that

$$\begin{aligned} \partial_1 \dots \partial_{2q} \left(\exp \left(-\frac{1}{2} f \right) \right) &= \exp \left(-\frac{1}{2} f \right) \sum_{\substack{A \subset \{1, \dots, 2q\} \\ |A| \text{ even}}} (-1)^{\frac{|A|}{2}} \left(\sum_{B \in \Pi(A)} \prod_{(k,l) \in B} \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right) \\ &\quad \times \prod_{k \in A^c} \left(\sum_{\substack{l=1 \\ k \neq l}}^{2q} s_l \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right). \end{aligned}$$

We deduce, using the mutual independence of the family X_1, \dots, X_{2q} and the estimate 3.31,

$$\begin{aligned} \tilde{\Delta}_\lambda^{(q)} &\leq C \mathbb{E}_{\mathbf{X}} \left[\sum_{\substack{A \subset \{1, \dots, 2q\} \\ |A| \text{ even}}} \left(\sum_{B \in \Pi(A)} \prod_{(k,l) \in B} \left| \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right| \right) \prod_{k \in A^c} \left(\sum_{\substack{l=1 \\ k \neq l}}^{2q} \left| \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right| \right) \right] \\ &\leq C \sum_{\substack{A \subset \{1, \dots, 2q\} \\ |A| \text{ even}}} \left(\sum_{B \in \Pi(A)} \prod_{(k,l) \in B} \mathbb{E}_{\mathbf{X}} \left[\left| \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right| \right] \right) \mathbb{E}_{\mathbf{X}} \left[\prod_{k \in A^c} \left(\sum_{\substack{l=1 \\ k \neq l}}^{2q} \left| \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right| \right) \right] \\ &\leq C \sum_{\substack{A \subset \{1, \dots, 2q\} \\ |A| \text{ even}}} \left(\|t\|^2 \frac{\eta(\lambda)}{\lambda} \right)^{\frac{|A|}{2}} \mathbb{E}_{\mathbf{X}} \left[\prod_{k \in A^c} \left(\sum_{\substack{l=1 \\ k \neq l}}^{2q} \left| \mathbb{E}_a \left[N_\lambda^{(k)} N_\lambda^{(l)} \right] \right| \right) \right]. \end{aligned} \quad (3.34)$$

To give a bound on the right-hand term and thus establish Lemma 3.2.3, we use the following lemma whose proof again relies on the decorrelation estimates of Lemma 3.2.4.

Lemma 3.4.2. *There is a constant C depending only on \mathcal{M} , K and q such that*

$$\mathbb{E}_{\mathbf{X}} \left[\prod_{k \in A^c} \left(\sum_{\substack{l=1 \\ l \neq k}}^{2q} |\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]| \right) \right] \leq C(1 + \|t\|^4)^{q - \frac{|A|}{2}} \left(\frac{\eta(\lambda)}{\lambda} \right)^{q - \frac{|A|}{2}}.$$

Proof of Lemma 3.4.2. Assume without loss of generality that $A^c = \{1, \dots, 2m\}$. We compute

$$\mathbb{E}_{\mathbf{X}} \left[\prod_{k=1}^{2m} \left(\sum_{\substack{l=1 \\ k \neq l}}^{2q} |\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]| \right) \right] = \sum_{\substack{l_1, \dots, l_{2m}=1 \\ l_k \neq k}}^{2q} \underbrace{\mathbb{E}_{\mathbf{X}} \left[\prod_{k=1}^{2m} |\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l_k)}]| \right]}_{\Delta_\ell}.$$

Now fix $\ell = (l_1, \dots, l_{2m})$. Consider the following graph G_ℓ with vertices in $\{1, \dots, 2q\}$: two vertices k and l are connected if the term $|\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]|$ appears into the expression Δ_ℓ . If the graph G_ℓ is disconnected, we can use independence of the random variables X_1, \dots, X_{2q} , and we are left to show the aforementioned bound for connected graphs. Thanks to Weyl law, there is a constant C such that

$$|\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]| \leq C\|t\|^2,$$

and we can assume (up to bounding one of the terms in the product) that G_ℓ is a tree (with $2m - 1$ edges). Suppose without loss of generality that 1 is a leaf of the tree attached to 2. Recalling the definition (3.9) of $\eta(\lambda)$, one has

$$\begin{aligned} \Delta_\ell &\leq C\|t\|^2 \mathbb{E}_{X_2, \dots, X_{2q}} \left[\mathbb{E}_{X_1} \left[|\mathbb{E}_a [N_\lambda^{(1)} N_\lambda^{(2)}]| \prod_{\substack{(k,l) \in G_\ell \\ (k,l) \neq (1,2)}} |\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]| \right] \right] \\ &\leq C\|t\|^4 O\left(\frac{\eta(\lambda)}{\lambda}\right) \mathbb{E}_{X_2, \dots, X_{2q}} \left[\prod_{\substack{(k,l) \in G_\ell \\ (k,l) \neq (1,2)}} |\mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l)}]| \right], \end{aligned}$$

after the estimate (3.31). Repeating the procedure leaf by leaf we obtain the bound

$$\mathbb{E}_{\mathbf{X}} \left[\prod_{k=1}^{2m} \left| \mathbb{E}_a [N_\lambda^{(k)} N_\lambda^{(l_k)}] \right| \right] = \|t\|^{4m} O \left(\left(\frac{\eta(\lambda)}{\lambda} \right)^{2m-1} \right) = \|t\|^{4m} O \left(\left(\frac{\eta(\lambda)}{\lambda} \right)^m \right).$$

We used the fact that $2m - 1 \geq m$, with equality when $m = 1$. That is, the worst case is attained for graphs G_l such that their sets of edges forms partitions into pairs, for instance when $G_\ell = \{(1, 2), (3, 4) \dots, (2m - 1, 2m)\}$. \square

Proof of Theorem 3.2.6

The proof of Theorem 3.2.6 is rather technical, and relies on the following Sobolev injection for a smooth domain Ω :

$$W^{d+1,1}(\Omega) \subset L_\infty(\Omega).$$

It allows us to bound the supremum norm by the $W^{d+1,1}$ Sobolev norm, which is interchangeable with the expectation under \mathbb{P}_a . We only detail the proof of the first assertion for simplicity. The second assertion is the generalization to the case $t \in \mathbb{R}^p$, and its proof follows the same lines.

Proof of Theorem 3.2.6. Let B_K be a ball containing the compact K . Let $t \mapsto h(t)$ be a non-negative symmetric function, and non-increasing on \mathbb{R}_+ . For any smooth function $f : B_K \times \mathbb{R} \rightarrow \mathbb{R}$ with $f(v, 0) = 0$,

$$\begin{aligned} \sup_{v \in K} \sup_{t \in \mathbb{R}} h(t) |f(v, t)| &\leq \sup_{t \in \mathbb{R}} h(t) \int_{[0,t]} \sup_{v \in B_K} |\partial_t f(v, s)| ds \\ &\leq \sup_{t \in \mathbb{R}} \int_{[0,t]} h(s) \sup_{v \in B_K} |\partial_t f(v, s)| ds \\ &\leq C \int_{-\infty}^{+\infty} h(t) \|\partial_t f(v, t)\|_{W^{d+1,1}(B_K)} dt. \end{aligned} \quad (3.35)$$

We set

$$f(v, t) = \left| \mathbb{E}_X [e^{itg_X^\chi(v)}] - e^{-\frac{t^2}{2}} \right|^{2q} \quad \text{and} \quad h(t) = \frac{1}{(1 + |t|^{2+\varepsilon})^{2q}}.$$

in (3.35). By Fubini theorem,

$$\mathbb{E}_a \left[\sup_{v \in K} \sup_{t \in \mathbb{R}} h(t) |f(v, t)| \right] \leq C \sum_{|\alpha| \leq d+1} \int_{-\infty}^{+\infty} h(t) \int_{B_K} \mathbb{E}_a [|\partial_\alpha \partial_t f(v, t)|] dv dt. \quad (3.36)$$

It remains to estimate the integrand. Using the derivative of the power function, we have

$$\partial_\alpha \partial_t f(v, t) = g(v, t) \left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right|^{2(q-d-2)},$$

for some function g to be explicited. Using Cauchy–Schwarz inequality

$$\mathbb{E}_a [|\partial_\alpha \partial_t f(v, t)|] \leq \sqrt{\mathbb{E}_a [g(v, t)^2]} \cdot \sqrt{\mathbb{E}_a \left[\left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right|^{4(q-d-2)} \right]}.$$

According to Theorem 3.2.2, there is a constant C independent of t and λ such that

$$\sqrt{\mathbb{E}_a \left[\left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right|^{4(q-d-2)} \right]} \leq C(1 + |t|^{4(q-d-2)}) \left(\frac{\eta(\lambda)}{\lambda} \right)^{q-d-2}.$$

We will show that for some polynomial P of degree m independent of q ,

$$\mathbb{E}_a [g(v, t)^2] \leq P(t). \quad (3.37)$$

We will establish this fact in the end of the proof. Injecting this into the expression (3.36), we obtain

$$\mathbb{E}_a \left[\sup_{v \in K} \sup_{t \in \mathbb{R}} h(t) f(v, t) \right] \leq C \left(\frac{\eta(\lambda)}{\lambda} \right)^{q-d-2} \int_{-\infty}^{+\infty} \frac{(1 + |t|^{\frac{m}{2}})(1 + |t|^{4(q-d-2)})}{(1 + |t|^{2+\varepsilon})^{2q}} dt,$$

and for q large enough, the integral is bounded. The end of the proof is the same as in the remark following Theorem 3.2.2. We have by Markov inequality and q large enough,

$$\begin{aligned} \mathbb{P}_a \left(\sup_{t \in \mathbb{R}} \sup_{v \in K} \frac{\left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right|}{1 + |t|^{2+\varepsilon}} > \lambda^\varepsilon \left(\frac{\eta(\lambda)}{\lambda} \right)^{1/2} \right) &\leq \left(\frac{\lambda}{\eta(\lambda) \lambda^{2\varepsilon}} \right)^q \mathbb{E}_a \left[\sup_{v \in K, t \in \mathbb{R}} h(t) f(v, t) \right] \\ &\leq C \frac{1}{\lambda^{2q\varepsilon}} \left(\frac{\eta(\lambda)}{\lambda} \right)^{d+2}. \end{aligned}$$

For q large enough the right-hand term is summable, and Borel–Cantelli lemma implies

the existence a constant $C(\omega)$ depending only on K and ε such that

$$\left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right| \leq C(\omega)(1 + |t|^{2+\varepsilon}) \frac{\sqrt{\eta(\lambda)}}{\lambda^{\frac{1}{2}-\varepsilon}}.$$

It remains to show the estimate (3.37). We have

$$\left| \mathbb{E}_X \left[e^{itg_\lambda^X(v)} \right] - e^{-\frac{t^2}{2}} \right|^{2q} = \mathbb{E}_{\mathbf{X}} \left[\prod_{k=1}^{2q} \left(e^{\pm itg_\lambda^{X_k}(v)} - e^{-\frac{t^2}{2}} \right) \right].$$

From this expression we deduce that

$$g(v, t) = \mathbb{E}_{\mathbf{X}} \left[F \left(t, \left(\partial_\alpha g_\lambda^{X_k} \right)_{\substack{|\alpha| \leq d' \\ 1 \leq k \leq 2q}} \right) \right],$$

with F a function bounded by a polynomial of degree $2(d+2)$ in its arguments. But the partial derivatives of g_λ are still Gaussian under \mathbb{P}_a , and the local Weyl law (see also the expression (3.38)) implies the existence of a universal constant C such that

$$\mathbb{E}_a \left[\left(\partial_\alpha g_\lambda^{X_k} \right)^{2p} \right] = \frac{(2p)!}{2^p p!} \mathbb{E}_a \left[\left(\partial_\alpha g_\lambda^{X_k} \right)^2 \right]^p \leq \frac{(2p)!}{2^p p!} C,$$

It implies that

$$\mathbb{E}_a \left[g(v, t)^2 \right] \leq P(t),$$

for some polynomial P in t whose degree is independent of q .

□

Proof of tightness estimates

Proof of Lemma 3.2.8. Let $x, y \in \mathcal{M}$. We have

$$\mathbb{E}_a \left[\partial_\alpha g_\lambda^x(u) \partial_\alpha g_\lambda^y(v) \right] = \frac{1}{K(\lambda) \lambda^{2|\alpha|}} \sum_{\lambda_n \leq \lambda} \left(\partial_\alpha (\varphi_n \circ \Phi_x) \left(\frac{u}{\lambda} \right) \right) \left(\partial_\alpha (\varphi_n \circ \Phi_y) \left(\frac{v}{\lambda} \right) \right).$$

Setting

$$x_u = \Phi_x \left(\frac{u}{\lambda} \right) \quad \text{and} \quad y_v = \Phi_y \left(\frac{v}{\lambda} \right),$$

and using the fact that $d \exp_x = \text{Id}$, we obtain

$$\mathbb{E}_a [\partial_\alpha g_\lambda^x(u) \partial_\alpha g_\lambda^y(v)] = \frac{1}{K(\lambda)\lambda^{2|\alpha|}} (\partial_{\alpha,\alpha} K_\lambda(x_u, y_v)) + O\left(\frac{1}{\lambda}\right).$$

Recalling that the kernel \mathcal{B}_d (resp. \mathcal{S}_d) is the \mathcal{C}^∞ scaling limit of the spectral projector we have

$$\mathbb{E}_a [\partial_\alpha g_\lambda^X(u) \partial_\alpha g_\lambda^Y(v)] = \partial_{\alpha,\alpha} [\mathcal{B}_d(\lambda \text{ dist}(x_u, y_v))] + O\left(\frac{\eta(\lambda)}{\lambda}\right).$$

We briefly describe the \mathcal{C}^∞ extension of decorrelation estimates given by Lemma 3.2.4. The proof is very similar and we refer to the proof of Lemma 3.2.4 for more details. Firstly, by the local Weyl law in the \mathcal{C}^∞ topology, we have uniformly on $x \in \mathcal{M}$ and $u \in B$,

$$\mathbb{E}_a \left[\left(\partial_\alpha g_\lambda^X(u) \right)^2 \right] = C_\alpha + O\left(\frac{\eta(\lambda)}{\lambda}\right) \quad \text{and} \quad C_\alpha = \partial_{\alpha,\alpha} \mathcal{B}_d(\|u - v\|)|_{u=v=0}. \quad (3.38)$$

Secondly, take X and Y two independent uniform random variables on \mathcal{M} , and $k \geq 1$. As in Lemma 3.2.4, we write

$$\mathbb{E}_X \left[\mathbb{E}_a [\partial_\alpha g_\lambda^X(u) \partial_\alpha g_\lambda^Y(v)]^{2k} \right] = I_1 + I_2,$$

with

$$I_1 = \int_{\text{dist}(x,Y) > \varepsilon} \mathbb{E}_a [\partial_\alpha g_\lambda^X(u) \partial_\alpha g_\lambda^Y(v)]^{2k} dx \quad \text{and} \quad I_2 = \int_{\text{dist}(x,Y) < \varepsilon} \mathbb{E}_a [\partial_\alpha g_\lambda^X(u) \partial_\alpha g_\lambda^Y(v)]^{2k} dx.$$

Using a similar argument as in Lemma 3.2.4, we deduce that uniformly on $u, v \in B$,

$$\mathbb{E}_X \left[\mathbb{E}_a [\partial_\alpha g_\lambda^X(u) \partial_\alpha g_\lambda^Y(v)]^{2k} \right] = O\left(\frac{\eta(\lambda)}{\lambda}\right). \quad (3.39)$$

Define

$$W_X : u \mapsto \partial_\alpha g_\lambda^X(u) \quad \text{and} \quad W_Y : u \mapsto \partial_\alpha g_\lambda^Y(u).$$

The joint process (W_X, W_Y) is Gaussian under \mathbb{P}_a . We fix $u, v \in B$ and set

$$\rho(u, v) = \frac{\mathbb{E}_a [W_X(u) W_Y(v)]^2}{\mathbb{E}_a [W_X(u)^2] \mathbb{E}_a [W_Y(v)^2]}.$$

A direct Gaussian computation shows that

$$\mathbb{E}[W_X(u)^{2p}] = \frac{(2p)!}{2^p p!} \mathbb{E} [W_X(u)^2]^p \quad \text{and} \quad \mathbb{E}[W_Y(v)^{2p}] = \frac{(2p)!}{2^p p!} \mathbb{E} [W_Y(v)^2]^p,$$

and

$$\begin{aligned} Q_p(\rho(u, v)) &:= \frac{\mathbb{E}[(W_X(u)W_Y(v))^{2p}]}{\mathbb{E}_a [W_X(u)^{2p}] \mathbb{E}_a [W_Y(v)^{2p}]} \\ &= \sum_{k=0}^p \frac{\binom{2p+2k}{2p} \binom{2p}{p+k}}{\binom{2p}{p}} \rho(u, v)^k (1 - \rho(u, v))^{p-k}. \end{aligned} \quad (3.40)$$

From identity (3.40) we compute

$$\mathbb{E}_a \left[\left(\mathbb{E}_X \left[\int_B W_X(u)^{2p} du \right] - \frac{(2p)!}{2^p p!} (C_\alpha)^p \right)^2 \right] = \left(\frac{(2p)!}{2^p p!} \right)^2 (\Delta_1 + \Delta_2), \quad (3.41)$$

with

$$\Delta_1 := \left(\mathbb{E}_X \left[\int_B \mathbb{E}_a [W_X(u)^2]^p du \right] - (C_\alpha)^p \right)^2,$$

and

$$\Delta_2 := \int_B \int_B \mathbb{E}_{X,Y} \left[\mathbb{E}_a [W_X(u)^2]^p \mathbb{E}_a [W_Y(v)^2]^p (Q_p(\rho(u, v)) - 1) \right] dudv.$$

From Equation (3.38) we have

$$\Delta_1 = O \left(\left(\frac{\eta(\lambda)}{\lambda} \right)^2 \right).$$

As for the term Δ_2 , we use the fact that $\mathbb{E}_a [W_X(u)^2]$ is bounded above and below by positive constants for λ large enough, from equation (3.38). We develop the polynomial

Q_p and we use equation (3.39) to obtain

$$\begin{aligned} \mathbb{E}_{X,Y} \left[\mathbb{E}_a \left[W_X(u)^2 \right]^p \mathbb{E}_a \left[W_Y(v)^2 \right]^p (Q_p(\rho(u,v)) - 1) \right] &\leq C \mathbb{E}_{X,Y} [|Q_p(\rho(u,v)) - 1|] \\ &\leq C \mathbb{E}_{X,Y} \left[\sum_{k=1}^p |p_k| \rho(u,v)^k \right] \\ &= O\left(\frac{\eta(\lambda)}{\lambda}\right). \end{aligned}$$

Since the estimate is uniform on u, v we deduce

$$\Delta_2 = O\left(\frac{\eta(\lambda)}{\lambda}\right).$$

Injecting this estimate into identity (3.41) we obtain

$$\mathbb{E}_a \left[\left(\mathbb{E}_X \left[\int_B W_X(u)^{2p} du \right] - \frac{(2p)!}{2^p p!} (C_\alpha)^p \right)^2 \right] = O\left(\frac{\eta(\lambda)}{\lambda}\right),$$

The quantity inside the square is a polynomial of degree at most $2p$ in the Gaussian random variables $(a_n)_{n \geq 0}$, and hence belongs to a finite fixed sum of Wiener chaos. The hypercontractivity property asserts that for such a polynomial, all the L^q norms for $q \geq 2$ are equivalent, which in our case implies that for every $q \geq 2$,

$$\mathbb{E}_a \left[\left(\mathbb{E}_X \left[\int_B W_X(u)^{2p} du \right] - \frac{(2p)!}{2^p p!} (C_\alpha)^p \right)^q \right] = O\left(\left(\frac{\eta(\lambda)}{\lambda}\right)^{q/2}\right).$$

For more details on Wiener chaos and hypercontractivity we refer the reader to the book [NP12]. Borel–Cantelli lemma implies the existence for every $\varepsilon > 0$ of a constant $C(\omega)$ independent of λ such that

$$\left| \mathbb{E}_X \left[\int_B W_X(u)^{2p} du \right] - \frac{(2p)!}{2^p p!} (C_\alpha)^p \right| \leq C(\omega) \left(\frac{\eta(\lambda)}{\lambda}\right)^{1/2-\varepsilon},$$

which in turn implies the existence of a constant $\tilde{C}(\omega)$ such that

$$\sup_{\lambda > 0} \mathbb{E}_X \left[\int_B |\partial_\alpha g_\lambda^X(u)|^{2p} du \right] \leq \tilde{C}(\omega).$$

□

Proof of Wassertein estimates

Proof of Lemma 3.3.5. Let ϕ be a function in $\mathcal{S}(\mathbb{R})$, supported on the compact $[-(M+1), M+1]$. Using Plancherel isometry we have (the constant C may change from line to line)

$$\begin{aligned}
 |\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X)]| &\leq \frac{1}{2\pi} \int_{\mathbb{R}} \left| \mathbb{E}[e^{itX_n}] - \mathbb{E}[e^{itX}] \right| |\widehat{\phi}(t)| dt \\
 &\leq \frac{C}{n^\alpha} \int_{\mathbb{R}} (1 + |t|^m) |\widehat{\phi}(t)| dt \\
 &\leq \frac{C}{n^\alpha} \int_{\mathbb{R}} (1 + |t|^{m+1}) |\widehat{\phi}(t)| \frac{1}{1 + |t|} dt \\
 &\leq \frac{C}{n^\alpha} \int_{\mathbb{R}} |\widehat{\phi}(t)| \frac{dt}{1 + |t|} + \frac{C}{n^\alpha} \int_{\mathbb{R}} |t|^{m+1} |\widehat{\phi}(t)| \frac{dt}{1 + |t|} \\
 &\leq \frac{C}{n^\alpha} \sqrt{\int_{\mathbb{R}} |\widehat{\phi}(t)|^2 dt} + \frac{C}{n^\alpha} \sqrt{\int_{\mathbb{R}} |t|^{2m+2} |\widehat{\phi}(t)|^2 dt} \quad (\text{Jensen}) \\
 &\leq \frac{C}{n^\alpha} \|\phi\|_2 + \frac{C}{n^\alpha} \|\phi^{(m+1)}\|_2 \quad (\text{Plancherel}) \\
 &\leq C \frac{\sqrt{M+1}}{n^\alpha} (\|\phi\|_\infty + \|\phi^{(m+1)}\|_\infty). \tag{3.42}
 \end{aligned}$$

By standard approximation argument, the inequality is true for every $\phi \in C^{m+1}(\mathbb{R})$ with support in $[-(M+1), M+1]$. Suppose now that ϕ does not have compact support. Let χ_M a smooth function with support in $[-(M+1), M+1]$ such that $\chi_M = 1$ on $[-M, M]$. Set $\phi_M = \phi \cdot \chi_M$. We write

$$|\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X)]| \leq |\mathbb{E}[\phi_M(X_n)] - \mathbb{E}[\phi_M(X)]| + \|\phi_\infty\| \mathbb{P}(X_n > M) + \|\phi_\infty\| \mathbb{P}(X > M).$$

From inequality (3.42) and Markov inequality applied to the function $x \mapsto |x|^\beta$,

$$|\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X)]| \leq C \frac{\sqrt{M+1}}{n^\alpha} (\|\phi_M\|_\infty + \|\phi_M^{(m+1)}\|_\infty) + C \frac{\|\phi_\infty\|}{M^\beta}.$$

Using Leibniz rule, we have

$$\|\phi_M\|_\infty \leq \|\phi\|_\infty \quad \text{and} \quad \|\phi_M^{(m+1)}\|_\infty \leq C_m \sup_{k \leq m+1} \|\phi^{(k)}\|_\infty.$$

Choosing

$$M = n^{\frac{\alpha}{\beta+1}}$$

and under the requirement that $M > 1$, we obtain

$$|\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X)]| \leq Cn^{-\frac{2\alpha\beta}{2\beta+1}} \sup_{k \leq m+1} \|\phi^{(k)}\|_\infty,$$

from which it follows that

$$\text{Wass}_{(m+1)}(X_n, X) \leq Cn^{-\frac{2\alpha\beta}{2\beta+1}}.$$

□

Proof of the geometric lemma

Proof of Lemma 3.3.8. Both sides are dimensionless and it suffices to prove the assertion for $a = 1$. We can assume that $\mathcal{H}^{d-1}(E \cap \partial D) = 0$, else we could find a segment S passing through one of the vertices and such that $\mathcal{H}^1(E \cap S \cap \partial D) > 0$, and in that case the result is true.

We will prove Lemma 3.3.8 by a probabilistic method. We denote $(A_j)_{1 \leq j \leq 2^d}$ the vertices of the cube. Let P be a point chosen uniformly randomly on the cube $[0, 1]^d$. Let $(A_j P)$ be the random line passing through the points A_j and P . We will in fact prove that

$$\mathbb{E} \left[\sum_{j=1}^{2^d} \text{Card} \{E \cap (A_j P) \cap D\} \right] \geq \frac{1}{2^d} \mathcal{H}^{d-1}(E \cap D), \quad (3.43)$$

which implies the result, since for some realization of P and some j we must have

$$\text{Card}(E \cap (A_j P) \cap D) \geq \frac{1}{2^{d+1}d} \mathcal{H}^{d-1}(E \cap D).$$

Since we assumed that $\mathcal{H}^{d-1}(E \cap \partial D) = 0$ we can suppose that $E \subset \overset{\circ}{D}$. Since the manifold E has bounded curvature, it is a doubling space and Vitali–Lebesgue covering theorem (see [Hei01, p. 4]) asserts that for all $r_0 > 0$, we can find a disjoint family of (relatively compact) geodesic balls $(E_{r_n})_{n \geq 0}$ in E such that the geodesic ball E_{r_n} has radius $r_n < r_0$, and such that

$$\mathcal{H}^{d-1} \left(E \setminus \left(\bigsqcup_{n \in \mathbb{N}} E_{r_n} \right) \right) = 0.$$

By linearity of both sides of (3.43) and monotone convergence, it is sufficient to prove the inequality (3.43) by replacing E with E_r , a small (relatively compact) geodesic ball of

radius $r < r_0$ centered at some point $x \in E$. For r sufficiently small, the geodesic ball E_r is comparable to a \mathbb{R}^{d-1} -ball. More precisely, set $B_r(x) = \exp_x^{-1}(E_r)$. Riemannian volume comparison theorems asserts that

$$\mathcal{H}^{d-1}(B_r(x)) = \mathcal{H}^{d-1}(E_r)(1 + o(r_0)),$$

and the estimate is uniform on E by the curvature bound assumption. We will prove that for some $j \in \{1, \dots, 2^d\}$,

$$\mathbb{E} \left[\sum_{j=1}^{2^d} \text{Card} \{E_r \cap (A_j P)\} \right] \geq \mathbb{E} [\text{Card} \{E_r \cap (A_j P)\}] \geq \frac{1}{2^d} \mathcal{H}^{d-1}(E_r).$$

Let \mathbf{n}_x denote a normal unit vector at x . A little geometry shows that we can choose j such that

$$|\langle \overrightarrow{A_j x}, \mathbf{n}_x \rangle| \geq \frac{1}{2}. \quad (3.44)$$

For r_0 small enough and uniformly on E , the hypersurface E_r is almost flat and the line $(A_j P)$ has at most one point of intersection with E_r . The opposite would imply that for some $y \in E_r$ the line $(A_j y)$ is tangent to E_r at some point y . That is $\langle \overrightarrow{A_j y}, \mathbf{n}_y \rangle = 0$. But it contradicts the inequality (3.44) and the continuity on E_r of the mapping

$$x \mapsto |\langle \overrightarrow{A_j x}, \mathbf{n}_x \rangle|.$$

The uniformity of E comes from the fact that the modulus of continuity of this application is controlled by the curvature of E . We deduce

$$\mathbb{E} [\text{Card} \{E_r \cap (A_j P)\}] = \mathbb{P} (\text{Card} \{E_r \cap (A_j P)\} \neq \emptyset).$$

Uniformly on x in E , we can find $r' = r + o(r_0)$ such that every line that passes through A_j and intersects the $d-1$ -dimensional ball $B_{r'}(x) \subset T_x E$, also passes through E_r . Indeed, the central projection of E_r onto $T_x E$ with center of projection A_j is almost a \mathbb{R}^{d-1} -ball and must contain a ball of radius $r' = r + o(r_0)$ (see Figure 3.3).

We deduce

$$\mathbb{P} (\text{Card} \{E_r \cap (A_j P)\} \neq \emptyset) \geq \mathbb{P} (\text{Card} \{B_{r'}(x) \cap (A_j P)\} \neq \emptyset).$$

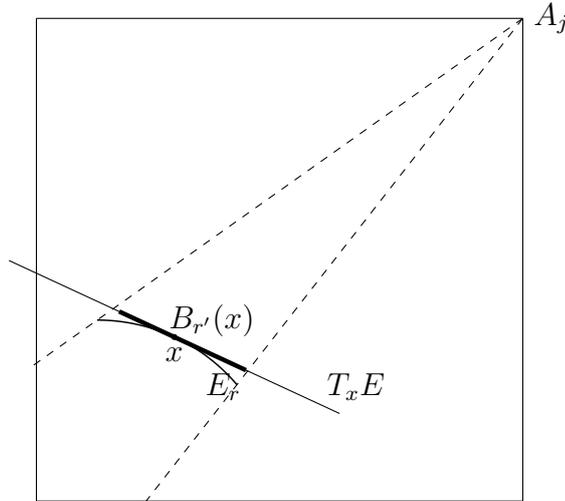


Figure 3.3 – Construction of $B_{r'}(x)$.

But the right hand side is easy to estimate. It is the volume of the cone in $[0, 1]^d$, based at A_j and generated by the ball $B_{r'}(x)$. The formula base \times height/ d gives

$$\begin{aligned} \mathbb{P}(\text{Card}\{B_{r'}(x) \cap (A_j P)\} \neq \emptyset) &\geq \frac{\mathcal{H}^{d-1}(B_{r'}(x))}{d} |\overrightarrow{A_j x}, \mathbf{n}_x| \\ &\geq \frac{1}{2d} \mathcal{H}^{d-1}(B_r(x))(1 + o(r_0)) \\ &\geq \frac{1}{2d} \mathcal{H}^{d-1}(E_r)(1 + o(r_0)). \end{aligned}$$

Patching up the above estimates we recover

$$\mathbb{E} \left[\sum_{j=1}^{2^d} \text{Card}\{E \cap (A_j P) \cap D\} \right] \geq \frac{1}{2d} \mathcal{H}^{d-1}(E \cap D)(1 + o(r)),$$

and letting r go to zero we deduce the result. □

VARIANCE FOR THE NUMBER OF ZEROS OF TRIGONOMETRIC POLYNOMIAL WITH DEPENDENT COEFFICIENTS

This chapter consists in the second submitted paper of this thesis [Gas21b], concerning the variance asymptotics for trigonometric polynomials with dependent coefficients.

4.1	Introduction	133
4.2	Kac–Rice formula for the second moment	136
4.2.1	Kac–Rice formula for the variance	137
4.2.2	Matrix notations	139
4.2.3	Factorization of the Kac density	139
4.3	Proof of the main theorems	143
4.3.1	Asymptotics of the Kac density	143
4.3.2	Proof of Theorem 4.1.2	146
4.3.3	Proof of Theorem 4.1.1	147

4.1 Introduction

The asymptotic behavior of the variance of the number of zeros of random trigonometric polynomials $\sum a_k \cos(kt) + b_k \sin(kt)$ with independent Gaussian coefficients has been established in [GW11]. Since then, the variances of numerous models have been studied: for instance, see [ADL16] for the analogous model $\sum a_k \cos(kt)$, or more recently [LP21] for random orthogonal polynomials on the real line. In this paper we make explicit the asymptotics of the variance of the number of zeros of random trigonometric polynomials with dependent coefficients.

This paper serves three purposes. Firstly, it is the natural continuation of [ADP19] and [APP21], in which the asymptotic behavior of the expected number of zeros of this model has been established. Secondly, we avoid the use of the explicit expression for the first and second order Kac densities for the number of zeros (see e.g. [LP21, Lemma 2.2]), but we rather exploit on one hand the intrinsic properties of the Kac density and on the other hand the convergence in distribution of the model of random trigonometric polynomials towards the sinc process. This point of view allows us to avoid heavy computations that usually comes with the Kac–Rice formula, in particular in the non-stationary case, and allows us to give a unified point of view for all the previously cited Gaussian models for which the variance asymptotics is known.

Let us now detail our model of random trigonometric polynomial with dependent coefficients. Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ be the one-dimensional torus, which can be identified with a segment of \mathbb{R} of length 2π . For $s, t \in \mathbb{T}$ we define the distance $\text{dist}(s, t) = d(s - t, 2\pi\mathbb{Z})$. For $s \in \mathbb{T}$ we define

$$X_n(s) = \frac{1}{\sqrt{n}} \sum_{k=1}^n a_k \cos(ks) + b_k \sin(ks), \quad (4.1)$$

with $(a_k)_k, (b_k)_k$ two independent stationary centered Gaussian processes with correlation function $\rho : \mathbb{Z} \rightarrow \mathbb{R}$. That is for $k, l \geq 0$, $\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \rho(k - l)$. Thanks to Bochner Theorem, the correlation function ρ is associated with a spectral measure μ on the torus \mathbb{T} via the relation

$$\rho(k) = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-iku} d\mu(u).$$

We denote by $Z_{X_n}(I)$ the number of zeros of X_n on a subinterval I of the torus \mathbb{T} . Under suitable conditions on the spectral measure μ , it has been shown in [ADP19] and [APP21] that the expectation of the number of zeros of the process X_n on \mathbb{T} behaves like $\frac{2}{\sqrt{3}}n$, as in the independent framework. The following theorem makes explicit the variance asymptotics

for the number of zeros of the process X_n on \mathbb{T} , as n grows to infinity.

Theorem 4.1.1. *We suppose that the spectral measure μ has a positive continuous density ψ with respect to the Lebesgue measure on the torus \mathbb{T} . Then there is an explicit positive constant γ_2 that does not depend on ψ , such that for any subinterval I of the torus,*

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{X_n}(I))}{n} = \text{length}(I) \gamma_2.$$

The universal constant $\gamma_2 \simeq 0.089$ is thus the same constant computed in [GW11] in the particular framework of independent Gaussian random variables. Note that this universality result is not unsurprising, given the non-universality behavior in the non-Gaussian framework (see [BCP19]).

Theorem 4.1.1 is in fact a corollary of the next Theorem 4.1.2, that puts forwards the principal ingredients necessary to obtain such a universal asymptotics for the variance. Let $(X_n)_n$ be a sequence of centered Gaussian random processes defined on an open subinterval I of \mathbb{T} or \mathbb{R} . We define for $n \in \mathbb{N}$ the process $Y_n = X_n(\cdot/n)$, defined on the open subinterval nI of the rescaled torus $n\mathbb{T}$ or \mathbb{R} endowed with its canonical distance. We denote r_n the covariance function of Y_n

$$\forall s, t \in nI, \quad r_n(s, t) := \mathbb{E}[Y_n(s)Y_n(t)].$$

Let also Y_∞ be a stationary Gaussian random process on \mathbb{R} with covariance function r_∞ , and ψ be a uniformly continuous function on I , bounded above and below by positive constants.

We denote by $C_0(\mathbb{R})$ the space of continuous functions that converges to zero in $\pm\infty$.

Theorem 4.1.2. *We suppose that the sequence of random processes $(X_n)_n$ satisfies the following two conditions.*

(H_1) *The sequence of processes $(X_n)_{n \in \mathbb{N}}$ is of class \mathcal{C}^2 on I and for $u, v \in \{0, 1, 2\}$, the following convergence holds uniformly for $x \in I$ and locally uniformly for $s, t \in \mathbb{R}$,*

$$\lim_{n \rightarrow +\infty} r_n^{(u,v)}(nx + s, nx + t) = \psi(x)r_\infty^{(u,v)}(s, t).$$

(H_2) *There is a positive function $g \in L^2(\mathbb{R}) \cap C_0(\mathbb{R})$ such that for $u, v \in \{0, 1\}$,*

$$\forall s, t \in nI, \quad |r_n^{(u,v)}(s, t)| \leq g(\text{dist}(s, t)).$$

Then there is an explicit positive constant γ_2 depending only on r_∞ , such that

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{X_n}(I))}{n} = \text{length}(I) \gamma_2.$$

This last Theorem 4.1.2 covers Theorem 4.1.1, as we will prove that the assumptions on the spectral measure μ in Theorem 4.1.1 ensure that the associated sequence of processes $(X_n)_n$ satisfies hypotheses (H_1) and (H_2) of Theorem 4.1.2.

In fact, Theorem 4.1.2 covers various previously known results about the asymptotics of variance of the number of zeros of general random trigonometric polynomials. For instance, Theorem 4.1.2 allows to make explicit the variance asymptotics for the number of zeros on any compact subinterval of $]0, \pi[$ of the process

$$\tilde{X}_n(s) = \frac{1}{\sqrt{n}} \sum_{k=0}^n a_k \cos(ks + \theta), \quad \theta \in \mathbb{R}, \quad (4.2)$$

with $(a_k)_{k \geq 0}$ a stationary sequence of centered random Gaussian variables whose spectral measure has a positive continuous density. Note that the variance asymptotics for the number of zeros on $[0, \pi]$ for this process was established in the case of Gaussian iid $(a_k)_{k \geq 0}$ in [ADL16]. More recently, the authors in [LP21] established the variance asymptotics for the number of zeros of sums of real orthogonal polynomials with iid Gaussian coefficients. Let $(P_n)_n$ be a sequence of real orthogonal polynomials with respect to a density measure

$d\phi$ supported on $[-1, 1]$. We define

$$X_n(s) = \frac{1}{\sqrt{n}} \sum_{k=0}^n a_k P_k(\cos(s)),$$

where $(a_n)_{n \geq 0}$ are iid centered Gaussian random variables. Under suitable conditions on the density ϕ , hypotheses (H_1) and (H_2) of Theorem 4.1.2 are satisfied with $\psi = 1/\phi$ and Y_∞ a process with sinc covariance function, and Theorem 4.1.2 implies [LP21, Cor. 1.3].

Section 4.2 of the paper is devoted to the study of the Kac density, that is the integrand in the Kac–Rice formula that gives the second moment of the number of zeros of a general one-dimensional Gaussian process. Under suitable regularity assumptions on the covariance function, we first remove the apparent singularity of the Kac density near the diagonal. We then show a factorization property of the second order Kac density.

In Section 4.3, we make explicit the asymptotics of the Kac density associated with a sequence of processes satisfying the hypotheses of Theorem 4.1.2. Hypothesis (H_1) implies in particular that as n grows to infinity, the sequence $(Y_n)_n$ has a limit proportional to the stationary process Y_∞ on \mathbb{R} . Together with the results of Section 4.2, we deduce that the Kac density associated with the process Y_n uniformly converges towards the Kac density of the process Y_∞ . This fact leads to the proof of Theorem 4.1.2. We then check that the model of trigonometric polynomials with dependent coefficients satisfying the hypotheses of Theorem 4.1.1, also satisfies the hypotheses of Theorem 4.1.2.

4.2 Kac–Rice formula for the second moment

In this section we give a short study of the Kac density for a general centered Gaussian process satisfying some natural regularity assumptions. We first remove the apparent singularity along the diagonal, which is the object of Lemma 4.2.2 and then prove a factorization property of the Kac density, whose result is contained in Corollary 4.2.6.

In the following, let Y be a centered Gaussian process defined on an interval I with covariance function r . We assume that the process Y has \mathcal{C}^2 sample paths and that the joint distribution of the Gaussian vector $(Y(s), Y(t))$ is nondegenerate for $s \neq t$. We denote by p_s the density of $Y(s)$ and by $p_{s,t}$ the density of the Gaussian vector $(Y(s), Y(t))$. We

will assume the existence of a constant M such that for all $s, t \in I$ and $u, v \leq 2$,

$$|r^{(u,v)}(s, t)| \leq M \quad \text{and} \quad \det \text{Cov}(Y(s), Y'(s)) \geq \frac{1}{M}. \quad (4.3)$$

4.2.1 Kac–Rice formula for the variance

We define

$$Z_Y(I) := \text{Card} \{s \in I \mid Y(s) = 0\},$$

the number of zeros of Y in the interval I . The following proposition gives an integral representation for the first and second moment of Z_Y , and a proof can be found in [AW09, Thm. 3.2].

Proposition 4.2.1 (Kac–Rice formula).

$$\mathbb{E}[Z_Y] = \int_I \rho_1(s) ds \quad \text{and} \quad \mathbb{E}[Z_Y^2] - \mathbb{E}[Z_Y] = \iint_{I^2} \rho_2(s, t) ds dt,$$

with

$$\begin{aligned} \rho_1(s) &= \mathbb{E}[|Y'(s)| \mid Y(s) = 0] p_s(0), \\ \rho_2(s, t) &= \mathbb{E}[|Y'(s)||Y'(t)| \mid Y(s) = Y(t) = 0] p_{s,t}(0, 0). \end{aligned}$$

The function ρ_1 (resp. ρ_2) is called the Kac density associated with the first (resp. second) moment of the number of zeros. We then have the integral representation

$$\begin{aligned} \text{Var}(Z_Y) &= \left(\mathbb{E}[Z_Y^2] - \mathbb{E}[Z_Y] - \mathbb{E}[Z_Y]^2 \right) + \mathbb{E}[Z_Y] \\ &= \left(\iint_{I^2} \rho_2(s, t) - \rho_1(s)\rho_1(t) ds dt \right) + \int_I \rho_1(s) ds. \end{aligned} \quad (4.4)$$

Note that the quantity $\rho_2(s, s)$ is ill-defined, since the Gaussian vector $(Y(s), Y'(s))$ is degenerate. The following lemma allows us to remove the nondegeneracy of the function $\rho_2(s, t)$ when s and t are close, in order to show that the second moment is indeed well-defined. The following procedure is standard, see for instance [AL21b] for a general treatment, or [AW09, Prop. 5.8]. The constant M is defined in (4.3).

Lemma 4.2.2. *There is a positive constant η depending only on M , such that for $|s - t| \leq \eta$,*

$$\rho_2(s, t) \leq M^{3/2}|t - s|.$$

Proof. We define for $s \neq t$ the quantities

$$Y[s, t] = \frac{Y(t) - Y(s)}{t - s} \quad \text{and} \quad Y[s, s, t] = \frac{Y[s, t] - Y'(s)}{t - s}.$$

And we extend them by continuity respectively by $Y'(s)$ and $Y''(s)/2$ when $s = t$. The mean value theorem and the uniform bounds (4.3) on $r^{(u,v)}$ imply that

$$\|\text{Cov}(Y(s), Y[s, t]) - \text{Cov}(Y(s), Y'(s))\| \leq M|t - s| \quad \text{and} \quad \mathbb{E}[Y[s, s, t]^2] \leq M.$$

The lower bound (4.3) on $\det \text{Cov}(Y(s), Y'(s))$ implies that for some positive constant η depending only on M , and $s, t \in I$ such that $|s - t| \leq \eta$,

$$\det \text{Cov}(Y(s), Y[s, t]) \geq \frac{1}{2M}.$$

The density $p_{s,t}(0, 0)$ of the vector $(Y(s), Y(t))$ then satisfies for $|s - t| \leq \eta$

$$p_{s,t}(0, 0) = \frac{1}{2\pi\sqrt{\det(\text{Cov}(Y(s), Y(t)))}} = \frac{1}{2\pi|t - s|\sqrt{\det \text{Cov}(Y(s), Y[s, t])}} \leq \frac{\sqrt{M}}{|t - s|}.$$

The conditional Gaussian density appearing in the Kac-Rice formula satisfies, by Cauchy-Schwarz inequality for conditional expectation

$$\begin{aligned} \mathbb{E}[|Y'(s)||Y'(t)| \mid Y(s) = Y(t) = 0] &= |t - s|^2 \mathbb{E}[|Y[s, s, t]||Y[t, t, s]| \mid Y(s) = Y[s, t] = 0] \\ &\leq |t - s|^2 \sqrt{\mathbb{E}[Y[s, s, t]^2] \mathbb{E}[Y[t, t, s]^2]} \\ &\leq M|t - s|^2, \end{aligned}$$

and the conclusion follows. □

Remark 4.2.3. Making explicit the convergence, one has in fact

$$\lim_{t \rightarrow s} \frac{\rho_2(s, t)}{|t - s|} = \frac{\mathbb{E}[(Y''(s))^2 \mid Y(s) = Y'(s) = 0]}{8\pi\sqrt{\det \text{Cov}(Y(s), Y'(s))}} = \frac{\det(\text{Cov}(Y(s), Y'(s), Y''(s)))}{8\pi(\det \text{Cov}(Y(s), Y'(s)))^{3/2}}.$$

4.2.2 Matrix notations

Before going further, we will need a few matrix notations for the next subsection. In the following, Ω (resp. Σ) are square matrices of size 2 (resp. 4). We write

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \quad \text{and} \quad \Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array} \right),$$

where for $i, j \in \{1, 2\}$, Ω_{ij} are real numbers and Σ_{ij} are square matrices of size 2. We define the diagonal and off diagonal matrices

$$\Omega^{\text{diag}} = \begin{pmatrix} \Omega_{11} & 0 \\ 0 & \Omega_{22} \end{pmatrix}, \quad \Omega^{\text{off}} = \begin{pmatrix} 0 & \Omega_{12} \\ \Omega_{21} & 0 \end{pmatrix},$$

and

$$\Sigma^{\text{diag}} = \left(\begin{array}{c|c} \Sigma_{11}^{\text{diag}} & \Sigma_{12}^{\text{diag}} \\ \hline \Sigma_{21}^{\text{diag}} & \Sigma_{22}^{\text{diag}} \end{array} \right), \quad \Sigma^{\text{off}} = \left(\begin{array}{c|c} \Sigma_{11}^{\text{off}} & \Sigma_{12}^{\text{off}} \\ \hline \Sigma_{21}^{\text{off}} & \Sigma_{22}^{\text{off}} \end{array} \right),$$

so that

$$\Omega = \Omega^{\text{diag}} + \Omega^{\text{off}} \quad \text{and} \quad \Sigma = \Sigma^{\text{diag}} + \Sigma^{\text{off}}.$$

If Ω_{11} (resp. Σ_{11}) is non-zero (resp. invertible) we define the Schur complements

$$\Omega_c := \Omega_{22} - \frac{\Omega_{12}\Omega_{21}}{\Omega_{11}} \quad \text{and} \quad \Sigma_c := \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}. \quad (4.5)$$

By row reduction, one has

$$\det(\Sigma) = \det(\Sigma_{11}) \det(\Sigma_c) \quad \text{and} \quad (\Sigma^{-1})_{22} = (\Sigma_c)^{-1}. \quad (4.6)$$

4.2.3 Factorization of the Kac density

In this section, we explicit the Kac densities ρ_1 and ρ_2 in terms of Gaussian covariance matrices and we prove a useful factorization property for the Kac density ρ_2 that will be used in the next section. We define

$$\Omega(s) := \text{Cov}(Y(s), Y'(s)) \quad \text{and} \quad \Sigma(s, t) = \text{Cov}(Y(s), Y(t), Y'(s), Y'(t)).$$

From these definitions and the notation of the previous subsection, one has

$$\det \Sigma^{\text{diag}}(s, t) = \det \Omega(s) \det \Omega(t) \quad \text{and} \quad (\Sigma^{\text{diag}}(s, t))_c = \begin{pmatrix} \Omega_c(s) & 0 \\ 0 & \Omega_c(t) \end{pmatrix}. \quad (4.7)$$

The following Lemma 4.2.4 relies the Schur complement of a matrix with Gaussian conditioning. A proof can be found in [AW09].

Lemma 4.2.4. *For $s \neq t$, one has*

$$\text{Law}((Y'(s)) \mid Y(s) = 0) \sim \mathcal{N}(0, \Omega_c(s)).$$

$$\text{Law}((Y'(s), Y'(t)) \mid Y(s) = Y(t) = 0) \sim \mathcal{N}(0, \Sigma_c(s, t)).$$

We define the function $\tilde{\rho}_1$ (resp. $\tilde{\rho}_2$) on the space of symmetric positive definite matrices of size 2 (resp. 4) as

$$\tilde{\rho}_1(\Omega) = \frac{1}{2\pi\sqrt{\det \Omega}} \int_{\mathbb{R}} |y| \exp\left(-\frac{y^2}{2\Omega_c}\right) dy$$

and

$$\tilde{\rho}_2(\Sigma) = \frac{1}{(2\pi)^2\sqrt{\det \Sigma}} \iint_{\mathbb{R}^2} |y_1||y_2| \exp\left(-\frac{1}{2} y^T (\Sigma_c)^{-1} y\right) dy_1 dy_2$$

The previous Lemma 4.2.4 and relations (4.7) then implies the following formulas

$$\rho_1(s) = \tilde{\rho}_1(\Omega(s)), \quad \rho_2(s, t) = \tilde{\rho}_2(\Sigma(s, t)) \quad \text{and} \quad \rho_1(s)\rho_1(t) = \tilde{\rho}_2(\Sigma^{\text{diag}}(s, t)). \quad (4.8)$$

When the random Gaussian vectors $(Y(s), Y'(s))$ and $(Y(t), Y'(t))$ are independent then one has directly from Gaussian conditioning the equality $\rho_2(s, t) = \rho_1(s)\rho_1(t)$. The following Lemma 4.2.5 and Corollary 4.2.6 makes explicit the error term between these two quantities.

Lemma 4.2.5. *Let K be a compact subset of the symmetric positive definite matrices of size 4. There is a constant C_K depending only on the compact K such that*

$$\forall \Sigma \in K, \quad \left| \tilde{\rho}_2(\Sigma) - \tilde{\rho}_2(\Sigma^{\text{diag}}) \right| \leq C_K \|\Sigma^{\text{off}}\|^2.$$

Proof. A straightforward computation shows that

$$\left((\Sigma^{\text{diag}})^{-1} \Sigma^{\text{off}} (\Sigma^{\text{diag}})^{-1} \right)^{\text{diag}} = \left((\Sigma^{\text{diag}})^{-1} \Sigma^{\text{off}} \right)^{\text{diag}} = 0. \quad (4.9)$$

In particular the trace of these matrices are zero and a Taylor expansion of the determinant function yields

$$\det \Sigma = \det \Sigma^{\text{diag}} + O(\|\Sigma^{\text{off}}\|^2). \quad (4.10)$$

The expansion of the inverse of a matrix and identity (4.6) yields

$$(\Sigma_c)^{-1} = \left(\Sigma^{\text{diag}} + \Sigma^{\text{off}} \right)_{22}^{-1} = \left((\Sigma^{\text{diag}})_c \right)^{-1} + \left((\Sigma^{\text{diag}})^{-1} \Sigma^{\text{off}} (\Sigma^{\text{diag}})^{-1} \right)_{22} + O(\|\Sigma^{\text{off}}\|^2).$$

By identity (4.9), there is an explicit coefficient h and a constant C_K such that

$$\left((\Sigma^{\text{diag}})^{-1} \Sigma^{\text{off}} (\Sigma^{\text{diag}})^{-1} \right)_{22} = \begin{pmatrix} 0 & h \\ h & 0 \end{pmatrix} \quad \text{with} \quad |h| \leq C_K \|\Sigma^{\text{off}}\|. \quad (4.11)$$

We now express the difference

$$\tilde{\rho}_2(\Sigma) - \tilde{\rho}_2(\Sigma^{\text{diag}}) = \frac{1}{4\pi^2} (R_1 + R_2),$$

with

$$R_1 = \left[\frac{1}{\sqrt{\det \Sigma}} - \frac{1}{\sqrt{\det \Sigma^{\text{diag}}}} \right] \left(\iint_{\mathbb{R}^2} |y_1| |y_2| \exp \left(-\frac{1}{2} y^T (\Sigma_c)^{-1} y \right) dy_1 dy_2 \right),$$

$$R_2 = \frac{1}{\sqrt{\det(\Sigma^{\text{diag}})}} \iint_{\mathbb{R}^2} |y_1| |y_2| \left[\exp \left(-\frac{1}{2} y^T (\Sigma_c)^{-1} y \right) - \exp \left(-\frac{1}{2} y^T (\Sigma^{\text{diag}})_c^{-1} y \right) \right] dy_1 dy_2.$$

Estimate (4.10) directly implies

$$R_1 = O \left(\|\Sigma_{\text{off}}(s, t)\|^2 \right). \quad (4.12)$$

A Taylor expansion of the exponential function and relations (4.6) and (4.11) yields

$$R_2 = \frac{2h}{\sqrt{\det(\Sigma^{\text{diag}})}} \iint_{\mathbb{R}^2} |y_1| |y_2| y_1 y_2 \exp \left(-\frac{1}{2} y^T (\Sigma^{\text{diag}})_c^{-1} y \right) dy_1 dy_2 + O(\|\Sigma^{\text{off}}\|^2).$$

By parity the double integral is zero, thus

$$R_2 = O\left(\|\Sigma_{\text{off}}(s, t)\|^2\right). \quad (4.13)$$

The Lagrange rest theorem implies that all the O appearing above are uniform for Σ in the compact set K . Gathering estimates (4.12) and (4.13), we deduce the existence of a constant C_K such that

$$\forall \Sigma \in K, \quad \left| \tilde{\rho}_2(\Sigma) - \tilde{\rho}_2(\Sigma^{\text{diag}}) \right| \leq C_K \|\Sigma^{\text{off}}\|^2.$$

□

Corollary 4.2.6. *There are positive constants ε and C depending only on the constant M defined in (4.3), such that for all $s, t \in I$ satisfying $\max_{u,v \in \{0,1\}} |r^{(u,v)}(s, t)| \leq \varepsilon$,*

$$\det \Sigma(s, t) \geq \frac{1}{2C^2} \quad \text{and} \quad |\rho_2(s, t) - \rho_1(s)\rho_1(t)| \leq C \left(\sup_{u,v \in \{0,1\}} |r^{(u,v)}(s, t)| \right)^2.$$

Proof. Given the equality (4.7), the upper bound assumption (4.3) and the regularity of the determinant, we deduce the existence of a constant C depending only on the constant M , such that

$$|\det \Sigma(s, t) - \det \Omega(s) \det \Omega(t)| \leq C \|\Sigma_{\text{off}}(s, t)\| = C \max_{u,v \in \{0,1\}} |r^{(u,v)}(s, t)|.$$

The lower bound assumption on the determinant (4.3) implies the existence of a positive number ε depending only on M such that for all $s, t \in I$ with $\max_{u,v \in \{0,1\}} |r^{(u,v)}(s, t)| \leq \varepsilon$ one has

$$\det \Sigma(s, t) \geq \frac{1}{2M^2}. \quad (4.14)$$

From now we fix two such real numbers s and t . The above lower bound (4.14) implies that the matrix $\Sigma(s, t)$ lives in a compact subset K_M of the space of symmetric positive definite matrices, that depends only on the constant M . The previous Lemma 4.2.5 applied to the matrix $\Sigma(s, t)$ and the relations (4.8) imply the existence of a constant C depending only

on M such that

$$|\rho_2(s, t) - \rho_1(s)\rho_1(t)| \leq C \left(\max_{u, v \in \{0, 1\}} |r^{(u, v)}(s, t)| \right)^2.$$

□

4.3 Proof of the main theorems

4.3.1 Asymptotics of the Kac density

Let $(X_n)_n$ be a sequence of processes on I satisfying the hypotheses of Theorem 4.1.2. We recall the definition for $n \in \mathbb{N}$ of the process $Y_n = X_n(\cdot/n)$ with covariance function r_n , and Y_∞ is a centered stationary Gaussian process with covariance function r_∞ . We define for $n \in \mathbb{N} \cup \{+\infty\}$

$$\Omega_n(s) := \text{Cov}(Y_n(s), Y'_n(s)) \quad \text{and} \quad \Sigma_n(s, t) = \text{Cov}(Y_n(s), Y_n(t), Y'_n(s), Y'_n(t)).$$

By stationarity, Ω_∞ is constant and the matrix $\Sigma_n(s, t)$ depends only on the difference $s - t$. The following Lemma 4.3.1 shows that the process Y_n satisfies the bound (4.3), uniformly for n large enough.

Lemma 4.3.1. *There is a constant C such that for $s, t \in nI$ and $u, v \in \{0, 1, 2\}$ one has*

$$|r_n^{(u, v)}(s, t)| \leq C.$$

For all parameter $\eta > 0$, there is a positive constant C_η and a rank n_0 such that for $n \geq n_0$ and for all $s, t \in nI$ with $\text{dist}(s, t) > \eta$,

$$\det \Sigma_n(s, t) \geq C_\eta.$$

Proof. The uniform convergence of hypothesis (H_1) on the process X_n ensures that the quantities $\mathbb{E}[Y_n(s)^2]$, $\mathbb{E}[Y'_n(s)^2]$ and $\mathbb{E}[Y''_n(s)^2]$ are bounded by some constant C . By Cauchy-Schwartz inequality, for $u, v \in \{0, 1, 2\}$ and $s, t \in nI$ one has $|r_n^{(u, v)}(s, t)| \leq C$. For $s \in \mathbb{R}$, the random variables $Y_\infty(s)$ and $Y'_\infty(s)$ are decorrelated by stationarity, thus

$$\det \Omega_\infty = -r_\infty(0)r''_\infty(0) > 0.$$

We deduce the following uniform convergence on $x \in I$

$$\lim_{n \rightarrow +\infty} \Omega_n(nx) = \psi(x)\Omega_\infty. \quad (4.15)$$

The function ψ is bounded from below by a positive constant C_ψ , and the convergence (4.15) implies the existence of rank n_0 independent of x such that

$$\forall n \geq n_0, \forall x \in I, \quad \det[\Omega_n(nx)] \geq \frac{C_\psi^2 \det \Omega_\infty}{2} > 0. \quad (4.16)$$

In particular, the function Y_n satisfies the bounds (4.3) for some constant M independent of n . Now the function g of hypothesis (H_2) in Theorem 4.1.2 decreases to zero at infinity. Given $\varepsilon > 0$, there is a constant T_ε such that for s, t satisfying $\text{dist}(s, t) > T_\varepsilon$, one has

$$\max_{u, v \in \{0, 1\}} |r_n^{(u, v)}(s, t)| \leq \varepsilon.$$

One can then apply Corollary 4.2.6, to deduce the existence of a constant T independent of n , such that for all s, t satisfying $\text{dist}(s, t) > T$, it holds that

$$\det \Sigma_n(s, t) \geq \frac{1}{2M^2}. \quad (4.17)$$

Let $\eta > 0$. Since the process Y_∞ is stationary and the support of its spectral measure has an accumulation point, then (see [AW09, Ex. 3.5]) the covariance matrix $\Sigma_\infty(s, t)$ is nondegenerate for $s \neq t$. By compactness, one can find a positive constant C_η such that for all $s, t \in nI$ with $\eta \leq \text{dist}(s, t) \leq T$, one has $\det \Sigma_\infty(s - t) \geq C_\eta$. The uniform convergence of $r_n^{(u, v)}$ towards $r_\infty^{(u, v)}$ and then implies that for n greater than some rank n_0 ,

$$\det \Sigma_n(s, t) \geq \frac{C_\eta}{2}. \quad (4.18)$$

Gathering (4.17) and (4.18) we deduce the second assertion. \square

For $n \in \mathbb{N} \cup \{+\infty\}$ we define $\rho_{1, n}$ and $\rho_{2, n}$ the Kac densities associated with the process Y_n . By stationarity, the function $\rho_{1, \infty}$ is a constant and the function $\rho_{2, \infty}$ depends only on the difference $s - t$. We deduce the following corollary.

Corollary 4.3.2. *Let $\eta > 0$. We have the following uniform convergences, uniformly in $x \in I$ and s, t in a compact set of \mathbb{R} with $\text{dist}(s, t) > \eta$.*

$$\lim_{n \rightarrow +\infty} \rho_{1,n}(nx) = \rho_{1,\infty} \quad \text{and} \quad \lim_{n \rightarrow +\infty} \rho_{2,n}(nx + s, nx + t) = \rho_{2,\infty}(s - t).$$

Proof. Hypothesis (H_1) and the previous Lemma 4.3.1 implies the following uniform convergences uniformly in $x \in I$ and s, t in a compact set of \mathbb{R} with $|s - t| > \eta$

$$\lim_{n \rightarrow +\infty} \Omega_n(nx) = \psi(x)\Omega_\infty \quad \text{and} \quad \lim_{n \rightarrow +\infty} \Sigma_n(nx + s, nx + t) = \psi(x)\Sigma_\infty(s - t). \quad (4.19)$$

The functions $\tilde{\rho}_1$ (resp. $\tilde{\rho}_2$) is continuous on the space of symmetric positive definite matrices, which combined with the convergence (4.19) directly implies the conclusion of the corollary. Note that by a change of variable, the limit does not depend on the function ψ . \square

The following Lemma 4.3.3 establishes a decay property for the Kac density, whose rate depends on the function g of hypothesis (H_2) of Theorem 4.1.2.

Lemma 4.3.3. *There is a constant C and a rank n_0 such that for $n \geq n_0$ and $s, t \in nI$,*

$$|\rho_{2,n}(s, t) - \rho_{1,n}(s)\rho_{1,n}(t)| \leq Cg^2(\text{dist}(s, t)).$$

Proof. According to Corollary 4.2.6 and the proof of Lemma 4.3.1, there are constants T and C independent of n such that when $\text{dist}(s, t) \geq T$ it holds that

$$|\rho_{2,n}(s, t) - \rho_{1,n}(s)\rho_{1,n}(t)| \leq Cg^2(\text{dist}(s, t)).$$

When $\text{dist}(s, t) \leq T$, the functions $\rho_{1,n}$ and $\rho_{2,n}$ are bounded by a constant independent of n , according to Lemmas 4.2.2 and 4.3.1. Since the function g is assumed to be continuous and positive, it is bounded below by a positive constant on compact subsets of \mathbb{R} and the conclusion follows. \square

4.3.2 Proof of Theorem 4.1.2

We identify the interval I with $]a, b[$ and we write

$$\begin{aligned} \text{Var}(Z_{X_n}(I)) &= \text{Var}(Z_{Y_n}(nI)) \\ &= n \int_a^b \rho_{1,n}(nx) dx \\ &\quad + n \int_a^b \int_{n(a-x)}^{n(b-x)} (\rho_{2,n}(nx, nx + \tau) - \rho_{1,n}(nx)\rho_{1,n}(nx + \tau)) d\tau dx. \end{aligned}$$

For the first term, Corollary 4.3.2 implies that

$$\lim_{n \rightarrow +\infty} \int_a^b \rho_{1,n}(nx) dx = |b - a| \rho_{1,\infty}.$$

For the second term, we write

$$\begin{aligned} R_n &= \int_a^b \int_{n(a-x)}^{n(b-x)} (\rho_{2,n}(nx, nx + \tau) - \rho_{1,n}(nx)\rho_{1,n}(nx + \tau)) d\tau dx \\ &\quad - |b - a| \int_{\mathbb{R}} (\rho_{2,\infty}(\tau) - \rho_{1,\infty}^2) d\tau. \end{aligned}$$

We fix two positive constants $\eta > 0$ and A . We split R_n into four parts, as

$$R_n = R_{1,n}^{A,\eta} - R_{2,n}^{A,\eta} + R_{3,n}^{A,\eta} - R_{3,\infty}^{A,\eta},$$

with

$$\begin{aligned} R_{1,n}^{A,\eta} &:= \int_a^b \int_{n(a-x)}^{n(b-x)} \mathbb{1}_{\eta \leq |\tau| \leq A} [(\rho_{1,n}(nx)\rho_{1,n}(nx + \tau) - \rho_{1,\infty}^2)] d\tau dx, \\ R_{2,n}^{A,\eta} &:= \int_a^b \int_{n(a-x)}^{n(b-x)} \mathbb{1}_{\eta \leq |\tau| \leq A} [(\rho_{2,n}(nx, nx + \tau) - \rho_{2,\infty}(\tau))] d\tau dx, \\ R_{3,n}^{A,\eta} &:= \int_a^b \int_{n(a-x)}^{n(b-x)} \mathbb{1}_{\{|\tau| \geq A\} \cup \{|\tau| \leq \eta\}} (\rho_{2,n}(nx, nx + \tau) - \rho_{1,n}(nx)\rho_{1,n}(nx + \tau)) d\tau dx, \\ R_{3,\infty}^{A,\eta} &:= |b - a| \int_{\mathbb{R}} \mathbb{1}_{\{|\tau| \geq A\} \cup \{|\tau| \leq \eta\}} (\rho_{2,\infty}(\tau) - \rho_{1,\infty}^2) d\tau. \end{aligned}$$

Corollary 4.3.2 directly implies that $R_{1,n}^{A,\eta}$ and $R_{2,n}^{A,\eta}$ converge towards 0 when n goes to infinity. The bound given by Lemma 4.3.3 implies that for some constant C independent

of A and η ,

$$|R_{3,n}^{A,\eta}| + |R_{3,\infty}^{A,\eta}| \leq C|b - a| \int_{\mathbb{R}} \mathbb{1}_{\{|\tau| \geq A\} \cup \{|\tau| \leq \eta\}} g^2(\tau) d\tau. \quad (4.20)$$

Gathering the estimates for $R_{3,n}^{A,\eta}$, $R_{3,n}^{A,\eta}$, $R_{3,n}^{A,\eta}$ and $R_{3,\infty}^{A,\eta}$, we deduce that for some constant C ,

$$\limsup_{n \rightarrow +\infty} |R_n| \leq C|b - a| \int_{\mathbb{R}} \mathbb{1}_{\{|\tau| \geq A\} \cup \{|\tau| \leq \eta\}} g^2(\tau) d\tau.$$

The function g is assumed to be square integrable. Letting A go to infinity and η go to zero, we deduce that $\lim_{n \rightarrow +\infty} R_n = 0$, from which follows the following convergence

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{X_n}(I))}{n} = \text{length}(I) \gamma_2,$$

where

$$\gamma_2 := \int_{\mathbb{R}} (\rho_{2,\infty}(\tau) - \rho_{1,\infty}^2) d\tau + \rho_{1,\infty}.$$

It remains to show the positivity of the constant C_∞ . The above proof shows in fact that

$$\gamma_2 = \lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_{Y_\infty}[0, n])}{n},$$

and it has been shown (see [AL21b]) that $C_\infty > 0$ for a large class of processes including Y_∞ .

4.3.3 Proof of Theorem 4.1.1

In the following, we consider the sequence of trigonometric Gaussian polynomials $(X_n)_n$ defined in (4.1). Assuming the hypotheses of Theorem 4.1.1, we show that this sequence of processes satisfies hypotheses (H_1) and (H_2) of Theorem 4.1.2, from which follows the conclusion of Theorem 4.1.1. Following [APP21], the next computation gives an integral

expression for the covariance function r_n of $Y_n = X_n(\cdot/n)$.

$$\begin{aligned}
 r_n(s, t) &:= \mathbb{E}[Y_n(s)Y_n(t)] \\
 &= \frac{1}{n} \sum_{k,l=1}^n \rho(k-l) \cos\left(\frac{ks-lt}{n}\right) \\
 &= \frac{1}{2\pi n} \int_0^{2\pi} \operatorname{Re} \left(\sum_{k,l=1}^n e^{-i(k-l)y} e^{\frac{iks-ilt}{n}} \right) d\mu(y) \\
 &= \cos\left(\frac{n+1}{2n}(s-t)\right) \frac{1}{2\pi} \int_0^{2\pi} K_n\left(\frac{s}{n}-y, \frac{t}{n}-y\right) \psi(y) dy, \tag{4.21}
 \end{aligned}$$

where K_n is the two points Fejér kernel

$$K_n(x, y) = \frac{1}{n} \frac{\sin\left(\frac{nx}{2}\right) \sin\left(\frac{ny}{2}\right)}{\sin\left(\frac{x}{2}\right) \sin\left(\frac{y}{2}\right)}.$$

In the case where $\rho(k-l) = \delta_{k,l}$, the measure μ is the normalized Lebesgue measure on $[-\pi, \pi]$. In that case, we denote by $r_{0,n}$ its covariance function, whose expression is given by

$$r_{0,n}(s, t) = \frac{1}{2n} \left[\frac{\sin\left(\left(\frac{2n+1}{2n}\right)(s-t)\right)}{\sin\left(\frac{s-t}{2n}\right)} - 1 \right]. \tag{4.22}$$

By assumption the spectral measure μ has a continuous and positive density ψ on \mathbb{T} . The following two lemmas show that the covariance function r_n satisfies hypotheses (H_1) and (H_2) of Theorem 4.1.2 with function $g(t) = C/(1+|t|^\alpha)$, for $1/2 < \alpha < 1$, and Y_∞ a stationary Gaussian process with sinc covariance function.

Lemma 4.3.4. *Let $u, v \geq 0$. Uniformly for $x \in \mathbb{T}$ and s, t in compact subsets of \mathbb{R} ,*

$$\lim_{n \rightarrow +\infty} r_n^{(u,v)}(nx+s, nx+t) = \psi(x)(-1)^v \operatorname{sinc}^{(u+v)}(s-t).$$

Proof. Let us first remark that the covariance function r_n is a trigonometric polynomial and can thus be extended to an analytic function on \mathbb{C} . We will prove that the conclusion of Lemma 4.3.4 holds when s and t belong to a compact subset of \mathbb{C} . By analyticity, it

suffices then to prove the lemma for $u = v = 0$. We have

$$r_n(nx + s, nx + t) = I_n^x(s, t) + \psi(x)r_{0,n}(nx + s, nx + t),$$

where

$$I_n^x(s, t) = \cos\left(\frac{n+1}{2n}(s-t)\right) \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n\left(\frac{s}{n} - y, \frac{t}{n} - y\right) [\psi(x+y) - \psi(x)] dy.$$

From expression (4.22), one has uniformly for $x \in \mathbb{T}$ and s, t in compact subsets of \mathbb{C} ,

$$r_{0,n}(nx + s, nx + t) = \text{sinc}(s - t) + O\left(\frac{1}{n}\right).$$

It remains to prove that the quantity $I_n^x(s, t)$ converges towards 0 uniformly on $x \in \mathbb{T}$ and s, t in compact sets of \mathbb{C} . Let $A > 1$ and $K = B(0, A - 1)$ the disk centered in 0 of radius $A - 1$ in \mathbb{C} . Denoting by ω_ψ the uniform modulus of continuity of the spectral density ψ , we have

$$\begin{aligned} |I_n^x(s, t)| &\leq \frac{1}{2\pi n^2} \int_{-n\pi}^{n\pi} \left| \frac{\sin\left(\frac{s-y}{2}\right) \sin\left(\frac{t-y}{2}\right)}{\sin\left(\frac{s-y}{2n}\right) \sin\left(\frac{t-y}{2n}\right)} \right| \omega_\psi\left(\frac{y}{n}\right) dy \\ &\leq R_1 + R_2, \end{aligned}$$

where

$$R_1 = \frac{1}{2\pi n^2} \int_{-A}^A \left| \frac{\sin\left(\frac{s-y}{2}\right) \sin\left(\frac{t-y}{2}\right)}{\sin\left(\frac{s-y}{2n}\right) \sin\left(\frac{t-y}{2n}\right)} \right| \omega_\psi\left(\frac{y}{n}\right) dy,$$

and

$$R_2 = \frac{1}{2\pi n^2} \int_{-n\pi}^{n\pi} \mathbb{1}_{\{|y| \geq A\}} \left| \frac{\sin\left(\frac{s-y}{2}\right) \sin\left(\frac{t-y}{2}\right)}{\sin\left(\frac{s-y}{2n}\right) \sin\left(\frac{t-y}{2n}\right)} \right| \omega_\psi\left(\frac{y}{n}\right) dy.$$

The term R_1 is bounded by

$$R_1 \leq \frac{1}{2\pi} \int_{-A}^A \omega_\psi\left(\frac{y}{n}\right) dy.$$

Since the spectral density is (uniformly) continuous on \mathbb{T} , the quantity R_1 converges towards zero as n goes to infinity, uniformly on $x \in \mathbb{T}$ and $s, t \in K$. For the term R_2 we

use the following inequalities, valid for $\operatorname{Re}(z) \in [-5\pi/6, 5\pi/6]$:

$$\frac{3}{5\pi} |\operatorname{Re}(z)| \leq |\sin(\operatorname{Re}(z))| \leq |\sin(z)|. \quad (4.23)$$

There is a rank n_0 depending only on the compact subset K such that, for all $n \geq n_0$, $s \in K$ and $y \in [-n\pi, n\pi]$,

$$-\frac{5\pi}{6} \leq \frac{\operatorname{Re}(s) - y}{2n} \leq \frac{5\pi}{6} \quad \text{and} \quad -\frac{5\pi}{6} \leq \frac{\operatorname{Re}(t) - y}{2n} \leq \frac{5\pi}{6}.$$

It follows from the series of inequalities (4.23) that there is a constant C such that for $n \geq n_0$,

$$\begin{aligned} R_2 &\leq \frac{\cosh(A)^2}{2\pi n^2} \int_{-n\pi}^{n\pi} \mathbb{1}_{\{|y| \geq A\}} \frac{1}{\left| \sin\left(\frac{s-y}{2n}\right) \right|} \frac{1}{\left| \sin\left(\frac{t-y}{2n}\right) \right|} \omega_\psi\left(\frac{y}{n}\right) dy \\ &\leq C \cosh(A)^2 \int_{-n\pi}^{n\pi} \mathbb{1}_{\{|y| \geq A\}} \frac{1}{|\operatorname{Re}(s) - y| |\operatorname{Re}(t) - y|} \omega_\psi\left(\frac{y}{n}\right) dy \\ &\leq C \cosh(A)^2 \int_{-\infty}^{\infty} \mathbb{1}_{\{|y| \geq A\}} \frac{1}{(|y| - A + 1)^2} \omega_\psi\left(\frac{y}{n}\right) dy. \end{aligned}$$

The quantity R_2 thus converges towards 0 as n goes to infinity, uniformly on $x \in \mathbb{T}$ and $s, t \in K$. \square

Lemma 4.3.5. *Let $u, v \geq 0$ and $0 < \alpha < 1$. There is a constant C such that*

$$\forall s, t \in n\mathbb{T}, \quad |r_n^{(u,v)}(s, t)| \leq \frac{C}{1 + \operatorname{dist}(s, t)^\alpha}.$$

Proof. Let $s, t \in n\mathbb{T}$. According to the previous Lemma 4.3.4, the function r_n and its derivatives are uniformly bounded, we can thus assume that $\operatorname{dist}(s, t) \geq 4$. By Cauchy integral formula, there is a constant C such that

$$|r_n^{(a,b)}(s, t)| \leq C \sup_{|w| \leq 1} \sup_{|z| \leq 1} |r_n(s + w, t + z)|.$$

Let w, z be complex numbers such that $|w| \leq 1$ and $|z| \leq 1$. Using the explicit formula

(4.21) for r_n we obtain

$$r_n(s+w, t+z) = \frac{1}{2\pi n} \cos\left(\frac{n+1}{2n}(s-t+w-z)\right) \int_{-\pi}^{\pi} \frac{\sin\left(\frac{ny+s+w}{2}\right) \sin\left(\frac{ny+t+z}{2}\right)}{\sin\left(\frac{y+\frac{s+w}{n}}{2}\right) \sin\left(\frac{y+\frac{t+z}{n}}{2}\right)} \psi(y) dy.$$

Using the fact that the cosine function is bounded on a horizontal complex strip,

$$|r_n(s+w, t+z)| \leq \frac{C\|\psi\|_{\infty}}{2\pi n} \int_{-\pi}^{\pi} \left| \frac{\sin\left(\frac{ny+s+w}{2}\right) \sin\left(\frac{ny+t+z}{2}\right)}{\sin\left(\frac{y+\frac{s+w}{n}}{2}\right) \sin\left(\frac{y+\frac{t+z}{n}}{2}\right)} \right| dy.$$

Let $\delta = \text{dist}(s, t)/2$. Up to translating s and t by $\pm 2\pi n$ and exchanging s and t , we can assume that $\delta = \frac{t-s}{2}$. We then make the change of variable $x = y + \frac{t+s}{2n}$ to obtain

$$\int_{-\pi}^{\pi} \left| \frac{\sin\left(\frac{ny+s+w}{2}\right) \sin\left(\frac{ny+t+z}{2}\right)}{\sin\left(\frac{y+\frac{s+w}{n}}{2}\right) \sin\left(\frac{y+\frac{t+z}{n}}{2}\right)} \right| dy = \int_{-\pi}^{\pi} \left| \frac{\sin\left(\frac{nx+\delta+w}{2}\right) \sin\left(\frac{nx-\delta+z}{2}\right)}{\sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right) \sin\left(\frac{x+\frac{-\delta+z}{n}}{2}\right)} \right| dx.$$

This last integral splits into two integrals I_1 and I_2 defined by

$$I_1 := \int_0^{\pi} \left| \frac{\sin\left(\frac{nx+\delta+w}{2}\right) \sin\left(\frac{nx-\delta+z}{2}\right)}{\sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right) \sin\left(\frac{x+\frac{-\delta+z}{n}}{2}\right)} \right| dx \quad \text{and} \quad I_2 := \int_{-\pi}^0 \left| \frac{\sin\left(\frac{nx+\delta+w}{2}\right) \sin\left(\frac{nx-\delta+z}{2}\right)}{\sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right) \sin\left(\frac{x+\frac{-\delta+z}{n}}{2}\right)} \right| dx.$$

Both term can be treated the exact same way. We have by Hölder inequality with $0 < \alpha < 1$,

$$I_1 \leq \left(\int_0^{\pi} \left| \frac{\sin\left(\frac{nx+\delta+w}{2}\right)}{\sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right)} \right|^{\frac{1}{1-\alpha}} dx \right)^{1-\alpha} \left(\int_0^{\pi} \left| \frac{\sin\left(\frac{nx-\delta+z}{2}\right)}{\sin\left(\frac{x+\frac{-\delta+z}{n}}{2}\right)} \right|^{\frac{1}{\alpha}} dx \right)^{\alpha}. \quad (4.24)$$

For the left integral in (4.24), we make use of the following inequalities, which are consequences of inequalities (4.23), and the fact that $|w| \leq 1$ and $\delta \geq 2$

$$\left| \sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right) \right| \geq \frac{3}{10\pi} \left(x + \frac{\delta + \text{Re}(w)}{n} \right) \geq \frac{3}{10\pi} \left(x + \frac{\delta}{2n} \right),$$

to get

$$\begin{aligned} \left(\int_0^\pi \left| \frac{\sin\left(\frac{nx+\delta+w}{2}\right)}{\sin\left(\frac{x+\frac{\delta+w}{n}}{2}\right)} \right|^{\frac{1}{1-\alpha}} dx \right)^{1-\alpha} &\leq C \left(\int_0^\infty \frac{dx}{\left(x + \frac{\delta}{2n}\right)^{\frac{1}{1-\alpha}}} \right)^{1-\alpha} \\ &\leq C \left(\frac{n}{\delta}\right)^\alpha. \end{aligned} \quad (4.25)$$

For the right integral in (4.24), we make the change of variable $y = nx - \delta + \operatorname{Re}(z)$ and we use the inequality

$$|x + iy| \leq C|\sin(x + iy)|,$$

valid for $x \in [-2\pi/3, 2\pi/3]$ and $y \in [-1/4, 1/4]$, to get

$$\begin{aligned} \left(\int_0^\pi \left| \frac{\sin\left(\frac{nx-\delta+z}{2}\right)}{\sin\left(\frac{x+\frac{-\delta+z}{n}}{2}\right)} \right|^{\frac{1}{\alpha}} dy \right)^\alpha &\leq n^{-\alpha} \left(\int_{\operatorname{Re}(z)-\delta}^{n\pi+\operatorname{Re}(z)-\delta} \left| \frac{\sin\left(\frac{y+i\operatorname{Im}(v)}{2}\right)}{\sin\left(\frac{y+i\operatorname{Im}(v)}{2n}\right)} \right|^{\frac{1}{\alpha}} dy \right)^\alpha \\ &\leq 2Cn^{1-\alpha} \left(\int_{-\infty}^\infty \left| \frac{\sin\left(\frac{y+i\operatorname{Im}(v)}{2}\right)}{y+i\operatorname{Im}(v)} \right|^{\frac{1}{\alpha}} dy \right)^\alpha \\ &\leq C'n^{1-\alpha}, \end{aligned} \quad (4.26)$$

where in the last inequality we used the fact that the integrand is uniformly bounded in a neighborhood of the origin, and that $\frac{1}{\alpha} > 1$ so the integrand is also integrable near $\pm\infty$. Plugging estimates (4.25) and (4.26) into inequality (4.24), we obtain for some constant C that

$$|r_n^{(u,v)}(nx + s, nx + t)| \leq \frac{C\|\psi\|_\infty}{2\pi n} (I_1 + I_2) \leq \frac{2C\|\psi\|_\infty}{2\pi n} \left(\frac{n}{\delta^\alpha}\right) \leq \frac{C}{\operatorname{dist}(s, t)^\alpha}.$$

□

CUMULANTS FOR THE ZEROS COUNTING MEASURE OF REGULAR GAUSSIAN PROCESSES

This chapter consists in the third submitted paper of this thesis [Gas21c], concerning the exact cumulant asymptotics for the zeros counting measure of regular Gaussian processes.

5.1 Introduction	154
5.1.1 Statement of the main results	155
5.1.2 Outline of the proof	164
5.2 Basics and notations	167
5.2.1 Partition and cumulants	167
5.2.2 Diagonal set and factorial power measure	170
5.2.3 Matrix notations	172
5.2.4 Divided differences	176
5.3 Kac–Rice formula for Gaussian processes	186
5.3.1 Kac density and cumulants of the zeros counting measure	187
5.3.2 Matrix representation of the Kac density and factorization property	192
5.3.3 Decay of the cumulant Kac density	199
5.4 Asymptotics of the cumulants of the zeros counting measure	205
5.4.1 Uniform non degeneracy of the covariance matrix	205
5.4.2 Asymptotics of the cumulants	210

5.1 Introduction

The study of the number of zeros of smooth Gaussian processes has a long history and is in particular motivated by the pioneering works of Kac and Rice, see e.g. [AW09] for a general introduction to this topic. The asymptotics for the expectation and the variance of the number of zeros of a stationary Gaussian process on an interval growing interval $[0, R]$ as R grows to infinity has been known since [Cuz76], where a central limit theorem (CLT) for the number of zeros is also proved. The variance asymptotics is there established using the celebrated Kac–Rice method and the CLT is proved using approximation by an m -dependent process.

With similar methods, the variance of the number of zeros of random Gaussian trigonometric polynomials with large degree has been studied in [GW11], as well as the associated CLT. Later on, the machinery of Wiener chaos expansion was then successfully used in order to compute the variance asymptotics as well as establishing CLTs for the number of zeros of various models of stochastic processes, see for instance [AL13; Arm+21; Do+21]. Central limit theorems for the number of real roots of random algebraic polynomials have also been investigated, see for example [NV21] and the references therein.

In the recent paper [Gas21a], focusing on the asymptotics of the Kac density rather than on the full integral Kac–Rice formula, the author managed to avoid some of the technical computations inherent to the use of Kac–Rice method. This allowed him to get a unifying point of view, make explicit the needed decorrelation estimates and then deduce the variance asymptotics for the number of zeros of many models of Gaussian processes. It has been then conjectured that the same heuristics could be applied to treat the asymptotics of the higher central moments of the number of zeros of a Gaussian process, which is the goal of the present paper.

Up to now, very few results about the asymptotics of higher central moments are known. The best result so far is the one by M. Ancona and T. Letendre [AL21a], where it has been proved that the p -th central moment, when properly rescaled, converges towards the p -th moment of a Gaussian random variable, under restrictive condition that the covariance function and their derivatives decreases faster than x^{-4p} . This last result then yields another proof of the CLT for the number of zeros by the method of moments, for the processes whose covariance function is in the Schwartz class of regular and rapidly decreasing functions. Note that the stationary sine process, i.e. the Gaussian process with sinc covariance function, which plays a central role in probability theory and mathematical

physics, is ruled out from their framework, due to the slow decay of the sinc kernel. In the more general context of point processes, higher moments of geometric statistics have also been studied under the hypothesis of fast decreasing correlation [BYY19].

In this paper, we prove the exact asymptotics of the higher central moments of the number of zeros of large class of Gaussian processes, under the only hypothesis, apart from regularity, that the covariance function as well as its derivatives are square integrable. Our results apply in particular for Gaussian trigonometric and orthogonal polynomials, as well as the stationary process with sinc kernel and other Gaussian stationary processes on the real line with possibly slow decaying kernels. We prove in fact a more general theorem by computing the exact asymptotics of the cumulants of linear statistics associated with the zeros counting measure of the underlying processes. The use of cumulants instead of central moments simplifies the rather intricate combinatorics involved when estimating higher order moments via the Kac–Rice method.

Our result in turn implies the convergence of associated moments of any order and thus a CLT, with an exact rate of convergence. As a corollary, we deduce a polynomial concentration of any order of the number of zeros, and by a Borel–Cantelli argument, the almost sure convergence of the number of zeros. Note that these last facts cannot be deduced from chaos expansion methods. More generally, in the context of linear statistics, we prove the almost sure equidistribution of the zeros set at the limit for a large class of smooth Gaussian processes.

5.1.1 Statement of the main results

Cumulants asymptotics and central limit theorems

In the following, all the random variables considered are defined on a common abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and \mathbb{E} will denote the associated expectation. In the sequel, W stands for a standard Gaussian random variable, i.e. centered with unit variance. We denote by $\kappa_p(Z)$ the p -th cumulant of a random variable Z , given by the expression

$$\kappa_p(Z) = \sum_{\mathcal{I} \in \mathcal{P}_p} (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \prod_{I \in \mathcal{I}} \mathbb{E} [Z^{|\mathcal{I}|}], \quad (5.1)$$

where the sum is indexed by the set \mathcal{P}_p of all the partitions of the finite set $\{1, \dots, p\}$. We refer to [Spe83; PT11] and the paragraph 5.2.1 below for more details on the cumulants of

a random variable. The following theorem describes the asymptotics of all the cumulants of the number of zeros of a Gaussian trigonometric polynomial with independent coefficients.

Theorem 5.1.1. *Let $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ be two iid sequences of standard Gaussian variables. Let Z_n be the number of zeros on $[0, 2\pi]$ of the process*

$$h_n(x) := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} a_k \cos(kx) + b_k \sin(kx).$$

For p a positive integer, there is an explicit constant γ_p such that

$$\lim_{n \rightarrow +\infty} \frac{\kappa_p(Z_n)}{n} = \gamma_p.$$

The constants γ_1 and γ_2 are positive. The above theorem implies in particular that

$$\lim_{n \rightarrow +\infty} \frac{\text{Var}(Z_n)}{n} = \gamma_2 \quad \text{and} \quad \forall p \geq 3, \quad \frac{\kappa_p(Z_n)}{n^{p/2}} = O\left(\frac{1}{n^{\frac{p}{2}-1}}\right). \quad (5.2)$$

Given the expression of the central moments in terms of cumulants and the fact that the cumulants of a Gaussian random variable are zero for $p \geq 3$, the asymptotics (5.2) imply in fact that for every positive integer p ,

$$\mathbb{E} \left[\left(\frac{Z_n - \mathbb{E}[Z_n]}{\sqrt{\text{Var}(Z_n)}} \right)^p \right] = \mathbb{E}[W^p] + O\left(\frac{1}{\sqrt{n}}\right). \quad (5.3)$$

Note that the exact asymptotics of the cumulants given by Theorem 5.1.1 is in nature stronger than the cruder bound given by (5.2) and thus the central moment asymptotics (5.3).

As a consequence, it implies a central limit theorem for the number of zeros, as well as a polynomial concentration to any order of the number of zeros around its mean.

Corollary 5.1.2. *As n goes to infinity, we have the convergence in distribution*

$$\frac{Z_n - \mathbb{E}[Z_n]}{\sqrt{\text{Var}(Z_n)}} \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, 1).$$

For all $p \geq 2$, there is a constant C_p such that for all integer n and positive constant ε ,

$$\mathbb{P}(|Z_n - \mathbb{E}[Z_n]| \geq n\varepsilon) \leq \frac{C_p}{(\sqrt{n}\varepsilon)^p}.$$

Note that the variance estimate in Equation (5.2) and the associated CLT were first established in [GW11] by the Kac–Rice method and in [AL13] by the Wiener chaos expansion. So far the exact asymptotics of the p -th central moment or cumulants of Z_n has never been computed for $p \geq 3$. Theorem 5.1.1 shows that it asymptotically behaves like the p -th moment of a Gaussian random variable, which is expected from the already existing central limit theorem for the random variable Z_n . The polynomial concentration of the number of zeros and a Borel–Cantelli argument implies the almost sure convergence

$$\lim_{n \rightarrow +\infty} \frac{Z_n}{n} = \gamma_0 \quad \text{a.s.},$$

a result already known from [AP21], using a derandomization method. Exponential concentration has been established in [NZ19] for this particular model but the proof is of very different nature and strongly use the trigonometric nature of the random process h_n . Our proof only uses the fact that the process is of class \mathcal{C}^∞ and is adaptable to many other models.

The error term in (5.3) is new and cannot be deduced from the Wiener chaos expansion. It implies in particular a rate of convergence towards the Gaussian random variable of order $1/\sqrt{n}$ for the moment metric. It is reminiscent of the Berry–Essen bound for more classical CLT.

The independence hypothesis above on the Gaussian random coefficients can be relaxed. Namely, we can extend the previous Theorem 5.1.1 to the case where the Gaussian sequences $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are independent and stationary.

Theorem 5.1.3. *Let $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ be two independent sequences of standard Gaussian variables, such that*

$$\mathbb{E}[a_k a_l] = \mathbb{E}[b_k b_l] = \rho(k - l).$$

We assume that the spectral measure associated with the correlation function ρ has a continuous positive density on the torus \mathbb{T} . Let Z_n be the number of zeros on $[0, 2\pi]$ of the process

$$h_n(x) := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} a_k \cos(kx) + b_k \sin(kx).$$

Then the conclusion of Theorem 5.1.1 and its Corollary 5.1.2 holds.

The expectation of the number of zeros in this model has been studied in [ADP19; APP21] and the variance in [Gas21a]. The above Theorem 5.1.3 gives the asymptotics of every cumulants and therefore, as discussed above in the independent case, it proves a central limit theorem for the number of zeros, which is a new result in this dependent framework, as well as concentration around the mean and a quantification of rate of convergence.

In another direction, one can replace the functions \cos and \sin by more general functions. A standard framework is then the following model of random orthogonal polynomials, for which we can give a similar statement.

Theorem 5.1.4. *Let $(a_k)_{k \geq 0}$ be an iid sequence of standard Gaussian variables. Let $(P_k)_{k \geq 0}$ be a sequence of orthogonal polynomials associated with a measure μ on the line, and let $[a', b']$ be an interval. We assume that the measure μ and the interval $[a', b']$ satisfies the hypotheses of [Do+21, Thm. 1.1]. Let Z_n be the number of zeros on $[a', b']$ of the process*

$$h_n(x) := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} a_k P_k(x).$$

Then the conclusion of Theorem 5.1.1 and its Corollary 5.1.2 holds.

The expectation, the variance and a central limit theorem for this model have been very recently studied with the method of chaos in [Do+21]. Here again, we extend this result by determining the asymptotics of higher cumulants and thus higher moments. As already

discussed after the previous statements above, from the cumulants asymptotics established in Theorem 5.1.4, we can also deduce concentration around the expected number of zeros and as well as a rate of convergence in the associated CLT, these two last results cannot be obtained with the Wiener chaos method.

At last, we extend known results about the number of zeros of a stationary Gaussian process on a growing interval, establishing in particular a CLT under the sole square integrability of the associated correlation function and its derivatives.

Theorem 5.1.5. *Let f be a stationary Gaussian process with C^∞ paths and covariance function r . For $R > 0$ we define Z_R to be the number of zeros on $[0, R]$ of the process f .*

— *If the covariance function r and its derivatives are in $L^2(\mathbb{R})$, then for all $p \geq 2$*

$$\lim_{R \rightarrow +\infty} \mathbb{E} \left[\left(\frac{Z_R - \mathbb{E}[Z_R]}{\sqrt{\text{Var}(Z_R)}} \right)^p \right] = \mathbb{E}[W^p].$$

— *If the covariance function r and its derivatives are in $L^q(\mathbb{R})$ for all $q > 1$, then for p a positive integer, there is an explicit constant γ_p such that*

$$\lim_{R \rightarrow +\infty} \frac{\kappa_p(Z_R)}{R} = \gamma_p.$$

As mentioned above, the CLT which is obtained from the above moments asymptotics by the method of moments is already known in the particular case of stationary Gaussian processes with covariance function belonging to the Schwartz class, see [AL21a]. Here the assumption on the decay of the correlation function is greatly relaxed and we only need to assume the square integrability of the covariance kernel as well as its derivatives.

As a particular and representative case, Theorem 5.1.5 englobes the example of the stationary Gaussian process f with sinc kernel, which is a completely new result. This process plays a central role in the study of determinantal point processes, and appears as the limit of the local statistics of various random models, from eigenvalues of random matrices to random integer partitions. For this particular process, the asymptotic of the expectation and the variance of Z_R , as well as a CLT were known since the pioneering

works of [Cuz76] and the references therein. But so far, the exact asymptotics of higher central moments or cumulants of Z_R remained unknown.

Under the stronger hypothesis that the covariance function and its derivatives are in L^p for all $p > 1$, we deduce a polynomial concentration around the mean to any order for the number of zeros, which appears to be a new result. We also deduce the exact asymptotic of the cumulants of any order. This integrability hypothesis in particular holds true for processes whose covariance functions r and their derivatives satisfy the bound

$$\forall x \in \mathbb{R}, \quad r^{(u)}(x) \leq \frac{C_u}{1 + |x|},$$

which is the case for a stationary Gaussian process with sinc covariance function.

A more general and unifying statement

In fact Theorems 5.1.1, 5.1.3, 5.1.4 and 5.1.5 are all corollaries of a single, more general statement given below. In order to state it, we need to introduce first a few notations that will be used for the rest of the paper.

Let U be a non-empty open interval of the real line \mathbb{R} or of the one-dimensional torus \mathbb{T} , endowed with their canonical distance $|\cdot|$. Let $n \in \mathbb{R}_+^* \cup \{+\infty\}$. If n is finite then nU is a non-empty open subset of \mathbb{R} or of the one-dimensional torus $n\mathbb{T}$ of length n . For $n = +\infty$ we use the convention $(+\infty)U = \mathbb{R}$. This setting allows us to give a unified exposition for processes defined on the torus (e.g. random trigonometric polynomials) and on the real line (e.g. the sinc process).

Let \mathbf{N} be an unbounded subset of \mathbb{R}_+^* and $\bar{\mathbf{N}} = \mathbf{N} \sqcup \{+\infty\}$. For each $n \in \bar{\mathbf{N}}$, we consider a centered Gaussian process f_n defined on nU , and we assume that the process f_∞ is a non degenerate stationary process on \mathbb{R} . Note that for $n \in \mathbf{N}$ the process $h_n = f_n(n \cdot)$ is a Gaussian process on U . For $n \in \bar{\mathbf{N}}$ and $s, t \in nU$ we define the covariance function

$$r_n(s, t) = \mathbb{E}[f_n(s)f_n(t)] \quad \text{and} \quad r_\infty(s - t) = r_\infty(s, t).$$

If the process f_n is of class $\mathcal{C}^k(U)$ for $k \geq 0$ then the covariance function r_n is also of class \mathcal{C}^k in each variable, and one has for $u, v \leq k$ and $x, y \in nU$

$$r_n^{(u,v)}(x, y) = \mathbb{E}[f_n^{(u)}(x)f_n^{(v)}(y)].$$

For $n \in \mathbf{N}$ we define

$$Z_n = \{x \in U \mid f_n(nx) = 0\} \quad \text{and} \quad \nu_n := \sum_{x \in Z_n} \delta_x, \quad (5.4)$$

the random counting measure on Z_n . Note that $(\nu_n)_{n \in \mathbf{N}}$ is a family of measures on U . Assume for now (it will be a consequence of Bulinskaya Lemma) that for each $n \in \mathbf{N}$ the set Z_n is almost surely locally finite. For a bounded function $\phi : U \rightarrow \mathbb{R}$, with compact support in U , we define the bracket

$$\langle \nu_n, \phi \rangle = \sum_{x \in Z_n} \phi(x).$$

For instance, if $U = \mathbb{R}$ and $\phi = \mathbb{1}_{[0,1]}$ then

$$\langle \nu_n, \phi \rangle = \text{Card} \{x \in [0, n] \mid f_n(x) = 0\}.$$

Note the Kac–Rice formula implies that in expectation, the counting measure has a density with respect to the Lebesgue measure. Integrability properties of this density shown in this article allows us to extend the domain of definition of tests functions ϕ to the space

$$L^{(\infty)}(U) = \bigcap_{1 \leq p < +\infty} L^p(U).$$

For $q \geq 1$ we define the two following hypotheses.

- $H_1(q)$: The sequence of processes $(f_n)_{n \in \bar{\mathbf{N}}}$ is of class $\mathcal{C}^q(U)$, and there is a uniformly continuous function ψ on U , bounded below and above by positive constants, such that for $u, v \leq q$, the following convergence holds uniformly for $x \in U$ and locally uniformly for $s, t \in \mathbb{R}$,

$$\lim_{n \rightarrow +\infty} r_n^{(u,v)}(nx + s, nx + t) = \psi(x) r_\infty^{(u,v)}(s, t). \quad (5.5)$$

- $H_2(q)$: There is a function g , even, bounded and going to zero near infinity, such that for $u, v \leq q$, $n \in \bar{\mathbf{N}}$ and $s, t \in nU$,

$$|r_n^{(u,v)}(s, t)| \leq g(s - t), \quad (5.6)$$

and for some positive constant ε the function g_ε is in $L^2(\mathbb{R})$, where

$$g_\varepsilon : x \mapsto \sup_{|u| \leq \varepsilon} g(x + u).$$

Theorem 5.1.6. *Let $p \geq 2$ and $q = 2p - 1$. We assume that the sequence of processes $(f_n)_{n \in \mathbb{N}}$ satisfies hypotheses $H_1(q)$ and $H_2(q)$ defined above. Then for every function $\phi \in L^\infty(U)$,*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\left(\frac{\langle \nu_n, \phi \rangle - \mathbb{E}[\langle \nu_n, \phi \rangle]}{\sqrt{\text{Var}(\langle \nu_n, \phi \rangle)}} \right)^p \right] = \mathbb{E}[W^p].$$

Assume moreover that $g_\varepsilon \in L^{\frac{p}{p-1}}(\mathbb{R})$. Then there is an explicit constant γ_p depending only on the process f_∞ , such that

$$\lim_{n \rightarrow +\infty} \frac{\kappa_p(\langle \nu_n, \phi \rangle)}{n} = \gamma_p \cdot \left(\int_U \phi(x)^p dx \right).$$

The assumption $H_1(q)$ characterizes the convergence of the family of processes $(f_n)_{n \in \mathbb{N}}$ towards a limit stationary process in \mathcal{C}^q norm. This hypothesis is natural and arises in many models. For instance the covariance function of random trigonometric polynomials converges towards the sinc function. The regularity of the process f_n ensures the well-definiteness of the p -th moment. The convergence towards a non-degenerate stationary process ensures the uniform non-degeneracy on the process f_n , as well as the explicit asymptotics for the cumulants.

The assumption $H_2(q)$ is much weaker than the one present in [AL21a], where the authors require a function g that decrease like x^{-4p} . Here we show that the asymptotics of higher moments is independent of the rate of decay of the covariance function, and must only satisfy some uniform square integrability condition.

Let us briefly now show that the unifying Theorem 5.1.6 indeed implies the collection of theorems of the previous subsection. First, Theorem 5.1.1 and 5.1.3 are a consequence of Theorem 5.1.6, by setting $U = \mathbb{T}$, $\mathbf{N} = \mathbb{N}^*$, $\phi = \mathbb{1}_{\mathbb{T}}$ and

$$f_n(x) = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} a_k \cos\left(\frac{kx}{n}\right) + b_k \sin\left(\frac{kx}{n}\right).$$

Let ψ be the spectral density of the correlation function ρ of the stationary Gaussian sequences $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$, which is assumed to be continuous and positive on \mathbb{T} . Assumptions $H_1(q)$ and $H_2(q)$ are proved for all $q > 0$ for this model in the paper [Gas21a] with limit process having sinc covariance function, and

$$g = \frac{C_\alpha}{1 + |x|^\alpha},$$

where the exponent α can be taken in $]1/2, 1[$. Note that Theorem 5.1.1 is a particular case of Theorem 5.1.3 with $\psi = 1$ (in that case, one can take $\alpha = 1$ above).

Similarly, Theorem 5.1.4 is a consequence of Theorem 5.1.6. Let μ be a measure with compact support on the real line. We set U a subinterval of \mathbb{R} such that μ has a positive continuous density on \overline{U} . It is proved in [Do+21] under mild assumption on the measure μ that for the model of random orthogonal polynomials with respect to the measure μ , the assumption $H_2(q)$ holds true for all $q > 0$ and

$$g(x) = \frac{C}{1 + |x|}.$$

Also, a slight modification of assumption $H_1(q)$ holds true for all $q > 0$ with limit process having sinc covariance function, and ψ the inverse of the density of the measure μ . This minor difference involves the equilibrium measure of the support of the measure μ , through some rescaling of the ambient space. The proof of Theorem 5.1.6 adapts verbatim to this setting. Note that if $\text{supp } \mu = [0, 1]$, then after a change of variable, the equilibrium measure is simply the Lebesgue measure on the torus \mathbb{T} and hypothesis $H_1(q)$ then exactly holds true.

At last, Theorem 5.1.5 is again a consequence of Theorem 5.1.6 with $U = \mathbb{R}$, $\mathbf{N} = \mathbb{R}_+^*$, $f_n = f$, and test function $\phi = \mathbb{1}_{[0,1]}$.

Asymptotics for the linear statistics

Let ν_∞ denotes the Lebesgue measure on the interval U . Theorem 5.1.6 implies a strong law of large number and a central limit theorem for the sequence of random measure $(\nu_n)_{n \in \mathbf{N}}$. The two following Corollaries 5.1.7 and 5.1.8 extend the results of [AL21a, Sec. 1.4] to our framework, and we refer to this paper for a more thorough discussion.

Corollary 5.1.7 (Law of large numbers). *Assume that the hypotheses $H_1(q)$ and $H_2(q)$ are satisfied for all $q \geq 1$, and either $\mathbf{N} = \mathbb{N}^*$, or $\mathbf{N} = \mathbb{R}_+^*$ and for $n \in \mathbb{R}_+^*$, $f_n = f_\infty$. Then we have the following almost-sure convergence for the vague topology*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \nu_n = \gamma_1 \nu_\infty \quad \text{a.s.} \quad .$$

Corollary 5.1.7 shows that zeros of the process $f_n(n \cdot)$ tend to be equidistributed on the set U as n goes to $+\infty$. When $\mathbf{N} = \mathbb{N}^*$, the proof follows from an application of the Borel–Cantelli Lemma. When $\mathbf{N} = \mathbb{R}_+^*$ and $\forall n \in \mathbb{R}_+^*$, $f_n = f_\infty$, we can apply the Borel–Cantelli Lemma to prove the almost sure convergence on a polynomial subsequence. It is then a standard fact that the monotonicity of Z_n ensures the almost sure convergence of the whole sequence.

Corollary 5.1.8 (Central limit theorem). *Assume that the hypotheses $H_1(q)$ and $H_2(q)$ are satisfied for all $q \geq 1$. Then we have the following convergence in distribution*

$$\forall \phi \in C_c(U), \quad \frac{\sqrt{n}}{\gamma_2} \left\langle \left(\frac{1}{n} \nu_n - \gamma_1 \nu_\infty \right), \phi \right\rangle \underset{n \rightarrow +\infty}{\sim} \mathcal{N} \left(0, \|\phi\|_2^2 \right).$$

Corollary 5.1.8 implies that the fluctuations around the mean of the counting measure ν_n is comparable to a Gaussian white noise.

5.1.2 Outline of the proof

Before giving a complete and detailed proof of Theorem 5.1.6, let us sketch its main ingredients and arguments. It mainly relies on a careful analysis of the Kac–Rice formula, which asserts that for a test function ϕ ,

$$\mathbb{E}[\langle \nu_n, \phi \rangle^p] = \int_{(nU)^p} \left(\prod_{i=1}^p \phi \left(\frac{x_i}{n} \right) \right) \rho_{p,n}(x_1, \dots, x_p) dx_1 \dots dx_p + \text{extra terms}, \quad (5.7)$$

where

$$\rho_{p,n}(x_1, \dots, x_p) := T_n(x_1, \dots, x_p) \mathbb{E} \left[\prod_{i=1}^p |f'_n(x_i)| \mid f_n(x_1) = \dots = f_n(x_p) = 0 \right], \quad (5.8)$$

and $T_n(x_1, \dots, x_p)$ is the density at zero of the Gaussian vector $(f_n(x_1), \dots, f_n(x_p))$. The extra terms appearing in Equation (5.7) are of combinatorial nature and can be treated the exact same way as the first term, so we will omit them in the following heuristics. The function $\rho_{p,n}$ is called the Kac density of order p associated with the process f_n . Observe that the function $\rho_{p,n}$ is ill defined when two of its arguments collapse. This issue is solved by using the technique of divided differences, that appeared in [Cuz75] and was developed in [AL21a]. Let us give an example with $p = 2$. The idea is to replace in (5.8) the quantity

$$f_n(x_1) = f_n(x_2) = 0 \quad \text{by} \quad f_n(x_1) = \frac{f_n(x_2) - f_n(x_1)}{x_2 - x_1} = 0.$$

If the variables x_1 and x_2 collapse, the second expression becomes $f_n(x_1) = f'_n(x_1) = 0$. The regularity of the process f_n given by the assumption $H_1(q)$ and the non-degeneracy of the limit process f_∞ ensures that the Gaussian vector $(f_n(x), f'_n(x))$ is non-degenerate and gives an alternative non-singular expression of the function $\rho_{2,n}$ near the diagonal. For higher integers p , the reasoning is the same. For each partition \mathcal{I} of the set $\{1, \dots, p\}$, we will thus give an alternative and non-singular expression of the density $\rho_{p,n}$, that extends by continuity on points (x_1, \dots, x_p) such that x_i and x_j are equals if i and j belong to the same cell of the partition \mathcal{I} . This procedure is explained in Section 5.2.4.

Now we develop the expression of the cumulant of order p as a function of the moments in order to get

$$\begin{aligned} \kappa_p(\langle \nu_n, \phi \rangle) &= \sum_{\mathcal{J}} (|\mathcal{J}| - 1)! (-1)^{|\mathcal{J}|-1} \prod_{J \in \mathcal{J}} \mathbb{E} [\langle \nu_n, \phi \rangle^{|J|}] \\ &= \int_{(nU)^p} \prod_{i=1}^p \phi\left(\frac{x_i}{n}\right) F_{p,n}(x_1 \dots, x_p) dx_1 \dots dx_p, + \text{extra terms}, \end{aligned} \quad (5.9)$$

where the sum indexed by \mathcal{J} runs over all the partitions of the set $\{1, \dots, p\}$, and with

$$F_{p,n}(x_1 \dots, x_p) = \sum_{\mathcal{J}} (|\mathcal{J}| - 1)! (-1)^{|\mathcal{J}|-1} \prod_{J \in \mathcal{J}} \rho_{|J|,n}(\underline{x}_J).$$

Now let \mathcal{I} be a partition of $\{1, \dots, p\}$, and assume that for i and j belonging to two different cells of the partition \mathcal{I} , the variable x_i and x_j are far from each other. Then the decay hypothesis $H_2(q)$ implies that the Gaussian random variable $f_n(x_i)$ and $f_n(x_j)$ are almost independent, and from the definition of the Kac density $\rho_{p,n}$ we deduce that for

$A \subset \{1, \dots, p\}$,

$$\rho_{|A|,n}((x_a)_{a \in A}) \simeq \prod_{I \in \mathcal{I}} \rho_{|A \cap I|,n}((x_a)_{a \in A \cap I}), \quad (5.10)$$

Note that the function $\rho_{p,n}$ depends on f_n only through the covariance matrix of the vector $(f_n(x_1), \dots, f_n(x_p), f'_n(x_1), \dots, f'_n(x_p))$. This matrix representation formula allows us to give a precise error term in (5.10), proportional to the square of the magnitude of $r_n^{(u,v)}(x_i, x_j)$, where i and j belong to different cells of the partition \mathcal{I} . We refer to Section 5.2.3 for matrix notations and to Section 5.3.2 for the matrix representation of the Kac Density.

The combinatoric properties of cumulants and (5.10) imply that

$$F_{p,n}(x_1 \dots, x_p) \simeq 0, \quad (5.11)$$

as soon as the variables $(x_i)_{1 \leq i \leq p}$ are clustered with respect to some partition \mathcal{I} with at least two cells. A refinement of Taylor expansion using graph theoretic arguments (see Section 5.3.3), gives a much more precise error in (5.11) than the approach taken in [AL21a]. We then show that far from the diagonal (x, \dots, x) the function $F_{p,n}$ is small and will have sufficiently nice integrability properties on $(nU)^p$ in order to show in (5.9) that for $p \geq 3$,

$$\lim_{n \rightarrow +\infty} \frac{1}{n^{p/2}} \int_{(nU)^p} \prod_{i=1}^p \phi\left(\frac{x_i}{n}\right) F_{p,n}(x_1 \dots, x_p) dx_1 \dots dx_p = 0.$$

Given the link between cumulants and central moments, this fact leads to the convergence of the central moment of order p to the central moment of a Gaussian random variable. If moreover, the function g_ε is in $L^{\frac{p}{p-1}}(\mathbb{R})$ then the function $F_{p,n}(0, x_2, \dots, x_p)$ is integrable on $(nU)^{p-1}$, uniformly for $n \in \bar{\mathbb{N}}$. This fact leads to the exact asymptotics of the p -th cumulant of the random variable $\langle \nu_n, \phi \rangle$.

Despite its apparent simplicity, the detailed proof is quite technical and the diversity of arguments used justifies the following section, which introduces several notions and associated notations for the rest of the paper. In particular, the notion of partition of a finite set plays a central role in this article. From a combinatoric point of view, it appears in the Kac–Rice formula when expressing moments of the factorial power counting measure in terms of moments of the usual power measure, but also from the interpretation of cumulants in the context of Möebius inversion in the lattice of partition. The interplay between these last two combinatoric facts leads to an elegant expression of the cumulants

of the zeros counting measure (given by Proposition 5.3.5), and simplifies the approach taken by the authors in [AL21a], where they computed directly the asymptotics of central moments.

A novelty of this paper is also the intensive use of the matrix representation of the Kac density which allows us to dissociate the probabilistic setting, and facts concerning pure matrix analysis. We believe that this approach, already taken by the author in [Gas21a] to treat the asymptotic of the variance, greatly simplifies the exposition of proofs using Kac–Rice formulas.

5.2 Basics and notations

We define a few notations that will be of use and simplify the exposition. In the following, A is a finite set. The letter a, b, \dots denote elements of A . The letters B, C, \dots denote subset of A . The letters $\mathcal{I}, \mathcal{J}, \dots$ denote subsets of the power set of A .

5.2.1 Partition and cumulants

Set theory

We denote by $|A|$ the cardinal of the set A and $\mathcal{P}(A)$ the power set of A . For a set E , we define E^A the product of $|A|$ copies of E . A generic element of E^A is denoted

$$\underline{x}_A = (x_a)_{a \in A}$$

to avoid any confusion when elements of A are also sets. For a function $f : E \rightarrow \mathbb{R}$ and $\underline{x}_A \in E^A$ we write

$$f(\underline{x}_A) = (f(x_a))_{a \in A}.$$

Let $\underline{\phi}_A = (\phi_a)_{a \in A}$ be functions from E to \mathbb{R} . We define

$$\underline{\phi}_A^{\otimes} : \underline{x}_A \mapsto \prod_{a \in A} \phi(x_a). \quad (5.12)$$

At last, we denote by

$$2A = \{1, 2\} \times A \quad (5.13)$$

The set $2A$ should be seen as the disjoint union of A and a copy of itself. For an element $x_A \in E^A$ we denote $\underline{x}_{A,A}$ the element $(x_A, x_A) \in E^{2A}$.

The lattice of partitions and cumulants

The material of this paragraph is very standard, we refer to [Spe83; PT11] for a nice introduction on this topic. We define \mathcal{P}_A as the set of partitions of A . The partition of A into singletons will be denoted \bar{A} . In the following, B is a subset of A and \mathcal{I} is a partition of A . For $a \in A$ we denote $[a]_{\mathcal{I}}$ the cell of \mathcal{I} in which the element a belongs, and \mathcal{I}_B the partition of B induced by the partition \mathcal{I} of A . For instance, if $\mathcal{I} = \{\{1, 2\}, \{3, 4\}, \{5\}\}$, $a = 1$ and $B = \{1, 2, 3\}$ then

$$[a]_{\mathcal{I}} = \{1, 2\} \quad \text{and} \quad \mathcal{I}_B = \{\{1, 2\}, \{3\}\}.$$

The partition \mathcal{I} induce a partition on the set $2A$ via the relation

$$\{2I \mid I \in \mathcal{I}\}, \tag{5.14}$$

and we will still denote by \mathcal{I} this partition.

The set \mathcal{P}_A has a natural structure of poset (partially ordered set). Given \mathcal{I} and \mathcal{J} two partition of A , we say that \mathcal{I} is *finer* than \mathcal{J} (or that \mathcal{J} is *coarser* than \mathcal{I}) and we denote it $\mathcal{I} \preceq \mathcal{J}$ (or $\mathcal{J} \succeq \mathcal{I}$), if

$$\forall I \in \mathcal{I}, \exists J \in \mathcal{J} \text{ such that } I \subset J.$$

We then have

$$\mathcal{I}_J = \{I \in \mathcal{I} \mid I \subset J\} \quad \text{and} \quad J = \bigsqcup_{I \in \mathcal{I}_J} I.$$

Note that there is a one-to-one correspondence between the set of partitions of A coarser than a partition \mathcal{I} , and the set of partition of \mathcal{I} , given by the application

$$\mathcal{J} \mapsto \{\mathcal{I}_J \mid J \in \mathcal{J}\}. \tag{5.15}$$

Following this observation, we denote $[I]_{\mathcal{J}}$ the cell of \mathcal{J} in which the set I is included. For

instance, if $\mathcal{I} = \{\{1\}, \{2, 3\}, \{4\}, \{5\}\}$ and $\mathcal{J} = \{\{1, 2, 3\}, \{4, 5\}\}$ then $\mathcal{I} \preceq \mathcal{J}$ and

$$[\{2, 3\}]_{\mathcal{J}} = \{1, 2, 3\} \quad \text{and} \quad \mathcal{I}_{\{4, 5\}} = \{\{4\}, \{5\}\}.$$

Note that two partitions \mathcal{I} and \mathcal{J} have a greatest lower bound and a least upper bound for this partial order, which turns (\mathcal{P}_A, \preceq) into a finite lattice. Let $(m_B)_{B \subset A}$ and $(\kappa_B)_{B \subset A}$ two families of numbers. In our case of interest, the Möebius inversion on this particular lattice takes the form

$$\left(\forall B \subset A, \quad m_B = \sum_{\mathcal{I} \in \mathcal{P}_B} \prod_{I \in \mathcal{I}} \kappa_I \right) \quad \text{iff} \quad \left(\forall B \subset A, \quad \kappa_B = \sum_{\mathcal{I} \in \mathcal{P}_B} (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \prod_{I \in \mathcal{I}} m_I \right). \quad (5.16)$$

We will make use of the following cancellation property of the cumulants.

Lemma 5.2.1. *Let $(m_B)_{B \subset A}$ and $(\kappa_B)_{B \subset A}$ be two families of numbers related by one of the equivalent formulas in (5.16). Assume the existence of a partition $\mathcal{I} \neq \{A\}$ such that*

$$\forall B \subset A, \quad m_B = \prod_{I \in \mathcal{I}} m_{I \cap B}.$$

Then

$$\kappa_A = 0.$$

Proof. See [Spe83]. □

If $(X_a)_{a \in A}$ is a family of random variables, we can define for a subset B of A

$$m_B((X_b)_{b \in B}) = \mathbb{E} \left[\prod_{b \in B} X_b \right] \quad \text{and} \quad \kappa_B((X_b)_{b \in B}) = \sum_{\mathcal{I} \in \mathcal{P}_B} (-1)^{|\mathcal{I}|-1} (|\mathcal{I}| - 1)! \prod_{I \in \mathcal{I}} \mathbb{E} \left[\prod_{i \in I} X_i \right]. \quad (5.17)$$

The quantity $m_B((X_b)_{b \in B})$ (resp. $\kappa_B((X_b)_{b \in B})$) is the joint moment (resp. cumulant) of the family of random variables $(X_b)_{b \in B}$. The previous Lemma 5.2.1 translates in the following property for the cumulant. If there is a partition \mathcal{I} with at least two cells, such that the collection random variables $(X_i)_{i \in I}$ for $I \in \mathcal{I}$ are mutually independent, then the joint cumulant of the family $(X_a)_{a \in A}$ is zero.

The joint cumulants are a convenient tool in the Gaussian framework, since for a Gaussian vector $(X_a)_{a \in A}$, the joint cumulant $\kappa_A((X_a)_{a \in A})$ cancels as soon as $|A| \geq 3$. Reciprocally, a random variable X such that $\kappa_p(X, \dots, X) = 0$ for all $p \geq 3$ is Gaussian.

5.2.2 Diagonal set and factorial power measure

We will see in Section 5.3 that the Kac–Rice formula gives an integral formula for the p -th factorial power measure of the zero set of a Gaussian process. The expression of the Kac density degenerates near the diagonal and it motivates the introduction of a few notations for the diagonal of a set and factorial power measure. In the following, A is a finite set and (E, d) is a metric space.

Diagonal set and diagonal inclusion

We define the (large) diagonal of E^A the subset

$$\Delta := \Delta^A = \left\{ \underline{x}_A \in E^A \mid \exists a, b \in A \text{ with } a \neq b \text{ and } x_a = x_b \right\}.$$

Let \mathcal{I} be a partition of the set A . We define

$$\Delta_{\mathcal{I}} = \left\{ \underline{x}_A \in E^A \mid x_a = x_b \iff [a]_{\mathcal{I}} = [b]_{\mathcal{I}} \right\}.$$

From this definition, one has the following decomposition of the space E

$$E^A = \bigsqcup_{\mathcal{I} \in \mathcal{P}_A} \Delta_{\mathcal{I}}, \quad \Delta = \bigsqcup_{\substack{\mathcal{I} \in \mathcal{P}_A \\ \mathcal{I} \neq \bar{A}}} \Delta_{\mathcal{I}} \quad \text{and} \quad E^A \setminus \Delta = \Delta_{\bar{A}},$$

where \bar{A} is the partition of A in singletons. We also define

$$\Delta_{\mathcal{I}^+} := \bigsqcup_{\mathcal{J} \preceq \mathcal{I}} \Delta_{\mathcal{J}} = \left\{ \underline{x}_A \in E^A \mid x_a = x_b \implies [a]_{\mathcal{I}} = [b]_{\mathcal{I}} \right\}.$$

Enlargement of the diagonal set

We fix a number $\eta \geq 0$ and $\underline{x}_A \in E^A$. We define the graph $G_{\eta}(\underline{x}_A)$ with set vertices A , and where two vertices a and b are connected by an edge if $d(x_a, x_b) \leq \eta$. Denote by $\mathcal{I}_{\eta}(\underline{x}_A)$ the partition of A induced by the connected components of $G_{\eta}(\underline{x}_A)$. It allows us to define the subset

$$\Delta_{\mathcal{I}, \eta} = \left\{ \underline{x}_A \in E^A \mid \mathcal{I}_{\eta}(\underline{x}_A) = \mathcal{I} \right\}. \quad (5.18)$$

If $\eta = 0$ then $\Delta_{\mathcal{I},0} = \Delta_{\mathcal{I}}$. In the case where $\eta > 0$ we have $\overline{\Delta_{\mathcal{I},\eta}} \subset \Delta_{\mathcal{I}^+}$. As in the case $\eta = 0$, we also have

$$E^A = \bigsqcup_{\mathcal{I} \in \mathcal{P}_A} \Delta_{\mathcal{I},\eta}.$$

The fundamental property of this construction is the following. Let $a, b \in A$ and $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$. If $[a]_{\mathcal{I}} = [b]_{\mathcal{I}}$ then

$$d(x_a, x_b) \leq |A|\eta,$$

and if $[a]_{\mathcal{I}} \neq [b]_{\mathcal{I}}$ then

$$d(x_a, x_b) \geq \eta.$$

The factorial power measure

We define the diagonal inclusion

$$\begin{aligned} \iota_{\mathcal{I}} : E^{\mathcal{I}} &\longrightarrow E^A \\ \underline{x}_{\mathcal{I}} &\longrightarrow (x_{[a]_{\mathcal{I}}})_{a \in A}. \end{aligned}$$

For instance, if $\mathcal{I} = \{\{1, 3\}, \{2\}\}$ then $\iota_{\mathcal{I}}(x, y) = (x, y, x)$. A direct consequence of this definition is that the mapping $\iota_{\mathcal{I}}$ is a bijection between $E^{\mathcal{I}} \setminus \Delta$ and $\Delta_{\mathcal{I}}$.

Let Z be a locally finite subset of the metric space E . We set $\nu := \sum_{x \in Z} \delta_x$ the counting measure on Z , and

$$\nu^A = \sum_{x \in Z^A} \delta_x \quad \text{and} \quad \nu^{[A]} = \sum_{x \in Z^A \setminus \Delta} \delta_x.$$

The measure ν^A (resp. $\nu^{[A]}$) is the power (resp. factorial power) measure of the measure ν . Both measures are linked by the following lemma.

Lemma 5.2.2. *With the notations as above, one has*

$$\nu^A = \sum_{\mathcal{I} \in \mathcal{P}_A} \iota_{\mathcal{I}*} \nu^{[\mathcal{I}]}.$$

Proof. We have

$$\sum_{x \in Z^A} \delta_x = \sum_{\mathcal{I} \in \mathcal{P}_A} \left(\sum_{x \in Z^A \cap \Delta_{\mathcal{I}}} \delta_x \right).$$

Using the fact that the mapping $\iota_{\mathcal{I}}$ is a bijection between $E^{\mathcal{I}} \setminus \Delta$ and $\Delta_{\mathcal{I}}$, one gets

$$\sum_{x \in Z^A \cap \Delta_{\mathcal{I}}} \delta_x = \sum_{y \in Z^{\mathcal{I}} \setminus \Delta} \delta_{\iota_{\mathcal{I}}(y)} = \iota_{\mathcal{I}*} \nu^{[\mathcal{I}]}.$$

□

5.2.3 Matrix notations

The Kac density (see Section 5.3 and Lemma 5.3.6) is expressed in term of the covariance matrix of the underlying Gaussian process and its derivatives. This fact allows us to consider the Kac density as a function defined on the set of positive definite matrices, evaluated in some covariance matrix related to our underlying Gaussian process. To this end, we introduce a few useful notations

Basic matrix notations

In the following, we define $\mathcal{M}_A(\mathbb{R})$, $\mathbb{S}_A(\mathbb{R})$ and $\mathbb{S}_A^+(\mathbb{R})$ respectively the sets of square, symmetric and symmetric positive definite matrices acting on the space \mathbb{R}^A equipped with its canonical basis. If B is another finite set, we define $\mathcal{M}_{A,B}(\mathbb{R})$ the set of matrices from \mathbb{R}^A to \mathbb{R}^B . The open subset of matrices in $\mathcal{M}_{A,B}(\mathbb{R})$ with maximal rank is denoted $\mathcal{M}_{A,B}^*(\mathbb{R})$. For a matrix $\Gamma \in \mathcal{M}_{A,B}(\mathbb{R})$ we define

$$\|\Gamma\| = \sup_{i,j} |\Gamma_{i,j}|.$$

Given a matrix $\Sigma \in \mathcal{M}_{2A,2B}(\mathbb{R})$, we write

$$\Sigma = \left(\begin{array}{c|c} \Sigma^{11} & \Sigma^{12} \\ \hline \Sigma^{21} & \Sigma^{22} \end{array} \right), \quad \Sigma^{11}, \Sigma^{12}, \Sigma^{21}, \Sigma^{22} \in \mathcal{M}_{A,B}(\mathbb{R}).$$

Let $\Sigma \in \mathcal{M}_{2A}(\mathbb{R})$. If the matrix Σ^{11} is invertible, we define the matrix $\Sigma^c \in \mathcal{M}_A(\mathbb{R})$ to be the Schur complement of Σ^{11} in Σ :

$$\Sigma^c = \Sigma^{22} - \Sigma^{21}(\Sigma^{11})^{-1}\Sigma^{12}.$$

This matrix arises from the identity

$$\left(\begin{array}{c|c} \text{Id} & 0 \\ \hline -\Sigma^{21}(\Sigma^{11})^{-1} & \text{Id} \end{array} \right) \left(\begin{array}{c|c} \Sigma^{11} & \Sigma^{12} \\ \hline \Sigma^{21} & \Sigma^{22} \end{array} \right) = \left(\begin{array}{c|c} \Sigma^{11} & \Sigma^{12} \\ \hline 0 & \Sigma^c \end{array} \right).$$

In particular

$$\det(\Sigma) = \det(\Sigma^{11}) \det(\Sigma^c). \quad (5.19)$$

If $\Sigma \in \mathbb{S}_{2A}^+(\mathbb{R})$ then $\Sigma^c \in \mathbb{S}_A^+(\mathbb{R})$ and

$$(\Sigma^c)^{-1} = (\Sigma^{-1})^{22}. \quad (5.20)$$

Covariance matrix and Gaussian conditioning

Let $\underline{X}_A = (X_a)_{a \in A}$ and $\underline{Y}_A = (Y_a)_{a \in A}$ two sequences of jointly centered Gaussian vectors. We assume that the Gaussian vector $(\underline{X}_A, \underline{Y}_A)$ is non-degenerate. We define

$$\Sigma^{11} = \text{Cov}(\underline{X}_A), \quad \Sigma^{22} = \text{Cov}(\underline{Y}_A), \quad \Sigma^{12} = \text{Cov}(\underline{X}_A, \underline{Y}_A),$$

and

$$\Sigma := \text{Cov}[(\underline{X}_A, \underline{Y}_A)] = \left(\begin{array}{c|c} \Sigma^{11} & \Sigma^{12} \\ \hline {}^T \Sigma^{12} & \Sigma^{22} \end{array} \right).$$

Lemma 5.2.3. *We have*

$$\text{Law}(\underline{Y}_A | \underline{X}_A = 0) \sim \mathcal{N}(0, \Sigma^c).$$

Proof. We define the Gaussian vector

$$\underline{Y}_A^c = \underline{Y}_A - {}^T \Sigma^{12} (\Sigma^{11})^{-1} \underline{X}_A.$$

Then

$$\text{Cov}(\underline{X}_A, \underline{Y}_A^c) = 0 \quad \text{and} \quad \text{Cov}(\underline{Y}_A^c) = \Sigma^c.$$

Since decorrelation implies independence for Gaussian vectors, we have the following

equality of conditional distributions

$$\text{Law}(\underline{Y}_A | \underline{X}_A) \sim \mathcal{N}(T\Sigma^{12}(\Sigma^{11})^{-1}\underline{X}_A, \Sigma^c) \quad \text{and} \quad \text{Law}(\underline{Y}_A | \underline{X}_A = 0) \sim \mathcal{N}(0, \Sigma^c).$$

□

Compactness in matrix sets

The following lemmas give equivalent conditions to being compact in several matrix spaces.

Lemma 5.2.4. *A set K is relatively compact in $\mathbb{S}_A^+(\mathbb{R})$ if and only if one can find positive constants c_K and C_K such that*

$$\forall \Sigma \in K, \quad \det \Sigma \geq c_K, \quad \|\Sigma\| \leq C_K.$$

Lemma 5.2.5. *Let B be a subset of A . A set K is relatively compact in $\mathcal{M}_{B,A}^*(\mathbb{R})$ if and only if one can find positive constants c_K and C_K such that*

$$\forall Q \in K, \quad \det Q^T Q \geq c_K, \quad \|Q\| \leq C_K.$$

Lemma 5.2.6. *Let K be a relatively compact subset of $\mathcal{M}_{B,A}^*(\mathbb{R}) \times \mathbb{S}_A^+(\mathbb{R})$. Then the set*

$$\{Q\Sigma^T Q \mid (Q, \Sigma) \in K\}$$

is relatively compact in $\mathbb{S}_B^+(\mathbb{R})$.

Proof of the lemmas. The proof of Lemma 5.2.4 is a direct consequence of the continuity of the determinant. For Lemma 5.2.5, note that a matrix Q of $\mathcal{M}_{B,A}(\mathbb{R})$ is of maximal rank if and only if $\det Q^T Q > 0$. The conclusion follows again from the continuity of the determinant. As for Lemma 5.2.6, let $\Sigma \in \mathbb{S}_A^+(\mathbb{R})$ and $Q \in \mathcal{M}_{B,A}^*(\mathbb{R})$. A direct computation shows that the matrix $Q\Sigma^T Q$ is positive definite. The conclusion then follows from the continuity of the application $(Q, \Sigma) \mapsto Q\Sigma^T Q$. □

Block diagonal matrix with respect to a partition

Let B and C be subsets of A , and $\Gamma \in \mathcal{M}_{B,C}(\mathbb{R})$. For I and J subset of A we define

$$\Gamma_{I,J} = (\Gamma_{i,j})_{i \in I \cap B, j \in J \cap C} \quad \text{and} \quad \Gamma_I = \Gamma_{I,I}. \quad (5.21)$$

Now let $\Sigma \in \mathcal{M}_{2B,2C}(\mathbb{R})$. We define similarly

$$\Sigma_{I,J} = \left(\begin{array}{c|c} \Sigma_{I,J}^{11} & \Sigma_{I,J}^{12} \\ \hline \Sigma_{I,J}^{21} & \Sigma_{I,J}^{22} \end{array} \right) \quad \text{and} \quad \Sigma_I = \Sigma_{I,I}.$$

For a partition \mathcal{I} of the set A and a matrix $\Gamma \in \mathcal{M}_{B,C}(\mathbb{R})$ we define $\Gamma_{\mathcal{I}}$ to be the block diagonal matrix with blocks $(\Gamma_I)_{I \in \mathcal{I}}$. Similarly, for a matrix $\Sigma \in \mathcal{M}_{2B,2C}(\mathbb{R})$ we define

$$\Sigma_{\mathcal{I}} = \left(\begin{array}{c|c} \Sigma_{\mathcal{I}}^{11} & \Sigma_{\mathcal{I}}^{12} \\ \hline \Sigma_{\mathcal{I}}^{21} & \Sigma_{\mathcal{I}}^{22} \end{array} \right).$$

Note that if $\Sigma \in \mathbb{S}_{2A}^+(\mathbb{R})$ such that $\Sigma = \Sigma_{\mathcal{I}}$, then

$$(\Sigma^c)_{\mathcal{I}} = (\Sigma_{\mathcal{I}})^c, \quad (5.22)$$

but the equality is not true in all generality.

Power product space

We introduce a technical matrix space, that will be central in the alternative expression of the cumulant Kac density in Section 5.3.2. We define the sets

$$\widetilde{\mathcal{M}}_A(\mathbb{R}) := \prod_{B \subset A} (\mathcal{M}_B(\mathbb{R}) \times \mathcal{M}_{2B,2A}(\mathbb{R})) \quad \text{and} \quad \widetilde{\mathcal{M}}_A^*(\mathbb{R}) := \prod_{B \subset A} (\mathcal{M}_B(\mathbb{R}) \times \mathcal{M}_{2B,2A}^*(\mathbb{R})). \quad (5.23)$$

The space $\widetilde{\mathcal{M}}_A^*(\mathbb{R})$ is an open subset of $\widetilde{\mathcal{M}}_A(\mathbb{R})$. An element $(M^B, Q^B)_{B \subset A}$ of $\widetilde{\mathcal{M}}_A(\mathbb{R})$ will be denoted $(\widetilde{M}, \widetilde{Q})$, with

$$\widetilde{M} = (M^B)_{B \subset A} \quad \text{and} \quad \widetilde{Q} = (Q^B)_{B \subset A}.$$

If C is a subset of A , we denote $(\widetilde{M}_C, \widetilde{Q}_C)$ the element of $\widetilde{\mathcal{M}}_C(\mathbb{R})$ defined by

$$\widetilde{M}_C = \left(M_C^B\right)_{B \subset C} \quad \text{and} \quad \widetilde{Q}_C = \left(Q_C^B\right)_{B \subset C}.$$

At last, if \mathcal{I} is a partition of A , we define

$$\widetilde{M}_{\mathcal{I}} = \left(M_{\mathcal{I}}^B\right)_{B \subset A} \quad \text{and} \quad \widetilde{Q}_{\mathcal{I}} = \left(Q_{\mathcal{I}}^B\right)_{B \subset A}.$$

5.2.4 Divided differences

We now introduce the notion of divided differences. Classically used in interpolation theory, we use it in order to give a non degenerate expression of the Kac density near the diagonal. This approach was first taken in [Cuz75] in order to give a necessary and sufficient condition for the finiteness of the p -th moment of the number of zeros on an interval, and has been extensively used in [AL21a]. The results of this section is standard material about interpolation, but we will recall basic properties of divided differences.

Definition and basic properties

Definition

Let f be a regular function defined on an open interval U of \mathbb{R} or \mathbb{T} . We use the notations of Section 5.2.2. In particular, we consider Δ the large diagonal of U^A . Let A be a finite set and $\underline{x}_A \in U^A \setminus \Delta$. We define $\mathbb{R}^A[X]$ the space of polynomials of degree $|A| - 1$. The polynomial

$$L[f; \underline{x}_A] : x \mapsto \sum_{a \in A} f(x_a) \prod_{b \neq a} \frac{x - x_b}{x_a - x_b}$$

interpolates the function f at the point $(x_a)_{a \in A}$. It is the only polynomial in $\mathbb{R}^A[X]$ with this property, since the difference of two such polynomials cancels at least $|A|$ times and thus must be zero. The application

$$\pi_{\underline{x}_A} : f \mapsto L[f; \underline{x}_A]$$

is a projector onto the space $\mathbb{R}^A[X]$, whose kernel is the space of functions that cancels on \underline{x}_A . We then define the divided difference of f as the coefficient of degree $|A| - 1$ in

$L[f; \underline{x}_A]$. More explicitly,

$$f[\underline{x}_A] = \sum_{a \in A} f(x_a) \prod_{b \neq a} \frac{1}{x_a - x_b}.$$

For instance,

$$f[x] = f(x), \quad f[x, y] = \frac{f(x) - f(y)}{x - y}, \quad f[x, y, z] = \frac{f(x)(z - y) + f(y)(x - z) + f(z)(y - x)}{(x - y)(y - z)(z - x)},$$

and so on. The following lemma is an analogous of Taylor expansion theorem, in the context of divided differences.

Lemma 5.2.7. *For $a \in A$, one has*

$$L[f; \underline{x}_A](x) = L[f; \underline{x}_{A \setminus \{a\}}](x) + f[\underline{x}_A] \prod_{b \neq a} (x - x_b),$$

and

$$f(x) = L[f; \underline{x}_A](x) + f[\underline{x}_A, x] \prod_{a \in A} (x - x_a).$$

Proof. For $a \in A$, the polynomial

$$x \mapsto L[f; \underline{x}_A](x) - f[\underline{x}_A] \prod_{b \neq a} (x - x_b)$$

interpolates the points $(x_b)_{b \neq a}$ and is of degree $|A| - 2$. By uniqueness of the interpolating polynomial, it coincides with the polynomial $L[f; \underline{x}_{A \setminus \{a\}}]$. Hence the first statement. An application of this formula with interpolating points $A \cup \{x\}$ yields the second statement. \square

Continuity property of the divided differences

Recall that the function f is assumed to be regular. Define C_A to be the standard simplex of dimension $|A| - 1$:

$$C_A = \left\{ \underline{t}_A \in \mathbb{R}^A \mid \sum_{a \in A} t_a = 1 \right\}.$$

We can equip the simplex C_A with the induced Lebesgue measure dm . We then have the following integral representation for the divided differences, which is analogous to the integral rest in Taylor expansion in the context of divided differences.

Lemma 5.2.8 (Hermite–Genocchi formula). *We have*

$$f[\underline{x}_A] = \int_{C_A} f^{(|A|-1)} \left(\sum_{a \in A} t_a x_a \right) dm(\underline{t}_A).$$

In particular, the application $\underline{x}_A \mapsto f[\underline{x}_A]$ continuously extends to the whole space U^A .

Proof. See [Boo05]. □

This proposition allows us to extend by continuity the functions

$$\underline{x}_A \mapsto f[\underline{x}_A] \quad \text{and} \quad \underline{x}_A \mapsto L[f; \underline{x}_A],$$

from $U^A \setminus \Delta$ to the whole space U^A . For instance, if $x_a = y$ for all $a \in A$, then

$$f[\underline{x}_A] = \frac{f^{(|A|-1)}(y)}{(|A|-1)!} \quad \text{and} \quad L[f; \underline{x}_A](x) = \sum_{j=0}^{|A|-1} \frac{f^{(j)}(y)}{j!} (x-y)^j. \quad (5.24)$$

This expression coincides with the Taylor polynomial of order $|A| - 1$ of the function f at the point y . The continuity property of the divided differences allows us to extend Lemma 5.2.7 and 5.2.8 by taking \underline{x}_A in the whole space U^A and $x \in \mathbb{R}$.

Divided difference of a polynomial

In this section, P denotes a polynomial. The definition of the divided differences implies that the quantity $P[\underline{x}_A]$ is a rational fraction in the variables \underline{x}_A . From the Hermite–Genocchi formula, we can extend the definition of the divided difference to the whole complex space \mathbb{C}^A by analyticity. It implies the following lemma.

Lemma 5.2.9. *The quantity $P[\underline{x}_A]$ is a polynomial expression of the coefficients \underline{x}_A .*

For instance, if $P(x) = x^4$ then

$$P[x, y] = x^3 + x^2y + xy^2 + y^3.$$

If $\deg P < |A|$ then $L[P; \underline{x}_A, x] = P$, and the coefficient of degree $|A|$ of this polynomial is zero, which implies that $P[\underline{x}_A, x] = 0$. In the following we assume that $\deg P \geq |A|$.

Lemma 5.2.10. *The polynomial $x \mapsto P[\underline{x}_A, x]$ is of degree $\deg P - |A|$, and the leading coefficient this polynomial is the leading coefficient of the polynomial P .*

Proof. From the definition of the divided differences,

$$P[\underline{x}_A, x] \underset{x \rightarrow +\infty}{\simeq} \frac{P(x)}{\prod_{a \in A} (x - x_a)},$$

which implies that the polynomial $x \mapsto P[\underline{x}_A, x]$ is of degree $\deg P - |A|$, and its leading coefficient is the leading coefficient of P . \square

Iterated divided differences

Let B be a subset of A and $\underline{x}_B \in U^B$. We define the function

$$f^{[\underline{x}_B]} : x \mapsto f[\underline{x}_B, x]. \quad (5.25)$$

Lemma 5.2.11. *Let $\underline{x}_A \in U^A$. Then*

$$L[f; \underline{x}_A][\underline{x}_B, \cdot] = L[f^{[\underline{x}_B]}, \underline{x}_{A \setminus B}],$$

and

$$f[\underline{x}_A] = f^{[\underline{x}_B]}[\underline{x}_{A \setminus B}].$$

Let $x \in U^A$. Then

$$f[\underline{x}_B, x] = L[f; \underline{x}_A][\underline{x}_B, x] + f[\underline{x}_A, x] \prod_{a \in A \setminus B} (x - x_a).$$

Proof. Consider the two polynomials

$$P_1 = L[f^{[\underline{x}_B]}, \underline{x}_{A \setminus B}] \quad \text{and} \quad P_2 = L[f; \underline{x}_A]^{[\underline{x}_B]}.$$

They both interpolate the points $\underline{x}_{A \setminus B}$ at values $(f[\underline{x}_B, x_a])_{a \in A \setminus B}$. The polynomial P_1 is in $\mathbb{R}^{A \setminus B}[X]$, and $L[f; \underline{x}_A]$ is in $\mathbb{R}^A[X]$. By Lemma 5.2.10, the polynomial P_2 is also in

$\mathbb{R}^{A \setminus B}[X]$. By uniqueness of the interpolating polynomial, these two polynomials are equal, hence the first statement.

The coefficient of degree $|A \setminus B| - 1$ in P_1 is $f^{[x_B]}[x_{A \setminus B}]$, and the one in P_2 is $f[x_A]$ according to lemma 5.2.10. By the previous equality these two coefficients are equal, which yields the second formula. The last formula is a direct application of Lemma 5.2.7 applied to the function $f^{[x_B]}$. \square

Matrix viewpoint of the divided differences

In order to describe the divided differences of a Gaussian process and the induced transformation on the covariance matrix, we rewrite the operation of taking divided differences from a matrix viewpoint. From now on, we equip A with an arbitrary total order \leq and we introduce the notation

$$a \mid A = \{b \in A \mid b \leq a\}.$$

Thus, $\underline{x}_{a \mid A} = (x_b)_{b \leq a}$.

Basis of polynomials adapted to the divided difference

For $\underline{x}_A \in \mathbb{R}^A$ we define the polynomial

$$P_{\underline{x}_A}^a : x \mapsto \prod_{b < a} (x - x_b).$$

For any subset B of A , the family $(P_{\underline{x}_B}^b)_{b \in B}$ is a basis of the space $\mathbb{R}^B[X]$ and we will always equip the space $\mathbb{R}^B[X]$ with this basis.

Remark 5.2.12. The family $(P_{\underline{x}_A}^a)_{a \in A}$ is a family of monic polynomials of increasing degrees. Thus, if we equip A with another total order, the underlying transformation matrix is of determinant ± 1 , depending on the signature of the permutation, and it depends continuously on the quantities $(x_a - x_b)_{a, b \in A}$. It justifies in the following why the order can be chosen arbitrarily.

A direct induction based on Lemma 5.2.7 implies the following lemma.

Lemma 5.2.13. *Let $\underline{x}_A \in U^A$. Then*

$$L[f; \underline{x}_A] = \sum_{a \in A} f[\underline{x}_{a|A}] P_{\underline{x}_A}^a.$$

The finite differences of f are thus the coefficients of the Lagrange interpolation polynomial in the basis $(P_{\underline{x}_A}^a)_{a \in A}$. We define

$$f_A[\underline{x}_A] = (f[\underline{x}_{a|A}])_{a \in A} \quad \text{and} \quad M(\underline{x}_A) = (P_{\underline{x}_A}^b(x_a))_{a, b \in A}.$$

Then Lemma 5.2.13 rewrites matrixially

$$f(\underline{x}_A) = M(\underline{x}_A) f_A[\underline{x}_A]. \quad (5.26)$$

For instance,

$$\begin{pmatrix} f(x) \\ f(y) \\ f(z) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & y-x & 0 \\ 1 & z-x & (z-x)(z-y) \end{pmatrix} \begin{pmatrix} f[x] \\ f[x, y] \\ f[x, y, z] \end{pmatrix}.$$

The matrix $M(\underline{x}_A)$ is lower triangular, thus

$$\det M(\underline{x}_A) = \prod_{a \in A} P_{\underline{x}_A}^a(x_a) = \prod_{a < b} (x_b - x_a). \quad (5.27)$$

Divided differences with respect to a partition

In the following, \mathcal{I} is a partition of the set A . We define

$$f_{\mathcal{I}}[\underline{x}_A] = (f_I[\underline{x}_I])_{I \in \mathcal{I}} = (f[\underline{x}_{i|I}])_{i \in I, I \in \mathcal{I}}$$

We can perform the divided difference independently on each cell of the partition. That is, we can write for all $I \in \mathcal{I}$

$$f_I(\underline{x}_I) = M(\underline{x}_I) f_I[\underline{x}_I] \quad \text{and} \quad f(\underline{x}_A) = M^{\mathcal{I}}(\underline{x}_A) f_{\mathcal{I}}[\underline{x}_A]; \quad (5.28)$$

where $M^{\mathcal{I}}(\underline{x}_A)$ is the block diagonal matrix with blocks $(M(\underline{x}_I))_{I \in \mathcal{I}}$. For instance, if

$\mathcal{I} = \{\{1, 2\}, \{3, 4\}\}$ then

$$\begin{pmatrix} f(w) \\ f(x) \\ f(y) \\ f(z) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x-w & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & z-y \end{pmatrix} \begin{pmatrix} f[w] \\ f[w, x] \\ f[y] \\ f[y, z] \end{pmatrix}.$$

From Equation (5.27),

$$\det M^{\mathcal{I}}(\underline{x}_A) = \prod_{I \in \mathcal{I}} \det M(\underline{x}_I) = \prod_{I \in \mathcal{I}} \prod_{\substack{i, j \in I \\ i < j}} (x_j - x_i).$$

Lemma 5.2.14. *Let $\underline{x}_A \in \Delta_{\mathcal{I}, \eta}$. There is a constant $C(\eta)$ such that*

$$\|M^{\mathcal{I}}(\underline{x}_A)\| \leq C(\eta).$$

Proof. Let $a, b \in A$. If $[a]_{\mathcal{I}} \neq [b]_{\mathcal{I}}$, then $(M^{\mathcal{I}}(\underline{x}_A))_{a,b} = 0$. Else, there is $I \in \mathcal{I}$ such that $a, b \in I$. Then

$$(M^{\mathcal{I}}(\underline{x}_A))_{a,b} = P_{\underline{x}_I}^b(x_a) = \prod_{\substack{c < b \\ c \in I}} (x_a - x_c).$$

For $a, c \in I$, one has $|x_a - x_c| \leq |A|\eta$. The conclusion follows. \square

For $\underline{x}_A \in U^A$ we consider the mapping

$$\begin{aligned} \pi_{\underline{x}_A}^{\mathcal{I}} : \mathbb{R}^A[X] &\longrightarrow \prod_{I \in \mathcal{I}} \mathbb{R}^I[X] \\ P &\longrightarrow (L[P; \underline{x}_I])_{I \in \mathcal{I}}. \end{aligned}$$

It is well defined for $\underline{x}_A \in U^A$, since a polynomial is infinitely differentiable. For a subset I of A , we equip $\mathbb{R}^I[X]$ with the basis of polynomials $(P_{\underline{x}_I}^i)_{i \in I}$. Let $Q^{\mathcal{I}}(\underline{x}_A)$ be the matrix of the application $\pi_{\underline{x}_A}^{\mathcal{I}}$ in that basis. For instance, when $\mathcal{I} = \{\{1, 2\}, \{3\}\}$ then

$$\begin{pmatrix} f[x] \\ f[x, y] \\ f[z] \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & z-x & (z-x)(z-y) \end{pmatrix} \begin{pmatrix} f[x] \\ f[x, y] \\ f[x, y, z] \end{pmatrix}.$$

Let $\underline{x}_A \in U^A \setminus \Delta$. Then $M^{\mathcal{I}}(\underline{x}_A)$ is invertible and from Equation (5.26) and (5.28), one has

$$f_{\mathcal{I}}[\underline{x}_A] = [M^{\mathcal{I}}(\underline{x}_A)]^{-1} M(\underline{x}_A) f_A[\underline{x}_A],$$

and thus

$$Q^{\mathcal{I}}(\underline{x}_A) = [M^{\mathcal{I}}(\underline{x}_A)]^{-1} M(\underline{x}_A).$$

It implies that

$$\left| \det Q^{\mathcal{I}}(\underline{x}_A) \right| = \prod_{\substack{\{I,J\} \subset \mathcal{I} \\ I \neq J}} \prod_{i \in I} \prod_{j \in J} |x_i - x_j|. \quad (5.29)$$

By continuity with respect to \underline{x}_A of the application $\pi_{\underline{x}_A}^{\mathcal{I}}$ this formula remain true for $\underline{x}_A \in \mathbb{R}^A$. We deduce that the application $\pi_{\underline{x}_A}^{\mathcal{I}}$ is invertible for $\underline{x}_A \in \Delta_{\mathcal{I}^+}$.

Now let \mathcal{J} be another partition of A , finer than the partition \mathcal{I} . For $\underline{x}_A \in U^A$ we consider the mapping

$$\begin{aligned} \pi_{\underline{x}_A}^{\mathcal{I},\mathcal{J}} : \prod_{I \in \mathcal{I}} \mathbb{R}^I[X] &\longrightarrow \prod_{J \in \mathcal{J}} \mathbb{R}^J[X] \\ (P_I)_{I \in \mathcal{I}} &\longrightarrow (L[P_{J_I}; \underline{x}_J])_{J \in \mathcal{J}}. \end{aligned}$$

Restricted to $\mathbb{R}^I[X]$, the application $\pi_{\underline{x}_A}^{\mathcal{I},\mathcal{J}}$ coincides with $\pi_{\underline{x}_I}^{\mathcal{J}_I}$. Let $Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A)$ be the matrix of the application $\pi_{\underline{x}_A}^{\mathcal{I},\mathcal{J}}$, so that

$$f_{\mathcal{J}}[\underline{x}_A] = Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A) f_{\mathcal{I}}[\underline{x}_A]. \quad (5.30)$$

Lemma 5.2.15.

$$\left| \det Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A) \right| = \prod_{\substack{I \in \mathcal{I} \\ J_1 \neq J_2}} \prod_{\{J_1, J_2\} \subset \mathcal{J}_I} \prod_{j_1 \in J_1} \prod_{j_2 \in J_2} |x_{j_1} - x_{j_2}|.$$

In particular, the application $\pi_{\underline{x}_A}^{\mathcal{I},\mathcal{J}}$ is invertible in the case where $\underline{x}_A \in \Delta_{\mathcal{J}^+}$.

Proof. We have from Equation (5.29)

$$\left| \det Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A) \right| = \prod_{I \in \mathcal{I}} \left| \det Q^{\mathcal{J}_I}(\underline{x}_I) \right| = \prod_{\substack{I \in \mathcal{I} \\ J_1 \neq J_2}} \prod_{\{J_1, J_2\} \subset \mathcal{J}_I} \prod_{j_1 \in J_1} \prod_{j_2 \in J_2} |x_{j_1} - x_{j_2}|.$$

And this expression does not cancel when $\underline{x}_A \in \Delta_{\mathcal{J}^+}$. □

Divided difference on a subset

In this section, B is a subset of A and \mathcal{I} is a partition of A . From Equation (5.26) and (5.28), one has

$$f(\underline{x}_B) = M_{B,A}(\underline{x}_A) f_A[\underline{x}_A] \quad \text{and} \quad f(\underline{x}_B) = M_{B,A}^{\mathcal{I}}(\underline{x}_A) f_{\mathcal{I}}[\underline{x}_A],$$

where for a matrix M , the matrix $M_{B,A}$ is defined in (5.21). For $\underline{x}_A \in U^A$ we consider the mapping

$$\begin{aligned} \pi_{\underline{x}_A}^{\mathcal{I},B} : \prod_{I \in \mathcal{I}} \mathbb{R}^I[X] &\longrightarrow \prod_{I \in \mathcal{I}} \mathbb{R}^{I \cap B}[X] \\ (P_I)_{I \in \mathcal{I}} &\longrightarrow (L[P_I; \underline{x}_{I \cap B}])_{I \in \mathcal{I}}, \end{aligned} \tag{5.31}$$

and let $Q^{\mathcal{I},B}(\underline{x}_A)$ be the matrix of this application.

Lemma 5.2.16. *The set*

$$\left\{ Q^{\mathcal{I},B}(\underline{x}_A) \mid \underline{x}_A \in \Delta_{\mathcal{I},\eta} \right\}$$

is relatively compact in $\mathcal{M}_{B,A}^(\mathbb{R})$.*

Proof. Let $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$. The matrix $Q^{\mathcal{I},B}(\underline{x}_A)$ is block diagonal with respect to the partition \mathcal{I} , with blocks $(Q^{\{I\},B \cap I}(\underline{x}_I))_{I \in \mathcal{I}}$. For $I \in \mathcal{I}$, the quantities $(x_i - x_j)_{i,j \in I}$ are bounded by $|A|\eta$. According to remark 5.2.12, we can change the order on I and assume that $B \cap I = b|I$ for some $b \in B \cap I$. In that case, one has

$$Q^{\{I\},B \cap I}(\underline{x}_I) = \left(\text{Id}_{B \cap I} \mid 0 \right),$$

and the conclusion follows. □

Doubling divided differences

In the paragraph, we consider the divided difference on the set $2A$ defined in (5.13). We equip the set $2A$ with the lexicographic order inherited from the order on $\{1, 2\}$ and the arbitrary order on A . Note that

$$(f[x], f[x, x]) = (f(x), f'(x)).$$

The interest of doubling divided differences is to consider simultaneous interpolation of the function f and f' on prescribed points $(\underline{x}_a)_{a \in A}$. This part coincides with the classical Hermite interpolation.

Let \mathcal{I} be a partition of A (and thus of $2A$, following the notation (5.14)). As a consequence of Definition (5.31), one has

$$f_{\mathcal{I}B}[\underline{x}_{B,B}] = Q^{\mathcal{I},2B}(\underline{x}_{A,A})f_{\mathcal{I}}[\underline{x}_{A,A}]. \quad (5.32)$$

We define

$$\widetilde{M}^{\mathcal{I}}(\underline{x}_A) = (M^{\mathcal{I}B}(\underline{x}_B))_{B \subset A} \quad \text{and} \quad \widetilde{Q}^{\mathcal{I}}(\underline{x}_A) = (Q^{\mathcal{I},2B}(\underline{x}_{A,A}))_{B \subset A}. \quad (5.33)$$

Recall the definition of $\widetilde{\mathcal{M}}_A(\mathbb{R})$ and $\widetilde{\mathcal{M}}_A^*(\mathbb{R})$ in Paragraph 5.2.3. One has the following key proposition.

Proposition 5.2.17. *Let $\underline{x}_A \in \mathbb{R}^A$. Then*

$$(\widetilde{M}^{\mathcal{I}}(\underline{x}_A), \widetilde{Q}^{\mathcal{I}}(\underline{x}_A)) \in \widetilde{\mathcal{M}}_A(\mathbb{R}),$$

$$\widetilde{M}^{\mathcal{I}}(\underline{x}_A) = \widetilde{M}_{\mathcal{I}}^{\mathcal{I}}(\underline{x}_A) \quad \text{and} \quad \widetilde{Q}^{\mathcal{I}}(\underline{x}_A) = \widetilde{Q}_{\mathcal{I}}^{\mathcal{I}}(\underline{x}_A),$$

$$\forall B \subset A, \quad \forall I \in \mathcal{I}, \quad M_I^B = M^{I \cap B} \quad \text{and} \quad Q_I^B = Q^{I \cap B}.$$

Moreover, the set

$$\{(\widetilde{M}^{\mathcal{I}}(\underline{x}_A), \widetilde{Q}^{\mathcal{I}}(\underline{x}_A)) \mid \underline{x}_A \in \Delta_{\mathcal{I},\eta}\}$$

is relatively compact in $\widetilde{\mathcal{M}}_A^*(\mathbb{R})$.

Proof. The first three assertions directly follow from the definition of $\widetilde{M}^{\mathcal{I}}$ and $\widetilde{Q}^{\mathcal{I}}$. As for the second proposition, let $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$. According to Lemma 5.2.14, the coefficients of the matrix $M^{\mathcal{I}B}(\underline{x}_B)$ are bounded by a constant $C(\eta)$. We also have $\underline{x}_{A,A} \in \Delta_{\mathcal{I},\eta}$. Lemma 5.2.16 applies, and the set of matrices $(Q^{\mathcal{I},2B}(\underline{x}_{A,A}))_{\underline{x}_A \in \Delta_{\mathcal{I},\eta}}$ is relatively compact in $\mathcal{M}_{2B,2A}^*(\mathbb{R})$. The conclusion follows. \square

Divided differences of a Gaussian process

At last we describe the covariance matrix of the divided difference vector of a Gaussian

process. The integral representation given by the Hermite–Genocchi formula gives a convenient expression for the coefficients of the covariance matrix.

Let f be a Gaussian process of class $\mathcal{C}^{|A|-1}$ on the interval U . We denote r the covariance function of f , which is differentiable $|A| - 1$ times in each variable. We define

$$\Sigma^{\mathcal{I}}(\underline{x}_A) = \text{Cov}(f_{\mathcal{I}}[\underline{x}_A]). \quad (5.34)$$

Lemma 5.2.18. *Let $I, J \in \mathcal{I}$, $a \in I$ and $b \in J$. Then*

$$\Sigma^{\mathcal{I}}(\underline{x}_A)_{ab} = \int_{C_{a|I} \times C_{b|J}} r^{(|a|I|-1, |b|J|-1)} \left(\sum_{i \in a|I} s_i x_i, \sum_{j \in b|J} t_j x_j \right) dm(\underline{s}_{i|I}) dm(\underline{t}_{i|J}).$$

Proof. The Hermite Genocchi formula 5.2.8 asserts that

$$\begin{aligned} \Sigma^{\mathcal{I}}(\underline{x}_A)_{ab} &= \mathbb{E}[f[\underline{x}_{a|I}] f[\underline{x}_{b|J}]] \\ &= \int_{C_{a|I} \times C_{b|J}} \mathbb{E} \left[f^{(|a|I|-1)} \left(\sum_{i \in a|I} s_i x_i \right) f^{(|b|J|-1)} \left(\sum_{j \in b|J} t_j x_j \right) \right] dm(\underline{s}_{a|I}) dm(\underline{t}_{b|J}) \\ &= \int_{C_{a|I} \times C_{b|J}} r^{(|a|I|-1, |b|J|-1)} \left(\sum_{i \in a|I} s_i x_i, \sum_{j \in b|J} t_j x_j \right) dm(\underline{s}_{a|I}) dm(\underline{t}_{b|J}). \end{aligned}$$

□

5.3 Kac–Rice formula for Gaussian processes

For now on, A is a finite set and f is a centered Gaussian process defined on an interval U of \mathbb{R} or \mathbb{T} , with covariance function r . We assume for this section that the process f is of class $\mathcal{C}^{2|A|-1}$, and for every partition \mathcal{I} of A and $\underline{y}_{\mathcal{I}} \in U^{\mathcal{I}} \setminus \Delta$, the following non-degeneracy condition holds

$$\det \left[\text{Cov} \left((f^{(k)}(y_I))_{0 \leq k \leq 2|I|-1} \right)_{I \in \mathcal{I}} \right] > 0. \quad (5.35)$$

It has been shown for instance in [AW09] that this condition ensures the finiteness of the $|A|$ -th moment of the number of zeros of a Gaussian process on a bounded interval.

5.3.1 Kac density and cumulants of the zeros counting measure

In this section we give the expression of the p -th factorial moment and cumulant of the zeros counting measure. The first step is to lift the degeneracy of the Kac–Rice formula near the diagonal.

Non-degeneracy of the Kac density near the diagonal

We apply the method of divided differences to lift the degeneracy of the Kac density and give an alternative formula near the diagonal. The material in this section is quite standard, see for instance [AW09] and [AL21a]. Only Lemma 5.3.4 is new and allows us to express the Kac density as a function of a non degenerate Gaussian vector.

We fix for the rest of this paragraph a partition \mathcal{I} of A .

Lemma 5.3.1. *The Gaussian vectors $f_{\mathcal{I}}[\underline{x}_A]$ and $f_{\mathcal{I}}[\underline{x}_{A,A}]$ are non degenerate for $\underline{x}_A \in \Delta_{\mathcal{I}^+}$.*

Proof. We prove first the non degeneracy of $f_{\mathcal{I}}[\underline{x}_A]$. The process f is of class $\mathcal{C}^{|A|-1}$, and the Gaussian vector $f_{\mathcal{I}}[\underline{x}_A]$ is well-defined. Let $\underline{x}_A \in \Delta_{\mathcal{I}^+}$. By definition of the set $\Delta_{\mathcal{I}^+}$, there is a partition \mathcal{J} finer than the partition \mathcal{I} and such that $\underline{x}_A \in \Delta_{\mathcal{J}}$. We write $\underline{x}_A = \iota_{\mathcal{J}}(\underline{y}_{\mathcal{J}})$ for some $\underline{y}_{\mathcal{J}} \in \mathbb{R}^{\mathcal{J}} \setminus \Delta$. For $J \in \mathcal{J}$, we have from equation (5.24)

$$f_{\mathcal{J}}[\underline{x}_A] = \left(\frac{f^{(k)}(y_J)}{k!} \right)_{\substack{0 \leq k \leq |J|-1 \\ J \in \mathcal{J}}},$$

which is non-degenerate by the hypothesis (5.35). Now from Equation (5.30),

$$f_{\mathcal{I}}[\underline{x}_A] = Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A) f_{\mathcal{J}}[\underline{x}_A].$$

According to Lemma 5.2.15, the matrix $Q^{\mathcal{I},\mathcal{J}}(\underline{x}_A)$ is invertible when $\underline{x}_A \in \Delta_{\mathcal{J}}$, which implies that the Gaussian vector $f_{\mathcal{I}}[\underline{x}_A]$ is non degenerate for $\underline{x}_A \in \Delta_{\mathcal{I}^+}$.

Now for the Gaussian vector $f_{\mathcal{I}}[\underline{x}_{A,A}]$, the process f is of class $\mathcal{C}^{2|A|-1}$ and the Gaussian vector $f_{\mathcal{I}}[\underline{x}_{A,A}]$ is well-defined. Moreover, if $\underline{x}_A \in \Delta_{\mathcal{I}^+}$ then $\underline{x}_{A,A} \in \Delta_{\mathcal{I}^+}$ and we can apply the previous case to the set $2A$ to deduce the non degeneracy of the Gaussian vector $f_{\mathcal{I}}[\underline{x}_{A,A}]$. \square

We define the random set

$$Z = \{x \in U \mid f(x) = 0\}.$$

By Bulinskaya lemma (see [AW09]) and the assumption on f , the subset Z is almost surely a closed discrete subset of U and we can define the random measure ν to be the counting measure of Z . Kac–Rice formula (see [AW09, Thm. 3.2] and [AL21a, Prop. 3.6]) asserts that for a positive function $\Phi : U^A \rightarrow \mathbb{R}$, one has, following the notations of Section 5.2.2,

$$\mathbb{E}[\langle \nu^{[A]}, \Phi \rangle] = \int_{U^A} \rho(\underline{x}_A) \Phi(\underline{x}_A) d\underline{x}_A, \quad (5.36)$$

with

$$\rho(\underline{x}_A) := \rho_{|A|}(\underline{x}_A) = \frac{N(\underline{x}_A)}{D(\underline{x}_A)},$$

where

$$N(\underline{x}_A) = \mathbb{E} \left[\prod_{a \in A} |f'(x_a)| \mid \forall a \in A, f(x_a) = 0 \right] \quad \text{and} \quad D(\underline{x}_A) = \sqrt{\det [2\pi \text{Cov}(f(x_A))]}.$$

The function ρ is only defined for $\underline{x}_A \in U^A \setminus \Delta$. Along the diagonal Δ , the function N is ill-defined and the function D cancels. The first step consists in giving an alternative non singular formula for ρ in a neighborhood of the diagonal Δ . Let $\underline{x}_A \in \Delta_{\mathcal{I}^+}$. We define

$$D^{\mathcal{I}}(\underline{x}_A) = \sqrt{\det [2\pi \text{Cov}(f_{\mathcal{I}}[\underline{x}_A])]}, \quad N^{\mathcal{I}}(\underline{x}_A) = \mathbb{E} \left[\prod_{I \in \mathcal{I}} \prod_{i \in I} |f[\underline{x}_I, x_i]| \mid f_{\mathcal{I}}[\underline{x}_A] = 0 \right],$$

and

$$\rho^{\mathcal{I}}(\underline{x}_A) = \frac{N^{\mathcal{I}}(\underline{x}_A)}{D^{\mathcal{I}}(\underline{x}_A)}.$$

Lemma 5.3.1 implies that the three quantities above are well defined.

Remark 5.3.2. If $\mathcal{I} = \overline{A}$, then $f_{\overline{A}}(\underline{x}_A) = f(\underline{x}_A)$ and $f[x_a, x_a] = f'(x_a)$, which implies that

$$\rho(\underline{x}_A) = \rho^{\overline{A}}(\underline{x}_A).$$

One has the following relations.

Lemma 5.3.3. *Let \mathcal{J} be a finer partition than \mathcal{I} . Then for $\underline{x}_A \in \Delta_{\mathcal{J}^+}$ one has*

$$\begin{aligned} D^{\mathcal{J}}(\underline{x}_A) &= \left| \det Q^{\mathcal{I}, \mathcal{J}}(\underline{x}_A) \right| D^{\mathcal{I}}(\underline{x}_A), \\ N^{\mathcal{J}}(\underline{x}_A) &= \left(\det Q^{\mathcal{I}, \mathcal{J}}(\underline{x}_A) \right)^2 N^{\mathcal{I}}(\underline{x}_A), \\ \rho^{\mathcal{J}}(\underline{x}_A) &= \left| \det Q^{\mathcal{I}, \mathcal{J}}(\underline{x}_A) \right| \rho^{\mathcal{I}}(\underline{x}_A). \end{aligned}$$

It implies that the function $\rho^{\mathcal{J}}$ can be extended by continuity from $\Delta_{\mathcal{J}^+}$ to $\Delta_{\mathcal{I}^+}$ via these relations. By taking $\mathcal{I} = \{A\}$ and $\mathcal{J} = \bar{A}$, it implies that the function the function ρ can be extended by continuity to the whole space U^A . Moreover, one has

$$\det Q^{\{A\}, \bar{A}}(\underline{x}_A) = \det M(\underline{x}_A) = \prod_{a < b} (x_b - x_a),$$

thus the function ρ vanishes on the diagonal Δ .

Proof. Let $\underline{x}_A \in \Delta_{\mathcal{J}^+}$. From Equation (5.30),

$$f_{\mathcal{J}}[\underline{x}_A] = Q^{\mathcal{I}, \mathcal{J}}(\underline{x}_A) f_{\mathcal{I}}[\underline{x}_A].$$

We deduce that

$$\begin{aligned} D^{\mathcal{J}}(\underline{x}_A) &= \sqrt{\det [2\pi \text{Cov}(f_{\mathcal{J}}[\underline{x}_A])]} \\ &= \left| \det Q^{\mathcal{I}, \mathcal{J}}(\underline{x}_A) \right| D^{\mathcal{I}}(\underline{x}_A). \end{aligned} \tag{5.37}$$

The quantities $N_A^{\mathcal{I}}(\underline{x}_A)$ and $N_A^{\mathcal{J}}(\underline{x}_A)$ are well defined. The Gaussian vectors $f_{\mathcal{I}}[\underline{x}_A]$ and $f_{\mathcal{J}}[\underline{x}_A]$ are equals up to a linear invertible change of variable, and they cancels simultaneously. In other words, one has

$$\begin{aligned} N^{\mathcal{J}}(\underline{x}_A) &= \mathbb{E} \left[\prod_{J \in \mathcal{J}} \prod_{j \in J} |f[\underline{x}_J, x_j]| \left| f_{\mathcal{J}}[\underline{x}_A] = 0 \right. \right] \\ &= \mathbb{E} \left[\prod_{J \in \mathcal{J}} \prod_{j \in J} |f[\underline{x}_J, x_j]| \left| f_{\mathcal{I}}[\underline{x}_A] = 0 \right. \right] \\ &= \mathbb{E} \left[\prod_{I \in \mathcal{I}} \prod_{J \in \mathcal{J}_I} \prod_{j \in J} |f[\underline{x}_J, x_j]| \left| f_{\mathcal{I}}[\underline{x}_A] = 0 \right. \right]. \end{aligned}$$

Let $I \in \mathcal{I}$, $J \in \mathcal{J}_I$ and $j \in J$. From Lemma 5.2.11, conditionally on $f_I[\underline{x}_I] = 0$ one has

$$f[\underline{x}_J, x_j] = \left(\prod_{i \in I \setminus J} (x_i - x_j) \right) f[\underline{x}_I, x_j],$$

from which we deduce

$$\begin{aligned} N^{\mathcal{J}}(\underline{x}_A) &= \left(\prod_{I \in \mathcal{I}} \prod_{J \in \mathcal{J}_I} \prod_{j \in J} \prod_{i \in I \setminus J} |x_i - x_j| \right) \mathbb{E} \left[\prod_{I \in \mathcal{I}} \prod_{i \in I} |f[\underline{x}_I, x_i]| \mid f_{\mathcal{I}}[\underline{x}_A] = 0 \right] \\ &= \left(\prod_{I \in \mathcal{I}} \prod_{\substack{J_1, J_2 \subset \mathcal{J}_I \\ J_1 \neq J_2}} \prod_{j_1 \in J_1} \prod_{j_2 \in J_2} |x_{j_1} - x_{j_2}| \right) N^{\mathcal{I}}(\underline{x}_A). \end{aligned} \quad (5.38)$$

We deduce the alternative expression for $\rho^{\mathcal{J}}$ from (5.37) and (5.38). \square

When the points $(\underline{x}_a)_{a \in A}$ collapse on the diagonal $\Delta_{\mathcal{I}}$ the vector $(f[\underline{x}_I, \underline{x}_i])_{I \in \mathcal{I}, i \in I}$ becomes degenerate, it makes unpractical the analysis of regularity of the function $\rho^{\mathcal{I}}$ in a neighborhood of the diagonal $\Delta_{\mathcal{I}}$. The following lemma gives another expression of the quantity $N^{\mathcal{I}}(\underline{x}_A)$ that depends fully on a non degenerate Gaussian vector. Recall the definition of the function $f^{[\underline{x}_B]}$ for a subset B of A in (5.25).

Lemma 5.3.4. *One has*

$$N^{\mathcal{I}}(\underline{x}_A) = \mathbb{E} \left[\prod_{I \in \mathcal{I}} \prod_{i \in I} \left| \left(M(\underline{x}_I) f_I^{[\underline{x}_I]}[\underline{x}_I] \right)_i \right| \mid f_{\mathcal{I}}[\underline{x}_A] = 0 \right].$$

Proof. Let $I \in \mathcal{I}$. According to formula (5.26),

$$f^{[\underline{x}_I]}(\underline{x}_I) = M(\underline{x}_I) f_I^{[\underline{x}_I]}[\underline{x}_I].$$

The conclusion follows from the definition of $N^{\mathcal{I}}(\underline{x}_A)$. \square

Expression of the cumulants of the zeros counting measure

We are now ready to give the expression of the cumulant of order $|A|$ of the linear statistics associated to zeros counting measure. Let $(\phi_a)_{a \in A}$ be a collection of bounded

functions with compact support. We define

$$\kappa_A(\nu)(\underline{\phi}_A) = \kappa\left(\langle \nu, \phi_a \rangle_{a \in A}\right),$$

the joint cumulant of the family of random variables $(\langle \nu, \phi_a \rangle)_{a \in A}$. We define the *cumulant Kac density* associated with the set A to be the function

$$\begin{aligned} F_A : \mathbb{R}^A &\longrightarrow \mathbb{R} \\ \underline{x}_A &\longrightarrow \sum_{\mathcal{J} \in \mathcal{P}_A} (|\mathcal{J}| - 1)! (-1)^{|\mathcal{J}| - 1} \prod_{J \in \mathcal{J}} \rho(\underline{x}_J). \end{aligned} \quad (5.39)$$

The following Proposition 5.3.5 express the cumulant of order $|A|$ of the linear statistics associated to zeros counting measure. It is key step in towards proof of Theorem 5.1.6, and reveals the elegant interplay between the factorial power counting measure and the combinatorics of cumulants.

Proposition 5.3.5. *We have*

$$\kappa_A(\nu)(\underline{\phi}_A) = \sum_{\mathcal{I} \in \mathcal{P}_A} \int_{U^{\mathcal{I}}} \left(\prod_{I \in \mathcal{I}} \prod_{i \in I} \phi_i(x_I) \right) F_{\mathcal{I}}(\underline{x}_{\mathcal{I}}) d\underline{x}_{\mathcal{I}}.$$

Proof. We have, using the expression of cumulants in terms of moments given by (5.17) and the notation (5.12)

$$\kappa_A(\nu)(\underline{\phi}_A) = \sum_{\mathcal{I} \in \mathcal{P}_A} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}| - 1} \prod_{I \in \mathcal{I}} \mathbb{E} \left[\langle \nu^I, \underline{\phi}_I^{\otimes} \rangle \right].$$

The link between the power measure and factorial power measure given by Lemma 5.2.2 implies that

$$\mathbb{E} \left[\langle \nu^I, \underline{\phi}_I^{\otimes} \rangle \right] = \sum_{\mathcal{J} \in \mathcal{P}_I} \mathbb{E} \left[\langle \nu^{[\mathcal{J}]}, \underline{\phi}_I^{\otimes} \circ \iota_{\mathcal{J}} \rangle \right].$$

The bijection given by (5.15) then implies

$$\begin{aligned} \kappa_A(\nu)(\underline{\phi}_A) &= \sum_{\mathcal{I} \in \mathcal{P}_A} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}| - 1} \sum_{\mathcal{J} \preceq \mathcal{I}} \prod_{I \in \mathcal{I}} \mathbb{E} \left[\langle \nu^{[\mathcal{J}]}, \underline{\phi}_I^{\otimes} \circ \iota_{\mathcal{J}} \rangle \right] \\ &= \sum_{\mathcal{J} \in \mathcal{P}_A} \sum_{\mathcal{I} \succeq \mathcal{J}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}| - 1} \prod_{I \in \mathcal{I}} \mathbb{E} \left[\langle \nu^{[\mathcal{J}]}, \underline{\phi}_I^{\otimes} \circ \iota_{\mathcal{J}} \rangle \right] \end{aligned}$$

Kac–Rice formula then asserts that

$$\mathbb{E} \left[\left\langle \nu^{[\mathcal{I}]}, \underline{\phi}_{\mathcal{I}}^{\otimes} \circ \iota_{\mathcal{I}} \right\rangle \right] = \int_{U^{\mathcal{I}}} \left(\prod_{\substack{J \in \mathcal{J} \\ J \subset \mathcal{I}}} \prod_{j \in J} \phi_j(x_J) \right) \rho(\underline{x}_{\mathcal{I}}) d\underline{x}_{\mathcal{I}},$$

from which we deduce that

$$\begin{aligned} \kappa_A(\nu)(\underline{\phi}_A) &= \sum_{\mathcal{J} \in \mathcal{P}_A} \int_{U^{\mathcal{J}}} \left(\prod_{J \in \mathcal{J}} \prod_{j \in J} \phi_j(x_J) \right) \sum_{\mathcal{I} \supseteq \mathcal{J}} (|\mathcal{I}| - 1)! (-1)^{|\mathcal{I}|-1} \prod_{I \in \mathcal{I}} \rho(\underline{x}_{\mathcal{I}}) d\underline{x}_{\mathcal{J}} \\ &= \sum_{\mathcal{J} \in \mathcal{P}_A} \int_{U^{\mathcal{J}}} \left(\prod_{J \in \mathcal{J}} \prod_{j \in J} \phi_j(x_J) \right) F_{\mathcal{J}}(\underline{x}_{\mathcal{J}}) d\underline{x}_{\mathcal{J}}, \end{aligned}$$

where the last equality follows from the bijection given by (5.15). \square

For instance if $|A| = 2$ then the second order cumulant coincides with the variance and

$$\kappa_A(\nu)(\underline{\phi}_A) = \left(\int_{U^2} \phi_1(x) \phi_2(y) [\rho_2(x, y) - \rho_1(x) \rho_1(y)] dx dy \right) + \left(\int_U \phi_1(x) \phi_2(x) \rho_1(x) dx \right).$$

5.3.2 Matrix representation of the Kac density and factorization property

In this section we prove a matrix representation for the Kac density and the cumulant Kac density. It allows us to dissociate the analysis of the covariance matrix of divided differences associated with the Gaussian process f , and of the Kac density seen as a functional of the covariance matrix.

Matrix representation of the Kac density

We define the mapping

$$\begin{aligned} \tilde{\rho} : \mathcal{M}_A(\mathbb{R}) \times \mathbb{S}_{2A}^+(\mathbb{R}) &\longrightarrow \mathbb{R} \\ M \times \Sigma &\longmapsto \frac{1}{\sqrt{\det[(2\pi)^2 \Sigma]}} \int_{\mathbb{R}^A} \prod_{a \in A} |(M \underline{x}_A)_a| \exp \left(-\frac{{}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A}{2} \right) d\underline{x}_A. \end{aligned}$$

Recall from definition (5.34) that

$$\Sigma^{\mathcal{I}}(\underline{x}_A) = \text{Cov}(f_{\mathcal{I}}[\underline{x}_A]) \quad \text{and} \quad \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) = \text{Cov}(f_{\mathcal{I}}[\underline{x}_{A,A}]).$$

The following lemma gives an alternative expression of ρ as a function of the covariance matrix $\Sigma^{\mathcal{I}}(\underline{x}_{A,A})$, and the matrix of divided differences $M^{\mathcal{I}}(\underline{x}_A)$ defined in (5.28).

Lemma 5.3.6. For $\underline{x}_A \in \Delta_{\mathcal{I}^+}$,

$$\rho(\underline{x}_A) = \left| \det M^{\mathcal{I}}(\underline{x}_A) \right| \tilde{\rho} \left(M^{\mathcal{I}}(\underline{x}_A), \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) \right).$$

Proof. Note first that Remark 5.3.2 and Lemma 5.3.3 implies that

$$\rho(\underline{x}_A) = \left| \det Q^{\mathcal{I},\bar{A}}(\underline{x}_A) \right| \rho^{\mathcal{I}}(\underline{x}_A) = \left| \det M^{\mathcal{I}}(\underline{x}_A) \right| \rho^{\mathcal{I}}(\underline{x}_A).$$

Let $I \in \mathcal{I}$. In virtue of Lemma 5.2.11, one has

$$f_I^{[\underline{x}_I]}[\underline{x}_I] = (f[\underline{x}_I, x_{i|I}])_{i \in I} \quad \text{and thus} \quad (f_I[\underline{x}_I], f_I^{[\underline{x}_I]}[\underline{x}_I]) = f_{2I}[\underline{x}_{I,I}].$$

Following the notations of Section 5.2.3 it implies that

$$\Sigma^{\mathcal{I}}(\underline{x}_{A,A})^{11} = \text{Cov}(f_{\mathcal{I}}[\underline{x}_A]).$$

From Equation (5.19), one has

$$\det \left[(2\pi)^2 \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) \right] = \det \left[2\pi \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^{11} \right] \det \left[2\pi \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^c \right].$$

Using the alternative expression of $N^{\mathcal{I}}$ given by Lemma 5.3.4 and the conditional formula of Lemma 5.2.3, we deduce

$$\begin{aligned} \rho^{\mathcal{I}}(\underline{x}_A) &= \frac{1}{\sqrt{\det \left[2\pi \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^{11} \right]}} \mathbb{E} \left[\prod_{I \in \mathcal{I}} \prod_{i \in I} \left| (M(\underline{x}_I) f_I^{[\underline{x}_I]}[\underline{x}_I])_i \right| \middle| f_{\mathcal{I}}[\underline{x}_A] = 0 \right] \\ &= \frac{1}{\sqrt{\det \left[(2\pi)^2 \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) \right]}} \int_{\mathbb{R}^A} \prod_{a \in A} \left| (M^{\mathcal{I}}(\underline{x}_A) \underline{x}_A)_a \right| \exp \left(-\frac{{}^T \underline{x}_A (\Sigma^{\mathcal{I}}(\underline{x}_{A,A})^c)^{-1} \underline{x}_A}{2} \right) d\underline{x}_A \\ &= \tilde{\rho} \left(M^{\mathcal{I}}(\underline{x}_A), \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) \right). \end{aligned}$$

The conclusion follows. \square

Lemma 5.3.7. *Let $\underline{x}_A \in \Delta_{\mathcal{I}^+}$. For a subset B of A ,*

$$\rho(\underline{x}_B) = \left| \det M^{\mathcal{I}_B}(\underline{x}_B) \right| \tilde{\rho} \left(M^{\mathcal{I}_B}(\underline{x}_B), Q^{\mathcal{I}, 2B}(\underline{x}_{A,A}) \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^T Q^{\mathcal{I}, 2B}(\underline{x}_{A,A}) \right).$$

Proof. Recall that the partition \mathcal{I}_B of B is the partition induced by the partition \mathcal{I} on the subset B . If $\underline{x}_A \in \Delta_{\mathcal{I}^+}$ then $\underline{x}_B \in \Delta_{\mathcal{I}_B^+}$. We can thus apply the previous Lemma 5.3.6 to get

$$\rho(\underline{x}_B) = \left| \det M^{\mathcal{I}_B}(\underline{x}_B) \right| \tilde{\rho} \left(M^{\mathcal{I}_B}(\underline{x}_B), \Sigma^{\mathcal{I}_B}[\underline{x}_B] \right).$$

From Equation (5.32),

$$f_{\mathcal{I}_B}[\underline{x}_{B,B}] = Q^{\mathcal{I}, 2B}(\underline{x}_{A,A}) f_{\mathcal{I}}[\underline{x}_{A,A}],$$

thus

$$\Sigma^{\mathcal{I}_B}[\underline{x}_B] = Q^{\mathcal{I}, 2B}(\underline{x}_{A,A}) \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^T Q^{\mathcal{I}, 2B}(\underline{x}_{A,A}).$$

□

Given two open subsets Ω_1 and Ω_2 of finite dimensional vector spaces, we define the function space $\mathcal{C}^{0,\infty}(\Omega_1, \Omega_2)$ to be the set of functions from $\Omega_1 \times \Omega_2$ to \mathbb{R} , that are infinitely differentiable with respect to the second argument and such that the partial derivatives (with respect to the second argument) are continuous.

Lemma 5.3.8. *The application $\tilde{\rho}$ belongs to $\mathcal{C}^{0,\infty}(\mathcal{M}_A(\mathbb{R}), \mathbb{S}_{2A}^+(\mathbb{R}))$.*

Proof. Let

$$h(\Sigma, \underline{x}_A) = \frac{1}{\sqrt{\det [(2\pi)^2 \Sigma]}} \exp \left(-\frac{{}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A}{2} \right).$$

The function $\Sigma \mapsto h(\Sigma, \underline{x}_A)$ is infinitely differentiable on $\mathbb{S}_{2A}^+(\mathbb{R})$ and its partial derivatives are also exponentially decreasing with respect to the variable \underline{x}_A . By differentiability under the integral, it implies that $\tilde{\rho}$ belongs to $\mathcal{C}^{0,\infty}(\mathcal{M}_A(\mathbb{R}), \mathbb{S}_{2A}^+(\mathbb{R}))$. □

Factorization of the Kac density and error term

In this section, we show that the function $\tilde{\rho}$ satisfies a nice factorization property. This is a rigorous statement of the approximation (5.10) stated in introduction. For the rest of

this section, \mathcal{I} is a partition of the set A , M is a matrix of $\mathcal{M}_A(\mathbb{R})$ and Σ is a matrix of $\mathbb{S}_{2A}^+(\mathbb{R})$ such that

$$M = M_{\mathcal{I}} \quad \text{and} \quad \Sigma = \Sigma_{\mathcal{I}}.$$

Lemma 5.3.9.

$$\tilde{\rho}(M, \Sigma) = \prod_{I \in \mathcal{I}} \tilde{\rho}(M_I, \Sigma_I).$$

Proof. Since the matrix Σ is block diagonal with respect to the partition \mathcal{I} ,

$$\sqrt{\det [(2\pi)^2 \Sigma]} = \prod_{I \in \mathcal{I}} \sqrt{\det [(2\pi)^2 \Sigma_I]}.$$

Similarly, for $\underline{x}_A \in \mathbb{R}^A$,

$$\exp\left(-\frac{{}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A}{2}\right) = \prod_{I \in \mathcal{I}} \exp\left(-\frac{{}^T \underline{x}_I ((\Sigma_I)^c)^{-1} \underline{x}_I}{2}\right).$$

The matrix $M_{\mathcal{I}}$ is also block diagonal with respect to the partition \mathcal{I} and

$$\prod_{a \in A} |(M \underline{x}_A)_a| = \prod_{I \in \mathcal{I}} \prod_{i \in I} |(M_I \underline{x}_I)_i|.$$

The conclusion follows from the definition of $\tilde{\rho}$. □

We want to describe the error term in Lemma 5.3.9 when the matrix Σ is not block diagonal with respect to the partition \mathcal{I} . We start with the following lemma.

Lemma 5.3.10. *Let H be a matrix of $\mathbb{S}_{2A}(\mathbb{R})$, such that $H_{\mathcal{I}} = 0$. Then*

$$\tilde{\rho}(M, \Sigma + H) = \tilde{\rho}(M, \Sigma) + O(\|H\|^2).$$

Proof. For such matrix H small enough, the matrix $\Sigma + H$ belongs to $\mathbb{S}_{2A}^+(\mathbb{R})$. The matrix Σ is block diagonal, whereas the matrix H satisfies $H_{\mathcal{I}} = 0$. It implies that

$$(\Sigma^{-1} H)_{\mathcal{I}} = (\Sigma^{-1} H \Sigma^{-1})_{\mathcal{I}} = 0 \quad \text{and thus} \quad \text{Tr}(\Sigma^{-1} H) = 0. \quad (5.40)$$

One has from identity (5.20)

$$\begin{aligned} [(\Sigma + H)^c]^{-1} &= [(\Sigma + H)^{-1}]^{22} \\ &= (\Sigma^c)^{-1} + H_\Sigma + O(\|H\|^2), \end{aligned}$$

where $H_\Sigma = [\Sigma^{-1}H\Sigma^{-1}]^{22}$. By (5.40), one has $(H_\Sigma)_\mathcal{I} = 0$. Differentiation under the integral sign gives

$$\tilde{\rho}(M, \Sigma + H) = \tilde{\rho}(M, \Sigma) + d\tilde{\rho}_{(M, \Sigma)}.H + O(\|H\|^2),$$

where

$$\begin{aligned} d\tilde{\rho}_{(M, \Sigma)}.H &= -\frac{1}{2\sqrt{\det [(2\pi)^2\Sigma]}} \int_{\mathbb{R}^A} \prod_{a \in A} |(M\underline{x}_A)_a| \left({}^T \underline{x}_A H_\Sigma \underline{x}_A \right) \exp \left(-\frac{{}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A}{2} \right) d\underline{x}_A \\ &= -\sum_{\substack{i, j \in A \\ [i]_\mathcal{I} \neq [j]_\mathcal{I}}} \frac{(H_\Sigma)_{ij}}{2\sqrt{\det [(2\pi)^2\Sigma]}} \int_{\mathbb{R}^A} \prod_{a \in A} |(M\underline{x}_A)_a| x_i x_j \exp \left(-\frac{{}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A}{2} \right) d\underline{x}_A. \end{aligned}$$

For each i, j of the sum we make the change of variable

$$\forall a \in A, \quad y_a = \begin{cases} -x_a & \text{if } [a]_\mathcal{I} = [i]_\mathcal{I} \\ x_a & \text{if } [a]_\mathcal{I} \neq [i]_\mathcal{I} \end{cases}.$$

Since M and Σ^c are block diagonal matrices, one has for $a \in A$ and $\underline{x}_A \in \mathbb{R}^A$

$$|(M\underline{y}_A)_a| = |(M\underline{x}_A)_a|, \quad {}^T \underline{y}_A (\Sigma^c)^{-1} \underline{y}_A = {}^T \underline{x}_A (\Sigma^c)^{-1} \underline{x}_A \quad \text{but} \quad y_i y_j = -x_i x_j.$$

Thus

$$d\tilde{\rho}_{(M, \Sigma)}.H = -d\tilde{\rho}_{(M, \Sigma)}.H = 0,$$

and the conclusion follows. \square

We can now state the following proposition that gives the error in Lemma 5.3.9 when the matrix Σ is not block diagonal with respect to the partition \mathcal{I} . Note that the following Lemma gives a quadratic error in the matrix coefficients of Σ , where in the somehow analogous Proposition [AL21a, prop. 6.43] only proves a square root error. The difference resides in Lemma 5.3.4, which allows us to bypass the lack of regularity of Gaussian integrals near the boundary of the cone of symmetric definite matrices.

Corollary 5.3.11. *Let B be a subset of A and K be a compact subset of $\mathcal{M}_B(\mathbb{R}) \times \mathcal{M}_{2B,2A}^*(\mathbb{R}) \times \mathbb{S}_{2A}^+(\mathbb{R})$. Then there is a constant C_K such that, for all $(M, Q, \Sigma) \in K$ such that $M = M_{\mathcal{I}}$ and $Q = Q_{\mathcal{I}}$, one has*

$$\left| \tilde{\rho}(M, Q\Sigma^T Q) - \prod_{I \in \mathcal{I}} \tilde{\rho}(M_I, Q_I \Sigma_I^T Q_I) \right| \leq C_K \sup_{\substack{I, J \in \mathcal{I} \\ I \neq J}} \|\Sigma_{I, J}\|^2.$$

Proof. Let $\Pi = Q\Sigma^T Q$. Lemma 5.2.6 asserts that the couple (M, Π) lives in a compact set of $\mathcal{M}_B(\mathbb{R}) \times \mathbb{S}_{2B}^+(\mathbb{R})$. From Lemma 5.3.10, one has

$$|\tilde{\rho}(M, \Pi) - \tilde{\rho}(M, \Pi_{\mathcal{I}})| = O(\|\Pi - \Pi_{\mathcal{I}}\|^2).$$

By Lemma 5.3.8, the application $\tilde{\rho}$ belongs to $\mathcal{C}^{0, \infty}(\mathcal{M}_B(\mathbb{R}) \times \mathbb{S}_{2B}^+(\mathbb{R}))$. Lagrange rest formula asserts the existence of a constant C_K such that

$$|\tilde{\rho}(M, \Pi) - \tilde{\rho}(M, \Pi_{\mathcal{I}})| \leq C_K \|\Pi - \Pi_{\mathcal{I}}\|^2 \leq C_K \sup_{\substack{I, J \in \mathcal{I} \\ I \neq J}} \|\Pi_{I, J}\|^2.$$

Since $Q = Q_{\mathcal{I}}$ and $\|Q\| \leq C_K$ for some constant C_K , we deduce

$$\|\Pi_{I, J}\| = \|Q_I \Sigma_{I, J}^T Q_J\| \leq C_K \|\Sigma_{I, J}\|.$$

Finally the conclusion follows from Lemma 5.3.9 since

$$\tilde{\rho}(M, \Pi_{\mathcal{I}}) = \prod_{I \in \mathcal{I}} \tilde{\rho}(M_I, \Pi_I) = \prod_{I \in \mathcal{I}} \tilde{\rho}(M_I, Q_I \Sigma_I^T Q_I).$$

□

Matrix representation of the cumulant Kac density

Similarly to the Kac density, we can derive a matrix representation for the cumulant Kac density defined in (5.39). Note that the divided differences do not behave well when we consider them on a subfamily of interpolations points $(\underline{x}_A)_{a \in A}$. It explains why we introduced in Paragraph 5.2.3 the somehow complicated set $\widetilde{\mathcal{M}}_A^*(\mathbb{R})$. We introduce the

function \tilde{F}_A defined by

$$\begin{aligned} \tilde{F}_A : \tilde{\mathcal{M}}_A^*(\mathbb{R}) \times \mathbb{S}_{2A}^+(\mathbb{R}) &\longrightarrow \mathbb{R} \\ (\tilde{M}, \tilde{Q}) \times \Sigma &\longrightarrow \sum_{\mathcal{J} \in \mathcal{P}_A} (|\mathcal{J}| - 1)! (-1)^{|\mathcal{J}|-1} \prod_{J \in \mathcal{J}} \left| \det M^J \right| \tilde{\rho}(M^J, Q^J \Sigma^T Q^J). \end{aligned}$$

Let \mathcal{I} be a partition of A . The following lemma gives an alternative expression to the function F_A when $\underline{x}_A \in \Delta_{\mathcal{I}+}$.

Lemma 5.3.12. *For $\underline{x}_A \in \Delta_{\mathcal{I}+}$ one has*

$$F_A(\underline{x}_A) = \tilde{F}_A \left((\tilde{M}^{\mathcal{I}}(\underline{x}_A), \tilde{Q}^{\mathcal{I}}(\underline{x}_A)), \Sigma^{\mathcal{I}}(\underline{x}_{A,A}) \right).$$

Proof. One has

$$F_A(\underline{x}_A) = \sum_{\mathcal{J} \in \mathcal{P}_A} (|\mathcal{J}| - 1)! (-1)^{|\mathcal{J}|-1} \prod_{J \in \mathcal{J}} \rho(\underline{x}_J).$$

According to lemma 5.3.7, for a subset J of A one has

$$\rho(\underline{x}_J) = \left| \det M^{\mathcal{I}J}(\underline{x}_J) \right| \tilde{\rho} \left(M^{\mathcal{I}J}(\underline{x}_J), Q^{\mathcal{I}, 2J}(\underline{x}_{A,A}) \Sigma^{\mathcal{I}}(\underline{x}_{A,A})^T Q^{\mathcal{I}, 2J}(\underline{x}_{A,A}) \right).$$

The first statement follows from the definition (5.33) of $\tilde{M}^{\mathcal{I}}(\underline{x}_A)$ and $\tilde{Q}^{\mathcal{I}}(\underline{x}_A)$. \square

Corollary 5.3.11 translates directly into the following bound for the function \tilde{F}_A .

Lemma 5.3.13. *Let \mathcal{I} be a partition of A , with $\mathcal{I} \neq \{A\}$. Then there is a constant C_K such that for all $((\tilde{M}, \tilde{Q}), \Sigma) \in K$ with $\tilde{M} = \tilde{M}_{\mathcal{I}}$ $\tilde{Q} = \tilde{Q}_{\mathcal{I}}$, and*

$$\forall B \subset A, \forall I \in \mathcal{I}, \quad M_I^B = M^{I \cap B} \quad \text{and} \quad Q_I^B = Q^{I \cap B},$$

one has

$$\left| \tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma) \right| \leq C_K \sup_{\substack{I, J \in \mathcal{I} \\ I \neq J}} \|\Sigma_{I, J}\|^2.$$

Proof. The factorization property of the function $\tilde{\rho}$ given by Lemma 5.3.9 and the cumulant cancellation property given by Lemma 5.2.1 imply that

$$\tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma_{\mathcal{I}}) = 0.$$

The error term given by Corollary 5.3.11 translates directly for the function \tilde{F}_A to the desired estimate. \square

5.3.3 Decay of the cumulant Kac density

The goal of the following section is to improve the quadratic bound given by Lemma 5.3.13. We will do so, thanks to a refinement of Taylor expansion for regular functions that cancel on given affine subspaces. The next key Lemma 5.3.21 bounds the function \tilde{F}_A by a sum over a collection of graphs. We recall first a few definitions and propositions from graph theory.

Graph setting

A (finite) graph G is a couple $(E(G), V(G))$, where $E(G)$ is the set of vertices of the graph G and $V(G)$ the collection of edges of G . For our purposes, a graph G has no loop, but two different edges can have the same endpoints. The multiplicity of an edge $\{a, b\}$ is the number of edges in the graph that are equal to $\{a, b\}$. The edge connectivity of a graph G is the largest integer k such that the graph G remains connected when any subset of $(k - 1)$ edges are removed from the graph G . We define \mathcal{G}_A to be the set of graphs with set of vertices A and edge connectivity 2.

Let \mathcal{I} be a partition of A and let G be a graph with set of vertices A . We define the graph $G_{\mathcal{I}}$ on the set of vertices \mathcal{I} to be the quotient graph with respect to the partition \mathcal{I} . That is, the multiplicity of the edge $\{I, J\}$ (with $I \neq J$) of $G_{\mathcal{I}}$ is the number of edges $\{i, j\}$ in G (with multiplicity) such that $\{I, J\} = \{[i]_{\mathcal{I}}, [j]_{\mathcal{I}}\}$.

Lemma 5.3.14. *Let $H \in \mathcal{G}_{\mathcal{I}}$. There is $G \in \mathcal{G}_A$ such that*

$$H = G_{\mathcal{I}}.$$

Proof. For each $I \in \mathcal{I}$, we replace in H the vertex I by the cycle $(i)_{i \in I}$. The neighbors of I are arbitrary linked to vertices of this cycle. The obtained graph with set of vertices A satisfies $G_{\mathcal{I}} = H$ and will also have edge connectivity 2. \square

An *ear* of a graph G is a path in G such that its internal vertices all have degree two. Note that a cycle is a particular instance of an ear. An *ear decomposition* of the graph

G is a union (P_1, \dots, P_k) such that P_1 is a cycle, and for $i \geq 2$, P_i is an ear such that its endpoints belong to $\cup_{j < i} P_j$. We state the following standard fact for 2-edge connected graphs (see [Whi32]). The proof is a simple induction on the number of ears.

Lemma 5.3.15. *A 2-edge connected graph admits an ear decomposition. The number of ears is necessary the circuit rank of the graph G . Moreover, the starting cycle can be chosen arbitrarily among the cycles of G .*

It implies the following lemma.

Lemma 5.3.16. *Let G be a 2-edge connected graph. There is a family $(T_a)_{a \in A}$ of spanning trees of G such that for every edge $e \in E(G)$, one can find an element $a_e \in A$ such that e is not an edge of the spanning tree T_{a_e} .*

Proof. Let P_1 be a largest cycle in G , with vertices B , and (P_1, \dots, P_k) be an ear decomposition of G . For $i \geq 1$, we define E_i the set of edges of the path P_i . One has $|B| \geq |E_i|$, so that one can find a surjection $\tau_i : B \twoheadrightarrow E_i$.

For $a \notin B$, we define T_a to be any spanning tree of G . For $a \in B$, we define T_a to be the graph G where we removed, in each path E_i , the edge $\tau_i(a)$. The number k is the circuit rank of the graph G . By construction, the graph T_a is connected, so it must be a spanning tree of the graph G . Every edge $e \in E_i$ is the image of some element $a_e \in B$ by the surjection τ_i , so that the edge e does not belong to the tree T_{a_e} . \square

Crossed Taylor formula

In this paragraph we prove an enhancement of the Taylor remainder estimates for regular functions that cancel on affine subspaces. A simple observation of this phenomenon is the following. Assume that $F(x, y)$ is a regular function such that in a neighborhood of zero,

$$|F(x, y)| \leq |x| \quad \text{and} \quad |F(x, y)| \leq |y|.$$

Then for some constant C , one has in a neighborhood of zero that

$$|F(x, y)| \leq C|xy|,$$

which improves by a square factor the trivial bound $\sqrt{|xy|}$. We wish to extend this observation to the more complicated function \tilde{F}_A that satisfies the bounds given by Lemma

5.3.13 for several partitions \mathcal{I} of A . In the following, we give a general statement for this phenomenon.

Let Ω be an open subset of a finite dimensional vector space $V \simeq \mathbb{R}^E$, F an infinitely differentiable function on Ω , and \underline{y}_E be a vector in V . We fix an integer $d \in \mathbb{N}$. The following lemma states equivalent conditions for a regular function F to cancel on an affine subspace with order of cancellation equal to d .

Lemma 5.3.17. *Let B be a subset of E . Then the three following conditions are equivalent.*

1. *For every compact subset K of Ω , there is a constant C_K such that,*

$$\forall \underline{x}_E \in K, \quad |F(\underline{x}_E)| \leq C_K \left(\sup_{b \in B} |x_b - y_b| \right)^d.$$

2. *For every multi-index $\underline{n}_B \in \mathbb{N}^B$ with $|\underline{n}_B| = d$, there exists a function $H_{\underline{n}_B}$ bounded in $C^\infty(\Omega)$ and such that*

$$\forall \underline{x}_E \in \Omega, \quad F(\underline{x}_E) = \sum_{|\underline{n}_B|=d} (\underline{x}_B - \underline{y}_B)^{\underline{n}_B} H_{\underline{n}_B}(\underline{x}_E).$$

3. *For all $\underline{w}_E \in \Omega$ such that $\underline{w}_B = \underline{y}_B$, for every multi-index $\underline{m}_B \in \mathbb{N}^B$ with $|\underline{m}_B| < d$,*

$$\partial^{\underline{m}_B} F(\underline{w}_E) = 0.$$

Proof. We can assume that the Ω is a product of open intervals. The general case follows by a partition of unity. The implication (2) \Rightarrow (1) follows from the boundedness of the functions $H_{\mathcal{C}}$ on any compact subset K of Ω . The implication (1) \Rightarrow (3) is a consequence of the uniqueness of the polynomial approximation given by Taylor expansion. The implication (3) \Rightarrow (2) is a direct consequence of Taylor expansion with integral remainder of the function F on the segment between point \underline{x}_E and $(\underline{x}_{E \setminus B}, \underline{y}_B)$. \square

Now we extend the previous Lemma 5.3.17 to a collection \mathcal{B} of (not necessarily disjoint) subsets of E . For a fix positive integer d we define

$$\mathcal{C}_{\mathcal{B}} = \left\{ \underline{n}_E \in \{0, \dots, d\}^E \mid \forall B \in \mathcal{B}, |\underline{n}_B| \geq d \right\}. \quad (5.41)$$

For instance, if $\mathcal{B} = \{\{1, 2\}, \{2, 3\}, \{1, 3\}\}$ and $d = 2$, then

$$\mathcal{C}_{\mathcal{B}} = \{(2, 2, 0), (0, 2, 2), (2, 0, 2), (1, 1, 1), \dots\}.$$

Proposition 5.3.18. *Assume that for every $B \in \mathcal{B}$, the function F satisfies the equivalent statements of Lemma 5.3.17. Then there exists finitely many non-zero functions $(H_{\underline{n}_E})_{\underline{n}_E \in \mathcal{C}_{\mathcal{B}}}$, bounded in $\mathcal{C}^\infty(\Omega)$ and such that*

$$F(\underline{x}_E) = \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B}}} (\underline{x}_E - \underline{y}_E)^{\underline{n}_E} H_{\underline{n}_E}(\underline{x}_E).$$

Proof. Once again, we can assume that the Ω is a product of open intervals. The proof is an induction on the size of the set \mathcal{B} . If $\mathcal{B} = \{B\}$, this is exactly the hypothesis on F (second characterization in Lemma 5.3.17). Now let $B \in \mathcal{B}$ and suppose that the lemma is true for the family $\mathcal{B} \setminus \{B\}$. We have

$$F(\underline{x}_E) = \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B} \setminus \{B\}}} (\underline{x}_E - \underline{y}_E)^{\underline{n}_E} H_{\underline{n}_E}(\underline{x}_E).$$

Let $\underline{w}_E \in \Omega$ such that $\underline{w}_B = \underline{y}_B$. For every multi-index $\underline{m}_B \in \mathbb{N}^B$ with $|\underline{m}_B| < d$,

$$\begin{aligned} \partial^{\underline{m}_B} F(\underline{w}_E) &= \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B} \setminus \{B\}}} \partial^{\underline{m}_B} \left((\cdot - \underline{y}_E)^{\underline{n}_E} H_{\underline{n}_E}(\cdot) \right) (\underline{w}_E) \\ &= \sum_{\substack{\underline{n}_E \in \mathcal{C}_{\mathcal{B} \setminus \{B\}} \\ \underline{n}_B \leq \underline{m}_B}} (\underline{w}_{E \setminus B} - \underline{y}_{E \setminus B})^{\underline{n}_{E \setminus B}} \frac{\underline{m}_B!}{(\underline{m}_B - \underline{n}_B)!} \partial^{(\underline{m}_B - \underline{n}_B)} H_{\underline{n}_E}(\underline{w}_E) \\ &= 0, \end{aligned}$$

according to the third characterization in Lemma 5.3.17. Let $\underline{w}_E = (\underline{x}_{E \setminus B}, \underline{y}_B)$. For $\underline{n}_E \in \mathcal{C}_{\mathcal{B}}$ and $\underline{x}_E \in \Omega$ we define the quantity

$$\widetilde{H}_{\underline{n}_E}(\underline{x}_E) = H_{\underline{n}_E}(\underline{x}_E) - \sum_{|\underline{m}_B| < d - |\underline{n}_B|} \frac{(\underline{x}_B - \underline{y}_B)^{\underline{m}_B}}{(\underline{m}_B)!} \partial^{(\underline{m}_B)} H_{\underline{n}_E}(\underline{w}_E).$$

If $|\underline{n}_B| \geq d$, then $H_{\underline{n}_E} = \widetilde{H}_{\underline{n}_E}$ and $\underline{n}_E \in \mathcal{C}_{\mathcal{B}}$. If $|\underline{n}_B| < d$, then by Taylor expansion with

integral remainder, there exists functions $(H_{\underline{n}_E + \underline{p}_B})_{|\underline{p}_B| = d - |\underline{n}_B|}$ such that

$$\widetilde{H}_{\underline{n}_E}(\underline{x}_E) = \sum_{|\underline{p}_B| = d - |\underline{n}_B|} (\underline{x}_B - \underline{y}_B)^{\underline{p}_B} H_{\underline{n}_E + \underline{p}_B}(\underline{x}_E).$$

Now we compute

$$\begin{aligned} F(\underline{x}_E) &= F(\underline{x}_E) - \sum_{|\underline{m}_B| < d} \frac{(\underline{x}_B - \underline{y}_B)^{\underline{m}_B}}{\underline{m}_B!} \partial^{\underline{m}_B} F(\underline{w}_E) \\ &= \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B} \setminus \{B\}}} (\underline{x}_E - \underline{y}_E)^{\underline{n}_E} \left(H_{\underline{n}_E}(\underline{x}_E) - \sum_{\substack{|\underline{m}_B| < d \\ \underline{m}_B \geq \underline{n}_B}} \frac{(\underline{x}_B - \underline{y}_B)^{\underline{m}_B - \underline{n}_B}}{(\underline{m}_B - \underline{n}_B)!} \partial^{(\underline{m}_B - \underline{n}_B)} H_{\underline{n}_E}(\underline{w}_E) \right) \\ &= \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B} \setminus \{B\}}} (\underline{x}_E - \underline{y}_E)^{\underline{n}_E} \left(H_{\underline{n}_E}(\underline{x}_E) - \sum_{|\underline{m}_B| < d - |\underline{n}_B|} \frac{(\underline{x}_B - \underline{y}_B)^{\underline{m}_B}}{(\underline{m}_B)!} \partial^{(\underline{m}_B)} H_{\underline{n}_E}(\underline{w}_E) \right) \\ &= \sum_{\underline{n}_E \in (\mathcal{C}_{\mathcal{B} \setminus \{B\}} \cap \mathcal{C}_B)} (\underline{x}_E - \underline{y}_E)^{\underline{n}_E} H_{\underline{n}_E}(\underline{x}_E) + \sum_{\underline{n}_E \in (\mathcal{C}_{\mathcal{B} \setminus \{B\}} \setminus \mathcal{C}_B)} \sum_{|\underline{p}_B| = d - |\underline{n}_B|} (\underline{x}_B - \underline{y}_B)^{\underline{n}_E + \underline{p}_B} H_{\underline{n}_E + \underline{p}_B}(\underline{x}_E) \end{aligned}$$

One then have $|\underline{n}_E + \underline{p}_B| \geq d$ and thus the multi-index $\underline{n}_E + \underline{p}_B$ belongs to $\mathcal{C}_{\mathcal{B}}$ and the conclusion follows. \square

The previous Proposition 5.3.18 directly implies the following corollary.

Corollary 5.3.19. *If the function F satisfies the hypotheses of Proposition 5.3.18, then one can find a constant C_K such that for all \underline{x}_E in K ,*

$$|F(\underline{x}_E)| \leq C_K \sum_{\underline{n}_E \in \mathcal{C}_{\mathcal{B}}} |\underline{x}_E - \underline{y}_E|^{\underline{n}_E}.$$

For instance, let $E = \{1, 2, 3\}$. Let F be an infinitely differentiable function such that for (x, y, z) in any compact subset K of \mathbb{R}^3 ,

$$|F(x, y, z)| \leq x^2 + y^2, \quad |F(x, y, z)| \leq y^2 + z^2 \quad \text{and} \quad |F(x, y, z)| \leq x^2 + z^2.$$

Then the function F satisfies the hypotheses of Proposition 5.3.18 with $\mathcal{B} = \{\{1, 2\}, \{2, 3\}, \{1, 3\}\}$ and $d = 2$. It implies the existence of a constant C_K such that for $(x, y, z) \in K$,

$$|F(x, y, z)| \leq C_K (x^2 y^2 + y^2 z^2 + x^2 z^2 + |xyz|).$$

Remark 5.3.20. Let Ω^1 be an open subset of a finite dimensional vector space and assume that $F \in \mathcal{C}^{0,\infty}(\Omega^1, \Omega)$ (this function space is defined before Lemma 5.3.8). Then Proposition 5.3.18 remains true if one replace $\mathcal{C}^\infty(\Omega)$ by $\mathcal{C}^{0,\infty}(\Omega^1, \Omega)$ and the proof is in all points similar.

We now apply the previous Corollary 5.3.19 to the function \tilde{F}_A .

Lemma 5.3.21. *Let \mathcal{I} be a partition of A and K be a compact subset of $\tilde{\mathcal{M}}_A(\mathbb{R}) \times \mathbb{S}_{2A}^+(\mathbb{R})$. Then there is a constant C_K such that for all $((\tilde{M}, \tilde{Q}), \Sigma) \in K$ with $\tilde{M} = \tilde{M}_{\mathcal{I}}$ $\tilde{Q} = \tilde{Q}_{\mathcal{I}}$, and*

$$\forall B \subset A, \forall I \in \mathcal{I}, \quad M_I^B = M^{I \cap B} \quad \text{and} \quad Q_I^B = Q^{I \cap B},$$

one has

$$|\tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma)| \leq C_K \sum_{G \in \mathcal{G}_{\mathcal{I}}} \prod_{\{I, J\} \in E(G)} \|\Sigma_{I, J}\|.$$

Proof. The proposition is trivial if $\mathcal{I} = \{A\}$, and we can assume that $\mathcal{I} \neq \{A\}$. The proof is an application of Corollary 5.3.19. To this end, we define for subsets B, C of A the set

$$B \circ C = \left\{ \{(k, b), (l, c)\} \mid k, l \in \{1, 2\}, \quad b \in B, c \in C \right\}.$$

Then the set $V = \mathbb{S}_{2A}(\mathbb{R})$, endowed with its canonical basis, can be identified with $\mathbb{R}^{A \circ A}$. For $\mathcal{J} \in \mathcal{P}_A$ with $\mathcal{J} \succeq \mathcal{I}$ we define

$$B_{\mathcal{J}} = \left\{ \{(k, a), (l, b)\} \in A \circ A \mid [a]_{\mathcal{J}} \neq [b]_{\mathcal{J}} \right\},$$

and

$$\mathcal{B}_{\mathcal{I}} = \left\{ B_{\mathcal{J}} \mid \mathcal{I} \preceq \mathcal{J} \prec \{A\} \right\}.$$

As a consequence of Lemma 5.3.8 the function \tilde{F}_A is in $\mathcal{C}^{0,\infty}(\tilde{\mathcal{M}}_A^*(\mathbb{R}), \mathbb{S}_{2A}^+(\mathbb{R}))$. Let \mathcal{J} be a partition such that $\mathcal{I} \preceq \mathcal{J} \prec \{A\}$. According to Lemma 5.3.13 we have

$$\begin{aligned} |\tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma)| &\leq C_K \sup_{\substack{I, J \in \mathcal{J} \\ I \neq J}} \|\Sigma_{I, J}\|^2 \\ &\leq C_K \sup_{\{(k, a), (l, b)\} \in B_{\mathcal{J}}} |\Sigma_{ab}^{kl}|^2. \end{aligned}$$

The function \tilde{F} then satisfies the hypotheses of Proposition 5.3.18 with $d = 2$ and family of subsets $\mathcal{B}_{\mathcal{I}}$, from which we deduce the existence of a constant C_K such that

$$\begin{aligned}
 \left| \tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma) \right| &\leq C_K \sum_{\underline{n} \in \mathcal{C}_{\mathcal{B}_{\mathcal{I}}}} \prod_{\substack{e \in A^\circ \\ e = \{(k,a), (l,b)\}}} |\Sigma^{kl}|^{n_e} \\
 &\leq C_K \sum_{\underline{n} \in \mathcal{C}_{\mathcal{B}_{\mathcal{I}}}} \prod_{\substack{e \in A^\circ \\ e = \{(k,a), (l,b)\}}} \|\Sigma_{[a]_{\mathcal{I}}, [b]_{\mathcal{I}}}\|^{n_e} \\
 &\leq C_K \sum_{\underline{n} \in \mathcal{C}_{\mathcal{B}_{\mathcal{I}}}} \prod_{\substack{I, J \in \mathcal{I} \\ I \neq J}} \|\Sigma_{I, J}\|^{|\underline{n}_{I \circ J}|}. \tag{5.42}
 \end{aligned}$$

To every multi-index $\underline{n} \in \mathcal{C}_{\mathcal{B}_{\mathcal{I}}}$, we can associate a graph $G^{\underline{n}}$ with set of vertices \mathcal{I} , and where the multiplicity of the edge $\{I, J\}$ is given by the number $|\underline{n}_{I \circ J}|$. It follows directly from the definition of the set $\mathcal{C}_{\mathcal{B}_{\mathcal{I}}}$ and the set $\mathcal{G}_{\mathcal{I}}$ that the graph $G^{\underline{n}}$ belongs to the set $\mathcal{G}_{\mathcal{I}}$. Following inequality (5.42), one has

$$\left| \tilde{F}_A((\tilde{M}, \tilde{Q}), \Sigma) \right| \leq C_K \sum_{G \in \mathcal{G}_{\mathcal{I}}} \prod_{\{I, J\} \in E(G)} \|\Sigma_{I, J}\|.$$

□

5.4 Asymptotics of the cumulants of the zeros counting measure

We are now in position to study the asymptotics of the cumulants of the zeros counting measure associated with a sequence of processes $(f_n)_{n \in \mathbb{N}}$. We first prove that the non degeneracy condition (5.35) holds uniformly for $n \in \mathbb{N}$, which allows us to translate the previous Lemma 5.3.21 to the cumulant Kac density $F_{A, n}$ associated with the sequence $(f_n)_{n \in \mathbb{N}}$.

5.4.1 Uniform non degeneracy of the covariance matrix

Up to now, we assumed that the generic process f that we considered satisfied the non degeneracy condition (5.35). For stationary Gaussian processes, this non-degeneracy condition is true on very mild assumptions on the process. For non-stationary processes there seems to be no simple conditions that ensures the validity of (5.35). Nevertheless

in our case of interest, we consider a sequence of Gaussian processes that converges in distribution towards a stationary Gaussian process and we are able to prove some uniform non-degeneracy condition in this setting.

In this subsection, A denotes a finite set and \mathcal{I} a partition of A . For $n \in \bar{\mathbb{N}}$, we consider f_n a Gaussian process defined on nU . We will use the notations introduced in Section 5.3. In particular we consider the quantities $\rho_{A,n}$, $F_{A,n}$, $\Sigma_n^{\mathcal{I}}(\underline{x}_A)$, etc. associated with the process f_n .

We assume for now that the sequence $(f_n)_{n \in \bar{\mathbb{N}}}$ satisfies hypotheses $H_1(q)$ and $H_2(q)$ defined in (5.5) and (5.6), with $q = |A| - 1$. In particular the quantity $\Sigma_n^{\mathcal{I}}(\underline{x}_A) = \text{Cov}((f_n)_{\mathcal{I}}[\underline{x}_A])$ is well defined. Since the function g of hypothesis $H_2(q)$ decreases to zero, then for $\varepsilon > 0$, one can find a constant T_ε such that

$$\{x \in \mathbb{R} \mid g(x) \geq \varepsilon\} \subset [-T_\varepsilon, T_\varepsilon]. \quad (5.43)$$

The main proposition of this section is the following.

Proposition 5.4.1. *Under the above setting, there is a compact set K_η of $\mathbb{S}_A^+(\mathbb{R})$ such that for all $n \in \bar{\mathbb{N}}$ large enough, and $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$, the matrix $\Sigma_n^{\mathcal{I}}(\underline{x}_A)$ lives in K_η .*

We prove first Proposition 5.4.1 for the limit stationary process f_∞ .

Lemma 5.4.2. *Under the above setting, there is a compact set K_η of $\mathbb{S}_A^+(\mathbb{R})$ such that for all $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$, the matrix $\Sigma_\infty^{\mathcal{I}}(\underline{x}_A)$ lives in K_η .*

Proof. According to Lemma 5.2.4, one must show the existence of positive constants C_η and c_η such that

$$\forall \underline{x}_A \in \Delta_{\mathcal{I},\eta}, \quad \|\Sigma_\infty^{\mathcal{I}}(\underline{x}_A)\| \leq C_\eta \quad \text{and} \quad \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) \geq c_\eta.$$

From the Hermite Genocchi formula 5.2.8 and Lemma 5.2.18, we observe that the coefficients of the matrix $\Sigma_\infty^{\mathcal{I}}(\underline{x}_A)$ are bounded by $\|g\|_\infty$. It remains to prove the uniform positiveness of $\det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A)$ on $\Delta_{\mathcal{I},\eta}$.

The covariance function of f_∞ decreases to zero by assumption. Since for Gaussian vectors, decorrelation implies independence, one see that the process f_∞ is weakly mixing,

which in turn implies ergodicity. By Maruyama theorem (see [Mar49]), the spectral measure μ_∞ of f_∞ has no atoms. It is then a standard fact that this observation implies the non-degeneracy condition (5.35), and Lemma 5.3.1 implies that the Gaussian vector $(f_\infty)_\mathcal{I}[\underline{x}_A]$ is also non-degenerate for $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$.

We prove the uniform lower bound for $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$ by induction on the size of the set A . If $|A| = 1$ it reduces to the fact that the process f_∞ is non-degenerate. Assume that the property is true for every strict subset B of A . For another partition \mathcal{J} of A such that $\mathcal{J} \succeq \mathcal{I}$, and $\varepsilon > 0$, we define the subset

$$K_{\mathcal{J},\varepsilon} = \left\{ \underline{x}_A \in U^A \mid [a]_{\mathcal{J}} = [b]_{\mathcal{J}} \Leftrightarrow |x_a - x_b| \leq T_\varepsilon \right\}.$$

We can assume that $T_\varepsilon \geq |A|\eta$. In that case, one has

$$\Delta_{\mathcal{I},\eta} \subset \bigsqcup_{\mathcal{J} \succeq \mathcal{I}} K_{\mathcal{J},\varepsilon}.$$

In the case $\mathcal{J} = \{A\}$, one has for all $a, b \in A$

$$|x_a - x_b| \leq T_\varepsilon.$$

The set $\Delta_{\mathcal{I},\eta} \cap K_{\{A\},\varepsilon}$ is not compact, but it compact by translation in the sens that it is compact if one fixes one of the coordinates. This compactness property plus the stationarity of the process f_∞ implies that one can find a positive constant $c_{\eta,\varepsilon}$ such that

$$\forall \underline{x}_A \in \Delta_{\mathcal{I},\eta} \cap K_{\{A\},\varepsilon}, \quad \det \Sigma^\mathcal{I}(\underline{x}_A) \geq c_{\eta,\varepsilon}.$$

Now assume that $\mathcal{J} \neq \{A\}$. If $\underline{x}_A \in K_{\mathcal{J},\varepsilon}$ then for $a, b \in A$ such that $[a]_{\mathcal{J}} \neq [b]_{\mathcal{J}}$, and $u, v \leq |A| - 1$,

$$|r_\infty^{(u,v)}(x_a - x_b)| \leq \varepsilon.$$

It implies that,

$$\sup_{\substack{I, J \in \mathcal{J} \\ I \neq J}} \|\Sigma_\infty^\mathcal{I}(\underline{x}_A)_{I,J}\| \leq \varepsilon,$$

and thus

$$\|\Sigma_\infty^\mathcal{I}(\underline{x}_A) - (\Sigma_\infty^\mathcal{I}(\underline{x}_A))_{\mathcal{J}}\| \leq \varepsilon.$$

Since the determinant is a smooth function of the matrix coefficients and the matrix

$\Sigma_\infty^{\mathcal{I}}(\underline{x}_A)$ is bounded, we deduce the existence of a constant C_η such that for $\underline{x}_A \in K_{\mathcal{J},\varepsilon}$,

$$\begin{aligned} \left| \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) - \det (\Sigma_\infty^{\mathcal{I}}(\underline{x}_A))_{\mathcal{J}} \right| &= \left| \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) - \prod_{J \in \mathcal{J}} \det \Sigma_\infty^{\mathcal{I}J}(\underline{x}_J) \right| \\ &\leq C_\eta \varepsilon, \end{aligned}$$

and thus

$$\det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) \geq \prod_{J \in \mathcal{J}} \det \Sigma_\infty^{\mathcal{I}J}(\underline{x}_J) - C_\eta \varepsilon.$$

For all $J \in \mathcal{J}$, the set J is a strict subset of A . Moreover, if $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$ then $\underline{x}_J \in \Delta_{\mathcal{I}J,\eta}$. By induction hypothesis, one can find a positive constant c_η depending only on η such that $\det \Sigma_\infty^{\mathcal{I}J}(\underline{x}_J) \geq c_\eta$ when $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$. It implies that

$$\forall \underline{x}_A \in \Delta_{\mathcal{I},\eta} \cap K_{\mathcal{J},\varepsilon}, \quad \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) \geq (c_\eta)^{|\mathcal{J}|} - C\varepsilon.$$

Taking ε small enough and gathering the case $\mathcal{J} = \{A\}$ in and $\mathcal{J} \neq \{A\}$, the conclusion follows. \square

Proof of Proposition 5.4.1.

Proof. A reformulation of hypothesis $H_1(q)$ applied to the compact set $[-T_\varepsilon, T_\varepsilon]$ yields

$$\lim_{n \rightarrow +\infty} \sup_{\substack{s,t \in nU \\ |s-t| \leq T_\varepsilon}} \left| r_n^{(u,v)}(s,t) - \psi\left(\frac{s}{n}\right) r_\infty^{(u,v)}(s,t) \right| = 0.$$

The function ψ is uniformly continuous by hypothesis and we can define ω_ψ its uniform modulus of continuity. By hypothesis, there are positive constants c_ψ and C_ψ such that for all $x \in U$,

$$c_\psi \leq \psi(x) \leq C_\psi. \quad (5.44)$$

Let $n \in \bar{\mathbb{N}}$. If $s, t \in nU$ and $|t-s| > T_\varepsilon$ then hypothesis $H_2(q)$ implies that for $u, v \leq |A|-1$,

$$|r_n^{(u,v)}(s,t)| \leq \varepsilon. \quad (5.45)$$

Gathering (5.44) and (5.45), there is $n_\varepsilon \in \mathbb{N}$ such that for $n \geq n_\varepsilon$, and $s, t \in nU$

$$\left| r_n^{(u,v)}(s,t) - \psi\left(\frac{s}{n}\right) r_\infty^{(u,v)}(s,t) \right| \leq \varepsilon(1 + C_\psi). \quad (5.46)$$

Let $n \in \bar{\mathbb{N}}$ with $n \geq n_\varepsilon$ and $\underline{x}_A \in \Delta_{I,\eta}$. For $I, J \in \mathcal{I}$, $a \in I$ and $b \in J$ one has from Lemma 5.2.18

$$\Sigma_n^{\mathcal{I}}(\underline{x}_A)_{ab} = \int_{C_{a|I} \times C_{b|J}} r_n^{(|a|I|^{-1}, |b|J|^{-1})} \left(\sum_{i \in a|I} s_i x_i, \sum_{j \in b|J} t_j x_j \right) dm(\underline{s}_{a|I}) dm(\underline{t}_{b|J}).$$

Inequality (5.46) implies

$$\left| \Sigma_n^{\mathcal{I}}(\underline{x}_A)_{ab} - \psi \left(\frac{x_a}{n} \right) \Sigma_\infty^{\mathcal{I}}(\underline{x}_A)_{ab} \right| \leq \varepsilon(1 + C_\psi) + R_n(a, b), \quad (5.47)$$

where

$$\begin{aligned} R_n(a, b) &= \|g\|_\infty \int_{C_{a|I} \times C_{b|J}} \left| \psi \left(\frac{x_a}{n} \right) - \psi \left(\frac{\sum_{i \in a|I} s_i x_i}{n} \right) \right| dm(\underline{s}_{a|I}) dm(\underline{t}_{b|J}) \\ &\leq C \sup_{\underline{s}_{a|I} \in C_{a|I}} \left| \psi \left(\frac{x_a}{n} \right) - \psi \left(\frac{\sum_{i \in a|I} s_i x_i}{n} \right) \right|. \end{aligned}$$

For $i \in I$, one has $|x_a - x_i| \leq |A|\eta$. It implies that for any convex combination y of the variables $(x_i)_{i \in I}$ one also have $|x_a - y| \leq |A|\eta$. We deduce that

$$R_n(a, b) \leq C\omega_\psi \left(\frac{|A|\eta}{n} \right).$$

There is $n_{\eta,\varepsilon}$ such that for $n \geq n_{\eta,\varepsilon}$,

$$R_n(a, b) \leq \varepsilon,$$

and thus coming back to inequality (5.47),

$$\left| \Sigma_n^{\mathcal{I}}(\underline{x}_A)_{ab} - \psi \left(\frac{x_a}{n} \right) \Sigma_\infty^{\mathcal{I}}(\underline{x}_A)_{ab} \right| \leq \varepsilon(2 + C_\psi). \quad (5.48)$$

It implies the existence of a constant C_η such that for n large enough and $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$,

$$\left| \det \Sigma_n^{\mathcal{I}}(\underline{x}_A) - \left(\prod_{a \in A} \psi \left(\frac{x_a}{n} \right) \right) \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) \right| \leq C_\eta \varepsilon.$$

We deduce that

$$\det \Sigma_n^{\mathcal{I}}(\underline{x}_A) \geq c_\psi^{|A|} \det \Sigma_\infty^{\mathcal{I}}(\underline{x}_A) - C_\eta \varepsilon.$$

The conclusion follows from the previous Lemma 5.4.2 covering the stationary case, and taking ε small enough. \square

As a consequence of the previous Proposition 5.4.1, we deduce the following corollary about convergence of the Kac density associated with the process f_n .

Corollary 5.4.3. *Assume that the sequence of processes $(f_n)_{n \in \bar{\mathbb{N}}}$ satisfies hypotheses $H_1(q)$ and $H_2(q)$ defined in (5.5) and (5.6), with $q = 2|A| - 1$. Then there is a compact set K_η of $\mathbb{S}_{2A}^+(\mathbb{R})$ such that for all $n \in \bar{\mathbb{N}}$ large enough and $\underline{x}_A \in \Delta_{\mathcal{I}, \eta}$,*

$$\Sigma_n^{\mathcal{I}}(\underline{x}_{A,A}) \in K_\eta.$$

In particular we have the following convergence, uniformly for $x \in U$ and \underline{t}_A in compact subsets of \mathbb{R}^A

$$\lim_{n \rightarrow +\infty} \rho_n(nx + \underline{t}_A) = \rho_\infty(\underline{t}_A) \quad \text{and} \quad \lim_{n \rightarrow +\infty} F_{A,n}(nx + \underline{t}_A) = F_{A,\infty}(\underline{t}_A).$$

Proof. The first assertion is a direct application of Proposition 5.4.1 with the set $2A$, using the fact that $\underline{x}_{A,A} \in \Delta_{\mathcal{I}, \eta}$ when $\underline{x}_A \in \Delta_{\mathcal{I}, \eta}$. As for the second one, the proof of Proposition 5.4.1, and in particular equation (5.48), implies that for all partition \mathcal{I} of A , one has the following convergence, uniformly for $x \in U$ and \underline{t}_A in a compact subset of $\Delta_{\mathcal{I}, \eta}$

$$\lim_{n \rightarrow +\infty} \Sigma_n^{\mathcal{I}}(nx + \underline{t}_{A,A}) = \psi(x) \Sigma_\infty^{\mathcal{I}}(\underline{t}_{A,A}).$$

The conclusion follows from the alternative expression for ρ_n given by Lemma 5.3.6. Note that the function ρ_∞ does not depend on the function $\psi(x)$, by a change of variable. \square

5.4.2 Asymptotics of the cumulants

Let A be a finite set of cardinal p . We assume that the sequence of processes $(f_n)_{n \in \bar{\mathbb{N}}}$ satisfies hypotheses $H_1(q)$ and $H_2(q)$ defined in (5.5) and (5.6), with $q = 2p - 1$. We then choose $\eta = \frac{\varepsilon}{2p}$ where ε is the parameter of hypothesis $H_2(q)$, so that

$$g_\varepsilon(x) = \sup_{|u| \leq 2\eta p} g(x + u).$$

Decay of the cumulant Kac density

Let us now translate Lemma 5.3.21 to the cumulant Kac density $F_{A,n}$. The previous Corollary 5.4.3 ensures that the matrix $\Sigma^{\mathcal{I}}(\underline{x}_{A,A})$ lives in a compact subset of $\mathbb{S}_{2A}^+(\mathbb{R})$ when $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$ and n is large enough.

Lemma 5.4.4. *There is a constant C such that for all $\underline{x}_A \in (nU)^A$,*

$$|F_{A,n}(\underline{x}_A)| \leq C \sum_{G \in \mathcal{G}_A} \prod_{\{a,b\} \in E(G)} g_\varepsilon(x_a - x_b).$$

Proof. Let \mathcal{I} be a partition of A and $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$. According to Corollary 5.4.3, the matrix $\Sigma_n^{\mathcal{I}}(\underline{x}_{A,A})$, for n large enough depending only on η , lives in a compact subset of $\mathbb{S}_{2A}^+(\mathbb{R})$ depending only on the parameter η . By Proposition 5.2.17, the element $(\widetilde{M}^{\mathcal{I}}(\underline{x}_A), \widetilde{Q}^{\mathcal{I}}(\underline{x}_A))$ also lives in a compact subset of $\widetilde{\mathcal{M}}_A^*(\mathbb{R})$ that depends only on η . We can then apply Lemma 5.3.21 with $\Sigma = \Sigma_n^{\mathcal{I}}(\underline{x}_{A,A})$ and $(\widetilde{M}, \widetilde{Q}) = (\widetilde{M}^{\mathcal{I}}(\underline{x}_A), \widetilde{Q}^{\mathcal{I}}(\underline{x}_A))$. Given the representation formula for F_A given by Lemma 5.3.12, we deduce the existence of a constant C_η such that for all $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$,

$$|F_A(\underline{x}_A)| \leq C_\eta \sum_{G \in \mathcal{G}_{\mathcal{I}}} \prod_{\{I,J\} \in E(G)} \|(\Sigma_n^{\mathcal{I}}(\underline{x}_{A,A}))_{I,J}\|.$$

Let $H \in \mathcal{G}_{\mathcal{I}}$. According to Lemma 5.3.14, there is a graph $G \in \mathcal{G}_A$ such that $G_{\mathcal{I}} = H$. If we remove the edges $\{a, b\}$ of G such that $[a]_{\mathcal{I}} = [b]_{\mathcal{I}}$, then there is a bijection between the edges of G and the edges of H given by the mapping

$$\{a, b\} \longrightarrow \{[a]_{\mathcal{I}}, [b]_{\mathcal{I}}\}.$$

Let $\{a, b\}$ an be edge of the graph G .

— If $[a]_{\mathcal{I}} = [b]_{\mathcal{I}}$ then $|x_a - x_b| \leq A\eta$. We deduce that

$$0 < r_\infty(0) \leq g(0) \leq g_\varepsilon(x_a - x_b).$$

— If $[a]_{\mathcal{I}} \neq [b]_{\mathcal{I}}$ then from the Hermite-Genocchi formula and Lemma 5.2.18,

$$\|(\Sigma_n^{\mathcal{I}}(\underline{x}_{A,A}))_{I,J}\| \leq \sup_{|s| \leq 2\eta p} g(x_a - x_b + s) \leq g_\varepsilon(x_a - x_b).$$

We deduce that

$$\begin{aligned} \prod_{\{I,J\} \in E(H)} \|(\sum_n^{\mathcal{I}}(\underline{x}_{A,A}))_{I,J}\| &\leq \prod_{\substack{\{a,b\} \in E(G) \\ [a]_{\mathcal{I}} \neq [b]_{\mathcal{I}}}} g_\varepsilon(x_a - x_b) \\ &\leq C \prod_{\{a,b\} \in E(G)} g_\varepsilon(x_a - x_b). \end{aligned}$$

We deduce the existence of a constant C_η such that for $\underline{x}_A \in \Delta_{\mathcal{I},\eta}$,

$$|F_A(\underline{x}_A)| \leq C_\eta \sum_{G \in \mathcal{G}_A} \prod_{\{a,b\} \in E(G)} g_\varepsilon(x_a - x_b).$$

The inequality is true for every partition \mathcal{I} of A and the conclusion follows. \square

Convergence of the error term towards zero

Recall from Definition (5.4) that ν_n is the random counting measure of the zero set of the Gaussian process $f_n(n \cdot)$ defined on U . The previous Lemma 5.4.4 and the formula for the p -th cumulant given by Proposition 5.3.5 shows that the convergence of the cumulant reduces to the behavior of the quantity

$$I_n(G) = \int_{(nU)^A} \prod_{a \in A} \left| \phi_a \left(\frac{x_a}{n} \right) \right| \prod_{e=\{i,j\} \in E(G)} g_e(x_i - x_j) d\underline{x}_A, \quad (5.49)$$

where G is a 2-edge connected graph with set of vertices A and set of edges $E(G)$, $(\phi_a)_{a \in A}$ are functions in $L^{(\infty)}(U)$ and $(g_e)_{e \in E(G)}$ are even functions in $L^2 \cap L^\infty$.

The quantity $I_n(G)$, in the context of cumulants asymptotics, is somehow reminiscent of a theorem of Szegő (see [ALS10] and the references therein), where this kind of integral received a thorough treatment and Hölder bounds that depends on the structure of the graph were given. Nevertheless our setting is not exactly the same, and we were able to give a simple and self contained argument, that relies only on a basic interpolation inequality for Hölder norms, which proves an Hölder type bound for the quantity $I_n(G)$. The result is contained in the following lemma. Recall that $p = |A|$.

Lemma 5.4.5. *Assume that for all $e \in E(G)$, $g_e \in L^{\frac{p}{p-1}}$. Then for every $e \in E(G)$, there is a number $\alpha_e \geq p/(p-1)$ such that*

$$\frac{1}{n} |I_n(G)| \leq \left(\prod_{a \in A} \|\phi_a\|_p \right) \left(\prod_{e \in E(G)} \|g_e\|_{\alpha_e} \right).$$

Assume that $p \geq 3$ and $g_e \in L^2 \cap L^\infty$. Then

$$\lim_{n \rightarrow +\infty} \frac{1}{n^{p/2}} I_n(G) = 0.$$

Proof. Let $(T_a)_{a \in A}$ be the family of spanning trees of G given by Lemma 5.3.16. For fixed index $a \in A$, the linear mapping

$$\underline{x}_A \longmapsto (x_a, (x_b - x_c)_{\{b,c\} \in E(T_a)})$$

is volume preserving. For $e \notin T_a$ we bound the term $g_e(x_i - x_j)$ in $I_n(G)$ by $\|g_e\|_\infty$, and for $b \neq a$, the function ϕ_b by $\|\phi_b\|_\infty$ (this quantity may be infinite). By a change of variable, we get

$$I_n(G) \leq n \|\phi_a\|_1 \left(\prod_{b \neq a} \|\phi_b\|_\infty \right) \left(\prod_{e \in T_a} \|g_e\|_1 \right) \left(\prod_{e \notin T_a} \|g_e\|_\infty \right).$$

Since this inequality is true for all $a \in A$, one can interpolate these inequalities (see for instance [Ben+08, Thm. 5.5]) to obtain

$$I_n(G) \leq n \left(\prod_{a \in A} \|\phi_a\|_p \right) \left(\prod_{e \in E(G)} \|g_e\|_{p_e} \right), \quad \text{with} \quad \frac{1}{p_e} = \frac{1}{p} \sum_{a \in A} \mathbf{1}_{E(T_a)}(e). \quad (5.50)$$

Since for all $e \in E(G)$, there is a tree T_{a_e} that does not contain the edge e , one must have $p_e \geq \frac{p}{p-1}$, and the first part of the lemma follows. For the second part, note that we have the crude bound

$$|I_n(G)| \leq n^p \left(\prod_{a \in A} \|\phi_a\|_1 \right) \left(\prod_{e \in E(G)} \|g_e\|_\infty \right).$$

We interpolate this inequality with inequality (5.50) to get

$$\frac{1}{n^{p/2}} |I_n(G)| \leq \left(\prod_{a \in A} \|\phi_a\|_2 \right) \left(\prod_{e \in E(G)} \|g_e\|_{q_e} \right), \quad \text{with } q_e \geq 2. \quad (5.51)$$

It remains to show that the left hand side of (5.51) converges towards zero for $p \geq 3$. The Hölder bound given by (5.50) implies that by a density argument, we can assume that the functions $(g_e)_{e \in E(G)}$ are bounded and compactly supported. In that case, inequality (5.50) implies the convergence towards zero of the left hand side of (5.51) when $p \geq 3$. \square

In the following, $(\phi_a)_{a \in A}$ is a family of functions in $L^{(\infty)}(U)$. The previous Lemma 5.4.5 and the convergence of the Kac density given by Corollary 5.4.3 imply the following lemma.

Lemma 5.4.6. *For all $p \geq 3$,*

$$\lim_{n \rightarrow +\infty} \frac{1}{n^{p/2}} \int_{(nU)^A} \phi_A^{\otimes} \left(\frac{\underline{x}_A}{n} \right) F_{A,n}(\underline{x}_A) d\underline{x}_A = 0.$$

For all $p \geq 1$, if $g_\varepsilon \in L^{\frac{p}{p-1}}$ then

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \int_{(nU)^A} \phi_A^{\otimes} \left(\frac{\underline{x}_A}{n} \right) F_{A,n}(\underline{x}_A) d\underline{x}_A = \left(\int_U \prod_{a \in A} \phi_a(y) dy \right) \left(\int_{\mathbb{R}^{p-1}} F_{A,\infty}(0, \underline{x}) d\underline{x} \right).$$

Proof. According to Lemma 5.4.4, there is a constant C such that

$$\left| \int_{(nU)^A} \phi_A^{\otimes} \left(\frac{\underline{x}_A}{n} \right) F_{A,n}(\underline{x}_A) d\underline{x}_A \right| \leq C \sum_{G \in \mathcal{G}_A} I_n(G),$$

where $I_n(G)$ is defined in (5.49) with functions $g_e = g_\varepsilon$. The first part of the corollary is an immediate consequence of the first part of Lemma 5.4.5. As for the second part, the bound given by Lemma 5.4.5 implies that one can assume by a density argument that the functions $(\phi_a)_{a \in A}$ are continuous and compactly supported. In that case, pick $a_0 \in A$. We define $y_{a_0} = 0$ and we make the change of variables

$$x_{a_0} = ny \quad \text{and} \quad \forall a \in A \setminus \{a_0\}, \quad x_a = ny + y_a.$$

Then we have the following uniform convergence

$$\lim_{n \rightarrow +\infty} \phi_a \left(y + \frac{y_a}{n} \right) = \phi_a(y),$$

and according to Corollary 5.4.3,

$$\lim_{n \rightarrow +\infty} F_{A,n} \left(ny + \underline{y}_A \right) = F_{A,\infty}(\underline{y}_A),$$

from which the conclusion follows. □

Given the expression of cumulants given by Proposition 5.3.5 and the previous Lemma 5.4.6, we then deduce the following theorem concerning the convergence of cumulants associated with the linear statistics of the zeros counting measure of the sequence of processes $(f_n)_{n \in \bar{\mathbb{N}}}$. We define the Stirling number of the second kind

$$\left\{ \begin{matrix} p \\ k \end{matrix} \right\} := \text{Card} \{ \mathcal{I} \in \mathcal{P}_A \mid |\mathcal{I}| = k \}.$$

Theorem 5.4.7. *Let $p \geq 2$ and assume that the sequence of processes $(f_n)_{n \in \bar{\mathbb{N}}}$ satisfies the hypotheses $H_1(q)$ and $H_2(q)$ with $q = 2p - 1$. If $p \geq 3$ then*

$$\lim_{n \rightarrow +\infty} \frac{1}{n^{p/2}} \kappa_p(\langle \nu_n, \phi \rangle) = 0.$$

Moreover when $g_\varepsilon \in L^{\frac{p}{p-1}}$,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \kappa_p(\langle \nu_n, \phi \rangle) = \left(\int_U \phi^p(y) dy \right) \sum_{k=1}^p \left\{ \begin{matrix} p \\ k \end{matrix} \right\} \left(\int_{\mathbb{R}^{k-1}} F_{k,\infty}(0, \underline{x}) d\underline{x} \right).$$

Proof. Let $p \geq 3$. Recall from Proposition 5.3.5 that

$$\kappa_p(\langle \nu_n, \phi \rangle) = \sum_{\mathcal{I} \in \mathcal{P}_A} \int_{(nU)^\mathcal{I}} \left(\prod_{I \in \mathcal{I}} \phi \left(\frac{x_I}{n} \right)^{|I|} \right) F_{\mathcal{I},n}(\underline{x}_\mathcal{I}) d\underline{x}_\mathcal{I}.$$

According to the previous Lemma 5.4.6, one has

$$\frac{1}{n^{p/2}} |\kappa_p(\langle \nu, \phi \rangle)| \leq \sum_{\mathcal{I} \in \mathcal{P}_A} \frac{1}{n^{p/2}} \left| \int_{(nU)^{\mathcal{I}}} \left(\prod_{I \in \mathcal{I}} \phi \left(\frac{x_I}{n} \right)^{|I|} \right) F_{\mathcal{I},n}(x_{\mathcal{I}}) dx_{\mathcal{I}} \right| \xrightarrow{n \rightarrow +\infty} 0,$$

which proves the first assertion. As for the second assertion, it is again a consequence of Lemma 5.4.6, which implies that

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \kappa_p(\langle \nu_n, \phi \rangle) = \left(\int_U \phi^p(y) dy \right) \sum_{\mathcal{I} \in \mathcal{P}_A} \left(\int_{\mathbb{R}^{|\mathcal{I}|-1}} F_{|\mathcal{I}|,\infty}(0, \underline{x}) d\underline{x} \right).$$

□

The proof of the main Theorem 5.1.6 is a reformulation of the previous Theorem 5.4.7, with

$$\forall p \geq 1, \quad \gamma_p = \sum_{k=1}^p \left\{ \begin{matrix} p \\ k \end{matrix} \right\} \left(\int_{\mathbb{R}^{k-1}} F_{k,\infty}(\underline{x}) d\underline{x} \right).$$

In particular, one has

$$\gamma_1 = \frac{1}{\pi} \sqrt{\frac{-r''_{\infty}(0)}{r_{\infty}(0)}} \quad \text{and} \quad \gamma_2 = \gamma_1 + \int_{\mathbb{R}} F_{2,\infty}(0, x) dx.$$

It has been shown for instance in [Lac20] that under our assumptions on the process f_{∞} , the constant γ_2 is positive, from which follows the central limit theorem for the linear statistic associated with the zeros counting measure.

BIBLIOGRAPHY

- [ADL16] Jean-Marc Azaïs, Federico Dalmao, and José R. León, « CLT for the zeros of classical random trigonometric polynomials », in: *Ann. Inst. Henri Poincaré Probab. Stat.* 52.2 (2016), pp. 804–820, ISSN: 0246-0203, DOI: 10.1214/14-AIHP653, URL: <https://doi.org/10.1214/14-AIHP653>.
- [ADP19] Jürgen Angst, Federico Dalmao, and Guillaume Poly, « On the real zeros of random trigonometric polynomials with dependent coefficients », in: *Proc. Amer. Math. Soc.* 147.1 (2019), pp. 205–214, ISSN: 0002-9939, DOI: 10.1090/proc/14216, URL: <https://doi.org/10.1090/proc/14216>.
- [ÁF07] J. C. Álvarez Paiva and E. Fernandes, « Gelfand transforms and Crofton formulas », in: *Selecta Math. (N.S.)* 13.3 (2007), pp. 369–390, ISSN: 1022-1824, DOI: 10.1007/s00029-007-0045-5, URL: <https://doi.org/10.1007/s00029-007-0045-5>.
- [AL13] Jean-Marc Azaïs and José León, « CLT for crossings of random trigonometric polynomials », in: *Electronic Journal of Probability* 18.none (2013), pp. 1–17, DOI: 10.1214/EJP.v18-2403, URL: <https://doi.org/10.1214/EJP.v18-2403>.
- [AL21a] Michele Ancona and Thomas Letendre, « Roots of Kostlan polynomials: moments, strong law of large numbers and central limit theorem », in: *Ann. H. Lebesgue* 4 (2021), pp. 1659–1703, DOI: 10.5802/ahl.113, URL: <https://doi.org/10.5802/ahl.113>.
- [AL21b] Michele Ancona and Thomas Letendre, « Zeros of smooth stationary Gaussian processes », in: *Electron. J. Probab.* 26 (2021), Paper No. 68, 81, DOI: 10.1214/21-ejp637, URL: <https://doi.org/10.1214/21-ejp637>.
- [ALS10] Florin Avram, Nikolai Leonenko, and Ludmila Sakhno, « On a Szegő type limit theorem, the Hölder-Young-Brascamp-Lieb inequality, and the asymptotic theory of integrals and quadratic forms of stationary fields », in: *ESAIM Probab. Stat.* 14 (2010), pp. 210–255, ISSN: 1292-8100, DOI: 10.1051/ps:2008031, URL: <https://doi.org/10.1051/ps:2008031>.

-
- [AP21] Jürgen Angst and Guillaume Poly, « Variations on Salem–Zygmund results for random trigonometric polynomials: application to almost sure nodal asymptotics », in: *Electronic Journal of Probability* 26 (2021), pp. 1–36.
- [APP18] Jürgen Angst, Viet-Hung Pham, and Guillaume Poly, « Universality of the nodal length of bivariate random trigonometric polynomials », in: *Trans. Amer. Math. Soc.* 370.12 (2018), pp. 8331–8357, ISSN: 0002-9947, DOI: 10.1090/tran/7255, URL: <https://doi.org/10.1090/tran/7255>.
- [APP21] Jürgen Angst, Thibault Pautrel, and Guillaume Poly, « Real zeros of random trigonometric polynomials with dependent coefficients », in: (2021), URL: <https://arxiv.org/abs/2102.09653>.
- [Arm+19] Diego Armentano et al., « On the finiteness of the moments of the measure of level sets of random fields », 2019, URL: <https://arxiv.org/pdf/1909.10243v1>, submitted.
- [Arm+21] D. Armentano et al., « Central limit theorem for the number of real roots of Kostlan Shub Smale random polynomial systems », in: *Amer. J. Math.* 143.4 (2021), pp. 1011–1042, ISSN: 0002-9327, DOI: 10.1353/ajm.2021.0026, URL: <https://doi.org/10.1353/ajm.2021.0026>.
- [Arr+17] Benjamin Arras et al., « A new approach to the Stein-Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence », in: (2017), URL: <https://arxiv.org/abs/1605.06819>.
- [AW09] Jean-Marc Azaïs and Mario Wschebor, *Level sets and extrema of random processes and fields*, John Wiley & Sons, Inc., Hoboken, NJ, 2009, pp. xii+393, ISBN: 978-0-470-40933-6, DOI: 10.1002/9780470434642, URL: <https://doi.org/10.1002/9780470434642>.
- [BCP19] Vlad Bally, Lucia Caramellino, and Guillaume Poly, « Non universality for the variance of the number of real roots of random trigonometric polynomials », in: *Probab. Theory Related Fields* 174.3-4 (2019), pp. 887–927, ISSN: 0178-8051, DOI: 10.1007/s00440-018-0869-2, URL: <https://doi.org/10.1007/s00440-018-0869-2>.
- [Bel66] Yu. K. Belyaev, « On the Number of Intersections of a Level by a Gaussian Stochastic Process. I », in: *Theory of Probability & Its Applications* 11.1 (1966),

-
- pp. 106–113, DOI: 10.1137/1111006, eprint: <https://doi.org/10.1137/1111006>, URL: <https://doi.org/10.1137/1111006>.
- [Ben+08] Jonathan Bennett et al., « The Brascamp–Lieb inequalities: finiteness, structure and extremals », in: *Geometric and Functional Analysis* 17.5 (2008), pp. 1343–1415.
- [Ber02] M. V. Berry, « Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature », in: *J. Phys. A* 35.13 (2002), pp. 3025–3038, ISSN: 0305-4470, DOI: 10.1088/0305-4470/35/13/301, URL: <https://doi.org/10.1088/0305-4470/35/13/301>.
- [Ber77] M. V. Berry, « Regular and irregular semiclassical wavefunctions », in: *J. Phys. A* 10.12 (1977), pp. 2083–2091, ISSN: 0305-4470, URL: <http://stacks.iop.org/0305-4470/10/2083>.
- [Bér77] Pierre H. Bérard, « On the wave equation on a compact Riemannian manifold without conjugate points », in: *Math. Z.* 155.3 (1977), pp. 249–276, ISSN: 0025-5874, DOI: 10.1007/BF02028444, URL: <https://doi.org/10.1007/BF02028444>.
- [Bér85] Pierre Bérard, « Volume des ensembles nodaux des fonctions propres du laplacien », in: *Séminaire de théorie spectrale et géométrie* 3 (1985), pp. 1–9.
- [BG17] Vincent Beffara and Damien Gayet, « Percolation of random nodal lines », in: *Publ. Math. Inst. Hautes Études Sci.* 126 (2017), pp. 131–176, ISSN: 0073-8301, DOI: 10.1007/s10240-017-0093-0, URL: <https://doi.org/10.1007/s10240-017-0093-0>.
- [Bil95] Patrick Billingsley, *Probability and measure*, Third, Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1995, pp. xiv+593, ISBN: 0-471-00710-2.
- [BK13] D. Beliaev and Z. Kereta, « On the Bogomolny-Schmit conjecture », in: *J. Phys. A* 46.45 (2013), pp. 455003, 5, ISSN: 1751-8113, DOI: 10.1088/1751-8113/46/45/455003, URL: <https://doi.org/10.1088/1751-8113/46/45/455003>.
- [BM19] Jacques Benatar and Riccardo W. Maffucci, « Random waves on \mathbb{T}^3 : nodal area variance and lattice point correlations », in: *Int. Math. Res. Not. IMRN* 10 (2019), pp. 3032–3075, ISSN: 1073-7928, DOI: 10.1093/imrn/rnx220, URL: <https://doi.org/10.1093/imrn/rnx220>.

-
- [Boo05] Carl de Boor, « Divided differences », in: *Surv. Approx. Theory* 1 (2005), pp. 46–69.
- [Bou14] Jean Bourgain, « On toral eigenfunctions and the random wave model », in: *Israel J. Math.* 201.2 (2014), pp. 611–630, ISSN: 0021-2172, DOI: 10.1007/s11856-014-1037-z, URL: <https://doi.org/10.1007/s11856-014-1037-z>.
- [BS02] Eugene Bogomolny and Charles Schmit, « Percolation model for nodal domains of chaotic wave functions », in: *Physical Review Letters* 88.11 (2002), p. 114102.
- [BS07] Eugene Bogomolny and Charles Schmit, « Random wavefunctions and percolation », in: *Journal of Physics A: Mathematical and Theoretical* 40.47 (2007), p. 14033.
- [BS83] Richard A. Brualdi and Hans Schneider, « Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley », in: *Linear Algebra Appl.* 52/53 (1983), pp. 769–791, ISSN: 0024-3795, DOI: 10.1016/0024-3795(83)80049-4, URL: [https://doi.org/10.1016/0024-3795\(83\)80049-4](https://doi.org/10.1016/0024-3795(83)80049-4).
- [BW16] Jeremiah Buckley and Igor Wigman, « On the number of nodal domains of toral eigenfunctions », in: *Ann. Henri Poincaré* 17.11 (2016), pp. 3027–3062, ISSN: 1424-0637, DOI: 10.1007/s00023-016-0476-7, URL: <https://doi.org/10.1007/s00023-016-0476-7>.
- [BW18] Dmitry Beliaev and Igor Wigman, « Volume distribution of nodal domains of random band-limited functions », in: *Probab. Theory Related Fields* 172.1-2 (2018), pp. 453–492, ISSN: 0178-8051, DOI: 10.1007/s00440-017-0813-x, URL: <https://doi.org/10.1007/s00440-017-0813-x>.
- [BYY19] Bartłomiej Błaszczyszyn, Dhandapani Yogeshwaran, and Joseph E Yukich, « Limit theory for geometric statistics of point processes having fast decay of correlations », in: *The Annals of Probability* 47.2 (2019), pp. 835–895.
- [Cam19] Valentina Cammarota, « Nodal area distribution for arithmetic random waves », in: *Trans. Amer. Math. Soc.* 372.5 (2019), pp. 3539–3564, ISSN: 0002-9947, DOI: 10.1090/tran/7779, URL: <https://doi.org/10.1090/tran/7779>.

-
- [CH18] Yaiza Canzani and Boris Hanin, « C^∞ scaling asymptotics for the spectral projector of the Laplacian », in: *J. Geom. Anal.* 28.1 (2018), pp. 111–122, ISSN: 1050-6926, DOI: 10.1007/s12220-017-9812-5, URL: <https://doi.org/10.1007/s12220-017-9812-5>.
- [CH20] Yaiza Canzani and Boris Hanin, « Local universality for zeros and critical points of monochromatic random waves », English, in: *Commun. Math. Phys.* 378.3 (2020), pp. 1677–1712, ISSN: 0010-3616; 1432-0916/e.
- [CMR18] Simon Campese, Domenico Marinucci, and Maurizia Rossi, « Approximate normality of high-energy hyperspherical eigenfunctions », in: *J. Math. Anal. Appl.* 461.1 (2018), pp. 500–522, ISSN: 0022-247X, DOI: 10.1016/j.jmaa.2017.11.051, URL: <https://doi.org/10.1016/j.jmaa.2017.11.051>.
- [Cuz75] Jack Cuzick, « Conditions for finite moments of the number of zero crossings for Gaussian processes », in: *Ann. Probability* 3.5 (1975), pp. 849–858, ISSN: 0091-1798, DOI: 10.1214/aop/1176996271, URL: <https://doi.org/10.1214/aop/1176996271>.
- [Cuz76] Jack Cuzick, « A Central Limit Theorem for the Number of Zeros of a Stationary Gaussian Process », in: *The Annals of Probability* 4.4 (1976), pp. 547–556, ISSN: 00911798, URL: <http://www.jstor.org/stable/2243040>.
- [DF88] Harold Donnelly and Charles Fefferman, « Nodal sets of eigenfunctions on Riemannian manifolds », in: *Invent. Math.* 93.1 (1988), pp. 161–183, ISSN: 0020-9910, DOI: 10.1007/BF01393691, URL: <https://doi.org/10.1007/BF01393691>.
- [Die+20] Gauthier Dierickx et al., « Small scale CLTs for the nodal length of monochromatic waves », in: *arXiv preprint arXiv:2005.06577* (2020).
- [DNN20] Yen Do, Hoi Nguyen, and Oanh Nguyen, « Random trigonometric polynomials: universality and non-universality of the variance for the number of real roots », in: (2020), URL: <https://arxiv.org/abs/1912.11901>.
- [Do+21] Yen Do et al., *Central Limit Theorem for the number of real roots of random orthogonal polynomials*, 2021, arXiv: 2111.09015 [math.PR].

-
- [DS90] William Duke and Rainer Schulze-Pillot, « Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids », in: *Invent. Math.* 99.1 (1990), pp. 49–57, ISSN: 0020-9910, DOI: 10.1007/BF01234411, URL: <https://doi.org/10.1007/BF01234411>.
- [Fed69] Herbert Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969, pp. xiv+676.
- [Fla17] Hendrik Flasche, « Expected number of real roots of random trigonometric polynomials », in: *Stochastic Process. Appl.* 127.12 (2017), pp. 3928–3942, ISSN: 0304-4149, DOI: 10.1016/j.spa.2017.03.018, URL: <https://doi.org/10.1016/j.spa.2017.03.018>.
- [Gas21a] Louis Gass, « Almost-sure asymptotics for Riemannian random waves », in: *To appear in Bernoulli Journal* (2021), URL: <https://arxiv.org/abs/2005.06389>.
- [Gas21b] Louis Gass, *Variance of the number of zeros of dependent Gaussian trigonometric polynomials*, Mar. 2021, arXiv: 2103.08002 [math.PR].
- [Gas21c] Louis Gass, *Cumulants asymptotics for the zeros counting measure of real Gaussian processes*, 2021, arXiv: 2112.08247 [math.PR].
- [Gem72] Donald Geman, « On the variance of the number of zeros of a stationary Gaussian process », in: *Ann. Math. Statist.* 43 (1972), pp. 977–982, ISSN: 0003-4851, DOI: 10.1214/aoms/1177692560, URL: <https://doi.org/10.1214/aoms/1177692560>.
- [GW11] Andrew Granville and Igor Wigman, « The distribution of the zeros of random trigonometric polynomials », in: *Amer. J. Math.* 133.2 (2011), pp. 295–357, ISSN: 0002-9327, DOI: 10.1353/ajm.2011.0015, URL: <https://doi.org/10.1353/ajm.2011.0015>.
- [GW17] Damien Gayet and Jean-Yves Welschinger, « Betti numbers of random nodal sets of elliptic pseudo-differential operators », in: *Asian J. Math.* 21.5 (2017), pp. 811–839, ISSN: 1093-6106, DOI: 10.4310/AJM.2017.v21.n5.a2, URL: <https://doi.org/10.4310/AJM.2017.v21.n5.a2>.

-
- [Hei01] Juha Heinonen, *Lectures on analysis on metric spaces*, Universitext, Springer-Verlag, New York, 2001, pp. x+140, ISBN: 0-387-95104-0, DOI: 10.1007/978-1-4613-0131-8, URL: <https://doi.org/10.1007/978-1-4613-0131-8>.
- [HK78] Ernst Heintze and Hermann Karcher, « A general comparison theorem with applications to volume estimates for submanifolds », in: *Ann. Sci. École Norm. Sup. (4)* 11.4 (1978), pp. 451–470, ISSN: 0012-9593, URL: http://www.numdam.org/item?id=ASENS_1978_4_11_4_451_0.
- [Hör68] Lars Hörmander, « The spectral function of an elliptic operator », in: *Acta Math.* 121 (1968), pp. 193–218, ISSN: 0001-5962, DOI: 10.1007/BF02391913, URL: <https://doi.org/10.1007/BF02391913>.
- [Hux96] M. N. Huxley, *Area, lattice points, and exponential sums*, vol. 13, London Mathematical Society Monographs. New Series, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1996, pp. xii+494, ISBN: 0-19-853466-3.
- [Ing21] Maxime Ingremeau, « Local weak limits of Laplace eigenfunctions », in: *Tunis. J. Math.* 3.3 (2021), pp. 481–515, ISSN: 2576-7658, DOI: 10.2140/tunis.2021.3.481, URL: <https://doi.org/10.2140/tunis.2021.3.481>.
- [Ivi+06] A. Ivić et al., « Lattice points in large regions and related arithmetic functions: recent developments in a very classic topic », in: *Elementare und analytische Zahlentheorie*, vol. 20, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main, Franz Steiner Verlag Stuttgart, Stuttgart, 2006, pp. 89–128.
- [Ivr16] Victor Ivrii, « 100 years of Weyl’s law », in: *Bull. Math. Sci.* 6.3 (2016), pp. 379–452, ISSN: 1664-3607, DOI: 10.1007/s13373-016-0089-y, URL: <https://doi.org/10.1007/s13373-016-0089-y>.
- [Ivr80] V. Ja. Ivrii, « The second term of the spectral asymptotics for a Laplace-Beltrami operator on manifolds with boundary », in: *Funktsional. Anal. i Prilozhen.* 14.2 (1980), pp. 25–34, ISSN: 0374-1990.
- [Jub19] Benoît Jubin, « Intrinsic volumes of sublevel sets », 2019, URL: <https://arxiv.org/abs/1903.01592>, submitted.

-
- [Kac43] M. Kac, « On the average number of real roots of a random algebraic equation », in: *Bull. Amer. Math. Soc.* 49 (1943), pp. 314–320, ISSN: 0002-9904, DOI: 10.1090/S0002-9904-1943-07912-8, URL: <https://doi.org/10.1090/S0002-9904-1943-07912-8>.
- [Kee19] Blake Keeler, « A logarithmic improvement in the two-point Weyl law for manifolds without conjugate points », in: *arXiv preprint arXiv:1905.05136* (2019).
- [KKW13] Manjunath Krishnapur, Pär Kurlberg, and Igor Wigman, « Nodal length fluctuations for arithmetic random waves », in: *Ann. of Math. (2)* 177.2 (2013), pp. 699–737, ISSN: 0003-486X, DOI: 10.4007/annals.2013.177.2.8, URL: <https://doi.org/10.4007/annals.2013.177.2.8>.
- [KSW21] Zakhar Kabluchko, Andrea Sartori, and Igor Wigman, « Expected nodal volume for non-Gaussian random band-limited functions », in: *arXiv preprint arXiv:2102.11689* (2021).
- [Kun97] Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, vol. 24, Cambridge Studies in Advanced Mathematics, Reprint of the 1990 original, Cambridge University Press, Cambridge, 1997, pp. xiv+346, ISBN: 0-521-35050-6; 0-521-59925-3.
- [Lac20] Raphaël Lachieze-Ray, *Variance linearity for real Gaussian zeros*, 2020, arXiv: 006.10341 [math.PR].
- [Let16] Thomas Letendre, « Expected volume and Euler characteristic of random submanifolds », in: *J. Funct. Anal.* 270.8 (2016), pp. 3047–3110, ISSN: 0022-1236, DOI: 10.1016/j.jfa.2016.01.007, URL: <https://doi.org/10.1016/j.jfa.2016.01.007>.
- [LM16] Alexander Logunov and Eugenia Malinnikova, « Ratios of harmonic functions with the same zero set », in: *Geom. Funct. Anal.* 26.3 (2016), pp. 909–925, ISSN: 1016-443X, DOI: 10.1007/s00039-016-0369-4, URL: <https://doi.org/10.1007/s00039-016-0369-4>.
- [LM20] Alexander Logunov and Eugenia Malinnikova, « Review of Yau’s conjecture on zero sets of Laplace eigenfunctions », in: *Current developments in mathematics 2018*, Int. Press, Somerville, MA, [2020] ©2020, pp. 179–212.

-
- [Log18a] Alexander Logunov, « Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure », in: *Ann. of Math. (2)* 187.1 (2018), pp. 221–239, ISSN: 0003-486X, DOI: 10.4007/annals.2018.187.1.4, URL: <https://doi.org/10.4007/annals.2018.187.1.4>.
- [Log18b] Alexander Logunov, « Nodal sets of Laplace eigenfunctions: proof of Nadi-rashvili’s conjecture and of the lower bound in Yau’s conjecture », in: *Ann. of Math. (2)* 187.1 (2018), pp. 241–262, ISSN: 0003-486X, DOI: 10.4007/annals.2018.187.1.5, URL: <https://doi.org/10.4007/annals.2018.187.1.5>.
- [LP21] Doron S. Lubinsky and Igor E. Pritsker, « Variance of real zeros of random orthogonal polynomials », in: *J. Math. Anal. Appl.* 498.1 (2021), p. 124954, ISSN: 0022-247X, DOI: 10.1016/j.jmaa.2021.124954, URL: <https://doi.org/10.1016/j.jmaa.2021.124954>.
- [Lub09] Doron S. Lubinsky, « A new approach to universality limits involving orthogonal polynomials », in: *Ann. of Math. (2)* 170.2 (2009), pp. 915–939, ISSN: 0003-486X, DOI: 10.4007/annals.2009.170.915, URL: <https://doi.org/10.4007/annals.2009.170.915>.
- [Mar+16] Domenico Marinucci et al., « Non-universality of nodal length distribution for arithmetic random waves », in: *Geom. Funct. Anal.* 26.3 (2016), pp. 926–960, ISSN: 1016-443X, DOI: 10.1007/s00039-016-0376-5, URL: <https://doi.org/10.1007/s00039-016-0376-5>.
- [Mar49] Gisiro Maruyama, « The harmonic analysis of stationary stochastic processes », in: *Mem. Fac. Sci. Kyūsyū Univ. A* 4 (1949), pp. 45–106, ISSN: 0373-6385.
- [Mor98] Mitsuo Morimoto, *Analytic functionals on the sphere*, vol. 178, Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1998, pp. xii+160, ISBN: 0-8218-0585-1, DOI: 10.1090/mmono/178, URL: <https://doi.org/10.1090/mmono/178>.
- [MRW20] Domenico Marinucci, Maurizia Rossi, and Igor Wigman, « The asymptotic equivalence of the sample trispectrum and the nodal length for random spherical harmonics », in: *Ann. Inst. Henri Poincaré Probab. Stat.* 56.1 (2020), pp. 374–390, ISSN: 0246-0203, DOI: 10.1214/19-AIHP964, URL: <https://doi.org/10.1214/19-AIHP964>.

-
- [Non08] St'ephane Nonnenmacher, « Some open questions in "wave chaos" », in: *arXiv: Chaotic Dynamics* (2008).
- [NP12] Ivan Nourdin and Giovanni Peccati, *Normal approximations with Malliavin calculus*, vol. 192, Cambridge Tracts in Mathematics, From Stein's method to universality, Cambridge University Press, Cambridge, 2012, pp. xiv+239, ISBN: 978-1-107-01777-1, DOI: 10.1017/CB09781139084659, URL: <https://doi.org/10.1017/CB09781139084659>.
- [NPR19] Ivan Nourdin, Giovanni Peccati, and Maurizia Rossi, « Nodal statistics of planar random waves », in: *Comm. Math. Phys.* 369.1 (2019), pp. 99–151, ISSN: 0010-3616, DOI: 10.1007/s00220-019-03432-5, URL: <https://doi.org/10.1007/s00220-019-03432-5>.
- [NS09] Fedor Nazarov and Mikhail Sodin, « On the number of nodal domains of random spherical harmonics », in: *Amer. J. Math.* 131.5 (2009), pp. 1337–1357, ISSN: 0002-9327, DOI: 10.1353/ajm.0.0070, URL: <https://doi.org/10.1353/ajm.0.0070>.
- [NV21] Oanh Nguyen and Van Vu, « Random polynomials: Central limit theorems for the real roots », in: *Duke Math. J.* 170.17 (2021), pp. 3745–3813, ISSN: 0012-7094, DOI: 10.1215/00127094-2020-0089, URL: <https://doi.org/10.1215/00127094-2020-0089>.
- [NZ19] Hoi Nguyen and Ofer Zeitouni, *Exponential concentration for the number of roots of random trigonometric polynomials*, 2019, arXiv: 1912.12051 [math.PR].
- [Pau20] Thibault Pautrel, « New asymptotics for the mean number of zeros of random trigonometric polynomials with strongly dependent Gaussian coefficients », in: *Electronic Communications in Probability* 25 (2020), pp. 1–13.
- [PT11] Giovanni Peccati and Murad S. Taqqu, *Wiener chaos: moments, cumulants and diagrams*, vol. 1, Bocconi & Springer Series, A survey with computer implementation, Supplementary material available online, Springer, Milan; Bocconi University Press, Milan, 2011, pp. xiv+274, ISBN: 978-88-470-1678-1, DOI: 10.1007/978-88-470-1679-8, URL: <https://doi.org/10.1007/978-88-470-1679-8>.

-
- [Ric45] S. O. Rice, « Mathematical analysis of random noise », in: *Bell System Tech. J.* 24 (1945), pp. 46–156, ISSN: 0005-8580, DOI: 10.1002/j.1538-7305.1945.tb00453.x, URL: <https://doi.org/10.1002/j.1538-7305.1945.tb00453.x>.
- [Ros19] Maurizia Rossi, « Random nodal lengths and Wiener chaos », in: *Probabilistic methods in geometry, topology and spectral theory*, vol. 739, Contemp. Math. Amer. Math. Soc., Providence, RI, 2019, pp. 155–169, DOI: 10.1090/conm/739/14898, URL: <https://doi.org/10.1090/conm/739/14898>.
- [RW08] Zeév Rudnick and Igor Wigman, « On the volume of nodal sets for eigenfunctions of the Laplacian on the torus », in: *Ann. Henri Poincaré* 9.1 (2008), pp. 109–130, ISSN: 1424-0637, DOI: 10.1007/s00023-007-0352-6, URL: <https://doi.org/10.1007/s00023-007-0352-6>.
- [Slu91] Eric Slud, « Multiple Wiener-Itô integral expansions for level-crossing-count functionals », in: *Probab. Theory Related Fields* 87.3 (1991), pp. 349–364, ISSN: 0178-8051, DOI: 10.1007/BF01312215, URL: <https://doi.org/10.1007/BF01312215>.
- [Slu94] Eric V. Slud, « MWI representation of the number of curve-crossings by a differentiable Gaussian process, with applications », in: *Ann. Probab.* 22.3 (1994), pp. 1355–1380, ISSN: 0091-1798, URL: [http://links.jstor.org/sici?sici=0091-1798\(199407\)22:3%3C1355:MROTNO%3E2.0.CO;2-N&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(199407)22:3%3C1355:MROTNO%3E2.0.CO;2-N&origin=MSN).
- [Spe83] T. P. Speed, « Cumulants and partition lattices », in: *Austral. J. Statist.* 25.2 (1983), pp. 378–388, ISSN: 0004-9581.
- [SW16] Peter Sarnak and Igor Wigman, « Topologies of nodal sets of random band limited functions », in: *Advances in the theory of automorphic forms and their L-functions*, vol. 664, Contemp. Math. Amer. Math. Soc., Providence, RI, 2016, pp. 351–365, DOI: 10.1090/conm/664/13040, URL: <https://doi.org/10.1090/conm/664/13040>.
- [SZ54] R. Salem and A. Zygmund, « Some properties of trigonometric series whose terms have random signs », in: *Acta Math.* 91 (1954), pp. 245–301, ISSN: 0001-5962, DOI: 10.1007/BF02393433, URL: <https://doi.org/10.1007/BF02393433>.

-
- [Sze75] Gábor Szegő, *Orthogonal polynomials*, Fourth, American Mathematical Society Colloquium Publications, Vol. XXIII, American Mathematical Society, Providence, R.I., 1975, pp. xiii+432.
- [Tot00] Vilmos Totik, « Asymptotics for Christoffel functions for general measures on the real line », in: *J. Anal. Math.* 81 (2000), pp. 283–303, ISSN: 0021-7670, DOI: 10.1007/BF02788993, URL: <https://doi.org/10.1007/BF02788993>.
- [Wey11] H. Weyl, « Ueber die asymptotische Verteilung der Eigenwerte », in: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1911 (1911), pp. 110–117.
- [Whi32] Hassler Whitney, « Non-separable and planar graphs », in: *Trans. Amer. Math. Soc.* 34.2 (1932), pp. 339–362, ISSN: 0002-9947, DOI: 10.2307/1989545, URL: <https://doi.org/10.2307/1989545>.
- [Wig09] Igor Wigman, « On the distribution of the nodal sets of random spherical harmonics », in: *J. Math. Phys.* 50.1 (2009), pp. 013521, 44, ISSN: 0022-2488, DOI: 10.1063/1.3056589, URL: <https://doi.org/10.1063/1.3056589>.
- [Yau82] S. Yau, *Seminar on Differential Geometry*, Annals of Mathematics Studies, Princeton University Press, 1982, ISBN: 9780691082967, URL: <https://books.google.fr/books?id=Hm9fLdZGQY8C>.
- [Ylv68] Donald Ylvisaker, « A note on the absence of tangencies in Gaussian sample paths », in: *Ann. Math. Statist.* 39 (1968), pp. 261–262, ISSN: 0003-4851, DOI: 10.1214/aoms/1177698528, URL: <https://doi.org/10.1214/aoms/1177698528>.
- [Zel09] Steve Zelditch, « Real and complex zeros of Riemannian random waves », in: *Spectral analysis in geometry and number theory*, vol. 484, Contemp. Math. Amer. Math. Soc., Providence, RI, 2009, pp. 321–342, DOI: 10.1090/conm/484/09482, URL: <https://doi.org/10.1090/conm/484/09482>.
- [Zel10] Steve Zelditch, « Recent developments in mathematical quantum chaos », in: *Current developments in mathematics, 2009*, Int. Press, Somerville, MA, 2010, pp. 115–204.
- [Zel97] Steve Zelditch, « Fine structure of Zoll spectra », in: *J. Funct. Anal.* 143.2 (1997), pp. 415–460, ISSN: 0022-1236, DOI: 10.1006/jfan.1996.2981, URL: <https://doi.org/10.1006/jfan.1996.2981>.

Titre : Sur la géométrie des ensembles nodaux aléatoires

Mot clés : Mathématiques ; Probabilités ; Géométrie aléatoire

Résumé : Dans cette thèse, nous nous intéressons aux ensembles nodaux aléatoires, c'est-à-dire au lieu d'annulation de fonctions à valeurs réelles, dépendantes également d'un paramètre aléatoire. Notre principal modèle d'intérêt est celui des ondes aléatoires riemanniennes. L'étude de ce modèle, ainsi que de son ensemble nodal, est motivé par de célèbres conjectures physiques et mathématiques, comme la conjecture de Berry. Nous montrons dans une première partie un résultat de convergence presque-sûre concernant la mesure nodale, dans la limite de haute énergie. Ce résultat se base sur l'approche Salem–Zygmund, et vient renforcer les résultats de convergence en moyenne connus pour la mesure nodale.

En dimension un, l'étude des ondes aléatoires riemanniennes se réduit à celle des zéros des polynômes trigonométriques aléatoires. Elle s'inscrit dans la théorie plus générale des zéros de processus stochastiques unidimensionnels, elle aussi riche en applications : télécommunication, traitement du signal, etc. Nous montrons dans une seconde partie l'asymptotique exacte des moments pour la mesure de comptage aléatoire d'une large classe de processus gaussiens, et en particulier pour le modèle des polynômes trigonométriques aléatoires. Ce résultat, basé sur une analyse fine de la combinatoire liée aux zéros d'un processus gaussien, vient unifier et renforcer les théorèmes centraux limites déjà existants pour ces différents modèles.

Title: On the geometry of random nodal sets

Keywords: Mathematics ; Probability ; Random geometry

Abstract: In this thesis, we are interested in the nodal sets of random functions, that is in the vanishing locus of real functions, also depending on a random parameter. Our main model of interest is the so-called Riemannian random waves model. It is motivated by celebrated conjectures in both physics and mathematics, such as the Berry conjecture. We show in a first part the almost sure convergence of the nodal measure, in the high energy limit. This result is based on the Salem–Zygmund approach, and reinforces the previously known convergence in mean for the nodal measure.

In dimension one, this study reduces to the study of the zeros set of random trigonometric polynomials. It is a part of the more general theory of zeros of one dimensional stochastic processes, also rich in applications: telecommunication, signal processing, etc. We show in a second part the exact asymptotics of moments for the random counting measure of a large class of Gaussian processes, and in particular for the model of random trigonometric polynomials. This result is based on a fine analysis of the combinatorics related to the zeros of a Gaussian process. It unifies and strengthens the already existing central limit theorems for these different models.