Topology of random algebraic and analytic hypersurfaces

Damien Gayet
Master 2 lectures - 2018-2019
Université Grenoble Alpes

December 13, 2019

Abstract

These lecture notes are an expanded version of courses I gave for graduate students at the Université de Grenoble in the second semester of 2018-2019. I present some old and recent results on the topology and geometry of real random algebraic hypersurfaces, beginning with the average of the number of real roots of real polynomials of one variable. The measure is mainly the very natural so-called Kostlan or complex Fubini-Study measure. In higher dimensions I compute the average of the volume of the vanishing locus of a random polynomial of degree d in the real projective space, and of a random linear sum of eigenfunctions of the Laplacian over a compact Riemannian manifold. Then, I explain that every topology arises with uniform positive probability as part of the random hypersurface of degree d in any prescribed ball $B(x, 1/\sqrt{d})$. Finally, I present a link between percolation and the vanishing locus of random Gaussian analytic functions on the real plane, if the measure is the local rescaled measure given by the Kostlan polynomials. I tried to give the whole proofs. I did not try to give general results for Gaussian fields or even random functions, but I try to separate what is general and what is proper to the algebraic world.

Contents

Chapter 1

Introduction

Roughly speaking, algebraic geometry is the study of the vanishing locus of polynomials. The general question is

Question 1.0.1 *Can we describe for a given degree d, the possible geometries or topologies of the vanishing locus of all degree d polynomials?*

In this course, we will study algebraic geometry for fields equal to $\mathbb{R}$ or $\mathbb{C}$, and the ambient space will be either the affine spaces $\mathbb{R}^n$ or $\mathbb{C}^n$, or the projective spaces $\mathbb{R}P^n$ or $\mathbb{C}P^n$. The affine spaces are more intuitive, however the projective ones are much natural, since the space is compact and have simple symmetries.

1.1 Deterministic answers

1.1.1 Generalities

Notation: for any $n \geq 1$, any space M and any mapping $f : M \to \mathbb{R}^n$, let

$$Z(f) := f^{-1}(0).$$

Recall the following fundamental result in differential geometry:

Proposition 1.1.1 *([20, p. 29]) Let M^m a m-dimensional manifold and $f : M \to \mathbb{R}^n$ a C^1 function, such that*

$$\forall x \in Z(f),\ df(x) : T_xM \to \mathbb{R}^n \text{ is onto.}$$

Then $Z(f)$ is a submanifold of codimension n in M.

In particular, if $m = n$, $Z(f)$ is a discrete number of points without any accumulation. If moreover M is compact, then $Z(f)$ is the union of a finite number of points.

1.1.2 Roots (n=1)

In the complex case, the Gauss theorem asserts that the topological situation for $Z_{\mathbb{C}}(p) \subset \mathbb{C}$ is simple, where $p : \mathbb{C} \to \mathbb{C}$ is a polynomial:

Theorem 1.1.2 (Friedrich Gauss) *For $d \geq 1$, a degree d polynomial in one variable complex variable has always d roots (with multiplicity).*

Remark 1.1.3 1. *For any set Z of d distinct points, there exists a polynomial $p \in \mathbb{C}[z]$, such that $Z(p) = Z$.*

2. *When $p \in \mathbb{R}[x]$, the topology of $Z_{\mathbb{R}}(p)$, that is its number of points, is more... complex: for any $d \geq 1$, p can have between 0 (if d is even) or 1 (if d is odd) and d real roots, and every situation can be achieved.*

1.1.3 Homogeneous polynomials and projective spaces

Before going on with $n = 2$, we urge the reader not familiar with the projective spaces to read in the Annex the associated reminder 5.2. In projective spaces, we will deal with *homogeneous* polynomials

$$P \in \mathbb{K}^d_{hom}[X_0, \cdots X_n] := \{P \in \mathbb{K}[X_0, \cdots X_n] \mid \forall \lambda \in \mathbb{K}, X \in \mathbb{K}^{n+1}, P(\lambda X) = \lambda^d P(X).\}$$

This is easy to see that the set of degree $d$homogeneous monomials $\{X_0^{i_0} \cdots X_n^{i_n}, i_0 + \cdots i_n = d\}$ is a basis of $\mathbb{K}^d_{hom}[X_0, \cdots X_n]$. Why do we use homogeneous polynomials? Because their vanishing locus has a sense in the projective spaces, since

$$\forall t \in \mathbb{K}^*, P(X) = 0 \Leftrightarrow P(tX) = 0,$$

so that we can define $Z(P) \subset \mathbb{K}P^n$.

Note that $\mathbb{K}P^n \setminus \{X_0 = 0\}$ (removing one line) is diffeomorphic to $\mathbb{K}^n$ by the map

$$[X_0 : \cdots X_n] \mapsto (\frac{X_1}{X_0}, \cdots, \frac{X_n}{X_0}),$$

and in these coordinates on this chart, an homogeneous polynomial can be canonically transformed into a polynomial in n variables by

$$p(x_1, \cdots, x_n) := P(1 : X_1 : \cdots : X_n).$$

We see that $Z(P) \cap \{X_0 = 0\}^c \subset \mathbb{K}P^n$ is homeomorphic to $Z(p) \subset \mathbb{K}^n$. Conversely, given $p(x) \in \mathbb{R}[x_1, \cdots, x_n]$ and any degree d, we can *homogenize* p into

$$P(X_0, \cdots, X_n) := X_0^d p(\frac{X_1}{X_0}, \cdots, \frac{X_n}{X_0}).$$

By Proposition 1.1.1, if $P \in \mathbb{R}^d_{hom}$ (resp. $P \in \mathbb{C}^d_{hom}$) is a submersion on its vanishing locus, then $Z(P) \subset \mathbb{R}P^n$ (resp. $Z(P) \subset \mathbb{C}P^n$) is a (resp. complex) submanifold of dimension $n-1$, that is an *hypersuface.*

1.1.4 Algebraic curves (n=2)

In this case, when P is not singular and $\mathbb{K} = \mathbb{R}$, that is when $Z(P)$ is a submanifold of $\mathbb{R}P^2$, it is a finite number of topological circles, since a compact smooth manifold of dimension one is homeomorphic to a circle. These circles are its connected components. In $\mathbb{R}^2$, the connected components of $Z(p)$ can be topological lines (hence, going ton infinity) or circles.

1. $d = 1$ case: $Z(p)$ is a straight line in $\mathbb{R}^2$, so has one component. In $\mathbb{R}P^2$, by a linear change of variables, we can assume that $P = X_0$, so that its vanishing locus is one topological circle.

2. $d = 2$. In $\mathbb{R}^2$, $Z(p)$ is a *conic*, so has between 0 and two components. In $\mathbb{R}P^2$, there is 0 or one component. In general, we denote:

$$b_0(P) := \#\{ \text{ connected components of } Z(P)\}.$$

The first general result in higher dimensions (that is $n \geq 2$) was proved in 1876:

Theorem 1.1.4 *(**Axel Harnack Theorem 1876 [15], [9, Theorem 11.6.2]**) For any degree $d \geq 1$,*

1. *$b_0 \leq b_{max}(d) := \frac{1}{2}(d-1)(d-2)+1$.*
2. *For any d, there exists P such that $b_0(P) = b_{max}(d)$.*

In complex, we have a different situation:

Theorem 1.1.5 *For any degree $d \geq 1$, if $P \in \mathbb{C}^2_{hom}[X_0, \cdots, X_2]$ is not singular, then $Z(P) \subset \mathbb{C}P^2$ is a connected compact surface of genus $b_{max}(d) - 1$.*

The latter theorem says that the whole topology of the complex vanishing locus $Z(P) \subset \mathbb{C}P^2$ is always the same, on the contrary to the $Z(P) \subset \mathbb{R}P^2$ when P is a real polynomial.

In $\mathbb{R}P^2$, 16th David Hilbert famous problem is the following:

Question 1.1.6 *(Hilbert 1900) Describe relative positions of ovals originating from a real algebraic curve [...].*

The ovals are the connected components of $Z(P)$. For instance, there cannot be more than $d/2$ components encircling each other. Indeed, by Bézout's theorem, $\#(Z(P) \cap Z(Q)) \leq deg(P)deg(Q)$, so that $Z(P)$ intersect at most d times a line. However the number of intersections of a line passing by the inner of the smallest oval must intersect $Z(P)$ at least 2 times the number of circles, so that $N \leq d/2$.

1.1.5 Higher dimensions.

Theorem 1.1.7 *([37, p. 263], see also [24, Corollary 3]) For any real homogeneous polynomial P of degree d in $n+1$ real variables,*

$$b_0(Z(P)) \leq \frac{d}{d-1}(d^n - 1).$$

More precisely, $\sum_i b_i(Z(P), \mathbb{R}) \leq d(2d-1)^n$ where b_i is the $i-th$ Betti number, that is $b_i = \dim H_i(Z(P), \mathbb{R})$. In fact,

$$\sum_i b_i(Z_{\mathbb{C}}(P), \mathbb{R}) \leq \sum_i b_i(Z_{\mathbb{C}}(P), \mathbb{R}).$$

Again, the complex situation is far more clearer:

Theorem 1.1.8 *([37, p. 264]), [14, p. 157] The complex hypersurface $Z(P) \in \mathbb{C}P^n$ is always connected, and always simply connected for $n \geq 3$. In fact any two smooth complex hypersurfaces $Z(P)$ and $Z(Q)$ are isotopic, so that they are homeomorphic. The sum of the Betti numbers is a polynomial in d of degree n.*

The reader can check that these bounds are coherent with the Gauss and Harnack inequalities. For $n = 3$, $Z(P)$ is a union of real compact surfaces. The former bound implies that

$$(2b_0 + b_1)(Z_{\mathbb{R}}(P)) \leq 8d^3.$$

In particular, $b_0 \leq 4d^3$.

We have an equality versus a bound for the volume of $Z(P)$.

Theorem 1.1.9 *For any homogeneous polynomial of degree d,*

1. *if P is complex then $Vol Z_{\mathbb{C}}(P) = d$ (Wirtinger) ;*
2. *if P is real then $Vol Z_{\mathbb{R}}(P) \leq d\frac{Vol(\mathbb{R}P^n)}{Vol(\mathbb{C}P^n)}$ (Cauchy-Crofton).*

1.2 Probabilistic questions

We sum up the results exposed in the latter section.

1. in the complex projective case, the topology and the volume of $Z_{\mathbb{C}}(P)$ depend only on the degree of the polynomial, *not on the particular polynomial itself*, as long as P vanishes transversally;
2. in the real projective case, already with the $d = 2$ case, we see that the answer for the topology or the volume depend on the particular polynomial.
3. However, we saw that there is an upper bound for the Betti numbers (resp. the volume) of $Z_{\mathbb{R}}(P)$ by the the complex hypersurface $Z_{\mathbb{C}}(P)$.

In conclusion,

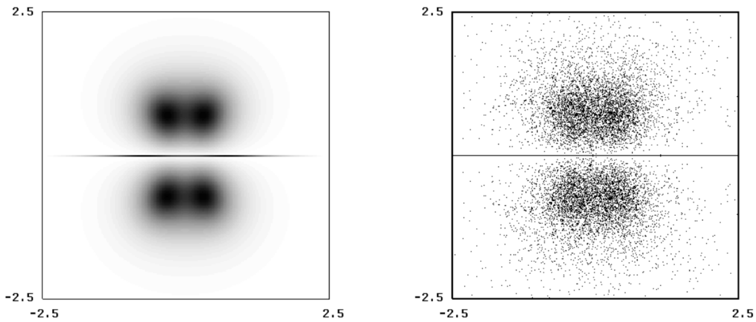
1. as far is the global topology (and even the global volume) of $Z(P)$ is concerned, there is no need to take at random polynomials.
2. On the contrary, It is very natural to take real polynomials at random and look at the statistics of the topology of $Z_{\mathbb{R}}(P)$, beginning with the mean number of real roots.
3. In fact, even in the complex case, there are very interesting observable to look at, for instance, the spatial distribution of the roots (in dimension 1) or in general $Z(P)$. The main general result in this case is [34].

We will be interested in three questions:

1. What is the average volume of an algebraic hypersurface?
2. What is its mean topological type?
3. Are there large connected components of these hypersurfaces, that is larger than the natural scale?

1.3 Other models

This section is very short, because the author does not want to enter into the bibliographical maelstrom.



20 LARRY A. SHEPP AND ROBERT J. VANDERBEI

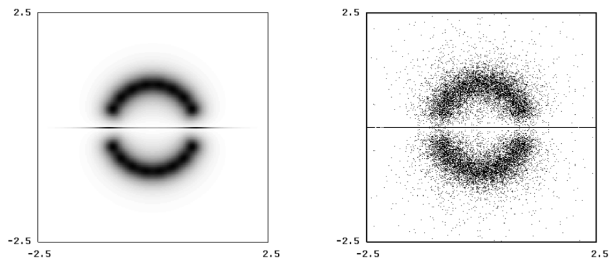


FIGURE 4. $n = 10$.

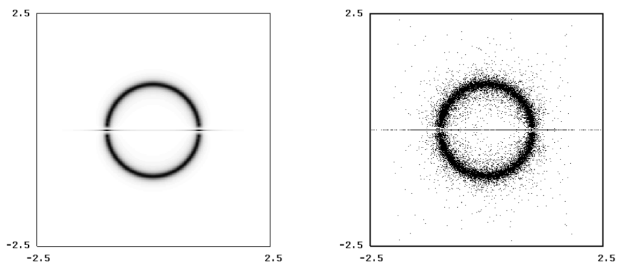


FIGURE 5. $n = 36$.

Figure 1.1: Repartition of roots of degree $d = 4, 10, 36$ of random Kac real polynomials.

1.3.1 Sections of ample line bundles

Everything that is proved in sections 2 and 3.2 has been in fact proved in a far more general situation. Namely, let M be a compact complex manifold equipped with a holomorphic line bundle $L \to X$ equipped itself with an hermitian product h with positive curvature equal to ω. This implies that ω is a Kähler manifold. If $c : X \to X$ is a anti-holomorphic involution, and $c_L : L \to L$ compatible with c. Since L has a positive curvature, then $H^0(X, L^{\otimes d})$ the set of holomorphic sections has a dimension which grows like d^n. The real part of it, that is

$$\mathbb{R}H^0(X, L^{\otimes d}) := \{s \in H^0(X, L^{\otimes d}), c_L \circ s = s \circ c\}$$

has the same real dimension.

Example 1.3.1 *The main example, and the only one that will be treated here is $M = \mathbb{C}P^n$ and c is the conjugation, if $L = O(1)$ the hyperplane bundle, then c_L can be chosen to be the natural conjugation. In this case, $\mathbb{R}H^0(X, L^{\otimes d}) = \mathbb{R}^d_{hom}[X_0, \cdots, X_n]$.*

1.3.2 Eigenfunctions of the Laplacian.

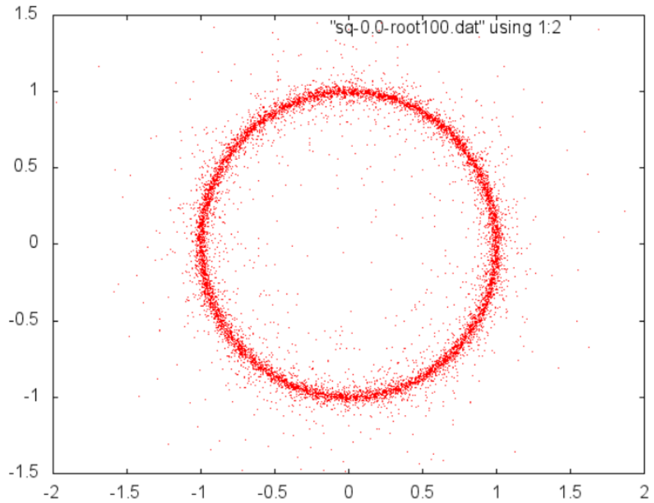
On (M, g) a compact Riemannian manifold, the Laplacian has a discrete infinite number of eigenvalues. As a random space we can take the sum of eigenspaces with eigenvalues less than a growing parameter L. We handle this kind of model for the mean volume in paragraph 2.3.1.

1.3.3 Gaussian fields on $\mathbb{R}^n$

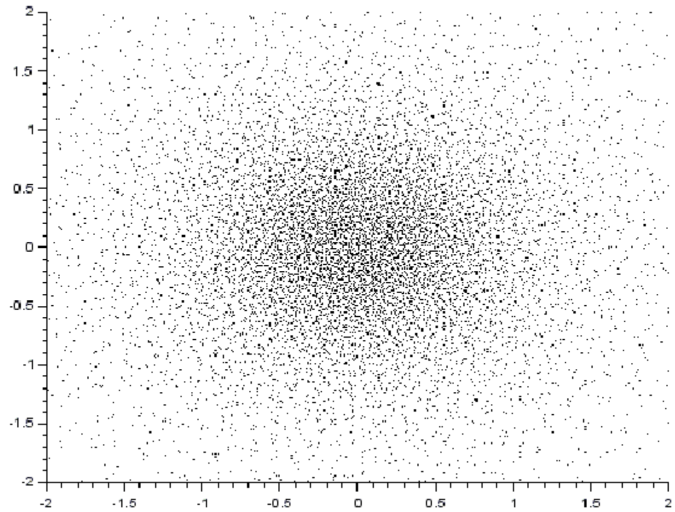
There is a lot of work for Gaussian fields $f : \mathbb{R}^n \to \mathbb{R}$. Even if $\mathbb{R}^n$ has a trivial geometry, the vanishing locus of f has not. See the books [1], [3] for results on the Euler Characteristics for instance. In these lectures, we will assume that the covariance function of f depends only on the distance between points. This allows us to use tools of percolation theory for the study of long connected components of the vanishing locus of f in $\mathbb{R}^2$, see chapter 4.

1.3.4 Complex hypersurfaces

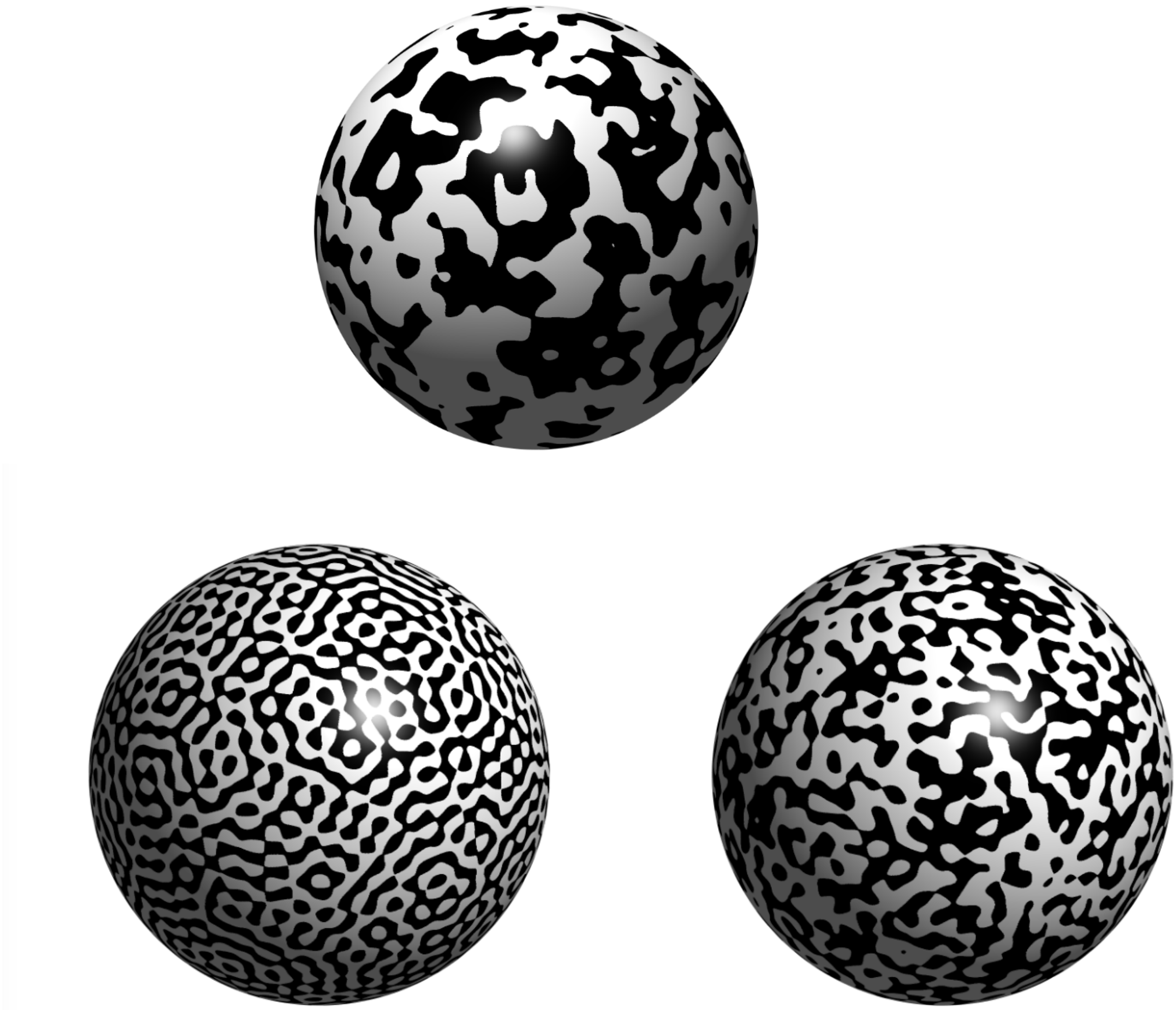
It is very natural to study the *complex* vanishing locus of a polynomial or a holomorphic section. As said before, [34] is one of the main reference. Other past results are referenced inside.



Repartition of roots of degree $d = 100$ random Kac complex polynomials.



Repartition of roots of degree $d = 100$ random for complex Fubini-Study or (Kostlan) polynomials.



Repartition of nodal lines for Kostlan random polynomials ($d = 300$), eigenfunctions ($d = 80$) and real Fubini-Study real polynomials ($d = 80$). Images Alex Barnett.

Chapter 2

Mean volume of random hypersurfaces

We saw in the introduction that the volume of $Z_{\mathbb{C}}(P) \subset \mathbb{C}P^n$ is always the same, whereas $Z_{\mathbb{R}} \subset \mathbb{R}P^n$ depends on P. In this section we estimate the average of this volume in the real case. Since for $n = 1$, the volume equals the number of roots of P, it is natural to begin with this case. Besides, this is an historical case.

2.1 Roots of random polynomials

2.1.1 Kac model

In 1943, Marek Kac's studied the following question:

Question 2.1.1 *Let*

$$p = \sum_{k=0}^{d} a_k x^k$$

be a random polynomial with independent coefficients following the same normal law :

$$\forall k \in \{0, \cdots, d\}, \ a_k \sim N(0, 1).$$

What is the mean number of real roots?

Note that by Gauss theorem, the number of *complex* roots is always d with probability one.

Notation. We will write in general, for any manifold M^n, $U \subset M$ an open set in M, and $f : U \to \mathbb{R}^n$,

$$N(f, U) := \#\{x \in U, \ f(x) = 0\} \in \mathbb{N} \cup \{+\infty\}$$

and $N = N(f) := N(f, M)$ when ambiguities are absent. Here we are interested in the case $n = 1$. Let us compute two simple cases.

1. For $d = 1$, $p = a_0 + a_1 x$ has one real root iff $a_1 \neq 0$ which happens with probability one, so
$$\mathbb{E}(N) = 1.$$
2. For $d = 2$, $p = a_0 + a_1 x + a_2 x^2$. The two complex roots are real iff
$$\Delta := a_1^2 - 4a_0 a_2 \geq 0,$$

so that

$$\mathbb{E}(N) = \int_{\Delta>0} 2e^{-\frac{1}{2}\|a\|^2} \frac{da}{\sqrt{2\pi}^3}$$

with $\|a\|^2 := \sum_{i=0}^2 a_i^2$ and $da = \prod_{i=0}^3 da_i$. Try to compute this.

Theorem 2.1.2 (Marek Kac 1943 [17]) *For a random polynomial of degree d as before,*

$$\mathbb{E}(N(p)) \sim_{d\to+\infty} \frac{2}{\pi}\log d.$$

2.1.2 Kac-Rice formula

The latter was originally proved by Kac by proving the following general formula:

Theorem 2.1.3 *(Kac-Rice formula ([3, Theorem 3.2])Let $I \subset \mathbb{R}$ be a segment, $f : I \to \mathbb{R}$ a random Gaussian field such that almost surely, f is C^1 on I, and for any $x \in I$, $\operatorname{Var} f(x) \neq 0$. Then,*

$$\mathbb{E}\big(N(f, I)\big) = \int_I \mathbb{E}\big(|f'(x)| \mid f(x) = 0\big)\phi_{f(x)}(0)dx.$$

Here, for any random Gaussian vector X, $\phi_X(x)$ denotes the density of X at $x \in \mathbb{R}$, see the probability toolbox for the Gaussian case. We emphasize that this formula does not depend on the law of $f(x)$. However when f is a random Gaussian field, it gives a very nice explicit computation:

Theorem 2.1.4 *(Kac 1943, Kostlan 1993) Under the same hypotheses than Theorem 2.1.3, assume moreover that f is centered, that is $\mathbb{E}(f(x)) = 0$ for any $x \in I$. Then, if $e : I^2 \to \mathbb{R}$ denote the two-point correlation function of f, then*

$$\mathbb{E}(N(f, I)) = \frac{1}{\pi}\int_{x\in I} \sqrt{\partial^2_{x,y} \log e(x, y)_{|x=y}}dx.$$

Proof. Since $f(x)$ is a centered Gaussian field,

$$\forall u \in \mathbb{R},\ \phi_{f(x)}(u) = e^{-\frac{1}{2}\frac{u^2}{\operatorname{Var}(f(x))}} \frac{du}{\sqrt{\operatorname{Var}(p(x))}\sqrt{2\pi}}.$$

For $u = 0$, this gives, using the correlation function,

$$\phi_{f(x)}(0) = \frac{du}{\sqrt{e(x, x))}\sqrt{2\pi}}.$$

We apply the regression formula given by Proposition 5.1.12 this to $X = (X_1, X_2) = (f'(x), f(x))$. By Theorem 5.1.10,

$$\operatorname{Cov}(f'(x), f'(y)) := \mathbb{E}(f'(x)f'(y)) = \frac{\partial^2}{\partial x\partial y}\mathbb{E}(f(x)f(y)) = \partial^2_{x,y}e,$$

in particular

$$\operatorname{Var}(f'(x)) = \partial^2_{x,y}e_{|x=y}.$$

and similarly

$$\operatorname{Cov}(f'(x), f(x)) = \mathbb{E}(f'(x)f(x)) = \frac{\partial}{\partial x}\mathbb{E}(f(x)f(y))_{|x=y} = \partial_x e_{|x=y}.$$

By Proposition 5.1.12,

$$\mathbb{E}(|f'(x)| \mid f(x) = 0) = \mathbb{E}(|X_3|),$$

where

$$X_3 \sim N\big(0, \partial^2_{xy} e_{|x=y} - \partial_x e^2_{|x=y} e(x,x)^{-1}\big).$$

Since for a centered $X \sim N(0, \Sigma)$,

$$\mathbb{E}(|X|) = \int_{x\in\mathbb{R}} |x| e^{-\frac{1}{2}\frac{x^2}{\Sigma}} \frac{dx}{\sqrt{\Sigma}\sqrt{2\pi}} = 2\sqrt{\Sigma}\int_{\mathbb{R}+} u e^{-\frac{1}{2}u^2} \frac{du}{\sqrt{2\pi}} = \sqrt{\frac{2\Sigma}{\pi}}$$

This implies

$$\mathbb{E}\big(N(f,I)\big) = \frac{1}{\pi}\int_{x\in I} \sqrt{\frac{e(x,x)\partial^2_{xy} e_{|x=y} - \partial_x e^2_{|x=y}}{e^2(x,x)}} dx.$$

□

2.1.3 Kostlan polynomials

There is another, more natural in fact than Kac polynomials, random model for polynomial, the *complex Fubini-Study*, or *Kostlan* measure:

$$p(x) = \sum_{k=0}^{d} a_k \sqrt{\binom{d}{k}} x^k,$$

where the $(a_k)_k$ are still independent and follow the same centered normal law. We will see later why this is a good measure to choose, which is not really clear at first sight! For the moment, we see that the 2-point correlatione equals:

$$e(x,y) = \sum_{k=0}^{d} (1+xy)^d.$$

Theorem 2.1.5 *(Kostlan [19], Shub-Smale [35] 1993) For Kostlan polynomials,*

$$\forall d \geq 0,\ \mathbb{E}(b_0(P)) = \sqrt{d}.$$

We apply now theorem 2.1.4 for Kostlan polynomials and then for the Kac.

Proof of Theorem 2.1.5.. The Kostlan polynomials are in fact easier than Kac's one. For it,

$$\sqrt{\partial^2_{xy} \log e(x,y)_{|x=y}} = \frac{\sqrt{d}}{1+x^2},$$

so that Theorem 2.1.4 gives

$$\mathbb{E}(N(p_{Kostlan})) = \frac{\sqrt{d}}{\pi}\int_{\mathbb{R}} \frac{1}{1+x^2} dx = \sqrt{d}.$$

We will recover this result in section 2.2 with a more geometric point of view. □

Proof of Theorem 2.1.2.. Now for the Kac polynomials, $e(x,y) = \frac{1-(xy)^{d+1}}{1-xy}$, so that

$$\sqrt{\partial^2_{xy} \log e(x,y)_{|x=y}} = \sqrt{\frac{1}{(1-x^2)^2} - (d+1)^2 \frac{x^{2d}}{1-t^{2d+2})^2}},$$

finir □

2.1.4 Sketchy proof of Kac-Rice formula

We begin by a deterministic relation which expresses N as an integral, and which has its own interest.

Lemma 2.1.6 *Let $I \subset \mathbb{R}$ be a closed interval, $f : I \to \mathbb{R}$ C^1, $f \neq 0$ on ∂I and $f(x) = 0$ implies $f'(0) = 0$ (transversality). Then,*

$$N(f, I) = \lim_{\delta \to 0} \frac{1}{2\delta} \int_{x \in I, |f(x)| \leq \delta} |f'(x)| dx.$$

Proof. The transversality assumption implies (why?) that there exists a finite number of zeros of f, $a < x_1 < \cdots < x_N < b$. For $\delta > 0$ small enough, $f^{-1}([-\delta, \delta])$ is a union of disjoint intervals $(J_k \ni x_k)_{k \in \{0, \cdots N\}}$ (why?). Then,

$$\begin{aligned} \int_{x \in I, |f(x)| \leq \delta} |f'(x)| dx &= \sum_{i=0}^{N} \int_{J_k} |f'(x)| dx && (2.1.1) \\ &= \sum_{i=0}^{N} \operatorname{sgn}(f'(x_k)) [f]_{\partial J_k} && (2.1.2) \\ &= \sum_{i=0}^{N} 2\delta = 2N\delta, && (2.1.3) \end{aligned}$$

hence the result. □

Proof of Kac-Rice. We have, admitting that every inversion of integral and limits are allowed,

$$E(N(p, I)) = \lim_{\delta \to 0} \frac{1}{2\delta} \int_I \mathbb{E}\big(\mathbf{1}_{\{|f(x)| < \delta\}} |f'(x)|\big) dx. \qquad (2.1.4)$$

Now for any $x \in I$, $f(x)$ is a random variable which law can be denoted by $d\mu_{f(x)}$, so that

$$\Pr(f(x) \in [u, v]) = \int_u^v d\mu_{f(x)}(s).$$

Using formula **??** we obtain

$$\begin{aligned} \mathbb{E}\big(1_{|f| < \delta} |f'(x)|\big) &= \int_{(u,v) \in \mathbb{R}^2} \mathbf{1}_{|p| < \delta} |f'(x)| p_{(f(x), f'(x))}(u, v) du dv && (2.1.5) \\ &= \int_{u \in \mathbb{R}} \mathbf{1}_{|p| < \delta} \Big(\int_{v \in \mathbb{R}} |f'(x)| p_{[f'(x) | f(x)]}(v|u) dv \Big) p_{f(x)}(u) du && (2.1.6) \\ &= \int_{\mathbb{R}} \mathbb{E}_{f'(x)} \big(|f'(x)| \mid f(x) = u \big) dp_{f(x)}(u) \mathbf{1}_{|p| < \delta} du && (2.1.7) \end{aligned}$$

so that

$$\lim_{\delta \to 0} \frac{1}{2\delta} \mathbb{E}\big(1_{|p| < \delta} |f'(x)|\big) = \mathbb{E}_{f'(x)} \big(|f'(x)| \mid f(x) = 0 \big) d\mu_{f(x)}(0). \qquad (2.1.8)$$

□

2.2 Algebraic hypersurfaces

The aim of this paragraph is to prove the following theorem:

Theorem 2.2.1 *(see [19]) If P of degree d is chosen at random with the Kostlan measure, then*

$$\forall d,\ \mathbb{E}\big(\operatorname{Vol}(Z_{\mathbb{R}}(P))\big) = \sqrt{d}\operatorname{Vol}(\mathbb{R}P^{n-1}).$$

For $n = 1$, this gives $\mathbb{E}(N(P)) = \sqrt{d}$ since $\mathbb{R}P^0 = \{1\}$, which was already proved by theorem 2.1.5. In section 5.3 we recall the definition of the Riemannian volume. There exists an estimate for the variance:

Theorem 2.2.2 *(Letendre [23] 2018) The variance of the volume of algebraic hypersufaces satisfies the estimate*

$$\operatorname{Var}\big(\operatorname{Vol}(Z_{\mathbb{R}}(P))\big) \sim_{d\to\infty} C_n d^{1-\frac{n}{2}},$$

where C_n is a universal constant. Besides, for any $\epsilon > 0$,

$$\Pr\Big(|\operatorname{Vol}(Z) - \mathbb{E}\operatorname{Vol}| > \sqrt{d}^{1-n/2+\epsilon}\Big) \leq C_\epsilon d^{-\epsilon}$$

We will not prove this theorem.

2.2.1 General Kac-Rice formula for volumes

For Theorem 2.2.1, we begin with a general formula for computing the expectation of the volumes of random hypersurfaces in Riemannian manifolds:

Theorem 2.2.3 *(Kac-Rice formula for volumes [3, Theoreom 6.8]) Let (M, g) be a Riemannian manifold and $f : M \to \mathbb{R}$ a Gaussian field on M, such that almost surely, f is C^1 and regular, and $\operatorname{Var}(f(x)) > 0$ for any $x \in M$. Then, for any integrable open subset $U \subset M$,*

$$\mathbb{E}\big(\operatorname{Vol}(f^{-1}(0)\cap U)\big) = \int_U \mathbb{E}\big(\|\nabla f(x))\| \mid f(x) = 0\big)\phi_{f(x)}(0) dvol_M(x).$$

We will use for he proof of this theorem the famous *co-area formula*, which is a kind of wonder formula which allows to compute an integral using the geometry given by levels of a given function. This generalizes the Fubini theorem, where the function equals a particular coordinate.

Theorem 2.2.4 *(Coarea Formula [10, Exercise III.12 (c)] Let M be a smooth Riemannian manifolds of dimension m, $U \subset M$ be any measurable subset of M, and $f : M \to \mathbb{R}^n$ be a smooth function with $n \leq m$. Then, for any continuous and bounded $g : M \to \mathbb{R}$, one has*

$$\int_U g\sqrt{\det df \circ df^*} dvol_M = \int_{y\in\mathbb{R}^m}\int_{f^{-1}(y)\cap U} g_{|f^{-1}(y)} dvol_{f^{-1}(y)} dy.$$

Recall that if $f : (E, g) \to (E', f)$ is a linear map between spaces equipped with scalar products, $f^* : E' \to E$ is defined by

$$\forall u' \in E', u \in E,\ g'(u', f(u)) = g(f^*(u'), u).$$

This implies that for two ONB B and B', $Mat(f^*, B', B) = Mat(f, B, B')^t$.

Exercise 2.2.5 *Prove that it is well defined.*

Example 2.2.6 *Let us give a simple example for $M = I \subset \mathbb{R}$ and $N = \mathbb{R}$ with the standard metric. Here $f^{-1}(y)$ is a discrete number of points, $df = f'(x)dx$ and $df^* = f'(x)dx$, so that $\sqrt{\det df \circ df^*} = |f'(x)|$. so that*

$$\int_I g|f'(x)|dx = \int_{\mathbb{R}} \sum_{x\in f^{-1}(y)} g(x)dy.$$

Note that if f vanishes transversally at its zeros, choosing g as $\mathbf{1}_{\{|f|\le\delta\}}$, we obtain

$$\int_{I\cap\{|f|\le\delta\}} |f'(x)|dx = \int_{[-\delta,\delta]} \#\{x \in I \cap f^{-1}(y)\}dy.$$

If the set of zeros of f is finite, f is transverse, then for δ is small enough, the right hand side equals $2\delta N(f, I)$, and we recover Lemma 2.1.6.

Exercise 2.2.7 *Volume of the spheres $\mathbb{S}^n$. Using the function $f(x) = \|x\|^2$ and the coarea formula, show that*

$$\mathrm{Vol}(\mathbb{S}^n) = (\sqrt{2\pi})^n\Big(\int_{\mathbb{R}} t^n e^{-\frac{1}{2}t^2}\frac{dt}{\sqrt{2\pi}}\Big)^{-1}.$$

and $\mathrm{Vol}(\mathbb{R}P^n) = \frac{1}{2}\mathrm{Vol}(\mathbb{S}^n)$.

The following proposition is the higher dimensional equivalent of Proposition 2.1.6

Proposition 2.2.8 *Let $f : M \to \mathbb{R}$ a C^1. Then, for any measurable $U \subset M$,*

$$\mathrm{Vol}(Z(f) \cap U) = \lim_{\delta\to 0}\frac{1}{2\delta}\int_{x\in U, |f(x)|\le\delta} \|\nabla f(x)\| dvol_M(x).$$

Proof. We apply the coarea formula given by Proposition 2.2.4. Note that by definition of the gradient of f,

$$df(x)(v) = \langle\nabla f(x), v\rangle_{T_xM},$$

and by definition of the adjoint,

$$df(x)(v) \times t = \langle v, df(x)^*(t)\rangle,$$

so that

$$df(x)^* = \nabla f(x)dt,$$

and

$$df(x) \circ df(x)^* = \|\nabla f(x)\|^2 dt.$$

This implies

$$\sqrt{\det df \circ df^*} = \|\nabla f(x)\|.$$

Consequently by Proposition 2.2.4, for any bounded continuous g and any measurable $U \subset M$, and in fact by any characteristic function of any measurable subset,

$$\int_U g\|\nabla f(x)\| dvol_M(x) = \int_{\mathbb{R}}\int_{x\in f^{-1}(y)\cap U} g(x)dvol_{|f^{-1}(y)}(x)dy.$$

Choose $g = g_\delta := \mathbf{1}_{[-\delta,\delta]} \circ f$. Then,

$$\int_{U\cap\{|f|\le\delta\}} \|\nabla f(x)\| dvol_M(x) = \int_{-\delta}^{\delta} \mathrm{Vol}(f^{-1}(y))dy.$$

Now, $\frac{1}{2\delta}\int_{-\delta}^{\delta}\mathrm{Vol}(f^{-1}(y))dy \to \mathrm{Vol}(f^{-1}(0))$ gives the result. $\square$

2.2.2 Computation

Proof of Theorem 2.2.3. Taking the average in f in the volume formula given by Proposition 2.2.8 gives, if we do not care for the inversion of limits and integral:

$$\begin{aligned}
\mathbb{E}(\mathrm{Vol}(f^{-1}(0))) &= \mathbb{E}\lim_{\delta\to 0}\frac{1}{2\delta}\int_{U\cap\{|f|\le\delta\}}\|\nabla f(x)\|dvol_M(x) &(2.2.1)\\
&= \lim_{\delta\to 0}\frac{1}{2\delta}\int_U \mathbb{E}\big(\mathbf{1}_{\{|f|\le\delta\}}\|\nabla f(x)\|\big)dvol_M(x) &(2.2.2)\\
&= \lim_{\delta\to 0}\frac{1}{2\delta}\int_U\int_{u=-\delta}^{\delta}\mathbb{E}\big(\|\nabla f(x)\|\mid f(x)=u\big)p_{f(x)}(u)dudvol_M(x) &(2.2.3)\\
&= \int_U \mathbb{E}\big(\|\nabla f(x))\|\mid f(x)=0\big)p_{f(x)}(0)dvol_M(x). &(2.2.4)
\end{aligned}$$

□

Let us apply Theorem 2.2.3 for random algebraic hypersurfaces in $\mathbb{R}P^n$. Recall that $P\in\mathbb{R}^d_{hom}[X_1,\cdots,X_n]$. I emphasize that the computation holds for far more general situations, see [22] for hypersurfaces which are the vanishing loci of random sums of eigenfunctions of the Laplace operator on a compact manifold.

Proof of Theorem 2.2.1. Firstly, since the integrand is invariant under isometries, we have for $x=[1:0\cdots:0]$,

$$\mathbb{E}(\mathrm{Vol}(Z(p))=\mathbb{E}\big(\|\nabla f(x))\|\mid f(x)=0\big)p_{f(x)}(0)\,\mathrm{Vol}(\mathbb{R}P^n),$$

and $f=\frac{P}{X_0^d}$ (locally f and P have the same vanishing locus).

The law of $f(x)$ is

$$e^{-\frac{1}{2}\frac{u^2}{e(x,x)}}\frac{du}{\sqrt{e(x,x)}\sqrt{2\pi}},$$

so that

$$p_{f(x)}(0)=\frac{1}{\sqrt{e(x,x)}\sqrt{2\pi}}.$$

Let $(x_1,\cdots,x_n)$ be local coordinates, such that $(\partial_{x_1},\cdots,\partial_{x_n})$ form an orthonormal basis of $T_x\mathbb{R}P^n$. Then, in this basis, $\nabla f(x)$ writes $X_1:=(\partial_i f)_{i=1,\cdots,n)}$ and $\|\nabla f(x)\|=\|X_1\|$. The random vector X_1 has covariance :

$$\Sigma_1=((\partial^2_{x_i,y_j}e(x,y)_{|x=y})_{i,j=1,\cdots,n}).$$

The random vector $X_2=f(x)$ has covariance $e(x,x)$. Lastly, the covariance

$$\mathrm{Cov}(X_1,X_2)=((\partial_{x_i}e(x,y)_{|x=y})_{i=1,\cdots,n}).$$

By Proposition **??**, the conditioned vector $(X_1|X_2=0)$ follows the normal law

$$N(0,\Sigma_1)-e^{-1}(x,x)\,\mathrm{Cov}(X_1,X_2)^t\,\mathrm{Cov}(X_1,X_2)).$$

So,

$$\mathbb{E}\big(\mathrm{Vol}(Z(p))\big)=\mathbb{E}\big(\|X_1\|\big)\frac{1}{\sqrt{2\pi e(x,x)}}\mathrm{Vol}(\mathbb{R}P^n).$$

Since $\mathbb{R}P^n$ is locally isometric to the sphere $\mathbb{S}^n$, we can assume that $x = (1, 0, \cdots, 0) \in \mathbb{R}^{n+1}$ and we can choose the local coordinates

$$(y_1, \cdots, y_n) \mapsto (\sqrt{1 - \|y\|^2}, y_1, \cdots, y_n).$$

We see that these coordinates satisfy the former orthonormal condition. Now, the affine coordinates $x = (x_1, \cdots x_n)$ on $\mathbb{R}P^n$ equal

$$x = \frac{y}{\sqrt{1 - \|y\|^2}}.$$

This implies that at $x = [1 : 0 : \cdots : 0]$, the affine coordinates satisfy the orthonormal condition too. We have

$$f(x) = f([1 : 0 : \cdots : 0]) = a_{d0\cdots 0}\sqrt{d^n},$$

so that $e(x, x) = d^n$. Moreover,

$$df(x) = a_{(d-1)0\cdots 1\cdots 0} d^{n/2}\sqrt{d} dx_k,$$

so that $\mathbb{E}\|X_1\| = d^{n/2}\sqrt{d}\mathbb{E}(\|a\|)$ with $a = (a_k)_{k=1\cdots n}$,

$$\Sigma(a) = I_n,$$

and $\mathrm{Cov}(X_1, X_2) = 0$, so that, using polar coordinates,

$$\begin{aligned} \mathbb{E}(\|a\|) &= \int_{a \in \mathbb{R}^n} \|a\| e^{-\frac{1}{2}\|a\|^2} \frac{da}{\sqrt{2\pi}^n} && (2.2.5) \\ &= \int_0^\infty r e^{-\frac{1}{2}r^2} \frac{r^{n-1} dr}{\sqrt{2\pi}^n} \mathrm{Vol}(\mathbb{S}^{n-1}), && (2.2.6) \end{aligned}$$

and finally

$$\mathbb{E}(\mathrm{Vol}(Z(P))) = \sqrt{d} \frac{1}{\sqrt{2\pi}^{n+1}} \mathrm{Vol}(\mathbb{R}P^n) \mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^\infty r^n e^{-\frac{1}{2}r^2} dr.$$

By 5.3, this is

$$\sqrt{d}\, \mathrm{Vol}(\mathbb{R}P^n) \frac{\mathrm{Vol}(\mathbb{S}^{n-1})}{Vol(\mathbb{S}^n)} = \sqrt{d}\, \mathrm{Vol}(\mathbb{R}P^{n-1})$$

□

2.3 Nodal hypersurfaces

2.3.1 Covariant function and spectral kernel

We will apply the Kac-Rice formula 2.2.3 for a different class of functions, namely the random sum of eigenfunctions of the Laplacien. Let (M, g) be a compact smooth Riemannian manifold, and define

$$\Delta := d^* d : C^\infty(M, \mathbb{R}) \quad \to \quad C^\infty(M, \mathbb{R}) \qquad (2.3.1)$$

the Laplacian, where $d^* : \Omega^1(M, \mathbb{R}) \to C^\infty(M, \mathbb{R})$ is the adjoint of the differential $d : C^\infty(M, \mathbb{R}) \to \Omega^1(M, \mathbb{R})$. The first main fact is the following:

Theorem 2.3.1 *([10, Theorem III.9.1.]) The set of eigenvalues with C^2 eigenfunctions consists of a sequence*

$$0 \leq \lambda_1 < \cdots < \lambda_2 \cdots \uparrow +\infty,$$

and each associated eigenspace is finite dimensional. Eigenspaces belonging to distinct eigenvalues are orthogonal in $L^2(M)$, and $L^2(M)$ is the direct sum of all the eigenspaces. Furthermore, each eigenfunction is in $C^\infty(M)$.

Example 2.3.2 *If (M,g) is the round sphere $(\mathbb{S}^n, g_0)$, then the eigenvalues of the Laplacian are the $\lambda_d = d(d+n-1)$, and the associated eigenspace is*

$$E_{\lambda_d} = \{P_{|\mathbb{S}^n} \mid P \in \mathbb{R}^d_{hom}[X_0, \cdots, X_n] \text{ and } \Delta_{\mathbb{R}^{n+1}} P = 0\}.$$

This implies that $N_L := \dim \oplus_{\lambda \leq L} \to_{L\to\infty} +\infty$. In fact, we have more:

Theorem 2.3.3 *(Weyl 1911)*

$$N_L \sim_{L\to\infty} \frac{\operatorname{Vol} B^n}{(2\pi)^n} L^{n/2} \operatorname{Vol}(M).$$

Now, we can take at random a function f in the following way: choose an L^2-orthonormal basis $(\phi_i)_{i=0}^\infty$, where the ϕ_i's are eigenfunctions of the Laplacian ordered with their eigenvalues. Then,

$$f_L = \sum_{\lambda_i \leq L} a_i \phi_i,$$

where as usual the $(a_i)_i$ are independent normal law following the same $N(0,1)$. The two-point correlation equals

$$e_L(x,y) = \sum_{\lambda_i \leq L} \phi_i(x)\phi_i(y).$$

In fact, this correlation has a geometrical interpretation. Let

$$\pi_L : L^2(M) \to E_L$$

be the orthogonal projection (for the L^2-metric on the functions). Then, by definition,

$$\begin{aligned} \forall f \in L^2(M),\ \pi_L(f) &= \sum_{\lambda_i \leq L} \langle f, \phi_i \rangle \phi_i & (2.3.2)\\ \Leftrightarrow \forall x \in M,\ \pi_L(f)(x) &= \int_{y\in M} \sum_{\lambda_i \leq L} \phi_i(x)\phi_i(y) f(y) d\operatorname{Vol}_M(y), & (2.3.3) \end{aligned}$$

so that

$$\pi_L(f)(x) = \int_{y\in M} e_L(x,y) f(y) d\operatorname{Vol}_M(y).$$

In orther termes, $e_L(x,y)$ is the *spectral kernel* or *Schwartz kernel* associated to E_L. It happens that there exists a precise result due to Hörmander (after others) about this kernel:

Theorem 2.3.4 *([16, Theorem 5.1] Fix $x \in M$ and choose $(x_i)_i$ normal coordinates. Then, uniformly in x, for any $\alpha, \beta \in \mathbb{N}^n$ and $L \geq 1$,*

$$\begin{aligned} \frac{\partial^{|\alpha|}}{\partial x_0^{\alpha_0} \cdots \partial x_n^{\alpha_n}} \frac{\partial^{|\beta|}}{\partial x_0^{\beta_0} \cdots \partial x_n^{\beta_n}} e_L(x,y)_{|x=y} &= L^{\frac{n+|\alpha|+|\beta|}{2}} \frac{1}{(2\pi)^n} \int_{\|\xi\|\leq 1} (i\xi)^\alpha (-i\xi)^\beta d\xi \\ &\quad + o(L^{\frac{n+|\alpha|+|\beta|}{2}}) \end{aligned}$$

This implies three things :

1. When $\alpha = \beta = 0$,
$$\forall x \in M,\ e_L(x,x) \sim_{L\to\infty} \frac{1}{(2\pi)^n} L^{\frac{n}{2}} \operatorname{Vol}(\mathbb{B}),$$

2. which itself implies the Weyl theorem. Indeed, let us integrate this equation over M. Then, the left hand side gives
$$\int_M e_L(x,x) dvol_M = \sum_{\lambda_i \leq L} \int_M \phi_i(x)^2 dvol_M = \dim E_L,$$
while the right-hand side gives
$$\int_M \frac{1}{(2\pi)^n} L^{\frac{n}{2}} \operatorname{Vol}(\mathbb{B}) d \operatorname{Vol}_M = \frac{1}{(2\pi)^n} L^{\frac{n}{2}} \operatorname{Vol}(\mathbb{B}) \operatorname{Vol}(M).$$

3. For any i, $\partial_{x_j} e_L = o(L^{\frac{n}{2}+1})$;

4. For any i, j,
$$\partial^2_{x_j y_j} e_L \quad = \quad \delta_{ij} \frac{1}{(2\pi)^n} L^{\frac{n}{2}+1} \operatorname{Vol}(\mathbb{B}^n)(n+2) + o(L^{\frac{n}{2}+1}). \qquad (2.3.4)$$

Exercise 2.3.5 *Show it.*

Exercise 2.3.6 *Let $M = S^1 \subset \mathbb{R}^2$ be the unit circle parametrized by*
$$\theta \in [0, 2\pi] \mapsto e^{i\theta}.$$
The Laplacian Δ reads $\Delta f = -\frac{\partial^2 f}{\partial \theta^2}$ acting on C^2 2π-periodic functions.

1. *What are the eigenfunctions of Δ?*

2. *We define the L^2-scalar product on C^2 2π-periodic real functions par*
$$\langle f, g \rangle := \int_0^{2\pi} fg(\theta) \frac{d\theta}{2\pi}.$$
For any $L \geq 0$, find an ONB $(f_i)_{i \in A(L)}$ of eigenfunctions with eigenvalues less or equal to L.

3. *For any $N \in \mathbb{N}$, let $L = N^2$ and $f_L : S^1 \to \mathbb{R}$ the random Gaussian field*
$$f_L := a_0 + \sqrt{2} \sum_{k=1}^{N} a_k \cos(kx) + b_k \sin(ky),$$
where the a_i and b_j are independent and $a_i, b_j \sim N(0,1)$. Show that the covariance function $e_L : [0, 2\pi]^2 \to \mathbb{R}$ of f_L satisfies
$$\forall (x,y) \in [0,2\pi]^2, x \neq y \Rightarrow e_L(x,y) = \frac{\sin((N+\frac{1}{2})(x-y))}{\sin(\frac{x-y}{2})}.$$

2.3.2 Computation

We can now apply this situation to the computation of the mean volume.

Theorem 2.3.7 *(Bérard 1985 [8]) Under the hypotheses from above,*

$$\mathbb{E}(\mathrm{Vol}(Z(f_L)) \sim_{L\to\infty} \sqrt{\frac{L}{n+2}}\,\mathrm{Vol}(M)\frac{\mathrm{Vol}(\mathbb{S}^{n-1})}{\mathrm{Vol}\,\mathbb{S}^n}.$$

Remark 2.3.8 1. *Recall by Example 2.3.2 that in the case of the round sphere,*

$$\forall d \geq 1,\ E_{d(d+n-1)} \subset \mathbb{R}^d_{hom}[X_0, \cdots, X_n]_{|\mathbb{S}^n},$$

as for the Kostlan measure. But for the latter, $\mathbb{E}(\mathrm{Vol}(Z(f)) \sim_d C\sqrt{d}$, when for the former, $\mathbb{E}(\mathrm{Vol}(Z(f)) \sim_d Dd$.

2. *Hörmander theorem says a bit more:*

$$L^{-n/2}e_L(x, x + L^{n/2}h) \to_{L\to\infty} \frac{1}{(2\pi)^n}\int_{\|\xi\|\leq 1} e^{-i\langle\xi,h\rangle}d\xi.$$

This means that when we look at the neighborhood of a point at the scale $L^{-n/2}$, then the geometry of the nodal random hypersurface does not change with L. Moreover, since the kernel measures the dependency of the Gaussian field, we see that at distances far from $L^{n/2}$, the values are almost independent. This can explain heuristically the $\sqrt{L}$ term in the mean volume. Indeed, a ball $B(x, L^{-n/2})$ has a volume equal to L^{-n}, so that we cans if we cover M with almost $\mathrm{Vol}\, ML^n$ such disjoint balls. If we assume that the field is independent in these balls, then, since the natural scale is $L^{-n/2}$, the volume of $f^{-1}(0)$ is close to the volume of an $n-1$-ball in the small ball, so that

$$\mathrm{Vol}\, f^{-1}(0) \cap B(x, 1/L^{n/2}) \approx C' L^{(n-1)}$$

Proof. Here, as before, $X_1 = (\partial_i f)_{i=1,\cdots n}$ in the normal coordinates, $X_2 = f$ so that $\mathrm{Var}(X_2) = e_L(x,x)$. This implies

$$p_{X_2}(0) = \frac{1}{\sqrt{2\pi e_L(x,x)}}.$$

If

$$a := \frac{1}{(2\pi)^n}L^{\frac{n}{2}+1}\,\mathrm{Vol}(\mathbb{B}^n)(n+2),$$

then

$$\mathrm{Var}(X_1) = (\partial^2_{i,j}e_{L|x=y})_{i,j=1,\cdots,n} = a\mathbf{1}_n + o(L^{n+2}),$$

$\mathrm{Cov}(X_1, X_2) = o(L^{\frac{n}{2}+1})$. This implies that the variance matrix of $(X_1 \mid X_2)$ is

$$\Sigma_1 := \mathrm{Var}(X_1) - o(L^{n+2}L^{-n/2}) = a\mathbf{1}_n + o(L^{n+2}).$$

Applying Theorem 2.2.3 gives

$$\mathbb{E}(\mathrm{Vol}(Z(f_L))) = \int_M\int_{x\in\mathbb{R}^n}\|x\|e^{-\frac{1}{2}\langle\Sigma_1^{-1}x,x\rangle}\frac{dx}{(2\pi)^{n/2}\sqrt{\det\Sigma_1}}\frac{1}{\sqrt{2\pi e_L(x,x)}}dvol_M.$$

Making the change of variables $x = y\sqrt{a}$ and noting that

$$a = e_L(x,x)L(n+2)(1+o(L)),$$

so that

$$\begin{aligned}
\mathbb{E}(\mathrm{Vol}(Z(f_L))) &= \int_M \int_{x\in\mathbb{R}^n} \|y\|\sqrt{a}e^{-\frac{1}{2}(1+o(1)\langle y,y\rangle}a^{n/2}\frac{dy}{(2\pi)^{n/2}a^{n/2}(1+o(1))} \\
&\sim_{L\to+\infty} \sqrt{\frac{L}{n+2}}\,\mathrm{Vol}(M)\,\mathrm{Vol}(\mathbb{S}^{n-1})\int_0^{+\infty} r^n e^{-\frac{1}{2}r^2}\frac{dr}{(2\pi)^{n/2}} \\
&= \sqrt{\frac{L}{n+2}}\,\mathrm{Vol}(M)\frac{\mathrm{Vol}(\mathbb{S}^{n-1})}{\mathrm{Vol}\,\mathbb{S}^n}
\end{aligned}$$

□

Chapter 3

Topology of random algebraic hypersurfaces

Let $P \in \mathbb{R}^d_{hom}[X_0, \cdots, X_n]$ be a homogeneous polynomial of degree d. We can look at its vanishing locus $Z(P)$ in the real projective space $\mathbb{R}P^n$, or on the unit sphere of $\mathbb{R}^{n+1}$, but it doubles the locus since it is invariant under antipodal transformation.

3.1 The Kostlan or complex Fubini-Study measure

We use the following non-intuitive, but miraculous, scalar product :

$$\langle P, Q\rangle = \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} P(Z)\bar{Q}(Z)e^{-\|Z\|^2}|dZ|^2,$$

with

$$|dZ|^2 = \prod_{k=0}^{n} dX_k \otimes dY_k,$$

where $Z_k = X_k + iY_k$. Of course the constant is not the origin of the miracle, but the integration over the complex space instead the real one.

Lemma 3.1.1 *The monomials*

$$\left(\sqrt{\frac{(d+n)!}{i_0! \cdots i_n!}} X_0^{i_0} \cdots X_n^{i_n}\right)_{i_0+\cdots+i_n=d}$$

form an orthonormal basis of $(\mathbb{R}^d_{hom}[X_0, \cdots, X_n], \langle , \rangle)$.

Proof. Let us compute $\langle X^I, X^J\rangle = \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} Z^I \bar{Z}^J e^{-\sum_i |Z_i|^2} |dZ_1|^2 \otimes \cdots \otimes |dZ_n|^2$. By Fubini, this is

$$\begin{aligned}
\langle X^I, X^J\rangle &= \frac{1}{(d+n)!\pi^{n+1}} \prod_{k=0}^{n} \int_{\mathbb{C}} Z_k^{i_k} \bar{Z}_k^{j_k} e^{-|Z_k|^2} |dZ_k|^2 \\
&= \frac{1}{(d+n)!\pi^{n+1}} \prod_{k=0}^{n} \int_0^{+\infty} \int_0^{2\pi} \rho^{i_k+j_k} e^{i(i_k-j_k)} e^{-\rho^2} \rho d\rho d\theta \\
&= \frac{1}{(d+n)!\pi^{n+1}} \delta_{I,J} (2\pi)^{n+1} \prod_{k=0}^{n} \int_0^{+\infty} \rho^{2i_k} e^{-\rho^2} \rho d\rho \\
&= \delta_{I,J} \frac{i_0! \cdots i_n!}{(d+n)!}.
\end{aligned}$$

□

Our random polynomials write,

$$P = \sum_{i_0+\cdots+i_n=d} a_{i_0\cdots i_n}\sqrt{\frac{(d+n)!}{i_0!\cdots i_n!}}X_0^{i_0}\cdots X_n^{i_n}$$

with independent normal $a_{i_0\cdots i_n}$. However, any orthonormal basis $(P_i)_{i=0,\cdots n}$ could be used instead.

Lemma 3.1.2 *The scalar product is invariant under the action on $\mathbb{R}^d_{\text{hom}}[X]$ by the orthonormal group $O(n+1)$, that is for any $g\in O(n+1)$, any polynomials P,Q, $\langle P\circ g, Q\circ g\rangle = \langle P, Q\rangle$.*

Proof. One implicit affirmation of the lemma is that $\mathbb{R}^d_{\text{hom}}[X]$ is invariant under the action of $O(n+1)$. This is an **exercise**. Indeed, for any $g\in O(n+1)$.

$$\langle P\circ g, Q\circ g\rangle = \frac{1}{(d+n)!\pi^{n+1}}\int_{\mathbb{C}^{n+1}} P\circ g(Z)\bar{Q}\circ(Z)e^{-\|Z\|^2}|dZ|^2.$$

Using the change of variable $Z' = g(Z)$, we obtain

$$\langle P\circ g, Q\circ g\rangle = \frac{1}{(d+n)!\pi^{n+1}}\int_{\mathbb{C}^{n+1}} P(Z')\bar{Q}(Z')e^{-\|g^{-1}Z'\|^2}|\det dg^{-1}||dZ'|^2.$$

However, $\det dg^{-1} = \det g^{-1} = \pm$ and $\|g^{-1}Z'\|^2 = \|Z'\|^2$. □

3.2 All topologies happen

3.2.1 In a small ball

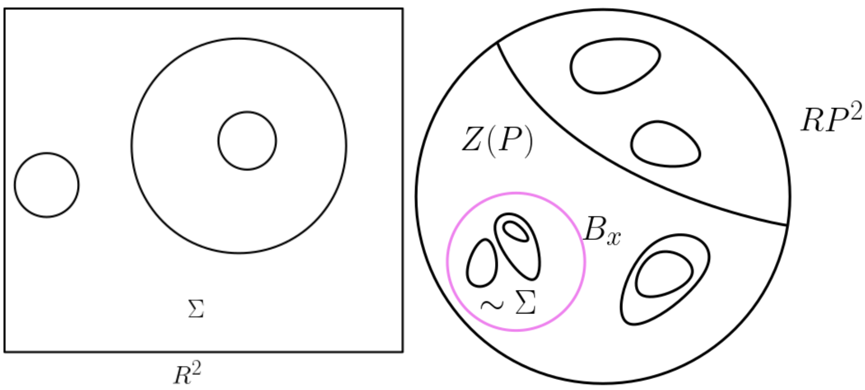


Figure 3.1: Left: a finite arrangement Σ of ovals in $\mathbb{R}^2$. Right: for any $x\in\mathbb{R}P^2$ there exists $c_\Sigma>0$, such that for any $d\gg 1$, $B(x,1/\sqrt{d})\cap Z(P)\sim(\Sigma,\mathbb{R}^2)$ with probability at least c_Σ. In a way, there is no random Hilbert problem.

Definition 3.2.1 *A smooth hypersurface $\Sigma\subset\mathbb{R}^n$ is said* algebraic *if this is the vanishing locus of some polynomial.*

Example 3.2.2 *Let C be any union of topological circles in $\mathbb{R}^2$. Then, it is isotopic to a union of geometrical circles, so that it is isotopic to an algebraic curve.*

Of course not any smooth hypersurface is not algebraic. In particular, it must be analytic, so that for instance in $\mathbb{R}^2$, any compact connected curve containing a straight segment is not algebraic. However, we have:

Theorem 3.2.3 (Herbert Seifert (1936) [32]) *Let $\Sigma \subset \mathbb{R}^n$ be any compact smooth hypersurface. Then, there exists a polynomial p and a diffeotopy of $\mathbb{R}^n$ sending Σ onto some connected components of $p^{-1}(0)$. The diffeotopy can be chosen as close as the identity as we want.*

It is not known which hypersufaces are diffeotopic to algebraic ones, see [9, Remark 14.1.1], and this reference for the state of art for these questions.

In the next theorem, for any $x \in \mathbb{R}P^n$, any $d, R > 0$, we denote by $B_{x,R}$ the ball $B(x, \frac{R}{\sqrt{d}})$ defined by the natural distance on $\mathbb{R}P^n$.

Theorem 3.2.4 (G.-Welschinger 2014 [12]) *Let $\Sigma \subset \mathbb{R}^n$ be any compact algebraic hypersurface, not necessarily connected, $R > 0$ and $x \in \mathbb{R}P^n$. Then, there exists $c > 0$,*

$$\forall d \gg 1,\ \Pr\Big((Z(P) \cap B_{x,R}, B_{x,R}) \sim_{diff} (\Sigma, B(0,1))\Big) > c_\Sigma.$$

The notation $(Z, B_x) \sim (\Sigma, B)$ means that there exists a diffeomorphism $\phi : B_x \to B$ such that $\phi(\partial B_x) = \partial B$ and $\phi(Z) = \Sigma$.

Remark 3.2.5

1. *If Σ is not algebraic, we can prove, using Seifert's theorem, the same with the following difference: there can be other components of $Z(P)$ in the ball that the ones associated to Σ. Indeed, if Σ in not algebraic, it can be a bit displaced such that it becomes* part of *the vanishing locus of some polynomial. The difference here are the other components.*

2. *If Σ is a union of circle as in example 3.2.2, then the Theorem says that for any $x \in \mathbb{R}P^n$, with uniform positive probability $Z(P) \cap D(x, 1/\sqrt{d})$ is diffeomorphic to the arrangement Σ. This means that in a way, there is no local random 16th Hilbert problem, since every arrangement arises locally, see Figure 3.1.*

3. *This bound is not surprising: for Kostlan polynomials, the kernel has natural scale $1/\sqrt{d}$ and converges after rescaling, so that we expect that in a ball of size $1/\sqrt{d}$, a universal geometry arises. In particular, the number of connected components, the Betti numbers and so on should be uniformly bounded. Since there are something like $\sqrt{d}^n$ such disjoint balls in $\mathbb{R}P^n$, we have the order $\sqrt{d}^n$.*

3.2.2 On the whole manifold

Denote by $N_\Sigma(P)$ the maximal number of disjoint balls b in $\mathbb{R}P^n$ such that

$$(b \cap Z(P), b) \sim_{diff} (\Sigma, \mathbb{B}^n).$$

We can now give a lower bound of the average number of apparitions of Σ in $Z(P)$:

Corollary 3.2.6 *Under the hypotheses of Theorem 3.2.4, there exists $c > 0$, such that*

$$\forall d \gg 1,\ \mathbb{E} N_\Sigma(P) \geq c\sqrt{d}^n.$$

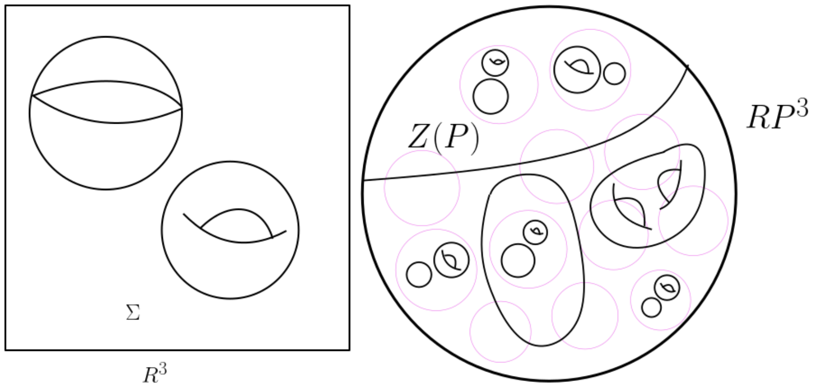


Figure 3.2: Left: $\Sigma \subset \mathbb{R}^3$ is the union of a torus and as sphere, which is an algebraic subset. Right: There exists $c > 0$ such that in average, Σ appears at least $c\sqrt{d}^3$ times in $Z(P)$. We can even assume that Σ appears in a set of disjoint balls $B(x, 1/\sqrt{d})$, represented here by violet spheres.

Remark 3.2.7 *In a the more general setting of holomorphic sections of an ample line bundle over a projective manifold M, see paragraph 1.3, the estimate writes*

$$\mathbb{E}N_\Sigma(P) \geq c_\Sigma \operatorname{Vol} \mathbb{R}M \sqrt{d}^n,$$

where c_Σ depends only on $\Sigma \subset \mathbb{R}^n$ and not on $\mathbb{R}M$.

Proof of Corollary 3.2.6. For any $x \in \mathbb{R}P^n$, denote by A_x the event that $(Z(P) \cap B_x, B_x) \sim (\Sigma, \mathbb{B})$. By Lemma 3.2.8 below, for any d large enough there exists a set $\Lambda_d \subset \mathbb{R}P^n$ such that $B_x \cap B_y = \emptyset$ for $\Lambda_d \ni x \neq y \in t\Lambda_d$, and such that

$$\#\Lambda_d \geq \sqrt{d}^n \frac{\operatorname{Vol}(M)}{2^{n+1} \operatorname{Vol} \mathbb{B}^n}.$$

Then,

$$\mathbb{E}N_\Sigma(P) \geq \mathbb{E} \sum_{x \in \Lambda_d} \mathbf{1}_{A_x} = \sum_{x \in \Lambda_d} \Pr(A_x).$$

By Theorem 3.2.4, there exists $c_\Sigma > 0$, such that for any d large enough independent of x, $\Pr(A_x) \geq c_\Sigma$. This concludes. □

Lemma 3.2.8 *Let (M, g) be a Riemannian smooth compact manifold. For any $\epsilon > 0$, denote by N_ϵ the maximal number of disjoint round balls of size ϵ in M. Then,*

$$\liminf_{\epsilon \to 0} \epsilon^n N_\epsilon \geq \frac{\operatorname{Vol}(M)}{2^n \operatorname{Vol} \mathbb{B}^n}.$$

Proof. Let Λ_ϵ be a maximal set of points on M, such that the balls $B(x, \epsilon)$ centered at points $x \in \Lambda_\epsilon$ are disjoint. Then $M \subset \cup_{x \in \Lambda_\epsilon} B(x, 2\epsilon)$. Indeed, if not, it contradicts the maximal character of Λ_ϵ. This implies

$$\operatorname{Vol} M \leq \sum_{x \in \Lambda_\epsilon} \operatorname{Vol}(B(x, 2\epsilon)).$$

But $\operatorname{Vol} B(x, 2\epsilon) \sim_{\epsilon \to 0} 2^n \operatorname{Vol}(\mathbb{B}^n)$. This can be seen choosing coordinates near x such that $d\operatorname{Vol}(x) = dx_1 \cdots dx_n$ at x, and since M is compact, the asymptotic is uniform in x, so that

$$\operatorname{Vol} M \leq \sum_{x \in \Lambda_\epsilon} 2^n \operatorname{Vol}(\mathbb{B}^n)\big(\epsilon + o(\epsilon)\big)^n,$$

hence the result. □

3.2.3 Betti numbers

Corollary 3.2.6 has itself an interesting corollary about Betti numbers:

Corollary 3.2.9 *For any $i \in \{0, \cdots n-1\}$, there exists $c > 0$,*

$$\forall d \gg 1,\ \mathbb{E}(b_i(Z(P))) \geq c\sqrt{d}^n.$$

For $i = 0$, this corollary was proved first by Lerario and Lunberg [21].

Proof. We first prove that for any $i \in \{0, \cdots n-1\}$, $S^i \times S^{n-i-1}$ can be embedded in $\mathbb{R}^n$ as a compact submanifold. Note first that $\mathbb{R} \times S^i$ embeds in $\mathbb{R}^{i+1}$ by $(t, x) \mapsto e^t x$. This implies that $\mathbb{R}^{n-i} \times S^i = \mathbb{R}^{n-i-1} \times \mathbb{R} \times S^i$ embeds in $\mathbb{R}^{n-i-1} \times \mathbb{R}^{i+1} = \mathbb{R}^n$. In particular, $S^{n-i-1} \times S^i$ embeds in $\mathbb{R}^n$.

Now by Künneth formula,

$$b_i(S^i \times S^{n-i-1}, \mathbb{R}) = \sum_{j=0}^{k} b_j(S^i, \mathbb{R}) b_{k-j}(S^{n-i-1}, \mathbb{R}).$$

If $k \geq 1$ then $b_j(S^k) = 0$ if $j \neq 0$ or k, and $b_0(S^k) = b_k(S^k) = 1$. And $b_0(S^0) = 2$. we obtain $b_i(\Sigma_i) \geq 1$ in any cases. A priori the embedding is not algebraic, but Seifert's theorem shows that a perturbation of it is a connected component of an algebraic hypersurface, so that Corollary 3.2.6 concludes. □

Exercise 3.2.10 *In fact, $S^{n-i-1} \times S^i$ embeds in $\mathbb{R}^n$ as an algebraic hypersurface not only as the perturbation of a connected component of an algebraic hypersurface. For this, use the polynomial*

$$q_i : (x, y) \in \mathbb{R}^{i+1} \times \mathbb{R}^{n-i-1} \mapsto (\|x\|^2 - 2)^2 + \|y\|^2 - 1 \in \mathbb{R}.$$

Question. What about an upper bound for Betti numbers? It does exist:

Theorem 3.2.11 *(G.-Welschinger [13]). For any $i \in \{0, \cdots, n-1\}$, there exists $C > 0$, such that*

$$\forall d,\ \mathbb{E}(b_i(Z(P))) \leq C\sqrt{d}^n.$$

Ideas of the proof. It holds on Morse theory. Let $p : \mathbb{R}P^n \to \mathbb{R}$ be a smooth Morse function, that is with non degenerate critical points. Then almost surely $p_P := p_{|Z(P)}$ is Morse. Morse theory implies that the number of critical points of given index $i \in \{0, \cdots, n-1\}$ (that is, the number of negative eigenvalues of the Hessian at the critical point) is bounded from above by the $b_i(Z(P))$. A Kac-Rice-type formula allows to compute the average of the number of critical points of p_P, and this number grows like $\sqrt{d}^n$, which gives the bound.

Question: Is there true that $\mathbb{E}(b_i)$ has an asymptotic? Answer: it is only known for $i = 0$. This is due to Nazarov and Sodin:

Theorem 3.2.12 *(Nazarov-Sodin 2016[26]) There exists $a > 0$, such that*

$$\mathbb{E}(b_0(Z(P)) \sim_{d\to\infty} a\sqrt{d}^n.$$

In fact, the result is far more general and holds for very general Gaussian fields on compact or non-compact manifolds, at least when their covariance function satisfies certain conditions. They also prove a concentration in probability. As far as concentration is concerned, there is another result in dimension 2 and dimension 1:

Theorem 3.2.13 *(G.-Welschinger [11]) For any degree d, define M the event that $Z(P)$ has the maximal number of connected components in $\mathbb{R}P^2$, that is $b_0(P) = \frac{1}{2}(d-1)(d-2)+1$, up to a linear term in d. Then, there exists a constant $c > 0$, such that*

$$\Pr(P \in M) \leq e^{-cd}.$$

In dimension 1, for any $\epsilon > 0$, there exists $c > 0$,

$$\Pr(b_0(P) \geq \sqrt{d}^{1+\epsilon}) \leq e^{-cd^\epsilon}.$$

These two last theorems use extensively the complex part $Z_{\mathbb{C}}(P)$ of the vanishing locus of P in $\mathbb{C}P^2$ and $\mathbb{C}P^1$.

Exercise 3.2.14 *Let $n \geq 1$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Gaussian field, where $f = (f_1, \cdots, f_n)$. The goal of this problem is, in a simple case, to estimate the number of zeros of f in an bounded open set of $\mathbb{R}^n$. We assume that there exists a Gaussian centered field $g : \mathbb{R}^n \to \mathbb{R}$, such that for any $i \in \{1, \cdots, n\}$, f_i is a copy of g, and all the f_i's are independent. We denote by $e : (\mathbb{R}^n)^2 \to \mathbb{R}$ the covariance function associated to g, and we assume that e is C^1.*

1. *For any $x \in \mathbb{R}^n$, write the variance matrix of the vector $f(x)$ as related to e.*

2. *For any $x \in \mathbb{R}^n$, compute the density $\phi_{f(x)}(0)$ as a function of e.*

3. *For any $x \in \mathbb{R}^n$, denote by $Df(x) \in M_n(\mathbb{R})$ the Jacobian matrix of f at x in the canonical basis. Write the variance matrix of the vector $Df(x)$ as a function of e and its derivative. In this question we consider that a matrix in $M_n(\mathbb{R})$ is a vector $(m_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n^2}$, however we keep the notations $1 \leq i,j \leq n$ instead of $1 \leq k \leq n^2$.*

4. *Write $Cov(f(x), Df(x))$.*

   Til the end of the problem we assume that there exists a smooth $\rho : \mathbb{R}^+ \to \mathbb{R}$ such that

   $$\forall x, y \in \mathbb{R}^n,\ e(x,y) = \rho(-\frac{1}{2}\|x-y\|^2).$$

5. *Write $Varf(x)$, $VarDf(x)$ and $Cov(f(x), Df(x))$ as a function of ρ and its derivatives.*

6. *Let $U \subset \mathbb{R}^n$ be an open subset and*

   $$N(f,U) := \#\{x \in U, f(x) = 0\}.$$

   Show that

   $$\mathbb{E}N(f,U) = cVol(U)^\alpha \rho(0)^\beta \rho'(0)^\gamma,$$

   with α, β, γ and c constants which depend only on n. Give the values of α, β and γ and write c as an integral over $M_n(\mathbb{R})$.

7. *Let $M \in M_n(\mathbb{R})$, and $C_1, \cdots, C_n$ its column vectors. Show that*

$$|\det M| = \prod_{i=1}^{n} \|C_i^{\perp}\|,$$

where for any $i \geq 2$, $C_i^{\perp}$ denotes the orthogonal projection onto $Vect^{\perp}(C_1, \cdots, C_{i-1})$, and $C_1^{\perp} = C_1$.

8. *Let $M \in M_n(\mathbb{R})$ be random, such that its coefficients are independent and follow a normal law $N(0,1)$.*

   (a) *Show that*

   $$\mathbb{E}(|\det M|) = \prod_{k=1}^{n} \mathbb{E}(\|X_k\|),$$

   where for any k, $X_k \in \mathbb{R}^k$ is a random vector which coefficients are independent and follow a normal law $N(0,1)$.

   (b) *Show that $\mathbb{E}(\|X_k\|) = \sqrt{2\pi}\frac{Vol(S^{k-1})}{Vol(S^k)}$.*

   (c) *Deduce $\mathbb{E}(|\det M|)$ as a function for $Vol(S^n)$.*

3.3 Proof of theorem 3.2.4

Proof of Theorem 3.2.4. By the invariance Lemma 3.1.2, it is enough to prove the theorem for $x = [1 : 0 : \cdots : 0]$, since the property is invariant under isometries of $\mathbb{R}P^n$. Then there is a local chart

$$\phi : [X_0 : \cdots : X_n] \mapsto (\frac{X_1}{X_0}, \cdots, \frac{X_n}{X_0}) = (x_1, \cdots, x_n).$$

Let $q \in \mathbb{R}[x_1, \cdots, x_n]$ be a polynomial such that some of its components are close to Σ. After an homothetic transformation and a translation, we can assume that $\Sigma \subset B(0,1)$. Let

$$q_d(x) := q(x\sqrt{d}).$$

Note that some of the components of $q_d = 0$ are diffeotopic in $B(0, 1/\sqrt{d})$ to $\Sigma/\sqrt{d}$. Then

$$Q_d := X_0^d q_d(\frac{X_1}{X_0}, \cdots, \frac{X_n}{X_0}) \in \mathbb{R}^d_{hom}[X_0, \cdots, X_n],$$

and its vanishing locus contains a copy of Σ in $\phi^{-1}(B(0, 1/\sqrt{d}))$.

Now define $H_Q := Q_d^{\perp}$ the orthogonal space to Q_d in $\mathbb{R}^d_{hom}[X], \langle , \rangle$. We would like to use a decomposition for our random polynomials adapted to Q_d an H_Q. However, we must compute the L^2-norm of Q.

Lemma 3.3.1

$$\|Q\| \sim_{d \to \infty} d^{-n/2} \|q\|_{\exp},$$

where $\|q\|^2_{\exp} := \frac{1}{\pi^n} \int_{\mathbb{C}^n} |q|^2(w) e^{-|w|^2} dw$.

Proof. We have

$$\|Q\|_2^2 = \int_{\mathbb{C}^{n+1}} |Z_0|^{2d} |q|^2 (\sqrt{d}\frac{Z'}{Z_0}) e^{-\|Z\|^2} dZ,$$

where $Z' = (Z_1, \cdots, Z_n)$. We use the change of variable $(W_0, w) = (Z_0, \frac{Z'}{Z_0})$, then $(w_0, w) = (W_0\sqrt{1+|w|^2}, w)$, and finally $y = \sqrt{d}w$ so that

$$\begin{aligned}
\|Q\|_2^2 &= \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} |W_0|^{2(d+n)} |q(\sqrt{d}w)|^2 e^{-|W_0|^2(1+\|w\|^2)} dW_0 dw \\
&= \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} |w_0|^{2(d+n)} |q(\sqrt{d}w)|^2 e^{-|w_0|^2} \frac{1}{(1+\|w\|^2)^{d+n+1}} dw_0 dw \\
&= d^{-n} \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} |w_0|^{2(d+n)} |q(w)|^2 e^{-|w_0|^2} \frac{1}{(1+\frac{1}{d}\|w'\|^2)^{d+n+1}} dw_0 dw' \\
&\sim_{d\to\infty} d^{-n} \frac{1}{(d+n)!\pi^{n+1}} \int_{\mathbb{C}} |z|^{2(d+n)} e^{-|z|^2} dz \int_{\mathbb{C}^n} |q|^2(w) e^{-|w|^2} dw \\
&= d^{-n} \|q\|^2_{\exp}.
\end{aligned}$$

□

We follow the barrier method of [25, p. 1343]. Since the random polynomial can be written in any fixed orthonormal basis, we can decompose our random polynomial P as

$$P = a\frac{Q_d}{\|Q_d\|} + R, \tag{3.3.1}$$

where $a \sim N(0,1)$ and $R \in Q^\perp$ is Gaussian random for the induced law on $Q^\perp$ and independent of a. More explicitly, $R = \sum_i b_i \phi_i$ for $(\phi_i)_i$ an ONB of $Q^\perp$ and the b_i are independent and independent of a. We want to prove that with uniform positive lower bound, R does not perturb too much the first term, such that the former still vanishes on a hypersurface diffeomorphic to Σ. Hence, we need to know when the vanishing locus of a perturbation of a function gives a diffeotopic perturbation of the vanishing locus of the function.

Proposition 3.3.2 *Let $f : \bar{B}(0,1) \to \mathbb{R}$ be a C^1-function such that there exists $a, \eta > 0$, such that :*

1. *(compact vanishing locus) $|f|_{\partial B(0,1)} \geq a$.*
2. *(transversality) $\forall x \in B(0,1),\ |f(x)| < a \Rightarrow \|df(x)\| > \eta$.*

Then for any $g \in C^1(\bar{B}(0,1))$ such that $|g|_\infty < a/2$ and $\|dg\|_{L^\infty(\mathbb{B})} < \eta/2$, $\big((f+g)^{-1}(0), \bar{B}(0,1)\big)$ is diffeotopic to $\big(f^{-1}(0), \bar{B}(0,1)\big)$.

The conclusion means that is there exists :

$$\psi : \bar{B}(0,1) \times [0,1] \to \bar{B}(0,1)$$

such that

1. ψ is continuous in the second variable
2. $\forall t \in [0,1],\ \psi(\cdot, t)$ is a diffeomorphism
3. $\psi(\cdot, 0) = Id$

4. $\psi(\cdot,1)\big(f^{-1}(0)\big) = (f+g)^{-1}(0)$.

Definition 3.3.3 *Let $k \geq 1$, M be a smooth manifold and $a, \eta > 0$. Then, a C^1 mapping $f : M \to \mathbb{R}^k$ is said to be (a,η)-transverse to 0, if whenever $|f(x)| \leq a$, then $df(x)$ has an right inverse whose norm is smaller than η^{-1}.*

We prove the more general :

Proposition 3.3.4 *Let $k \geq 1$, M be a smooth manifold, and $f : M \times [0,1] \to \mathbb{R}$ such that $\forall t \in [0,1], f_t := f(\cdot,t) : M \to \mathbb{R}$ is a submersion. Then, there is an isotopy $(\phi_t)_t$ from $Z(f_0)$ to $Z(f_t)$.*

Proof. The condition implies that f is a submersion, so that $Z(f) \subset M \times [0,1]$ is a smooth submanifold. Moreover, $\pi : Z(f) \to [0,1]$ is a submersion. Indeed,

$$\forall (x,t) \in Z(f), T_{x,t}Z(f) = \{(X,\tau) \in T_X M \times \mathbb{R}, -\partial_t f(x,t)\tau = df_t(x)(X)\}.$$

Let

$$X_t := -\frac{\nabla f_t(x)}{\|\nabla f_t\|^2}\partial_t f(x,t).$$

Then, X_t is a smooth vector field on M such that $(1, X_t) \in TZ(f)$. **finir** □

Proof of Proposition 3.3.2. □

We want to apply this to

$$f := a\frac{Q_d}{X_0^d\|Q_d\|} \text{ and } g := \frac{R}{X_0^d}$$

in our previous decomposition (3.3.1) of the random P and in the projective coordinates $(x_1, \cdots, x_n)$. Then, $f = a\|q\|_{\exp} d^{n/2} q(\sqrt{d}x)$. This implies that there is $c > 0$,

$$|f|_{|\partial B_x} \geq acd^{n/2}$$

since q does not vanish on $\partial B(0,1)$. Moreover,

$$df(x) = a\|q\|_{\exp} d^{n/2} d^{-1/2} dq(\sqrt{d}x).$$

since q vanishes transversally on $B(0,1)$, there exists $c' > 0$, such that

$$|f(x)| < ac'd^{n/2} \Rightarrow |df| > ac'd^{n/2}d^{-1/2}.$$

For $g := R/X_0$, we write it

$$g = p_1 + p_2 := \frac{1}{2}(g+f) + \frac{1}{2}(g-f).$$

Note that the law of $p_1 := g + f$ is the same of a general polynomial p, as well as for $p_2 := g - f$. We will use the trivial

$$\mathbb{E}\|g\| \leq \frac{1}{2}(\mathbb{E}\|p_1\| + \mathbb{E}\|p_2\|) \leq \mathbb{E}(\|p\|),$$

and similarly for the average of the derivative of g, where p is a general polynomial. Hence, it is enough too bound from above the norms of a general polynomial p. We have, for $p = p_1$ or p_2,

$$p = \sqrt{\frac{(d+n)!}{d!}}\big(a_{d0\cdots 0} + \sum_{k=1}^{n} a_{(d-1)0\cdots 1\cdots 0}\sqrt{d}x_k + O(\|\sqrt{d}x\|^2)\big).$$

For the C^0 bound, we write it

$$d^{n/2}\big(a_{d0\cdots 0} + O(\|\sqrt{d}x\|)\big),$$

so that

$$\mathbb{E}(|g|_\infty) \leq d^{n/2}(\mathbb{E}(|a|) + C) \leq C' d^{n/2}.$$

Now

$$dp = \sqrt{\frac{(d+n)!}{d!}}\sqrt{d}\big(\sum_{k=1}^{n} a_{(d-1)10\cdots 0}dx_k + O(\sqrt{d}x)\big),$$

so that on the ball $B(0, 1\sqrt{d})$,

$$\begin{aligned}\mathbb{E}(|dg|) &\leq d^{n/2}\sqrt{d}\sum\big(\mathbb{E}(|a_{(d-1)0\cdots 1\cdots 0}|) + C\big)\\ &\leq Cd^{n/2}\sqrt{d}.\end{aligned}$$

Lemma 3.3.5 *(Tchebychev Inequality) Let X be a non negative random variable. Then for any $M > 0$,*

$$\Pr(X > M) \leq \frac{\mathbb{E}(X)}{M}.$$

Proof.

$$\mathbb{E}(X) = \int_{\mathbb{R}} X d\mu(X) \geq \int_{X>M} X d\mu(X) \geq \int_{X>M} M d\mu(X) = M\Pr(X > M).$$

□

Now, for any $M > 0$, by what have been said,

$$\begin{aligned}\Pr(B_x \text{ contains a } \Sigma) &\geq \Pr\Big(|a|cd^{n/2} > Md^{n/2} \text{ and } |g|_\infty < \frac{M}{2}d^{n/2}\\ &\quad \text{ and } |a|cd^{n/2}d^{n/2}\sqrt{d} > M\sqrt{d} \text{ and } |g|_\infty < \frac{M}{2}d^{n/2}\sqrt{d}\Big)\\ &= \Pr\Big(|a|cd^{n/2} > Md^{n/2} \text{ and } |a|cd^{n/2}d^{n/2}\sqrt{d} > Md^{n/2}\sqrt{d}\Big)\\ &\quad \Pr\Big(|g|_\infty < \frac{M}{2}d^{n/2} \text{ and } |g|_\infty < \frac{M}{2}d^{n/2}\sqrt{d}\Big)\\ &\geq 2\int_{M/c}^{+\infty} e^{-\frac{1}{2}a^2}\frac{da}{\sqrt{2\pi}}(1 - 2C'/M - 2C/M).\end{aligned}$$

For M large enough, this is larger than $c_M > 0$ independently of d. □

Chapter 4

Percolation and random analytic nodal lines

4.1 Introduction

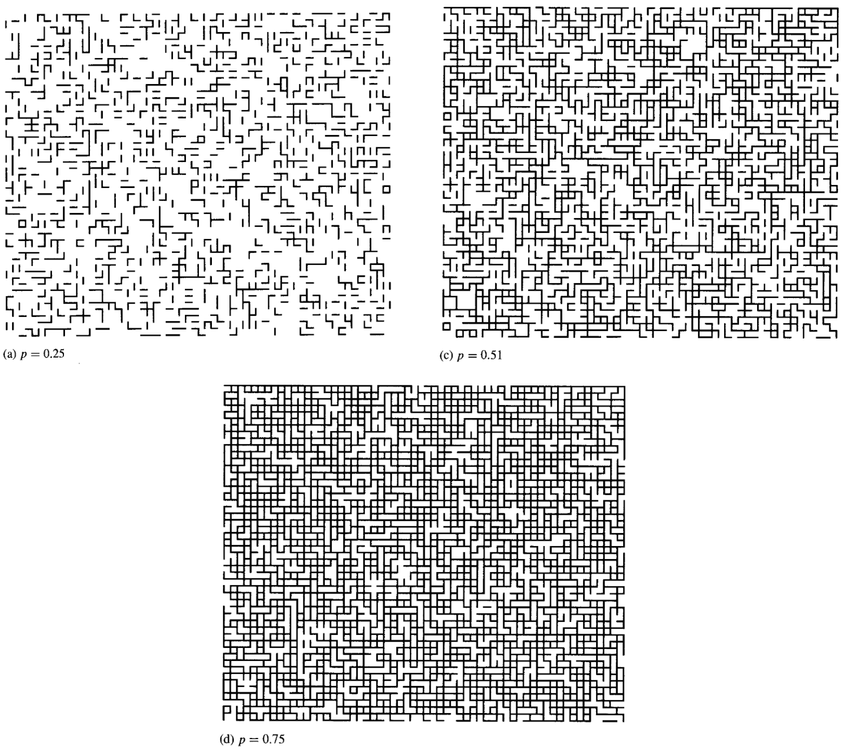


Figure 4.1: A realization of bond percolation for $p = 0.25$, $p = 0.51$ and $p = 0.75$ on the integer lattice. We see the existence of a large connected component in the two latter cases.

4.1.1 Bernoulli bond percolation

Let $\mathbb{Z}^2$ be the integer lattice over the real plane, and denote by $\mathcal{E}$ its set of edges, and by $\mathcal{V}$ its set of vertices. The so-called *Bernoulli bond percolation* is the following random geometrical model: With probability $p \in [0,1]$, color independently in black (open) every edges. We would like to know if there exists large connected components formed by the black set of edges. The question is a bit blur. Let us be more specific. Choose a rectangle $R \subset \mathbb{R}^2$. What is the probability that there is a black crossing of R in its length, that is a continuous arc in R through black edges from the left side to the right side of R? We will denote $\mathrm{Cross}(R)$ the event that there is such a crossing. In these lecture notes we will work at *criticality*, that is we make $p = 1/2$. We will explain later the sense of this term.

Theorem 4.1.1 (Russo [31], Seymour-Welsh [33] 1978) *Fix $R \subset \mathbb{R}^2$ a rectangle. Then, there exists $c > 0$, such that for*

$$n \geq 1, \Pr(\mathrm{Cross}(nR)) \geq c.$$

4.1.2 Site percolation

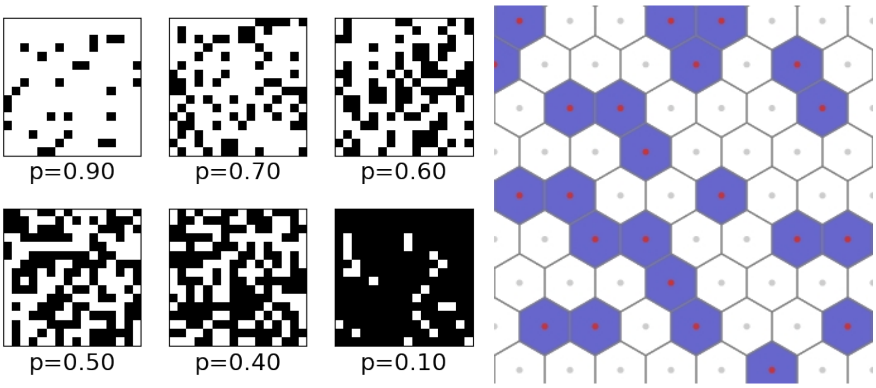


Figure 4.2: Left: examples of site percolation on the integer lattice in a box. Note that the dual faces are black iff the associate vertex is positive, so that an open edge correspond to two positive adjacent squares. Right: an example of site percolation on the triangular lattice, represented on the dual lattice, that is the hexagonal one.

In the results we present, the so-called *site percolation* will be preferred.

Definition 4.1.2 *Let $\subset \mathbb{R}^2$ be a planar lattice, $\mathcal{V}$ its set of vertices, and $\mathcal{E}$ its set of edges. Let $f : \mathcal{V} \to \{-1, 1\}$ be a random sign over the vertices. For any rectangle R, we say that* R is positively crossed *if there exists a continuous path in $\mathcal{E}$, such that for any edge in the path, $f = +1$ at its extremities.*

Often the site percolation is represented by the dual lattice: the positive vertices of $\mathcal{T}$ correspond to positive faces of $\mathcal{T}^*$, and an open edge correspond to two adjacent faces. For the triangular lattice, the dual lattice is the hexagonal lattice.

Theorem 4.1.3 (Russo [31], Seymour-Welsh [33] 1978) *Let $\mathcal{T}$ be a planar triangulation lattice invariant by orthogonal rotations and horizontal symmetry. Fix $R \subset \mathbb{R}^2$ a rectangle. Then, for the site Bernoulli site percolation on $\mathcal{T}$ at criticality $p = 1/2$, there exists $c > 0$, such that for*

$$n \geq 1, \Pr(\mathrm{Cross}(nR)) \geq c.$$

In fact, the theorem holds for more general lattices, like the triangular lattice. **le rÃľsultat gÃľnÃľral ?**

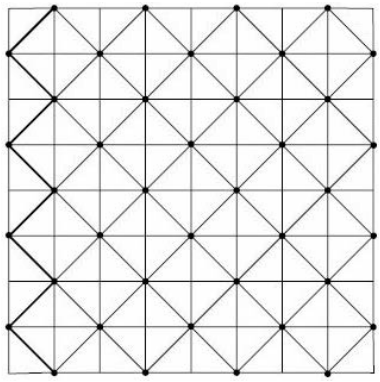


Figure 4.3: The Union-Jack lattice, an example of triangulation invariant under $\pi/2$-rotations, horizontal and vertical integer translations, and horizontal symmetries.

4.1.3 Continuous fields and percolation

It is very natural to ask the following question:

Question 4.1.4 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a random function, such that almost surely, f is smooth, $f^{-1}(0)$ is smooth, such that the measure of f is invariant under the symmetries of $\mathbb{R}^2$ and under the action $f \mapsto -f$. Does $f^{-1}(\mathbb{R}^+)$ satisfies a Russo-Seymour-Welsh property, that is: is it true that for any rectangle $R \subset \mathbb{R}^2$, there exists $c > 0$, such that*

$$\forall n \gg 1,\ \Pr(\mathrm{Cross}^+(nR)) > c?$$

Here $\mathrm{Cross}^+(R)$ is the event that a connected component of $f^{-1}(\mathbb{R}^+) \cap R$ intersects the two small sides of R.

Note that by the symmetry $f \mapsto -f$, the answer is identical for the negative crossings $\mathrm{Cross}^-(R)$. It is very natural to precise the question for random Gaussian fields. In this direction, a negative result was proved in 1996 :

Theorem 4.1.5 (Alexander 1996 [2]) *Assume that the correlation function of f depends only on the distance between the two points, and is non-negative. Then, almost surely, all the connected components of $f^{-1}(\mathbb{R}^+)$ are bounded.*

We will explain the following positive result:

Theorem 4.1.6 (Beffara-G. 2017 [4]) *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a Gaussian field whose correlation e function is smooth, depends only on the distance between points and is non negative. If there exists $C > 0$, such that $\forall e(x, y) \leq C\|x - y\|^{-D}$, with $D > 12$, then*

$$\forall n \geq 1,\ \Pr(\mathrm{Cross}^+(nR)) \geq c$$

and the same is true for nodal crossings, that is for $\mathrm{Cross}^0(nR)$, that is for the connected components of $f^{-1}(0)$, for n large enough.

In the original article, $D = 325$. Then, Belyaev and Muirhead proved it for $D = 16$ [5] and Rivera and Vanneuville proved it for $D = 4$ [30]. They all simplified the proof. Note that if the field is constant and almost surely not the function zero over $\mathbb{R}^2$, then almost surely it does not vanish, so there is no percolation of nodal lines. Even with the latter decorrelation, over smaller and smaller balls, the sign of the function f is constant and f does not vanish.

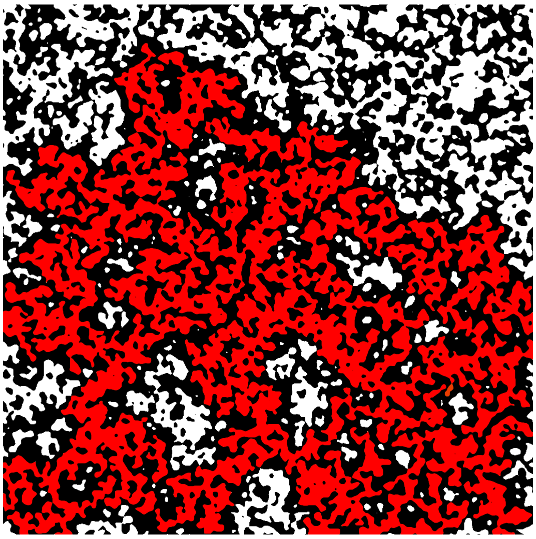


Figure 4.4: Percolation and Bargmann-Fock random functions: in black, the set $f^{-1}(\mathbb{R}^+)$ and in red, a connected component of $f^{-1}(\mathbb{R}^-)$ highlighted.

Example 4.1.7 *The Bargmann-Fock model. The kernel of the BF model is* $e(x,y) = \exp(-\|x-y\|^2)$*. It is positive and decreases exponentially fast, so that Theorem 4.1.6 holds.*

Let P be of degree d, and consider $\pi := \frac{P}{\|X\|^d}$, which vanishes on the same hypersurface as P and a well defined smooth function on $\mathbb{R}P^2$. If P is chosen at random for the Kostlan measure, denote by e_d the associate kernel for π.

Lemma 4.1.8 *The BF kernel is the rescaled limit of the kernel for Kostlan polynomials. More precisely, for any $x_0 \in \mathbb{R}P^n$, in the affine coordinates,*

$$e_d(x_0 + \frac{x}{\sqrt{d}}, x_0 + \frac{y}{\sqrt{d}}) \to \exp(-\|x-y\|^2).$$

Proof. By symmetry of the problem, we can assume that $x_0 = [1 : 0 \cdots : 0] = 0$ in the the affine coordinates. Then

$$\pi = \frac{p}{(1+\|x\|^2)^{d/2}},$$

where p is the affine polynomial associated to P in the chart $X_0 \neq 0$. The associate covariance function writes

$$e_d(x,y) = \frac{(1+\langle x,y\rangle)^d}{(1+\|x\|^2)^{d/2}(1+\|y\|^2)^{d/2}}$$

so that

$$e_d(\frac{x}{\sqrt{d}}, \frac{y}{\sqrt{d}}) =\to_{d\to\infty} e^{\frac{1}{2}(2\langle x,y\rangle - \|x\|^2 - \|y\|^2)} = e^{-\frac{1}{2}\|x-y\|^2}.$$

This Gaussian correlation function has two important features: it is non negative, and it decreases very fast. □

Remark 4.1.9 1. *In fact, the BF kernel limit is* very *universal: it holds for the general setting for holomorphic sections over projective manifolds, see paragraph 1.3.1. The 2-point correlation function for random holomorphic sections is the well-known* Bergman kernel *and is very important in complex analysis and complex geometry. The Szegö kernel used in [34] is essentially the same.*

2. *A natural question is: is there a similar result for Kostlan polynomials? That is, for any topological rectangle on $\mathbb{R}P^2$, is it true that there exists $c > 0$, such that for any degree d large enough, with probability at least c there exists a nodal line crossing this rectangle? The answer is yes, and has been proved by Dmitry Beliaev, Stephen Muirhead and Igor Wigman in [7].*

3. *What about taking upper level sets with constants different than zero? That is, what happens for the sets $f^{-1}([\epsilon, +\infty[)$? Alejandro Rivera and Hugo Vanneuville answered to this question [29]: for $\epsilon < 0$, with probability one there is one unique connected component of $f^{-1}([\epsilon, +\infty[)$, see figure 4.5*

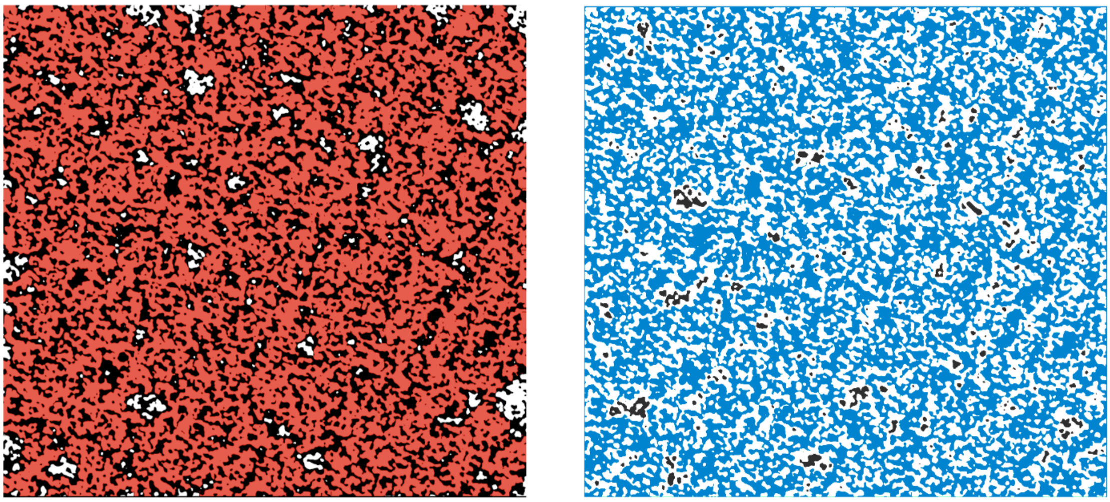


Figure 4.5: Bargmann-Fock random functions [29]. Left: the region $f \geq 0.1$ is colored in black, the small components of the region $f < 0.1$ are colored in white, and the giant component of the region $f < 0.1$ is colored in red. On the right hand side, the small components of the region $f \geq -0.1$ are colored in black, the giant component of the region $f \geq -0.1$ is colored in blue, and the region $f < -0.1$ is colored in white. The two pictures correspond to the same sample off.

4.2 A metatheorem

4.2.1 Statement

We will use a general result proved by Vincent Tassion in 2016, which gives in particular a proof of the Russo and Seymour-Welsh theorem 4.1.1, was written for Voronoi percolation, but holds in a very general situations:

Theorem 4.2.1 (Tassion [36]) *Let $f : \mathbb{R}^2 \to \{-1, 1\}$ be a random sign function over the plane, satisfying the following three conditions :*

- *(Symmetries) The measure is invariant under $f \mapsto -f$, under $\pi/2$-rotation and horizontal axis, and translations.*
- *(Crossing squares) There exists $c > 0$, such that for any horizontal square R,*

$$\Pr(\mathrm{Cross}^+(R)) \geq c.$$

- *(Positive correlation) For any two positive crossing events A and B in a bounded open set,* $\Pr(A \cap B) \geq \Pr A \Pr B$.

1. *Then for any rectangle R, there exists $c > 0$, such that*

$$\limsup_{n \to \infty} \Pr(\mathrm{Cross}^+(nR)) \geq c.$$

2. *If furthermore f satisfies the following additional condition:*

   - *(quasi-independency) For any crossing event A in the annulus $[-2n, 2n]^2 \setminus [-n, n]^2$ and B in $[-n \log n, n \log n]^2 \setminus [-3n, 3n]^2$,*

$$\left| \Pr(A \cap B) - \Pr(A) \Pr(B) \right| \to_{n \to 0} 0,$$

   then for any rectangle R, there exists $c > 0$,

$$\forall n \geq 1, \ \Pr(\mathrm{Cross}^+(nR)) \geq c.$$

4.2.2 The square crossings

Lemma 4.2.2 *In the following three percolation cases, the squares are uniformly crossed:*

1. *For the critical (that is $p = 1/2$) bond correlation on the lattice integer,*

2. *for the critical (that is $p = 1/2$) site percolation on a triangulation lattice invariant by $\pi/2$-rotation,*

3. *and for the continuous positive percolation for a continuous Gaussian field whose covariance depends only on the distance.*

Proof. For the first case, the probability is "morally" $1/2$. Indeed if there is no horizontal positive crossing, then there is a vertical negative crossing on the dual integer lattice, which is made of the centers of the faces of $\mathbb{Z}^2$, and of the translated of the edges of $\mathbb{Z}^2$ that link these vertices. I wrote "morally" because they are boundary effects that tend to zero when the size of the square grows to infinity.

For the second case, topology says that if there is no horizontal positive crossing, then there is a vertical negative crossing, see [18, pp. 52–53]. Note that this is false for the square lattice, which is not a triangulation. By the symmetries between $+$ and $-$ and the rotation, the crossing probability is $1/2$. Again, there are boundaries effects.

In the third case, the probability of crossing any square is exactly $1/2$, using the latter proof. □

4.2.3 The Fortuin Kasteleyn Ginibre condition

In this paragraph, we deal the positive correlation condition, in particular for the sign of a Gaussian vector. In fact, this condition is a particular case of the so-called *Fortuin-Kasteleyn-Ginibre* property, which is one of the most useful in percolation theory.

Definition 4.2.3 1. *A function $f : \mathbb{R}^N \to \mathbb{R}$ is called* increasing *if it is increasing in each of the coordinates.*

2. *An event A depending on the value of a vector $X \in \mathbb{R}^N$ is said* increasing *if $\mathbf{1}_A$ is increasing for the latter definition.*

3. *A random function* $X : \mathbb{R}^N \to \mathbb{R}$ *is said to* satisfy the FKG condition *if for any pair* $f : \mathbb{R}^N \to \mathbb{R}$ *of increasing functions,*

$$\mathrm{Cov}\,\big(f(X)g(X)\big) := \mathbb{E}\big(f(X)g(X)\big) - \mathbb{E}\big(f(X)\big)\mathbb{E}\big(g(X)\big) \geq 0.$$

Example 4.2.4 1. *If X satisfies the FKG condition and if A, B are two increasing events, then*

$$\Pr(A \cap B) \geq \Pr(A)\Pr(B).$$

2. *Let $(\mathcal{T}, \mathcal{V})$ be a lattice and its set of vertices, R be a rectangle, $N = \#R \cap \mathcal{V}$, $x = (x_1, \cdots, x_N)$ an enumeration of $R \cap \mathcal{V}$, $f : \mathbb{R}^N \to \{0,1\}$ be defined by $f(x) = 1$ if x produces a positive crossing for the site percolation of R, and $f(x) = 0$ in the contrary case. Then, f is an increasing function.*

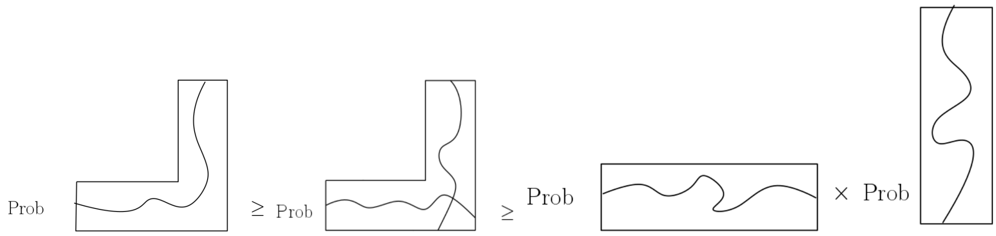


Figure 4.6: When the random coloring satisfies the FKG condition, the probability of crossing an "L" is greater than the squared probability of crossing one of the rectangles that make the L.

Proposition 4.2.5 *Let $(\mathcal{T}, \mathcal{V})$ be as in example 4.2.4, R and R' two rectangles, disjoint or not. Let $X = (X_1, \cdots, X_N)$ be a random coloring of $R \cap R' \cap \mathcal{V}$ satisfying the FKG condition. Then,*

$$\Pr(\mathrm{Cross}(R) \cap \mathrm{Cross}(R')) \geq \Pr(\mathrm{Cross}(R))\Pr(\mathrm{Cross}(R')).$$

Proof. Let $f : \mathbb{R}^N \to \{0,1\}$ (resp. g) be defined by $f(x) = 1$ (resp. $g(x) = 1$) if R (resp. R') is positively crossed. Then example 4.2.4 gives the result. □

Theorem 4.2.6 (Lauren Pitt 1982 [28]) *Let $X \in \mathbb{R}^N$ be a random Gaussian vector with a zero mean and non-singular variance matrix. Then, X satisfies FKG its covariance matrix has all its coefficient non negative.*

Corollary 4.2.7 *Bernoulli percolation for $p = 1/2$ satisfies the FKG condition.*

Proof. For any $p \in [0,1]$, let $Y \in \mathbb{R}$ be a normal law, so that $\Pr(Y \geq 0) = 1/2$. For any rectangle R, the sign at every vertex in $\mathcal{V} \cap R$ for Bernoulli can be seen as the sign of N independent copies of Y. Then, theorem 4.2.6 applies. □

As a consequence of the Tassion theorem 4.2.1 and Pitt theorem 4.2.6 we obtain a proof of the Russo-Seymour-Welsh for site percolation on lattices with good symmetries:

Proof of Theorem 4.1.3 and 4.1.1. The symmetry condition of Theorem 4.2.1 are insured by the hypotheses on the lattice and the probability 1/2. The square crossing condition is given by Lemma 4.2.2. Corollary 4.2.7 gives the positive correlation for the site percolation. The proof for bond percolation is similar. The last condition of quantitative dependency is trivial since this is Bernoulli percolation. □

As a consequence of the first part of Tassion theorem 4.2.1 and Pitt theorem 4.2.6 we obtain:

Corollary 4.2.8 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a centered Gaussian field whose correlation function depends only on the distance between points and is non negative, and such that f is C^0 almost surely. Then for any rectangle R, there exists $c > 0$, such that*

$$\limsup_{n\to\infty} \Pr(\mathrm{Cross}^+(nR)) \geq c.$$

Proof. The symmetry condition is induced by the hypothesis, and the square crossing by Lemma 4.2.2. The only thing to prove is the positive correlation. Pitt's theorem holds for finite dimension Gaussian vectors, but we prove now that it implies positive correlation for crossing events. Let U be an bounded subset of $\mathbb{R}^2$ and A, B two crossing events.

For any $n \in \mathbb{N}^*$ let $f_n : \frac{1}{n}\mathbb{Z}^2 \cap U \to \{-1, 1\}$ be the restriction of the sign of f over the refined lattice $\frac{1}{2n}\mathbb{Z}^2$. Let A_n and B_n be the crossing events for the associated site percolation, see definition 4.1.2.

Then, $A \subset \cup_k \cap_{n\geq k} A_n$. Indeed, fix $f \in A$ and $k \in \mathbb{N}^*$. By (uniform) continuity of f (on U), there exists $k \in N$ and a positive A-crossing continuous path $\gamma \subset f^{-1}(\mathbb{R}^+)$, such that its $\frac{2}{k}$−neighborhood lies in $f^{-1}(\mathbb{R}^+)$. This implies that $f \in A_n$ for any $n \geq k$. Similarly, if $f \in A^c$, then there is a "dual" negative crossing f and, for the same reasons, $A^c \subset \cup_k \cap_{n\geq k} A_n^c$. This implies that for any $f \in \Omega$, $\mathbf{1}_{A_n}(f) \to_{n\to\infty} \mathbf{1}_A(f)$ almost surely, so that $\Pr A_n \to \Pr A$ by the dominated convergence theorem, and similarly for B_n and B, and for $A_n \cap B_n$ and $A \cap B$.

Now, since $\tilde{A}_n$ and $\tilde{B}_n$ are increasing events, by Theorem 4.2.6,

$$\Pr(A \cap B) = \lim_{n\infty} \Pr(A_n \cap B_n) \geq \Pr(\tilde{A}_n)\Pr(\tilde{B}_n) \to \Pr(A \cap B).$$

□

If we want now to obtain the second part of theorem 4.2.1, we need to insure quantitative dependency. We will get a uniform lower bound for the probabilities of crossing instead of an inf limit and get theorem 4.1.6.

4.2.4 Proof of Pitt's theorem

First, we will give a general relation for Gaussian fields which will be useful for two applications : Pitt's theorem 4.2.6 and the independence theorem 4.3.5.

Proposition 4.2.9 *Let $S_0 = (\sigma^0_{ij})_{1\leq i,j\leq N}$ and $S_1 = (\sigma^1_{ij})_{1\leq i,j\leq N}$ be two non-singular covariance matrices and $X_0 \in \mathbb{R}^N$, $X_1 \in \mathbb{R}^N$ two associated Gaussian vectors. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a C^2 function, such that there exists N, $\max(|f(x)|, \|df(x)\|) = O_{\|x\|\to\infty}(\|x\|^N)$. Then,*

$$\mathbb{E}(f(X_1)) - \mathbb{E}(f(X_0)) = \sum_{1\leq i\leq j\leq N} \frac{1}{2^{\delta_{ij}}}(\sigma^1_{i,j} - \sigma^0_{i,j}) \int_0^1 \int_{x\in\mathbb{R}^N} \frac{\partial^2 f}{\partial x_i \partial x_j} \phi_{X_t}(x) dx dt,$$

where $x = (x_1, \cdots, x_n)$ and $\phi_{X_t}(x)$ is the density associated to the covariance $(1-t)S_0 + tS_1$.

Proof. For any $t \in [0,1]$ define the non negative covariance matrix

$$S_t := (1-t)S_0 + tS_1 = (\sigma^t_{ij})_{i,j} = \big((1-t)\sigma^0_{ij} + t\sigma^1_{ij}\big)_{i,j}$$

and X_t an associated Gaussian vector. Note that the hypothesis on the non-singularity implies that S_t is uniformly non-singular, that is the eigenvalues of S_t are uniformly bounded in t from below by a positive constant. We have, using the chain rule,

$$\begin{aligned}
\mathbb{E}(f(X_1)) - \mathbb{E}(f(X_0)) &= \int_0^1 \frac{d}{dt}\mathbb{E}f(X_t)dt \\
&= \int_0^1 \frac{d}{dt}\int_{x\in\mathbb{R}^N} f(x)\phi_{X_t}(x)dxdt \\
&= \sum_{i\le j}\int_0^1\int_{x\in\mathbb{R}^N} f(x)\frac{\partial}{\partial\sigma^t_{ij}}\phi_{X_t}(x)\frac{d\sigma^t_{i,j}}{dt}dxdt \\
&= \sum_{i\le j}(\sigma^1_{ij}-\sigma^0_{ij})\int_0^1\int_{x\in\mathbb{R}^N} f(x)\frac{\partial}{\partial\sigma^t_{ij}}\phi_{X_t}(x)dxdt.
\end{aligned}$$

Now we use a important relation, which proof is left as an exercise for the reader.

Lemma 4.2.10 *Let $\Sigma = (\sigma_{ij})_{1\le i,j\le N}$ be a covariance matrix, $X \in \mathbb{R}^N$ an associated Gaussian vector, and $\phi_X(x)$ its density. Then,*

$$\forall x \in \mathbb{R}^N,\ \forall 1 \le i \ne j \le N,\ \frac{\partial\phi(x)}{\partial\sigma_{ij}} = \frac{\partial^2\phi(x)}{\partial x_i\partial x_j} \quad \text{and } \forall 1\le i\le N,\ \frac{\partial\phi(x)}{\partial\sigma_{ii}} = \frac{1}{2}\frac{\partial^2\phi(x)}{\partial x_i^2}.$$

We can finish the proof or Proposition 4.2.9. Utilizing Lemma 4.2.10, we obtain

$$\mathbb{E}(f(X_1)) - \mathbb{E}(f(X_0)) = \sum_{i\le j}\frac{1}{2^{\delta_{ij}}}(\sigma^1_{ij}-\sigma^0_{ij})\int_0^1\int_{x\in\mathbb{R}^N} f(x)\frac{\partial^2\phi_{X_t}(x)}{\partial x_i\partial x_j}dxdt.$$

Since S_t is uniformly non-singular, the density decreases uniformly exponentially fast for $x \to \pm\infty$. Using two integrations by parts in the coordinates x_i and x_j, and using the bound condition for f, we obtain

$$\mathbb{E}(f(X_1)) - \mathbb{E}(f(X_0)) = \sum_{i\le j}\frac{1}{2^{\delta_{ij}}}(\sigma^1_{ij}-\sigma^0_{ij})\int_0^1\int_{x\in\mathbb{R}^N} \frac{\partial^2 f(x)}{\partial x_i\partial x_j}\phi_{X_t}(x)dxdt.$$

□

The first application of this Proposition 4.2.9 is the proof of Pitt's theorem.

Proof of Theorem 4.2.6. Assume first that X satisfies FKG. For any $(i,j) \in \{1,\cdots,N\}$, define $f(x) = x_i$ and $g(x) = x_j$. Then f and g are increasing and by hypothesis,

$$(\mathrm{Var}(X))_{i,j} = \mathbb{E}(x_i x_j) = \mathbb{E}(f(x)g(x)) \ge \mathbb{E}(f(x))\mathbb{E}(g(x)) = 0.$$

For the converse, let

$$S_1 := \mathrm{Var}(X,X) = \begin{pmatrix}\mathrm{Var}(X) & \mathrm{Var}(X)\\ \mathrm{Var}(X) & \mathrm{Var}(X)\end{pmatrix}$$

and

$$S_0 := \mathrm{Var}(X,Y) = \begin{pmatrix}\mathrm{Var}(X) & 0\\ 0(X) & \mathrm{Var}(X)\end{pmatrix},$$

where Y is an independent copy of X. Define

$$\begin{array}{rcl} F : (\mathbb{R}^N)^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & F(x,y) := f(x)g(y) \end{array}$$

Note that

$$\begin{array}{rcl} \mathbb{E}F(X_0) & = & \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ \text{and } \mathbb{E}F(X_1) & = & \mathbb{E}\big(f(X)g(X)\big). \end{array}$$

Since S_1 is singular, we cannot use directly Proposition 4.2.9. But for any $t < 1$, S_t is non-singular, so that we have

$$\begin{aligned} \mathbb{E}\big(f(X)g(X)\big) - \mathbb{E}(f(X))\mathbb{E}(g(X)) &= \lim_{\epsilon\to 0}\int_0^{1-\epsilon}\frac{d}{dt}\mathbb{E}(f(X_t)) \\ &= \lim_{\epsilon\to 0}\sum_{1\le i\le j\le 2N}(s_{1,ij}-s_{0,ij})\int_0^{1-\epsilon}\int_{x\in\mathbb{R}^{2N}}\frac{\partial^2 F}{\partial x_i\partial x_j}\phi_{X_t}(x)dxdt \\ &= \lim_{\epsilon\to 0}\sum_{i\in\{1,\cdots N\}j\in\{N+1,\cdots 2N\}}\mathrm{Var}(X)_{i(j-N)} \\ &\quad \int_0^{1-\epsilon}\int_{x\in\mathbb{R}^{2N}}\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_{j-N}}\phi_{X_t}(x)dxdt. \end{aligned}$$

Since f and g are increasing, their derivatives are non negative, and since moreover the coefficients of $\mathrm{Var}(X)$ are non negative, this gives the FKG inequality after passing through the limit $\epsilon \to 0$. $\square$

4.3 Proof of Theorem 4.1.6

4.3.1 The analytic continuation problem

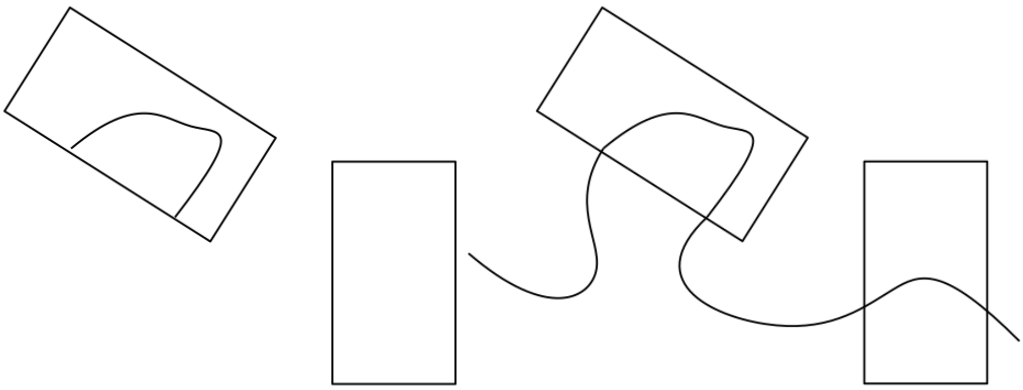


Figure 4.7: Knowing the sign of an analytic Gaussian field in a rectangle can give the information of its vanishing locus, and then by the continuation principle (right) information on another rectangle, as far as it is. This is a problem for the application of the second part of Theorem.

In order to apply the second part of Tassion's theorem 4.2.1, we must prove the dependency condition. However, there is a big problem for our functions, which are analytic:

the analytic continuation phenomenon. Indeed, knowing the sign of f on a whole rectangle where f vanishes gives its vanishing locus in the rectangle. If this a connected component of this vanishing locus crosses another rectangle, then by the ACP, we will know it, even if the other rectangle is very far. The solution for this problem which was chosen in [4] was to blur the information through discretization.

4.3.2 Discretization

Theorem 4.3.1 *(Improvement of [4, Theorem 1.6]) Let $f:\mathbb{R}^2\to\mathbb{R}$ be a centered Gaussian function with covariance function e satisfying $e(x,y)=k(\|x-y\|^2)$ for k C^1, with $k'(0)\neq 0$. Let $\mathcal{T}$ be a lattice in $\mathbb{R}^2$. Then, there exists $C_{\mathcal{T},f}>0$ such that for any fixed $\eta>0$, for any $n>1$, with probability at least $1-\frac{C_{\mathcal{T},f}}{n^\eta}$, the vanishing locus of f in the box $[-n,n]^2$ cuts at most one times every edge of $\frac{1}{n^{2+\eta}}\mathcal{T}$. In particular, with the same probability, any continuous crossing in rectangles in the box is a discrete crossing and vice versa.*

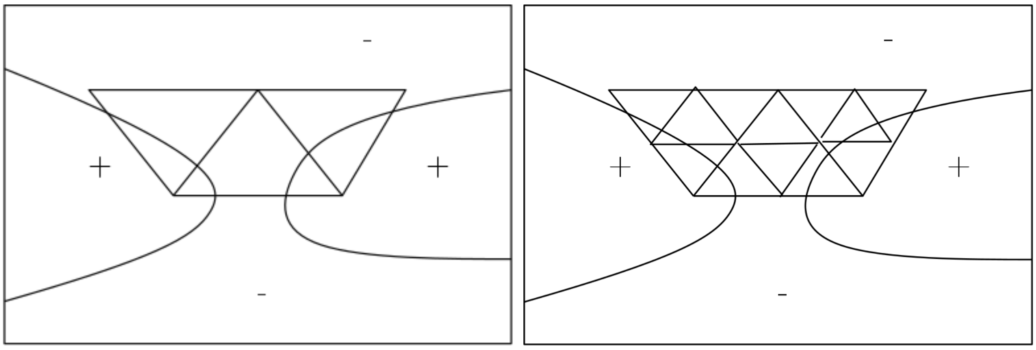


Figure 4.8: Discretization through a lattice. If the lattice is too coarse (left), the discretization is not trustful enough. The solution os too refine it, but the amount of information on the field increases and the dependency of two restrictions of the field to two disjoint rectangles increases.

In fact, the result in [4] is less precise (the power is 9 instead of 3) and its proof was pretty complicated. Here we follow the simpler proof giving better estimates of [5], see also [29]. The proof uses the following extended version of Kac-Rice formula;

Theorem 4.3.2 (Theorem 3.2 [3]) *Let f be a Gaussian field on an interval $I\subset\mathbb{R}$, such that almost surely, f is C^1 and that for any $x,y\in I$, $\operatorname{Cov}(f(x),f(y))$ is definite. Let N be the number of zeros of f on an interval I. Then,*

$$\mathbb{E}(N(N-1))=\int_{I^2}\mathbb{E}\big(|f'(x_1)||f'(x_2)|\;\big|\;f(x_1)=f(x_2)=0\big)\phi_{(f(x_1),f(x_2))}(0,0)dx_1dx_2.$$

We do not write down the proof and refer to the book [3].

Corollary 4.3.3 *[5, Proposition 7] Under the hypotheses of Theorem 4.3.2, and if $e(x,y)=k(x-y)$ with k C^2 and $k''(0)\neq 0$. Then, there exists a constant C depending only on $k''(0)$ such that*

$$\mathbb{E}(N(N-1))\leq C|I|^3.$$

Proof. Theorem 4.3.2 gives

$$\mathbb{E}(N(N-1)) \;=\; \int_{(x,y)\in I^2}\int_{(u,v)\in\mathbb{R}^2}|u||v|e^{-\frac12\langle\Sigma^{-1}(u,v),(u,v)\rangle}\frac{1}{2\pi\sqrt{\det\Sigma}}\phi_{(f(x),f(y))}(0,0)dxdy.$$

where

$$\Sigma = S_1 - C^t S_2^{-1} C,$$

with

$$S_2 := \mathrm{Var}\left(f(x), f(y)\right) = \begin{pmatrix} 1 & e(x,y) \\ e(x,y) & 1 \end{pmatrix}$$

is the covariance of the Gaussian vector $(f(x), f(y))$,

$$S_1 := \mathrm{Var}\left(f'(x), f'(y)\right) = \begin{pmatrix} \partial^2_{xy} e_{x=y} & \partial^2_{xy} e_{x,y} \\ \partial^2_{xy} e_{x,y} & \partial^2_{xy} e_{x=y} \end{pmatrix}.$$

is the variance matrix of he Gaussian vector $(f'(x), f'(y))$, and

$$C := \mathrm{Cov}\Big(\left(f(x), f(y)\right), \left(f'(x), f'(y)\right)\Big) = \begin{pmatrix} 0 & \partial_y e_{x,y} \\ \partial_x e_{x,y} & 0 \end{pmatrix}$$

where we used $\mathbb{E}(f(x)f'(x)) = \partial_x e(x,y)_{|x=y} = k'(0)\partial_x \|x-y\|^2_{|x=y} = 0$. The covariance matrix S_2 implies

$$\phi_{(f(x_1), f(x_2))}(0,0) = \frac{1}{\sqrt{1-e(x,y)^2} 2\pi}.$$

Now, using

$$e(x,y) = k(x-y) = 1 + \frac{1}{2} a u^2 + O(\|k^{(4)}\|_{L^\infty(I)} u^4),$$

with $a := k''(0)$, $u = x - y$,

$$k^2(u) = 1 + au^2 + O((k''(0)^4 + \|k^{(4)}\|_{L^\infty(I)}) u^4), ,$$

$$k' = au + O(3),\ k'^2 = a^2u^2 + O(4),$$

$$k'' = a + O(2),$$

$$\frac{k'^2}{1-k^2} = \frac{a^2u^2 + O(u^4)}{-au^2 + O(4)} = -a(1 + O(u^2))$$

$$\begin{aligned} S_1 - C^t S_2^{-1} C &= \begin{pmatrix} -k''(0) & -k''(x-y) \\ -k''(x-y) & -k''(0) \end{pmatrix} - \frac{1}{1-k^2} \begin{pmatrix} 0 & k' \\ -k' & 0 \end{pmatrix} \begin{pmatrix} 1 & -k \\ -k & 1 \end{pmatrix} \begin{pmatrix} 0 & -k' \\ k' & 0 \end{pmatrix} \\ &= \begin{pmatrix} -k''(0) & -k''(x-y) \\ -k''(x-y) & -k''(0) \end{pmatrix} - \frac{k'^2}{1-k^2} \begin{pmatrix} 1 & k \\ k & 1 \end{pmatrix} \\ &= \begin{pmatrix} O(2) & O(2) \\ O(2) & O(2) \end{pmatrix} \end{aligned}$$

Now,

Lemma 4.3.4 *Fix $n \geq 1$. Then there exists a constant $C > 0$, such that for any $X \in \mathbb{R}^n$ a centered Gaussian vector with variance matrix Σ,*

$$\mathbb{E}(\|X\|^k) \leq C \|\Sigma\|^{k/2}.$$

Proof. Since Σ is a positive real symmetric matrix, there exists an orthonormal basis $(e_i)_i$ of $\mathbb{R}^n$ of eigenvectors of Σ and associated positive eigenvalues $\lambda_1, \cdots, \lambda_n$, so that

$$\mathbb{E}\|X\|^k = \int_{(x_i)_i \in \mathbb{R}^n} \|(x_i)_i\| \prod_{i=1}^n e^{-\frac{1}{2}x_i^2/\lambda_i} dx_1 \cdots dx_n \frac{1}{\prod \sqrt{\lambda_i}(\sqrt{2\pi})^n}.$$

With the change of variables $u_i = x_i/\sqrt{\lambda_i}$, which gives

$$\mathbb{E}\|X\|^k = \int_{(u_i)_i \in \mathbb{R}^n} \|(u_i)_i\|^k (\max_i \lambda_i)^{k/2} \prod_{i=1}^n e^{-\frac{1}{2}u_i^2} du_1 \cdots du_n \frac{1}{(\sqrt{2\pi})^n} \leq C_n \|\Sigma\|^{k/2}.$$

□

In our case, since $\|S_1 - C^t S_2^{-1} C\| = O((x-y)^2)$, this gives

$$\begin{aligned} \mathbb{E}(N(N-1)) &\leq C \int_{I^2} O((x-y)^2) \frac{dxdy}{2\pi\sqrt{-2a(x-y)^2 + O(x-y)^4)}} \\ &\leq C' \int_{I^2} O((x-y)) dxdy \leq C'' |I|^3. \end{aligned}$$

□

Proof of Theorem 4.3.1. Fix $\epsilon > 0$ and ϵE an edge of $\epsilon\mathcal{T}$. The restriction of f on ϵE is a Gaussian field satisfying the hypotheses of Corollary 4.3.3. By Markov's inequality and Corollary 4.3.3,

$$\Pr(N(e) \geq 2) = \Pr(N(N-1) \geq 2) \leq \frac{1}{2}\mathbb{E}(N(N-1)) \leq C''' \epsilon^3.$$

On $[-n,n]^2$, there are at most Cn^2/ϵ^2 edges of $\epsilon\mathcal{T}$, so that

$$\Pr(\exists e \in [-n,n]^2 \mid N(e) \geq 2) \leq \sum_{e \in [-n,n]^2} \frac{1}{2}\mathbb{E}(N(N-1)) \leq C''' \frac{n^2}{\epsilon^2}\epsilon^3 \leq C'''' n^2 \epsilon.$$

Now, let us choose $\epsilon = \epsilon(n) = 1/n^{2+\eta}$. Then, with probability at last $1 - C'''' \frac{1}{n^\eta}$, the nodal lines of f on $[-n,n]^2$ crosses at most once every edge of $\frac{1}{n^{2+\eta}}\mathcal{T}$. □

4.3.3 Quantitative dependency

The discretization have temporally solved the analytic continuation problem. However the problem now we have to face is the following: by theorem 4.3.1, the number of points at which we look the sign of f increases in a rectangle R of size r as r^4. The more we know on the sign R, the more it gives informations on the sign on another rectangle R' and threats the independency condition of Tassion theorem.

The second application of Proposition 4.2.9 is an quantitative dependency theorem for events that depend only on the sign of the Gaussian field restricted to a finite number of points. In our case, this will be vertices of a lattice in a large box.

Theorem 4.3.5 *(Piterbarg 1982 [27, Theorem 1.1, p. 37]) Let A be an event that depends only on the signs of the coordinates of $X \in \mathbb{R}^N$, and $S_0 = (\sigma_{0,ij})_{1 \leq ij \leq N}$ and $S_1 = (\sigma_{1,ij})_{1 \leq i,j \leq N}$ two covariance matrices for $X_0 \in \mathbb{R}^N$ and $X_1 \in \mathbb{R}^N$, such that their diagonal elements equal 1. Then,*

$$|\Pr(X_0 \in A) - \Pr(X_1 \in A)| \leq \frac{1}{\pi} N^2 \max_{i<j} |\sigma_{ij}^1 - \sigma_{ij}^0| \sup_{t \in [0,1]} (1 - \sigma_{t,ij}^2)^{-1/2},$$

where for any $t \in [0,1]$, $\sigma_{t,ij}^2 := (1-t)\sigma_{0,ij} + t\sigma_{1,ij}$.

The left hand side compares the two Gaussian vectors for events depending on their signs. The given upper bound for it is the product which increases polynomially with the dimension of the vector, for us with the number of the sites where we look at the sign of the Gaussian field, and a second term which is bounded by the difference between the covariance matrices.

Remark 4.3.6 *There exists in fact an equality, see the proof below and [29] for a more profound analysis of it which gives a far better bound.*

Proof of Theorem 4.3.5. We use the notation of Proposition 4.2.9, that is $S_t := (1-t)S_0 + tS_1$ and X_t is a centered Gaussian vector associated to S_t. The proof of Proposition 4.2.9 gives, for $f(x) = \mathbf{1}_A$,

$$\begin{aligned}\Pr(X_1 \in A) - \Pr(X_0 \in A) &= \mathbb{E}(f(X_1)) - \mathbb{E}(f(X_0)) \\ &= \sum_{i<j}(\sigma^1_{ij} - \sigma^0_{ij}) \int_0^1 \int_{x\in A} \frac{\partial^2 \phi_{X_t}(x)}{\partial x_i \partial x_j} dx dt.\end{aligned}$$

Note that here f is not C^2, but it was used for the integration by parts in the sequence of the proof, so it is not necessary here. Now since A depends only on the signs of the x_i's, define

$$\forall x \in \mathbb{R}^N, \ \mathrm{sign}(x) := (\mathrm{sign}(x_i))_i \in \{-1,+1\}^N.$$

For any couple $i<j$ and any $\epsilon_i, \epsilon_j \in \{-1,+1\}$

$$Z^{\epsilon_i,\epsilon_j}(A) := \{x \in A \mid \mathrm{sign}(x_i) = \epsilon_i, \mathrm{sign}(x_j) = \epsilon_j\}.$$

The integration by part of $\phi_{X_t}(x)$ gives

$$\begin{aligned}\Pr(X_1 \in A) - \Pr(X_0 \in A) &= \int_0^1 \sum_{i<j}(\sigma_{1,i,j} - \sigma_{0,i,j}) \sum_{\epsilon_i,\epsilon_j} \int_{Z_A^{\epsilon_i,\epsilon_j}} \frac{\partial^2 \phi_t(x)}{\partial x_i \partial x_j} dx dt \\ &= \int_0^1 \sum_{i<j}(\sigma_{1,i,j} - \sigma_{0,i,j}) \sum_{\epsilon_i,\epsilon_j} \int_{Z_A^{\epsilon_i,\epsilon_j}} (-1)^{\epsilon_i+\epsilon_j} \phi_t(x)_{|x_i=x_j=0} \frac{dx}{dx_i dx_j} dt\end{aligned}$$

By section 5.1.4,

$$\phi_t(x)_{|x_i=x_j=0} = \phi_{X_t|(X_i^t, X_j^t)}(x|(0,0))\phi_{(X_i^t,X_j^t)}(0,0),$$

we get

$$\begin{aligned}\Pr(X_1 \in A) - \Pr(X_0 \in A) &= \int_0^1 \sum_{i<j}(\sigma_{1,i,j} - \sigma_{0,i,j}) \\ &\quad \sum_{\epsilon_i,\epsilon_j} (-1)^{\epsilon_i+\epsilon_j} \Pr(x \in Z_A^{\epsilon_i,\epsilon_j} \mid x_i = x_j = 0) \phi^t_{(X_i,X_j)}(0,0) dt.\end{aligned}$$

so that, using $\phi^t_{(X_i,X_j)}(0,0) = \frac{1}{2\pi}(1-\sigma^2_{t,ij})^{-1/2}$,

$$|\Pr(X_1 \in A) - \Pr(X_0 \in A)| \leq \frac{4}{2\pi} \frac{N(N-1)}{2} \max_{i<j} |\sigma_{1,i,j} - \sigma_{0,i,j}| \sup_{t\in[0,1]} (1-\sigma^2_{t,ij})^{-1/2}.$$

□

Note that we bound the probability in the sum by 1, which is a very crude bound. See [30] and [6] for better bounds.

The following corollary measures the dependency between the positive crossings of a discretized Gaussian field over two disjoint rectangles.

Corollary 4.3.7 *Let $\mathcal{V}$ be the set of vertices of a lattice in $\mathbb{R}^n$, $f:\mathbb{R}^n\to\mathbb{R}$ be a Gaussian field with covariance function e, such that $e(x,x)=1$, and U_0 and U_1 a pair of disjoint bounded open sets of $\mathbb{R}^n$. Then, there exists a constant C such that for any pair $(A_i)_{i=0,1}$ of events depending on the sign of f on $\mathcal{V}\cap(U_i)_{i=0,1}$,*

$$|\Pr\big(A_0,A_1) - \Pr(A_0)\Pr(A_1)\big) \leq \frac{1}{\pi}|(\mathcal{V}_0|+|\mathcal{V}_1|)^2 \max_{x\in\mathcal{V}_0, y\in\mathcal{V}_1} |e(x,y)|(1-e(x,y)^2)^{-1/2}.$$

Proof. Let

$$X_0 := (X,Y) := \Big((f(x))_{x\in\mathcal{V}_0}, (f(y))_{y\in\mathcal{V}_1}\Big).$$

Then, this Gaussian vector has a covariance matrix S_0 of the form

$$S_0 := \mathrm{Var}(X_0) = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y)^t & \mathrm{Var}(Y)\end{pmatrix}.$$

Let

$$S_1 = \begin{pmatrix} \mathrm{Var}(X) & 0 \\ 0 & \mathrm{Var}(Y)\end{pmatrix}.$$

If X_1 is the associated Gaussian vector, then $X' := (X_{1,i})_{1\leq i\leq|\mathcal{V}_0|}$ has the same law than X, and is independent of $X'' := (X_{1,i})_{|\mathcal{V}_0|+1\leq j\leq|\mathcal{V}_0|+|\mathcal{V}_1|}$, the latter having the same law than Y. By Theorem 4.3.5,

$$\begin{aligned}|\Pr(X_1\in A_0\cap A_1) - \Pr(X_0\in A_0\cap A_1)| \;\; &\leq \;\; \frac{1}{\pi}|\big(|\mathcal{V}_0|+|\mathcal{V}_1|\big)^2 \\ &\quad \max_{x\in\mathcal{V}_0, y\in\mathcal{V}_1} |e(x,y)|(1-e(x,y)^2)^{-1/2}.\end{aligned}$$

Now $\Pr(X_0\in A_0\cap A_1) = \Pr(A_0,A_1)$ and $\Pr(X_1\in A_0\cap A_1) = \Pr(A_0)\Pr(A_1)$, which concludes. □

Proof of the main Theorem 4.1.6 . In Corollary 4.2.8 we showed that the symmetry and positivity of correlations are satisfied for our continuous model. We must now prove the last quantitative dependence condition. Let $\mathcal{T}=(\mathcal{V},\mathcal{E})$ be the Union Jack lattice, see Figure 4.3. Fix $\eta>0$, such that $12+4\eta<D$. This is possible since $D>16$. Then, Theorem 4.3.1 shows that there exists $C>0$ such that for any $n\geq 2$, with probability at least $1-C(n\log n)^{-\eta}$, in the box $[-n\log n, n\log n]^2$, the continuous percolations happen simultaneously with the discrete site percolations on the vertices

$$V_n := \frac{1}{(n\log n)^{2+\eta}}\mathcal{V}.$$

Moreover by Corollary 4.3.7 applied to $U_0 := [-n\log n, n\log n]^2\backslash[-3n,3n]^2$, $U_1 := [-2n,2n]^2\backslash[-n,n]^2$ and

$$X_n = f_{|V_n\cap[-n\log n, n\log n]^2},$$

there exists a constant $C>0$ depending only on the kernel e such that for sign events A_0 on $\mathcal{V}_0$ (resp. A_1 on $\mathcal{V}_1$),

$$\max_{A_0,A_1} |\Pr\big(A_0,A_1) - \Pr(A_0)\Pr(A_1)| \leq C(n\log n)^4(n\log n)^{8+4\eta}n^{-D}.$$

Since this is bounded by $Cn^{12+4\eta-D}\log^{4n+4\eta}$, this max tends to 0 when n grows to infinity.

Now let $\epsilon > 0$, and N be such that

$$\forall n \geq N,\ C(n\log n)^4(n\log n)^{8+4\eta}n^{-D} + C(n\log n)^{-\eta} < \epsilon.$$

Then, by 4.3.1 and the last computation, the max of the sign events in Tassion theorem for the continuous percolation problem is no greater than ϵ for $n \geq N$, so that it converges to zero, which proves Theorem 4.1.6. □

4.4 Proof of Tassion's theorem

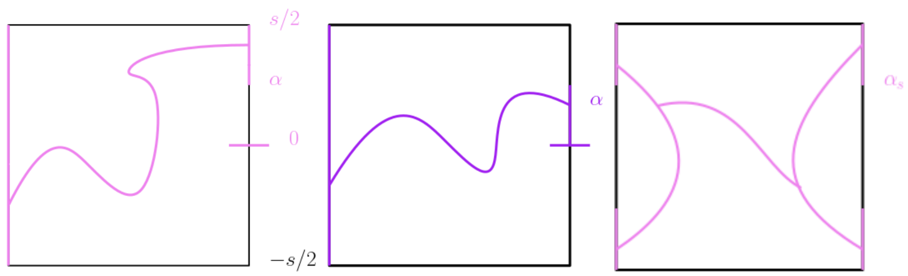


Figure 4.9: The events A^s_α, B^s_β and X^s_α

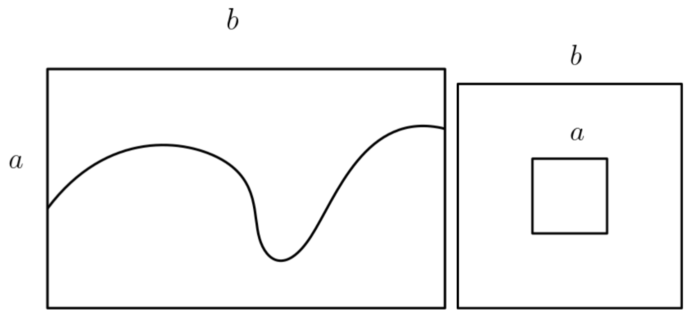


Figure 4.10: The event R^b_a and the event N^b_a.

For $0 < a < b$, $0 < \alpha < s/2$ Consider the events R^b_a, N^b_a, A^s_α, B^s_β and X^s_α defined by figure 4.10 and figure 4.9 respectively. Note that X^s_α can be achieved if we have 4 symmetric copies of A_α and a vertical crossing, so that

$$\Pr(X^s_\alpha) \geq \frac{1}{2}\Pr(A^s_\alpha)^4.$$

We will use a simple useful lemma:

Lemma 4.4.1 *There exists $C > 0$, such that for any $a < b$, $k \in \mathbb{N}^*$,*

$$\Pr(R_a^{b+k(b-a)}) > \frac{1}{2^k}\Pr(R^b_a)^k \text{ and if } b > 2a,\ \Pr(N^b_{2b-a}) \geq \Pr(R^b_a)^4.$$

Proof. This is a consequence of the FKG formula and figure 4.11. □

Proof of Theorem 4.2.1. Let $0 < \epsilon < 1/2$ (in the article of Tassion, $\epsilon = 1/3$). By FKG

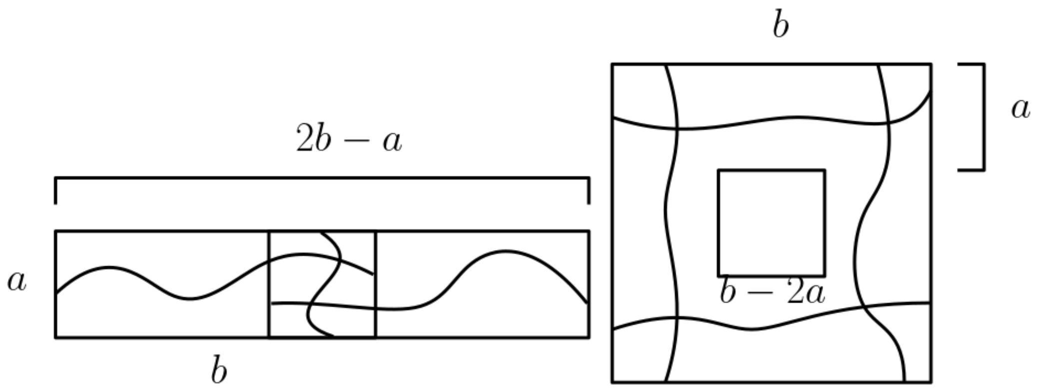


Figure 4.11: $\Pr(R_a^{2b-a}) \geq \frac{1}{2}(\Pr R_a^b)^2$ and $\Pr(N_{2b-a}^b) \geq \Pr(R_a^b)^4$.

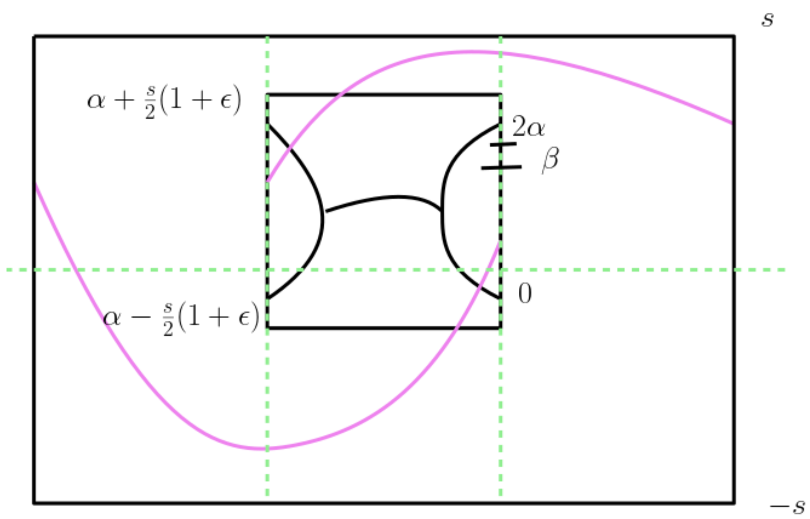


Figure 4.12: This is a horizontal crossing of the rectangle $(3-\epsilon)s \times 2s$ produced by 4 particular events translated or symmetrized from B_β^{2s} (two times) and $X_\alpha^{s(1+\epsilon)}$, if $\beta \leq 2\alpha$ and $\alpha < \frac{s}{2}(1-\epsilon)$.

inequality, we see on figure 4.12 that the probability of crossing a rectangle $(3-\epsilon)s \times 2s$ satisfies

$$\Pr(R_{2s}^{(3-\epsilon)s}) \geq \Pr(X_\alpha^s)\Pr(B_\beta^{2s})^2 \geq \frac{1}{2}\Pr(A_\alpha^s)^4 \Pr(B_\beta^{2s})^2$$

if $\beta \leq 2\alpha$ and $\alpha + \frac{s}{2}(1+\epsilon) < s$ which means $\alpha < \frac{s}{2}(1-\epsilon)$. We see that we must have both large probabilities for A and B. By horizontal symmetry, we have,

$$\Pr(A_\alpha^s) + \Pr(B_\alpha^s) \geq 1/4$$

Now define

$$\begin{aligned} \phi_s : [0, s/2] &\rightarrow [-1, 1] \\ \phi_s(\alpha) &:= \Pr(B_\alpha^s) - \Pr(A_\alpha^s). \end{aligned}$$

Then ϕ_s is non decreasing, $\phi_s(0) \leq 0$ since $B_0^s \subset A_0^s$ and $\phi_{s/2} \geq 1/4$ if the extremity α of $s/2 \times]\alpha, s/2]$ is not allowed in the definition of B_α^s. Since ϕ_s is continuous **à voir**, let α_s be such that $\phi(\alpha_s) = 1/8$. Then,

$$\min\left(\Pr(A^s_{\alpha_s}), \Pr(B^s_{\alpha_s})\right) \geq 1/8.$$

Now, by figure 4.12 we have

$$\alpha_s < \frac{s}{2}(1-\epsilon) \text{ and } \alpha_{2s} < 2\alpha_{s(1+\epsilon)} \Rightarrow \Pr(R_{2s}^{(3-\epsilon)s}) \geq \frac{1}{2^{19}}.$$

Now, fix $s_0 > 0$, $N \in \mathbb{N}^*$ and $\forall k \in \mathbb{N},\ s_k := s_0(\frac{2}{1+\epsilon})^k$.

• Assume first that

$$\exists k \in \{0, \cdots, N\}, \alpha_{s_k} \geq \frac{s_k}{2}(1-\epsilon).$$

Then $\alpha_{s_k} \geq s_k/4$ since $\epsilon < 1/2$ and in this case we construct large crossing without by figure 4.13. By this figure, we see that FKG and the condition implies that

$$\Pr(R_{s_k}^{\frac{3}{2}s_k}) \geq \Pr(X_{\alpha_{s_k}}^{s_k})^3 \geq \frac{1}{2^{25}}.$$

• Assume now that that

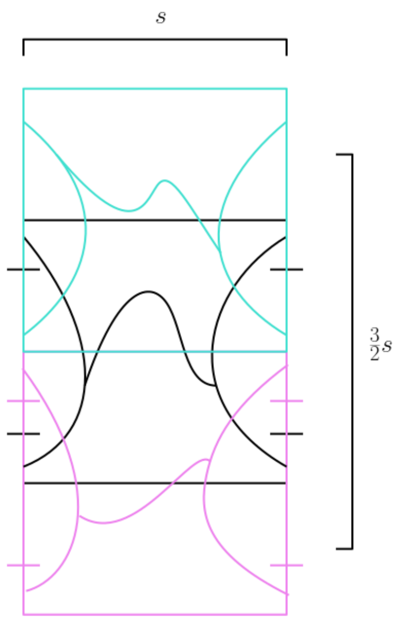


Figure 4.13: If $\alpha_s > s/4$, by FKG we can construct a crossing of the rectangle $\frac{3s}{2} \times s$ with three events $X^s_{s/4}$.

$$\forall k \in \{0, \cdots N\},\ \alpha_{2s_k} > 2\alpha_{s_k(1+\epsilon)} \text{ and } \alpha_{s_k} < \frac{s_k}{2}(1-\epsilon).$$

Recall that the latter condition is required for the small square being in the rectangle. Then,

$$\alpha_{(\frac{2}{1+\epsilon})^{N+1}} > 2^N \alpha_{s_0(1+\epsilon)}.$$

However, $\alpha_s \leq s$ implies $2^k\alpha_1 \leq (\frac{2}{1+\epsilon})^{k+1}$, so that $k \leq C_\epsilon$. That is, for s_1 large enough, there is a positive crossing of $s_1 R$. The first part of Theorem 4.2.1 is proved.

We prove now the second assertion of Theorem 4.2.1. □

We prove now the one-arm bound.

Theorem 4.4.2 *For any random centered Gaussian field $f : \mathbb{R}^2 \to \mathbb{R}$ such that $e(x,y) = k(\|x-y\|)$, with $e(x,y) \leq C/\|x-y\|^D$ with $D > 12$, there exists $\gamma > 0$ and $C > 0$ such that for any $1 \leq s < t$,*

$$\Pr(\textit{one arm in } A(s,t) \leq C(s/t)^{\gamma}.$$

Proof. Fix $1 \leq s < t$, and let $\mathcal{T}$ be a periodic symmetric triangulation. By Theorem 4.3.1, for any $\eta > 0$, there exist a constant $C > 0$, such that the probability of the continuous one arm event in $A(s,t)$ is bounded by the the one of the discrete one arm over the lattice $t^{-2+\eta}$ plus $Ct^{-\eta}$. The probability p of the latter event satisfies

$$p \;\leq\; \Pr\Big[\bigcap_{i\in\mathbb{N},\, s\leq 5^i\sqrt{st}\leq t/2} A^c_{5^i\sqrt{st}}\Big],$$

where A_s denotes the event that there is a positive circuit in the annulus $A(s,2s)$. Then, by Corollary 4.3.7,

$$\begin{aligned} p \;&\leq\; \Pr[\mathcal{A}^c_{\sqrt{st}}]\Pr\Big[\bigcap_{1\leq i\leq\lfloor\log_5(\frac{1}{2}\sqrt{\frac{t}{s}})\rfloor}\mathcal{A}^c_{5^i\sqrt{st}}\Big] + Ct^{4+4(2+\eta)}(\sqrt{st})^{-D}\\ &\leq\; \prod_{0\leq i\leq\lfloor\log_5(\frac{1}{2}\sqrt{\frac{t}{s}})\rfloor}\Pr[\mathcal{A}^c_{5^i\sqrt{st}}] + C't^{4+4(2+\eta)}(\sqrt{st})^{-D}\log t. \end{aligned}$$

By Theorem 4.1.6 and the FKG condition, there exists $c > 0$, such that $\Pr\mathcal{A}_s \geq c$ for the continuous model, so that $\Pr\mathcal{A}_s \geq c - t^{-\eta}$ for the discrete one. This implies

$$\begin{aligned} p \;&\leq\; C''(1-c+t^{-\eta})^{\log\sqrt{t/s}} + C''t^{4+4(2+\eta)}(\sqrt{st})^{-D}\log t\\ &\leq\; C''(s/t)^{\log(1-c+t^{-\eta}} + C''(s/t)^{D-12-4\eta}\log t. \end{aligned}$$

Fix η such that $D-12-4\eta > 0$ and t_0 such that $1-c+t_0^{-\eta} > 1-c/2$. We obtain the result for $t \leq t_0$, hence for all t. □

Chapter 5

Annex

5.1 Probability toolbox

5.1.1 Generalities.

Definition 5.1.1 *1. A probability space is a triplet* $(\Omega, \mathcal{F}, \mathbb{P})$*, where* Ω *is a set,* $\mathcal{F}$ *a* σ*-algebra, that is a non empty subset of* $\mathcal{P}(\Omega)$ *which is closed under complement and coutable unions, and* $\mathbb{P}$ *a measure over* $(\Omega, \mathcal{F})$*, that is a map :* $\mathbb{P} : \mathcal{F} \to [0, \infty]$ *satisfying* $\mathbb{P}(\emptyset) = 0$ *and* $\mathbb{P}$ *is additive for the union of disjoint sets.*

2. *If* Ω *is topological space, the Borelian* σ*-algebra is the* $\sigma-$*algebra generated by the open sets of* Ω*.*

3. *A random variable* $X : \Omega \to \mathbb{R}$ *is is a* $\mathcal{F}$*-measurable map, that is for any Borelian* $B \subset \mathbb{R}$*,* $B \in \mathcal{B}(\mathbb{R})$*,* $X^{-1}(B) \in \mathcal{F}$*. Note if* X *is* $(\Omega, \{\Omega, \emptyset\}, \mathbb{P})$*-measurable (that is we choose the trivial* σ*-algebra), then* X *is constant.*

4. *We say that events* $(A_1, \cdots, A_n) \in \mathcal{F}^n$ *are independent if*

$$\forall (I_k \subset \mathbb{R})_k,\ \Pr(A_1, \cdots, A_N) = \prod_{i=1}^{N} \Pr(A_k).$$

5. $(X_1, \cdots, X_N)$ *are independent variables iff*

$$\forall (I_k \subset \mathbb{R})_k,\ \Pr(X_1 \in I_1, \cdots, X_n \in I_N) = \prod_{i=1}^{N} \Pr(X_i \in I_k).$$

5.1.2 Gaussian vectors.

For $m \in \mathbb{R}$ and $\Sigma > 0$, a random variable X follows the normal law $N(m, \Sigma)$ iff

$$\forall t \in \mathbb{R},\ \Pr(X \leq t) = \int_{-\infty}^{t} e^{-\frac{1}{2}\frac{(x-m)^2}{\Sigma}} \frac{dx}{\sqrt{\Sigma}\sqrt{2\pi}}.$$

Exercise 5.1.2 *Verify that* $\Pr(X \in \mathbb{R}) = 1$*. You can use the classical*

$$\int_{\mathbb{R}} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}.$$

Definition 5.1.3 *A random vector $X = (X_i)_{i=1,\cdots,N} \in \mathbb{R}^N$ is a* non-degenerate Gaussian vector *iff there exist $m \in \mathbb{R}^N$ and a positive symmetric matrix $\Sigma \in Sym_N(\mathbb{R})$ such that for any Borelian $A \subset \mathbb{R}^N$,*

$$\Pr(X \in A) = \int_A e^{-\frac{1}{2}\langle \Sigma^{-1}(X-m),(X-m)\rangle} \frac{dx_1 \cdots dx_N}{\sqrt{(2\pi)^n \det \Sigma}}.$$

Σ is called the variance matrix *of X, and denoted by* $\mathrm{Var}(X)$.

Notation. For any Gaussian vector $X \in \mathbb{R}^N$, and any deterministic $t \in \mathbb{R}^N$, we denote by $\phi_X(x)$ the density of the law of X at x with respect to the Lebesgue measure, that is,

$$\Pr(X \in A) = \int_{x \in A} \phi_X(x) dx.$$

Proposition 5.1.4 *1. The coefficients of Σ can be computed by the formula*

$$\Sigma_{i,j} = \mathbb{E}\big((X_i - \mathbb{E}(X_i)(X_j - E(X_j)\big),$$

or $\Sigma = \mathbb{E}\big((X-m)(X-m)^t\big)$.

2. *If $X_1 \sim N(m_1, \Sigma_1) \in \mathbb{R}^{N_1}$ and $X_2 \sim N(m_2, \Sigma_2) \in \mathbb{R}^{N_2}$ are two non-degenerate Gaussian vectors, then the random vector (X_1, X_2) has variance*

$$\Sigma = \begin{pmatrix} \Sigma_1 & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_1, X_2)^t & \Sigma_2 \end{pmatrix},$$

where $\mathrm{Cov}(X_1, X_2) = \mathbb{E}\big((X_1 - m_1)(X_2 - m_2)^t\big)$.

3. *If $X_1 \sim N(m_1, \Sigma_1) \in \mathbb{R}^N$ and $X_2 \sim N(m_2, \Sigma_2) \in \mathbb{R}^N$ are two non-degenerate Gaussian vectors, then*

$$X_1 + X_2 \sim N(m_1 + m_2, \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\,\mathrm{Cov}(X_1, X_2)).$$

4. *If $\mathbb{R}^n \ni X \sim N(m, \Sigma)$ and $A \in M_{mn}(\mathbb{R})$ a matrix and $b \in \mathbb{R}^m$ is a deterministic vector, then*

$$\mathbb{R}^m \ni AX + b \sim N(Am + b, A\Sigma A^t).$$

We will see that X_1 is independent of X_2 iff $\mathrm{Cov}(X_1, X_2) = 0$. If $X_1 = X_2$, then $\mathrm{Cov}(X_1, X_2) = \mathrm{Var}(X_1)$.

Lemma 5.1.5 *For Gaussian vector $X \sim N(m, \Sigma)$, the density of X*

$$\phi_X(x) = \frac{1}{\sqrt{(2\pi)^N \det \Sigma}} \exp(-\frac{1}{2}\langle \Sigma^{-1}(x-m), (x-m)\rangle)$$

satisfies

$$\phi_X(x) = \int_{\lambda \in \mathbb{R}^N} e^{i\langle x-m, \lambda\rangle} e^{-\frac{1}{2}\langle \Sigma\lambda, \lambda\rangle} \frac{d\lambda}{(2\pi)^N}.$$

Proof. To prove this (a Fourier transform), we diagonalize Σ into an orthonormal basis and use the elementary

$$e^{-\frac{1}{2}x^2} = \int_{y \in \mathbb{R}} e^{ixy} e^{-\frac{1}{2}y^2} \frac{dy}{\sqrt{2\pi}}.$$

□

This equality allows to define Gaussian vectors with non negative covariance matrix, that is with Σ having a kernel. In this cases, any vector in the kernel correspond to **finir**.

5.1.3 Gaussian fields

The correlation function.

Definition 5.1.6 *Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space. A* Gaussian field *f over a topological space M a measurable function on $f : \Omega \times M \to \mathbb{R}$, where M and $\mathbb{R}$ are equipped with their Borelian σ-algebra, such that for any finite set of points $(x_1, \cdots, x_n) \in M^k$, the random vector $X_\omega := (f(\omega, x_i))_{i=1}^n$ is Gaussian.*

Example 5.1.7 *Let $(\phi_i)_{i=1,\cdots,N}$ be a finite set of functions $\phi_i : M \to \mathbb{R}$, and $(a_i)_{i=1\cdots n}$ a Gaussian vector. Then, $f(x) := \sum_{i=1}^n a_i \phi(x)$ is a Gaussian field over M.*

Definition 5.1.8 *The* two-point correlation function *associated to a random Gaussian function is the function*

$$\begin{array}{rcl} e : (\mathbb{R}^n)^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \mathrm{Cov}(f(x), f(y)) = \mathbb{E}\Big(\big(f(x) - \mathbb{E}f(x)\big)\big(f(y) - \mathbb{E}f(y)\big)\Big) \end{array}$$

In particular, $\mathrm{Var}(f(x)) = e(x,x)$, and $f(x)$ is independent of $f(y)$ iff $e(x,y) = 0$.

Example 5.1.9 *In example 5.1.7, the correlation function equals*

$$e(x,y) = \sum_{i=1}^N \mathrm{Cov}(a_i, a_j)\phi_i(x)\phi_j(x).$$

Assume that a_i are independent and follow the same normal law $N(0,1)$. Then,

$$e(x,y) = \sum_{i=1}^N \phi_i(x)\phi_i(y).$$

1. *For the Kac random polynomials defined by 1.0.1,*

$$e(x,y) = \sum_{k=0}^d x^k y^k = \frac{1 - (xy)^{d+1}}{1 - xy}.$$

2. *There is another, more natural in fact, random model for polynomial, the* complex Fubini-Study*, or* Kostlan *measure:*

$$p(x) = \sum_{k=0}^d a_k \sqrt{\binom{d}{k}} x^k,$$

where the $(a_k)_k$ are still independent and follow the same centered normal law. We will see later why this is a good measure to choose, which is not really clear at first sight! For the moment, let us compute the 2-point correlation:

$$e(x,y) = \sum_{k=0}^d (1 + xy)^d.$$

Theorem 5.1.10 (Kolmogorov theorem, see [26]) *Let $k \in \mathbb{N}^*$ and $f : \mathbb{R}^n \to \mathbb{R}$ be a Gaussian field with covariance e, such that e can be derivated in at least k times in x and k times in y, and that these derivatives are continuous. Then, almost surely f is C^{k-1}. Moreover, for any differential operator P, Pf is a Gaussian field whose kernel equals*

$$\mathbb{E}(Pf(x)Pf(y)) = P_x P_y e.$$

Theorem 5.1.11 *([1, Lemma 12.11.12]) Let M be a manifold and $f : M \to \mathbb{R}$ a Gaussian field, almost surely C^1. Then, almost surely f vanishes transversally.*

5.1.4 Conditional expectation.

Assume two real random variables X, Y are defined on the same probability space have a derivable joint density, that is

$$F(x,y) = \Pr(X \leq x, Y \leq y) = \int_{s \leq u, t \leq v} \phi_{X,Y}(u,v)dudv$$

is differentiable in x and y. Note that we can recover the law of X by

$$\Pr(X \leq x) = \Pr(X \leq x, Y \leq +\infty) = \int_{s \leq x} (\int_{v \in \mathbb{R}} \phi_{X,Y}(u,v)dv)du,$$

so that

$$\phi_X(u) = \int_{v \in \mathbb{R}} \phi_{X,Y}(u,v)dv.$$

Define the conditional density by

$$\phi_{X\big|Y}(x|y) = \{ \begin{array}{ccc} \frac{\phi_{X,Y}(x,y)}{\phi_Y(y)} & \text{if} & \phi_Y(y) > 0 \\ 0 & \text{if} & \phi_Y(y) = 0 \end{array}$$

and denote by $Z_y = (X \mid Y = y)$ the associate random variable. Then,

$$\begin{aligned} \lim_{\Delta \to 0} \Pr(X \leq x \mid Y \in [y \pm \Delta]) &= \lim_{\Delta \to 0} \frac{\Pr((X,Y) \in]-\infty, x] \times [y \pm \Delta])}{\Pr(Y \in [y \pm \Delta])} \\ &= \lim_{\Delta \to 0} \frac{\int_{-\infty}^{x} \phi_{X,Y}(u,v)\Delta + o(\Delta)du}{\phi_Y(y)\Delta + o(\Delta)} \\ &= \lim_{\Delta \to 0} \int_{-\infty}^{x} \phi_{X|Y}(u)du + o(1) \\ &= \int_{-\infty}^{x} \phi_{X|Y}(u|y) = \Pr(Z_y \leq x), \end{aligned}$$

which shows that the definition of Z_y is coherent with the intuitive definition given by de latter limit. Now

$$\begin{aligned} \mathbb{E}(f(X,Y)) &= \int_{(x,y) \in \mathbb{R}^2} f(x,y)\phi_{X,Y}(x,y)dxdy \\ &= \int_{(x,y) \in \mathbb{R}^2} f(x,y)\phi_{X|Y}(x|y)\phi_Y(y)dxdy \\ &= \int_{y \in \mathbb{R}} \Big(\int_{x \in \mathbb{R}} f(x,y)\phi_{X|Y}(x|y)dx \Big) \phi_Y(y)dy \\ &= \int_{y \in \mathbb{R}} \Big(\mathbb{E}(f \mid Y = y)\phi_Y(y)dy \end{aligned}$$

The following proposition in the Gaussian case is very useful, and is called the *regression formula* in books. It gives the law of Gaussian vector conditioned that another Gaussian vector has a particular value in terms of another unconditioned Gaussian vector.

Proposition 5.1.12 *(see [3, Proposition 1.2]) Let $X_1 \sim N(m_1, \Sigma_1) \in \mathbb{R}^{N_1}$ and $X_2 \sim N(m_2, \Sigma_2) \in \mathbb{R}^{N_2}$ be two Gaussian vectors, such that X_2 is non-degenerate. Then, for any $x_2 \in \mathbb{R}^{N_2}$ and any bounded function $f : \mathbb{R}^{N1} \to \mathbb{R}$,*

$$\mathbb{E}(f(X_1)|X_2 = x_2) = \mathbb{E}(f(X_3 + Cx_2)),$$

where

$$C = \mathrm{Cov}(X_1, X_2)\Sigma_2^{-1} \in M_{N_1,N_2}(\mathbb{R})$$

and $X_3 \in \mathbb{R}^{N_1}$ *is a Gaussian vector satisfying*

$$X_3 \sim N\big(m_1 - Cm_2, \Sigma_1 - \mathrm{Cov}(X_1, X_2)\Sigma_2^{-1}\,\mathrm{Cov}(X_1, X_2)^t.\big).$$

Remark 5.1.13 *1. Note that if* X_1 *and* X_2 *are independent,* $C = 0$, $X_3 = X_1$ *and* $\mathbb{E}(f(X_1)|X_2 = y) = \mathbb{E}(f(X_1))$ *: knowing something on* X_2 *has no influence on the result for* X_1.

2. On the converse, if $X_1 = X_2$*, then* $(f(X_1) \mid X_1 = x_2) = f(x_2)$*. Here* $C = I_{N_1}$*, and* $X_3 = 0$ *!*

Proof of Proposition 5.1.12. The trick is to find a matrix C such that $X_3 := X_1 - CX_2$ is independent of X_2. Since X_3 is Gaussian, independence is equivalent to $\mathrm{Cov}(X_3, X_2) = 0$, that is

$$\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(CX_2, X_2) = \mathbb{E}\Big((C(X_2 - m_2))(X_2 - m_2)^t)\Big) = CS_2,$$

so that

$$C = \mathrm{Cov}(X_1, X_2)S_2^{-1}.$$

From Proposition 5.1.4,

$$X_3 \sim N\big(m_1 - Cm_2, S_1 - \mathrm{Cov}(X_1, X_2)S_2^{-1}\,\mathrm{Cov}(X_1, X_2)^t.\big).$$

Now

$$\mathbb{E}(f(X_1) \mid X_2 = x_2) = \mathbb{E}(f(X_3 + CX_2) \mid X_2 = x_2) = \mathbb{E}(f(X_3 + Cx_2)).$$

□

Exercise 5.1.14 *For* $f(x)$ *a centered Gaussian field,*

$$\Pr(f(x) > 0 \mid f(y) > 0) = \frac{1}{2} + \frac{1}{\pi}\arcsin e(x, y).$$

5.2 Reminder for the projective spaces

5.2.1 Definition and basic properties

For $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $n \geq 1$, the *n-dimensional projective space* $\mathbb{K}P^n$ is defined by

$$\mathbb{K}P^n = \mathbb{K}^{n+1}/\sim,$$

where

$$(X_0, \cdots, X_n) \sim (Y_0, \cdots, Y_n) \Leftrightarrow \exists t \in \mathbb{K}^*, (Y_0, \cdots, Y_n) = t(X_0, \cdots, X_n).$$

In other terms, $\mathbb{K}^n$ is the set of lines in $\mathbb{K}^{n+1}$. Note that we have

$$\mathbb{R}P^n = \mathbb{S}^n/\sim,$$

where $\mathbb{S}^n := \{\|x\| = 1\}$, and if

$$\pi : \mathbb{K}^{n+1} \to \mathbb{K}P^n,$$

is the projection, then for any $[X] \in \mathbb{R}P^n$, $\pi^{-1}([X]) = \mathbb{R}X$ and

$$\pi^{-1}_{|\mathbb{S}^n}([X]) = \{\pm X\}.$$

For $\mathbb{K} = \mathbb{C}$, we have $\pi^{-1}([X]) = \mathbb{C}X$ and

$$\pi^{-1}_{|\mathbb{S}^{2n+1}}([X]) = \{tX | t \in \mathbb{C}, |t| = 1\}.$$

The set $\mathbb{K}P^n$ can be equipped with the topology induced by π, that is $U \subset \mathbb{K}P^n$ is open iff $\pi^{-1}(U) \in \mathbb{K}^{n+1}$ is open for the standard topology in the affine space.

Theorem 5.2.1 *[20, p. 60–61] The projective space $\mathbb{R}P^n$ (resp. $\mathbb{C}P^n$) is a smooth compact manifold of dimension n (resp. $2n$).*

Remark 5.2.2 *In fact, $\mathbb{C}P^n$ is a* complex manifold*, that is we can choose transition functions as holomorphic functions (locally defined on $\mathbb{C}^n$).*

5.2.2 Metric on $\mathbb{R}P^n$.

Recall that $\mathbb{R}P^n$ can be defined as the quotient of $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ by the projection $\pi : S^n \to \mathbb{S}^n / \sim = \mathbb{R}P^n$. Moreover, π is a local diffeomorphism. This allows to define a natural metric on $\mathbb{R}P^n$ by

$$\forall x \in \mathbb{R}P^n, \forall u, v \in T_x\mathbb{R}P^n, g(u, v) := \langle d\pi^{-1}(u), d\pi^{-1}(v)\rangle.$$

Here, $d\pi^{-1}$ is the differential of a local inverse π^{-1}. If we choose another and different inverse $\pi^{-1'}$, then $\pi^{-1'} = -\pi^{-1}$, so that $d\pi^{-1'} = -d\pi^{-1}$ and the metric defined by $\pi^{-1'}$ equals the first one.

5.3 Volumes and coarea formula

5.3.1 Riemannian volumes

We recall here some generalities for volumes in manifolds. If (M^n, g) is a smooth Riemannian manifold, that is a manifold equipped with a metric g, that is a scalar product g_x on the tangent space T_xM which is smooth in x. The definition of the latter means that in local coordinates $(x_i)_i$, for any pair of tangent vectors $v, w \in T_xM$, $v = \sum_{i=1}^n v_i\partial_{x_i}$ and $w = \sum_{i=1}^n w_i\partial_{x_i}$,

$$g(v, w) = \sum_{i,j} g_{ij}(x)v_iv_j$$

, where $g_{ij}(x)$ is a smooth function. Now, the *volume form* associated to g is defined in coordinates by the n-form

$$\sqrt{(\det(g_{ij}(x))_{ij})}dx_1 \wedge \cdots \wedge dx_n.$$

This definition does not depend on coordinates, since if (y_j) are other coordinates, then

$$(dx_i)_i = (\frac{\partial x_i}{\partial y_j})_{ij}(dy_j)_j,$$

so that by definition of the determinant,

$$dx_1 \wedge \cdots \wedge dx_n = \det(\frac{\partial x_i}{\partial y_j})_{ij}dy_1 \wedge \cdots \wedge dy_n.$$

Moreover, $\partial_{x_i} f = \sum_j \partial_{y_j} f \frac{\partial y_j}{\partial x_i}$ so that

$$\partial_{x_i} = \sum_j \partial_{y_j} \frac{\partial y_j}{\partial x_i},$$

so that the transformation matrix equals $P := (\frac{\partial y_j}{\partial x_i})_{ij}$. This implies

$$G := (g_{ij}(x))_{ij} = (g(\partial_{x_i}, \partial_{x_j}))_{ij} \tag{5.3.1}$$

$$= P^t \tilde{G} P, \tag{5.3.2}$$

where $\tilde{G}$ is the matrix of the scalar product in the basis $(\partial_{y_i})_i$, so that $\det G = \det^2 P \det \tilde{G}$, and

$$\sqrt{\det G} dx_1 \wedge \cdots \wedge dx_n = \sqrt{\det \tilde{G}} dy_1 \wedge \cdots \wedge dy_n.$$

We denote this intrinsic (that is, independent of the coordinates) form the *volume form* associated to the metric g and we dente it by $dvol_M$.

Note that if G is the matrix of a scalar product on $\mathbb{R}^n$, then $\sqrt{\det G}$ is the volume of the volume generated by the vectors ∂_{x_j}. In particular, if this basis is orthonormal, then $\sqrt{\det G} = 1$. This is why in a Riemannian manifold we define

$$\forall U \subset M, \ \mathrm{Vol}(U) := \int_U dvol_M(x).$$

5.3.2 The coarea formula

We will need the following useful equation:

Lemma 5.3.1 *For any linear map: $F : (E, g) \to (E', g')$ between euclidean spaces of finite dimension, such that $\dim E' \leq \dim E$ and F is onto. Then*

$$\sqrt{\det F \circ F^*} = |\det F_{|\ker F^\perp}|,$$

where in the right-hand-side, $\det$ is computed in orthormal basis.

Proof. Let $(e_i)_i$ be an ONB of $\ker F$ and $(e'_j)_j$ an ONG of $\ker F^\perp$, and $(f_k)_k$ an ONB of $\mathrm{im}\, F$. Then,

$$mat(F) = \begin{pmatrix} 0 & matF_{|\ker F^\perp} \end{pmatrix},$$

so that

$$matFF^* = \begin{pmatrix} 0 & matF_{\ker F^\perp} \end{pmatrix} \begin{pmatrix} 0 \\ matF^*_{|\ker F^\perp} \end{pmatrix} = matF_{|\ker F^\perp} matF^*_{|\ker F^\perp},$$

which gives the result since $\det matF^* = \det matF$. $\square$

Proof of Theorem 2.2.4. By Sard's theorem, since f is C^1, the Lebesgue measure of the set of critical points of f vanishes. Since for any such critical point x we have $\det df \circ df^* = 0$, we can assume that any x in both integrals are regular for f. Fix $x \in M$, and choose local coordinates $(x_i)_i$ based on an open subset $U \ni x$, such that

$$\langle \partial_{x_{m-n+1}}, \cdots, \partial_{x_m} \rangle \pitchfork \ker df(x).$$

For this fix first coordinates. This gives a basis of T_xM; find a linear change of variables to find a new basis that split in two parts that satisfies the condition, and apply this linear change of variables to the coordinates themselves. Define

$$\forall x \in U, \phi(x) := (f(x), x') \in \mathbb{R}^n \times \mathbb{R}^{m-n},$$

where $x' = (x_1, \cdots x_{m-n})$. Then

$$\ker d\phi(x) = \ker df(x) \cap \langle \partial_{x''} \rangle = \{0\},$$

where $x'' = (x_{m-n+1}, \cdots, x_m)$, so that by the local inversion theorem, ϕ is a local diffeomorphism. Note that for any $x \in M$,

$$\begin{aligned} v_x := \phi^{-1}(f(x), \cdot) : \mathbb{R}^{m-n} &\to M && (5.3.3)\\ x' &\mapsto \phi^{-1}(f(x), x') && (5.3.4) \end{aligned}$$

is in fact a parametrization of the fiber $f^{-1}(x)$, so that

$$\operatorname{im} dv_x(x') = T_{v_x(f(x),x')} f^{-1}(x) = \ker df(v(f(x), x')).$$

The volume form on $T_x f^{-1}(x)$ in the coordinates x' write

$$dvol_{f^{-1}(x)} = |\det dv_x| dx'$$

Decomposing $T_xM = \ker df(x)^{\perp} \oplus \ker df(x)$ gives

$$d\phi(x) = \begin{pmatrix} df_{|\ker df^{\perp}} & 0 \\ * & dv_x^{-1} \end{pmatrix},$$

so that in orthonormal basis,

$$|\det d\phi(x)| = |\det df_{|\ker df^{\perp}}||\det dv_x^{-1}|.$$

This implies

$$\begin{aligned} \int_U g\sqrt{\det df \circ df^*} dx &= \int_U g|\det d\phi(x)| \det dv_x(\phi(x)| dx && (5.3.5)\\ &= \int_{\phi(U)} g \circ \phi^{-1} |\det dv_x| dx' dx'' && (5.3.6)\\ &= \int_{\mathbb{R}^n} \int_{\phi(U)\cap f^{-1}(x)} g \circ \phi^{-1} |\det dv_x(x')| dx' dx'' && (5.3.7)\\ &= \int_{x''} \int_{\phi(U)\cap f^{-1}(x))} g dvol_{f^{-1}(x)} dx' dx'' && (5.3.8) \end{aligned}$$

Now, using a partition of unity compatible with the chart we have chosen, we obtain the result. □

5.4 Solutions of exercices

5.4.1 Exercice 5.1.14

We have

$$\Pr(f(x) > 0 \mid f(y) > 0) = 2\int_{a>0,b>0} e^{-\frac{1}{2}\langle \Sigma^{-1}(a,b),(a,b)\rangle} \frac{dadb}{2\pi\sqrt{\det \Sigma}}.$$

with

$$\Sigma = \begin{pmatrix} 1 & e(x,y) \\ e(x,y) & 1 \end{pmatrix}$$

so that

$$\Sigma^{-1} = (1-e(x,y)^2)^{-1} \begin{pmatrix} 1 & -e(x,y) \\ -e(x,y) & -1 \end{pmatrix}$$

and

$$\Pr(f(x) > 0 \mid f(y) > 0) = 2\int_{a>0,b>0} e^{-\frac{1}{2}\frac{a^2+b^2-2e(x,y)ab}{1-e(x,y)^2}} \frac{dadb}{2\pi(1-e(x,y)^2)^{1/2}}.$$

Using $u = r\sqrt{1-e(x,y)^2}\cos\theta$ and $a = r\sqrt{1-e(x,y)^2}\sin\theta$, this gives, with

$$duda = rdr(1-e(x,y)^2)d\theta,$$

$$\begin{aligned} \Pr(f(x) > 0 \mid f(y) > 0) &= 2\sqrt{1-e(x,y)^2}\int_0^{\pi/2}\int_0^{+\infty} e^{-\frac{1}{2}r^2(1-e(x,y)\sin(2\theta))} rdr\frac{d\theta}{2\pi} \\ &= \frac{1}{\pi}\sqrt{1-e(x,y)^2}\int_0^{\pi/2}\frac{1}{1-e(x,y)\sin(2\theta)}d\theta. \end{aligned}$$

Before going on, let us check that if $f(x)$ is independent of $f(y)$, then the result is $1/2$. In this case, $e(x,y)=0$ and we see that it is true. Let us go on: let $v = \tan\theta$. Then,

$$\begin{aligned} \Pr(f(x) > 0 \mid f(y) > 0) &= \frac{1}{\pi}\sqrt{1-e(x,y)^2}\int_0^{+\infty}(1-2e(x,y)\frac{v}{1+v^2}))^{-1}\frac{dv}{1+v^2} \\ &= \frac{1}{\pi}\sqrt{1-e^2}\int_0^{+\infty}(1+v^2-2ve)^{-1}dv \\ &= \frac{1}{\pi}\sqrt{1-e^2}\int_0^{+\infty}\left((v-e)^2+(1-e^2)\right)^{-1}dv \end{aligned}$$

Now with

$$w = \frac{v-e(x,y)}{\sqrt{1-e(x,y)^2}}$$

this gives

$$\begin{aligned} \Pr(f(x) > 0 \mid f(y) > 0) &= \frac{1}{\pi}\int_{-\frac{e(x,y)}{\sqrt{1-e(x,y)^2}}}^{+\infty}\frac{1}{1+w^2}dw = \frac{1}{2}+\frac{1}{\pi}\arctan\frac{e(x,y)}{\sqrt{1-e(x,y)^2}}) \\ &= \frac{1}{2}+\frac{1}{\pi}\arcsin e(x,y). \end{aligned}$$

Other method for the beginning. Using the conditional expectation. In this case, $\Sigma_1 = \Sigma_2 = 1$ and $\mathrm{Cov}(f(x), f(y)) = e(x,y)$. By the former proposition, for any $a \in \mathbb{R}$,

$$(f(x)|f(y) = a) \sim N(e(x,y)a, 1-e(x,y)^2).$$

We have

$$\begin{aligned}
\Pr(f(x)>0 \mid f(y)>0) &= \frac{\Pr(f(x)>0 \text{ and } f(y)>0)}{\Pr(f(y)>0)} \\
&= 2\int_0^{+\infty} \Pr(f(x)>0 \mid f(y)=a) e^{-\frac{1}{2}a^2}\frac{da}{\sqrt{2\pi}} \\
&= 2\int_0^{+\infty} \Big(\int_0^{+\infty} e^{-\frac{1}{2}\frac{(u-e(x,y)a)^2}{1-e(x,y)^2}} \frac{du}{\sqrt{1-e(x,y)^2}\sqrt{2\pi}}\Big) e^{-\frac{1}{2}a^2}\frac{da}{\sqrt{2\pi}} \\
&= 2\int_0^{+\infty} \Big(\int_0^{+\infty} e^{-\frac{1}{2}\frac{u^2-2e(x,y)ua+a^2}{1-e(x,y)^2}} \frac{du}{\sqrt{1-e(x,y)^2}\sqrt{2\pi}}\frac{da}{\sqrt{2\pi}}.
\end{aligned}$$

5.4.2 Exercise 2.3.6

1. Let $\lambda \in \mathbb{R}$, $\Delta f = \lambda f$. That means $f'' + \lambda f = 0$. For $\lambda < 0$, the solutions are $\mathrm{Vect}(e^{\sqrt{-\lambda}\theta}, e^{\sqrt{-\lambda}\theta})$, which has no non-trivial 2π-periodic. Hence $\lambda \geq 0$, and the solutions are $\mathrm{Vect}(\sin(\sqrt{\lambda}\theta), \cos(\sqrt{\lambda}\theta))$. By 2π-periodicity, $\lambda = k^2$ with $k \in \mathbb{N}$.

2. We have

$$\|\sin kx\|_2^2 = \|\cos kx\|_2^2 := \int_0^{2\pi} \cos^2(kx)\frac{d\theta}{2\pi} =: \int_0^{2\pi} \frac{1}{2}(1+\cos(2kx))\frac{d\theta}{2\pi} = \frac{1}{2}$$

   if $k \geq 1$, and 1 si $k = 0$. Moreover, these functions are orthogonal, so that $\{1, (\sqrt{2}\cos kx)_{k\geq 1}, (\sqrt{2}\sin kx)$ is an ONB made of eigenfunctions, and we choose $k \leq L$ for the answer.

3. We have

$$\begin{aligned}
e_L(x,y) &= 1 + 2\sum_{k=1}^{N} \cos(kx)\cos(ky) + \sin(kx)\sin(ky) \\
&= 1 + 2\sum_{k=1}^{N} \cos[k(x-y)] = \Re \sum_{k=-N}^{N} e^{ik(x-y)} = \\
&= e^{-iN(x-y)}\frac{e^{i(2N+1)(x-y)}-1}{e^{i(x-y)}-1} = \frac{\sin((N+\frac{1}{2})(x-y))}{\sin(\frac{x-y}{2})}
\end{aligned}$$

   if $e^{i(x-y)} \neq 1$ and in the other case. In the latter case, this is

5.4.3 Exercise 2.2.7

Let us use a classical trick that begins with

$$1 = \int_{\mathbb{R}^{n+1}} e^{-\frac{1}{2}\|x\|^2}\frac{dx}{(\sqrt{2\pi})^{n+1}}.$$

Define $f := \|x\|^2$ on $\mathbb{R}^{n+1}$. Then, $\nabla f(x) = 2x$, so that by the coarea formula

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}} e^{-\frac{1}{2}\|x\|^2}\frac{dx}{\sqrt{2\pi}^{n+1}}dx &= \int_{\mathbb{R}^+} e^{-\frac{1}{2}t^2}\frac{dx}{\sqrt{2\pi}^n}\int_{t\mathbb{S}^n} dvol\frac{1}{2\|x\|}dt \\
&= \int_{\mathbb{R}^+} e^{-\frac{1}{2}t^2}\frac{dx}{\sqrt{2\pi}^{n+1}}\mathrm{Vol}(\mathbb{S}^n)\frac{t^n}{2}dt
\end{aligned}$$

so that

$$\mathrm{Vol}(\mathbb{S}^n) = (\sqrt{2\pi})^n \Big(\int_{\mathbb{R}} t^n e^{-\frac{1}{2}t^2} \frac{dt}{\sqrt{2\pi}}\Big)^{-1}.$$

Since $\mathbb{R}P^n$ is locally isometric to $\mathbb{S}^n$ and the covering π is $2-1$, We obtain $\mathrm{Vol}(\mathbb{R}P^n) = \frac{1}{2}\mathrm{Vol}(\mathbb{S}^n)$. In particular, $Length(\mathbb{R}P^1) = \pi$.

5.4.4 Exercise 2.3.5.

We have

$$\partial^2_{x_j y_j} e_L = \delta_{ij} \frac{1}{(2\pi)^n} L^{\frac{n}{2}+1} \int_{\|\xi\|\leq 1} \xi_i^2 d\xi + o(L^{\frac{n}{2}+1}).$$

We have

$$\begin{aligned} n\int_{\|\xi\|\leq 1} \xi_i^2 d\xi &= \int_{\|\xi\|\leq 1} \|\xi\|^2 d\xi = \mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^1 r^2 r^{n-1} dr \\ &= \mathrm{Vol}(\mathbb{S}^{n-1})(n+2) = n\,\mathrm{Vol}(\mathbb{B}^n)(n+2). \end{aligned}$$

5.4.5 Exercise 2.2.5

This is well-defined because $u \mapsto g'(u', f(u))$ is a linear form on E, and because g is non-degenerate, $\phi : E \to E^*$ $\phi : v \mapsto (u \mapsto g(u,v)) \in E^*$ is a one-to- one map, so that $f^*(u') = \phi^{-1}(u \mapsto g'(u', f(u)))$.

5.4.6 Exercise 3.2.10

We see that $q_i(x,y) = 0$ implies $\|y\| \leq 1$ and $|\|x\|^2 - 2| \leq 1$, so that $\|x\|^2 \leq 3$ and $\|x\|^2 + \|y\|^2 \leq 4$, so that $q_i^{-1}(0) \subset B(0,2)$. We have $d_x q_i = 2(\|x\|^2 - 2)\sum_i x_i dx_i$ which is onto if $\|x\|^2 \neq 2$. If $\|x\|^2 = 2$, then $\|y\|^2 = 1$ and in this case $d_y q_i$ is onto. Now we prove that $q_i^{-1}(0) \sim S^i \times S^{n-i-1}$. For this, let

$$\begin{aligned} \varphi : \Sigma_i &\to \mathbb{R}^{i+1} \times \mathbb{R}^{n-i} && (5.4.1) \\ (x,y) &\mapsto (x', y') = \big(x, (y, Y)\big) := \Big(x, (y, \|x\|^2 - 2)\Big). && (5.4.2) \end{aligned}$$

Then clearly, φ is injective and smooth, and $y' \in \mathbb{S}^{n-1-i}$. For fixed y', $x \in S(0, \sqrt{2+Y})$ finir

5.4.7 Exercise 3.2.14

Let $n \geq 1$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Gaussian field, where $f = (f_1, \cdots, f_n)$. The goal of this problem is, in a simple case, to estimate the number of zeros of f in an bounded open set of $\mathbb{R}^n$. We assume that there exists a Gaussian centered field $g : \mathbb{R}^n \to \mathbb{R}$, such that for any $i \in \{1, \cdots, n\}$, f_i is a copy of g, and all the f_i's are independent. We denote by $e : (\mathbb{R}^n)^2 \to \mathbb{R}$ the covariance function associated to g, and we assume that e is C^1.

1. For any $x \in \mathbb{R}^n$, write the variance matrix of the vector $f(x)$ as related to e.
   We have $Varf(x) = (\mathbb{E}(f_i(x)f_j(x))_{ij}) = (\delta_{ij} e(x,x))_{ij}) = e(x,x)1_n$

2. For any $x \in \mathbb{R}^n$, compute the density $\phi_{f(x)}(0)$ as a function of e.
We have
$$\phi_{f(x)}(0) = \frac{1}{(2\pi e(x,x))^{n/2}}.$$

3. For any $x \in \mathbb{R}^n$, denote by $Df(x) \in M_n(\mathbb{R})$ the Jacobian matrix of f at x in the canonical basis. Write the variance matrix of the vector $Df(x)$ as a function of e and its derivative. In this question we consider that a matrix in $M_n(\mathbb{R})$ is a vector $(m_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n^2}$, however we keep the notations $1 \leq i,j \leq n$ instead of $1 \leq k \leq n^2$.
We have
$$\begin{aligned} VarDf(x) &= Var\big(Mat(df(x), B_{can})\big) \\ &= Var\big((\partial_j f_{i|x})_{i,j}\big) \\ &= (\mathbb{E}(\partial_j f_i(x)\partial_\ell f_k(x))_{1\leq i,j,k,\ell\leq n}) = (\delta_{ik}\partial^2_{x_j y_\ell} e_{|x=y})_{ij,k\ell}. \end{aligned}$$

4. Write $Cov(f(x), Df(x))$.
We have
$$Cov(f(x), Df(x)) = (\mathbb{E}f_i(x)\partial_\ell f_k(x))_{1\leq i,k,\ell\leq n} = (\delta_{ik}\partial_\ell e)_{1\leq i,k,\ell\leq n}.$$

Til the end of the problem we assume that there exists a smooth $\rho : \mathbb{R}^+ \to \mathbb{R}$ such that
$$\forall x, y \in \mathbb{R}^n,\ e(x,y) = \rho(-\frac{1}{2}\|x-y\|^2).$$

5. Write $Varf(x)$, $VarDf(x)$ and $Cov(f(x), Df(x))$ as a function of ρ and its derivatives.
We have $Varf(x) = \rho(0)1_n$**. Besides,**
$$\partial_{x_j} e = \rho'(-\frac{1}{2}\|x-y\|^2)(-(x_j - y_j))$$
and $\partial^2_{x_i y_\ell} e_{|x=y} = \delta_{i\ell}\rho'(0)$. **so that**
$$VarDf(x) = \rho'(0)1_{n^2}.$$
Lastly, $Cov(f(x), Df(x)) = 0_{n^3}$.

6. Let $U \subset \mathbb{R}^n$ be an open subset and
$$N(f,U) := \#\{x \in U, f(x) = 0\}.$$
Show that
$$\mathbb{E}N(f,U) = cVol(U)^\alpha \rho(0)^\beta \rho'(0)^\gamma,$$
with α, β, γ and c constants which depend only on n. Give the values of α, β and γ and write c as an integral over $M_n(\mathbb{R})$.
By Kac-Rice, we have
$$\mathbb{E}N(f,U) = \int_U \mathbb{E}(|\det Df(x)|\ |f(x) = 0)\phi_{f(x)}(0)dx,$$

where dx is the Lebesgue measure. Since $Cov(Df(x), f(x)) = 0$, the Gaussian vectors $Df(x)$ and $f(x)$ are independent, so that

$$\begin{aligned}\mathbb{E}(|\det Df(x)|\ |f(x)=0) &= \mathbb{E}(|\det Df(x)|)\\ &= \int_{M\in M_n(\mathbb{R})} |\det M| e^{-\frac{1}{2}\langle (VarDf(x))^{-1}M,M\rangle} d\mu(M),\end{aligned}$$

with

$$d\mu(M) = \frac{\prod dM_{ij}}{\sqrt{(2\pi)^{n^2}\det VarDf(x)}}.$$

Since $VarDf(x) = \rho'(0)1_{n^2}$, the change of variables $M = \sqrt{\rho'(0)}N$ gives

$$\mathbb{E}(|\det Df(x)|\ |f(x)=0) = \sqrt{\rho'(0)}^n \int_{N\in M_n(\mathbb{R})} |\det N| e^{-\frac{1}{2}\|N\|^2} \frac{\prod dN_{ij}}{\sqrt{(2\pi)^{n^2}}},$$

hence

$$\mathbb{E}N(f,U) = \Big(\frac{\rho'(0)}{2\pi\rho(0)}\Big)^{n/2} Vol(U) \int_{N\in M_n(\mathbb{R})} |\det N| e^{-\frac{1}{2}\|N\|^2} \frac{\prod dN_{ij}}{\sqrt{(2\pi)^{n^2}}}.$$

7. Let $M \in M_n(\mathbb{R})$, and $C_1, \cdots, C_n$ its column vectors. Show that

$$|\det M| = \prod_{i=1}^n \|C_i^\perp\|,$$

where for any $i \geq 2$, $C_i^\perp$ denotes the orthogonal projection onto $Vect^\perp(C_1, \cdots, C_{i-1})$, and $C_1^\perp = C_1$.

8. Let $M \in M_n(\mathbb{R})$ be random, such that its coefficients are independent and follow a normal law $N(0,1)$.

   (a) Show that

   $$\mathbb{E}(|\det M|) = \prod_{k=1}^n \mathbb{E}(\|X_k\|),$$

   where for any k, $X_k \in \mathbb{R}^k$ is a random vector which coefficients are independent and follow a normal law $N(0,1)$.

   (b) Show that $\mathbb{E}(\|X_k\|) = \sqrt{2\pi}\frac{Vol(S^{k-1})}{Vol(S^k)}$.
   We have

   $$\begin{aligned}\mathbb{E}(\|X_k\|) &= \int_{X\in\mathbb{R}^k} \|X\| e^{-\frac{1}{2}\|X\|^2} \frac{dX}{(2\pi)^{k/2}} &(5.4.3)\\ &= \int_0^\infty re^{-\frac{1}{2}r^2} r^{k-1} dr \frac{Vol(\mathbb{S}^{k-1})}{(2\pi)^{k/2}}. &(5.4.4)\end{aligned}$$

   Moreover

   $$\begin{aligned}1 &= \int_{x\in\mathbb{R}^{k+1}} e^{-\frac{1}{2}\|x\|^2} \frac{dx}{\sqrt{2\pi}^{k+1}}\\ &= Vol(S^k)\int r^n e^{-\frac{1}{2}r^2} \frac{dr}{\sqrt{2\pi}^{k+1}},\end{aligned}$$

   which gives the result.

   (c) Deduce $\mathbb{E}(|\det M|)$ as a function for $Vol(S^n)$.

5.4.8 Proof of Lemma 4.2.10

By Lemma 5.1.5,

$$\phi_X(x) = \int_{\lambda\in\mathbb{R}^N} e^{i\langle x-m,\lambda\rangle} e^{-\frac{1}{2}\langle\Sigma\lambda,\lambda\rangle} \frac{d\lambda}{(2\pi)^N},$$

so that for $i \neq j$, remembering that $\langle\Sigma\lambda,\lambda\rangle = \sum_{i,j}\sigma_{ij}\lambda_i\lambda_j$, and $\sigma_{ji} = \sigma_{ij}$,

$$\frac{\partial}{\sigma_{ij}}\phi_X(x) = \int_{\lambda\in\mathbb{R}^N} e^{i\langle x-m,\lambda\rangle}(-\frac{1}{2}(2\lambda_i\lambda_j))e^{-\frac{1}{2}\langle\Sigma\lambda,\lambda\rangle} \frac{d\lambda}{(2\pi)^N},$$

whereas for every i, j,

$$\frac{\partial}{\partial x_i\partial x_j}\phi_X(x) = \int_{\lambda\in\mathbb{R}^N} (-i\lambda_i)(-i\lambda_j)e^{i\langle x-m,\lambda\rangle}(-\frac{1}{2}(2\lambda_i\lambda_j))e^{-\frac{1}{2}\langle\Sigma\lambda,\lambda\rangle} \frac{d\lambda}{(2\pi)^N},$$

which gives the result for $i \neq j$.

Bibliography

[1] Robert J. Adler and Jonathan E. Taylor, *Random fields and geometry*, Springer Science & Business Media, 2009.

[2] Kenneth S. Alexander, *Boundedness of level lines for two-dimensional random fields*, Ann. Probab. **24** (1996), no. 4, 1653–1674.

[3] Jean-Marc Azaïs and Mario Wschebor, *Level sets and extrema of random processes and fields*, John Wiley & Sons, 2009.

[4] Vincent Beffara and Damien Gayet, *Percolation of random nodal lines*, Publications mathématiques de l'IHÉS **126** (2017), no. 1, 131–176.

[5] Dmitry Beliaev and Stephen Muirhead, *Discretisation schemes for level sets of planar gaussian fields*, Communications in Mathematical Physics **359** (2018), no. 3, 869–913.

[6] Dmitry Beliaev, Stephen Muirhead, and Alejandro Rivera, *A covariance formula for topological events of smooth gaussian fields*, arXiv preprint arXiv:1811.08169 (2018).

[7] Dmitry Beliaev, Stephen Muirhead, and Igor Wigman, *Russo-seymour-welsh estimates for the kostlan ensemble of random polynomials*, arXiv preprint arXiv:1709.08961 (2017).

[8] Pierre Bérard and Sylvestre Gallot, *Volume des ensemble nodaux des fonctions propes du laplacien*, Seminaire Equations aux derivees partielles (Polytechnique) **14** (1985), 1–9.

[9] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy, *Real algebraic geometry*, vol. 36, Springer Science & Business Media, 2013.

[10] Isaac Chavel, *Riemannian geometry: a modern introduction*, vol. 98, Cambridge university press, 2006.

[11] Damien Gayet and Jean-Yves Welschinger, *Exponential rarefaction of real curves with many components.*, Publ. Math., Inst. Hautes Étud. Sci. **113** (2011), 69–96.

[12] Damien Gayet and Jean-Yves Welschinger, *Lower estimates for the expected betti numbers of random real hypersurfaces*, Journal of the London Mathematical Society **90** (2014), no. 1, 105–120.

[13] Damien Gayet and Jean-Yves Welschinger, *Betti numbers of random real hypersurfaces and determinants of random symmetric matrices.*, J. Eur. Math. Soc. **18** (2016), no. 4, 733–772.

[14] Phillip Griffiths and Joseph Harris, *Principles of algebraic geometry*, John Wiley & Sons, 2014.

[15] Axel Harnack, *Ueber die Vieltheiligkeit der ebenen algebraischen Curven*, Mathematische Annalen **10** (1876), no. 2, 189–198.

[16] Lars Hörmander, *The spectral function of an elliptic operator*, Acta Math. **121** (1968), 193–218.

[17] Marek Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320.

[18] Harry Kesten, *Percolation theory for mathematicians*, vol. 423, Springer, 1982.

[19] Eric Kostlan, *On the distribution of roots of random polynomials*, pp. 419–431, Springer US, New York, NY, 1993.

[20] Jacques Lafontaine, *Introduction aux variétés différentielles*, EDP sciences, 2012.

[21] Antonio Lerario and Erik Lundberg, *Statistics on Hilbert's 16th problem*, International Mathematics Research Notices **2015** (2014), no. 12, 4293–4321.

[22] Thomas Letendre, *Expected volume and Euler characteristic of random submanifolds*, Journal of Functional Analysis **270** (2016), no. 8, 3047–3110.

[23] ———, *Variance of the volume of random real algebraic submanifolds*, Transactions of the American Mathematical Society (2018).

[24] John Milnor, *On the betti numbers of real varieties*, Proceedings of the American Mathematical Society **15** (1964), no. 2, 275–280.

[25] Fedor Nazarov and Mikhail Sodin, *On the number of nodal domains of random spherical harmonics*, American Journal of Mathematics **131** (2009), no. 5, 1337–1357.

[26] ———, *Asymptotic laws for the spatial distribution and the number of connected components of zero sets of Gaussian random functions.*, Zh. Mat. Fiz. Anal. Geom. **12** (2016), no. 3, 205–278.

[27] Vladimir Piterbarg, *Asymptotic methods in the theory of Gaussian processes and fields*, vol. 148, American Mathematical Soc., 2012.

[28] Loren D. Pitt, *Positively correlated normal variables are associated.*, Ann. Probab. **10** (1982), 496–499.

[29] Alejandro Rivera and Hugo Vanneuville, *The critical threshold for Bargmann-Fock percolation*, arXiv:1711.05012 (2017).

[30] ———, *Quasi-independence for nodal lines*, arXiv:1711.05009 (2017).

[31] Lucio Russo, *A note on percolation*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **43** (1978), no. 1, 39–48.

[32] Herbert Seifert, *Algebraische approximation von mannigfaltigkeiten*, Mathematische Zeitschrift **41** (1936), no. 1, 1–17.

[33] Paul D. Seymour and Dominic J. A. Welsh, *Percolation probabilities on the square lattice*, Annals of Discrete Mathematics, vol. 3, Elsevier, 1978, pp. 227–245.

[34] Bernard Shiffman and Steve Zelditch, *Distribution of zeros of random and quantum chaotic sections of positive line bundles*, Communications in Mathematical Physics **200** (1999), no. 3, 661–683.

[35] Michael Shub and Steve Smale, *Complexity of Bezout's theorem II volumes and probabilities*, Computational Algebraic Geometry (Boston, MA) (Frédéric Eyssette and André Galligo, eds.), Birkhäuser Boston, 1993, pp. 267–285.

[36] Vincent Tassion, *Crossing probabilities for Voronoi percolation*, The Annals of Probability **44** (2016), no. 5, 3385–3398.

[37] René Thom, *Sur l'homologie des variétés algébriques*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), 1965, pp. 255–265.