Lectures on Random Algebraic Geometry

Paul Breiding and Antonio Lerario

Contents

1 How many zeros of a polynomial are real?

How many zeros has a polynomial? The answer to this question is taught in a basic algebra course: it is equal to the degree of the polynomial. This is commonly known as Gauss' fundamental theorem algebra, and historically it laid the foundations for the subsequent development of what we know today as Algebraic Geometry.

However, Gauss' answer assumes that the question was stated as *How many complex zeros does a polynomial have?*. Yet, if the person, who asked that question, had in mind a real polynomial and real zeros, the answer is less clear. For instance, the polynomial $x^2 + ax + b$ has two real zeros, if $a^2 - 4b > 0$, it has one real zero, if $a^2 - 4b = 0$, and in the case $a^2 - 4b < 0$ it has no real zeros. The situation is depicted in Figure 1.1. This simple yet important example shows already that we can not give an answer to the above question in terms of the degree of the polynomial. Instead, we have to incorporate a list of algebraic equalities and inequalities. While for polynomials of degree 2 this was simple enough for us to understand, more complicated counting problems pose uncomparably harder challenges. Think of the number of 2-planes that tangentially touch a curve of the form $\{x \in \mathbb{R}^3 : q(x) = c(x) = 0\}$, where $q(x)$ is a polynomial of degree 2 and $c(x)$ is a polynomial of degree 3. Such curves are called genus-4-curves. The number of complex solutions of this counting problem is 120; see [12, Theorem 2.2]. But the algebraic constraints for the number of real solutions are so complicated, that it is very difficult just to prove which numbers of real solutions are possible [18]!

In this book we want to lay out an alternative perspective on enumerative problems like the ones above. Instead of computing a deterministic real picture, we want to know to understand its statistical properties. This thinking is not new: already in the 1930s and 1940s Littlewood, Offord [21] and Kac [13,14] considered real zeros of random polynomials. Later, in the 1950s, Wigner [25], Dyson [7] and others proposed using probabilistic methods for understanding models in theoretical physics. Ginibre [11] summarizes their motivation as follows.

> *"In the absence of any precise knowledge [...], one assumes a reasonable probability distribution [...], from which one deduces statistical properties [...]. Apart from the intrinsic interest of the problem, one may hope that the methods and results will provide further insight in the cases of physical interest or suggest as yet lacking applications."*

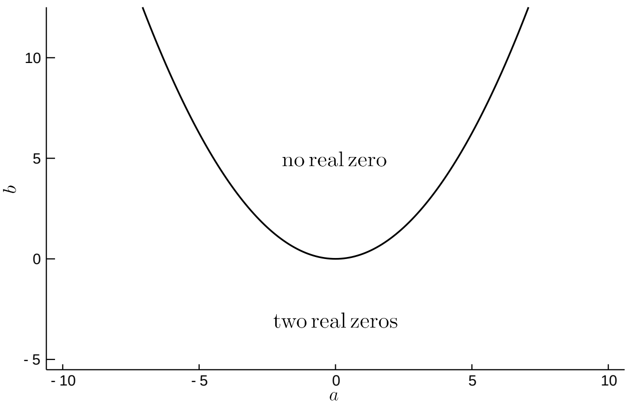

Figure 1.1: The configuration space for real zeros of the polynomial $f = x^2 + ax + b$. The blue curve $a^2 - 4b = 0$ is called the discriminant. If (a, b) is below the discriminant, then f has two real zeros. If it is above, it has no real zero. Polynomials on the discriminant have one real zero.

Although written in the context of statistical physics, Ginibre's words perfectly outline the ideas we wish to present with this book: we want to use tools from probability theory to understand the nature of algebraic–geometric objects.

Edelman and Kostlan [8] condense the probabilistic approach in the title of their seminal paper "How many zeros of a random polynomial are real?" (the answer is in Example 2 below). We chose the title of this introductory section as a homage of their work. Starting from their results, we explore in this book algebraic geometry from a probabilistic point of view. Our name for this new field of research is *Random Algebraic Geometry.*

Here is an illustrative example of what we have in mind: consider the degree 8 polynomial $f_\epsilon(x) = 1 + \epsilon_1 x + \epsilon_2 x^2 + \epsilon_3 x^3 + \epsilon_4 x^4 + \epsilon_5 x^5 + \epsilon_6 x^6 + \epsilon_7 x^7 + \epsilon_8 x^8$, where $\epsilon = (\epsilon_1, \ldots, \epsilon_8) \in \{-1, 1\}^8$. This polynomial can have $0, 2, 4, 6$ or 8 zeros, because complex zeros come in conjugate pairs. Instead of attempting to understand the equations separating the regions with a certain number of real solutions, we endow the coefficients of f_ϵ with a probability distribution. We assume that $\epsilon_1, \ldots, \epsilon_8$ are independent random variables with $\mathbb{P}\{\epsilon_i = 1\} = \frac{1}{2}$ for $1 \leq i \leq 8$, and we denote by $n(\epsilon)$ the random variable "number of real zeros of f_ϵ". Booth [3] showed that

$$\begin{aligned} &\mathbb{P}\{n(\epsilon) = 0\} = \frac{58}{2^8}, \quad \mathbb{P}\{n(\epsilon) = 2\} = \frac{190}{2^8}, \quad \mathbb{P}\{n(\epsilon) = 4\} = \frac{8}{2^8}, \quad \text{and} \\ &\mathbb{P}\{n(\epsilon) = 6\} = \mathbb{P}\{n(\epsilon) = 8\} = 0, \end{aligned}$$

which shows that f_ϵ has at most 4 zeros, and having more than 2 zeros is unlikely.

In Booth's example we have access to the full probability law. However, during this book we will encounter many situations in which computing the probability law is too ambitious. Instead, it is often feasible to compute or estimate the expected value of a random geometric property. For instance, in Booth's example the expected value of the number of roots is $\mathbb{E}\, n(\epsilon) = 1.609375$, and the variance is $\operatorname{Var} n(\epsilon) = 0.879$. Just based on this information we can conclude that having a large number of zeros is an exceptional property. Interestingly, many of the expected values we will meet obey what the "square-root law": the expected number of real solutions is roughly the square-root of the number of complex solutions. If this law holds, it immediately implies that instances, for which the number of real solutions equal the number of complex solutions, are *rarae aves.*

1.1 Reasonable probability distributions

In the quote of Ginibre it says "one assumes a reasonable probability distribution". He was probably thinking of physically meaningful distributions. But for us this means the following: suppose that $\mathcal{F}$ is a space of geometric objects endowed with a probability distribution, and that $X : \mathcal{F} \to \mathbb{R}^m$ is a random variable on $\mathcal{F}$. If X has symmetries, by which we mean that there is a group G acting on $\mathcal{F}$, such that $X(g.f) = X(f)$ for all g, then the probability distribution is reasonable, if it is invariant under G; that is $g.f \sim f$. This interpretation follows the *Erlangen program* by Felix Klein. In "A comparative review of recent researches in geometry" [15] Klein lays out a perspective on geometry based on a group of symmetries:

> *"Geometric properties are characterized by their remaining invariant under the transformations of the principal group."*

He writes that geometry should be seen as the following comprehensive problem.

> *"Given a manifoldness and a group of transformations of the same; to investigate the configurations belonging to the manifoldness with regard to such properties as are not altered by the transformations of the group."*

Therefore, reasonable probability distributions are distributions which respect geometry in Klein's sense. A reasonable probability distribution should not prefer one instance over another if they share the same geometry.

To illustrate this line of thought, we recall Booth's example within the framework above. The space of geometric objects $\mathcal{F}$ is the space of univariate polynomials of degree 8 with coefficients in $\{-1, 1\}$. The random variable $X(f)$ is the number of real zeros of the polynomial $f \in \mathcal{F}$. The group $G = \{-1, 1\}$ acts on $\mathcal{F}$ as $g.f(x) = 1 + \epsilon_1' x + \epsilon_2' x^2 + \epsilon_3' x^3 + \epsilon_4' x^4 + \epsilon_5' x^5 + \epsilon_6' x^6 + \epsilon_7' x^7 + \epsilon_8' x^8$, where $\epsilon_i' = \epsilon_i g^i$.

Since for all i we have $\epsilon_i g^i \in \{-\epsilon_i, \epsilon_i\}$ and since $\epsilon_i \sim -\epsilon_i$, we see that $gf \sim f$. In this sense, the distribution proposed by Booth is reasonable. In many cases the space $\mathcal{F}$ comes with the structure of a smooth manifold (e.g. a vector space, a Lie group or a homogeneous space) and in this case a "reasonable" probability distribution should be absolutely continuous with respect to Lebesgue measure (notice that the notion of sets of measure zero is well defined on a smooth manifold and independent of the possible choice of an actual measure).

1.2 Discriminants

Let us have a closer look at the picture in Figure 1.1. The discriminant $a^2 - 4b = 0$ divides the real (a, b)-plane into two components – one, where the number of real zeros is two, and one, where there are no real zeros. This is because Σ is a curve of codimension 1. The complex picture is different: here, the complex curve $\Sigma = \{(a, b) \in \mathbb{C}^2 : a^2 - 4b = 0\}$ is of *complex codimension one.* In particular, it is of real codimesion two, and $\mathbb{C}^2\backslash\Sigma$ is connected! This can be seen in Figure 1.2 and it is essentially the reason for why each polynomial of degree 2 outside Σ has two complex zeros: a function which is locally constant on a connected space is constant. We say that having two complex zeros is a *generic property.*

The propertiy of having two distinct complex zeros is invariant under scaling of the polynomials. In fact, being invariant under scaling holds for most properties considered in this book. The appropriate geometric framework for this is projective space.

Definition 1.1 (Complex projective Space)**.** The complex projective space $\mathbb{C}\mathrm{P}^n$ of dimension n is defined to be the set of lines through the origin in $\mathbb{C}^{n+1}$. That is, $\mathbb{C}\mathrm{P}^n := (\mathbb{C}^{n+1}\backslash\{0\})/\sim$, where $y \sim z$, if and only of there exists some $\lambda \in \mathbb{C}\backslash\{0\}$ with $y = \lambda z$. For a point $(z_0, z_1, \ldots, z_n) \in \mathbb{C}^{n+1}$ we denote by $[z_0, z_1, \ldots, z_n]$ its equivalence class in $\mathbb{C}\mathrm{P}^n$.

Here are a few more illustrative examples of generic properties. The first generalizes our introductory example to higher degrees.

Example 1. 1. A generic univariate polynomial $f \in \mathbb{C}[x]$ of degree d has d distinct zeros in $\mathbb{C}$ unless $\mathrm{Res}(f, f') = 0$.

2. The zero set $Z^{\mathbb{C}}(f) \subset \mathbb{C}\mathrm{P}^2$ of a generic $f \in \mathbb{C}[x_0, x_1, x_2]_{(d)}$ of degree d is homeomorphic to a genus $\frac{(d-1)(d-2)}{2}$ surface.
3. The zero set $Z^{\mathbb{C}}(f) \subset \mathbb{C}\mathrm{P}^3$ of a generic cubic $f \in \mathbb{C}[x_0, x_1, x_2, x_3]_{(3)}$ contains 27 complex lines.

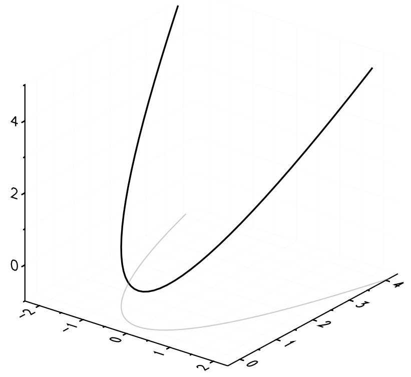

Figure 1.2: The picture shows the part of the complex discriminant $(a_1 + ia_2)^2 - 4(b_1 + ib_2) = 0$, where $a_1 = 2a_2$ As can be seen from the picture, the discriminant is of real codimension two. Because one can "go around" the discriminant without crossing it, a generic complex polynomial of degree 2 has two complex zeros.

For completing the terminology, and distinguishing it from projective space, we say that $\mathbb{C}^n$ is an n-dimensional *affine complex space*. The map

$$P : \mathbb{C}^{n+1} \backslash \{0\} \to \mathbb{C}\mathrm{P}^n \ , (z_0, z_1, \ldots, z_n) \mapsto [z_0, \ldots, z_n]$$

projects $n+1$-dimensional affine space onto n-dimensional projective space. On the other hand, the map $\psi : \mathbb{C}^n \to \mathbb{C}\mathrm{P}^n$, $(z_1, \ldots, z_n) \mapsto [1, z_1, \ldots, z_n]$ embeds n-dimensional affine space into n-dimensional projective space. One says that $\mathbb{C}^n$ is an *affine patch* of $\mathbb{C}\mathrm{P}^n$. Using this embedding we can define the zero sets in Example 1 to be in $\mathbb{C}\mathrm{P}^n$.

Such projective zero sets are defined by *homogeneous polynomials.* It is common to use the notation $f = \sum_{|\alpha|=d} f_\alpha \, x^\alpha$ for complex homogeneous polynomials of degree d in $n+1$ variables, where

$$\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1}, \quad z^\alpha = \prod_{i=0}^{n} z_i^{\alpha_i} \quad \text{and} \quad |\alpha| = \alpha_0 + \cdots + \alpha_n.$$

The space of homogeneous polynomials of degree d in $n+1$ many variables is

$$\mathbb{C}[x_0, \ldots, x_n]_{(d)} := \left\{ \sum_{|\alpha|=d} f_\alpha \, z^\alpha \;\middle|\; (f_\alpha) \in \mathbb{C}^N \right\}, \text{ where } N = \binom{n+d}{d},$$

and the projective space of polynomials is $P(\mathbb{C}[z_0, \ldots, z_n]_{(d)})$.

The complex projective zero set of k polynomials $f=(f_1,\dots,f_k)$, where the i-th polynomial is $f_i \in \mathbb{C}[z_0,\dots,z_n]_{(d_i)}$, is

$$Z^{\mathbb{C}}(f)=\{[z]\in \mathbb{C}\mathrm{P}^n : f_1(z)=0,\dots,f_k(z)=0\}, \tag{1.1}$$

Remark 1.2. A polynomial $f \in \mathbb{C}[z_0,\dots,z_n]_{(d)}$ is not a function on the complex projective space $\mathbb{C}\mathrm{P}^n$, but its zero set is still well defined. In fact, it is a section of an appropriate line bundle over the projective space (see Section 2.4) and the zero set of this section is the zero set of the polynomial.

In most cases the property we will be interested in are described by a list of numbers associated to elements of S. Let us re-interpret the statement from Example 1 using this language. If $S=P(\mathbb{C}[z_0,z_1]_{(d)})=\mathbb{C}\mathrm{P}^d$ is the projective space of complex polynomials of degree d, we might be interested in the number of zeroes of these polynomials. We can interpret this number as a map $\beta:\mathbb{C}\mathrm{P}^d\to\mathbb{C}$ given by

$$\beta : f \mapsto \#\{f=0\}.$$

This β is a constant map outside $\Sigma=\{\text{polynomials with multiple roots}\}$ (this is the classical discriminant hypersurface for space of polynomials).

The next definition gives a rigorous definition for genericity in our setting.

Definition 1.3 (Generic Properties)**.** Let S be a semialgebraic set. We say that a property β is *generic* for the elements of S if there exists a semialgebraic set $\Sigma\subset S$ of codimension at least one in S such that the property β is true for all elements in $S\backslash\Sigma$. We call the largest (by inclusion) such Σ the *discriminant* of the property β.

When working over the complex numbers most properties are generic in the sense that the discriminant is a proper complex algebraic set. Since proper complex algebraic sets in $\mathbb{C}\mathrm{P}^N$ do not disconnect the whole space, these properties are constant on an open dense set. This is a simple observation that we record in the next lemma.

Lemma 1.4. *Let $\Sigma\subsetneq\mathbb{C}\mathrm{P}^N$ be a proper algebraic subset. Then, $\mathbb{C}\mathrm{P}^N\backslash\Sigma$ is connected.*

Proof. Let $z_1,z_2\in\mathbb{C}\mathrm{P}^N\backslash\Sigma$. Choose a complex linear space $L\subset\mathbb{C}\mathrm{P}^N$ such that $z_1,z_2\in L$. Then, L intersects Σ either in finitely many points, or $L\subset\Sigma$. The latter is a contradiction to $z_1,z_2\notin\Sigma$, and so $L\cap\Sigma$ is finite. Thus $L\backslash\Sigma$ is path-connected and we find a real path from z_1 to z_2 that does not intersect Σ. □

The second point from Example 1 gives another occurence of this phenomenon. Let $\mathbb{C}\mathrm{P}^N = P(\mathbb{C}[z_0, \dots, z_n]_{(d)})$ be the projective space of polynomials and consider the property $\beta : \mathbb{C}\mathrm{P}^N \to \mathbb{C}^{2n+1}$ given by

$$\beta(f) = (b_0(Z^{\mathbb{C}}(f)), \dots, b_{2n}(Z^{\mathbb{C}}(f)))$$

(i.e. $\beta(f)$ is the list of the Betti numbers of the zero set of f in $\mathbb{C}\mathrm{P}^n$; this number does not depend on the representative of f that we pick, as a nonzero multiple of a polynomial has the same zero set as the original polynomial). The property β in this case takes a constant value on the complement of a complex discriminant $\Sigma \subset \mathbb{C}\mathrm{P}^N$. In the case $n = 2$ we have that

$$\beta(f) = (1, (d-1)(d-2), 1)$$

for all $f \in \mathbb{C}\mathrm{P}^n \backslash \Sigma$. A similar argument can be done for the third point in Example 1: the property "number of lines on the zero set of f" is constant outside a complex discriminant $\Sigma \subset P(\mathbb{C}[z_0, \dots, z_3]_{(3)})$ (in this case this discriminant coincides with the set of sections of an appropriate bundle which are not transversal to the zero section, we will discuss this better later, see Section 2.5.1).

As already briefly discussed in the beginning of this section, the topological reason for the existence of such strong generic properties over the complex numbers is Lemma 1.4. The formal statement goes under the name of *Thom's Isotopy Lemma* and we will prove it and discuss its implications later in Section 3.1.

1.3 Expected properties

Let us copy the notation from the preceding section to the real numbers.

Definition 1.5 (Real projective Space)**.** The real projective space $\mathbb{R}\mathrm{P}^n$ of dimension n is defined to be the set of lines through the origin in $\mathbb{R}^{n+1}$. That is, $\mathbb{R}\mathrm{P}^n := (\mathbb{R}^{n+1} \backslash \{0\}) / \sim$, where $y \sim z$, if and only of there exists some $\lambda \in \mathbb{R} \backslash \{0\}$ with $y = \lambda z$. For a point $(x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1}$ we denote by $[x_0, x_1, \dots, x_n]$ its equivalence class in $\mathbb{R}\mathrm{P}^n$.

Similar to before, we define the projection

$$P : \mathbb{R}^{n+1} \backslash \{0\} \to \mathbb{R}\mathrm{P}^n,\ (x_0, x_1, \dots, x_n) \mapsto [x_0, x_1, \dots, x_n]. \tag{1.2}$$

The space of real homogeneous polynomials is

$$\mathbb{R}[x_0, \dots, x_n]_{(d)} := \left\{ \sum_{|\alpha|=d} f_\alpha \, x^\alpha \;\middle|\; (f_\alpha) \in \mathbb{R}^N \right\},\ \text{where } N = \binom{n+d}{d}.$$

The projective space of real polynomials is $P(\mathbb{R}[x_0,\ldots,x_n]_{(d)})$. The real projective algebraic set of k polynomials $f=(f_1,\ldots,f_k)$ is

$$Z(f)=\{[x]\in\mathbb{R}\mathrm{P}^n : f_1(x)=0,\ldots,f_k(x)=0\}. \tag{1.3}$$

As we have seen, if the discriminant is a complex algebraic set, we have strong genericity over the complex numbers: as we noted above, the reason for this is Lemma 1.4, which says that the complex discriminant does not disconnect $\mathbb{C}\mathrm{P}^N$. However, if the discriminant is a real hypersurface, in general it might disconnect $\mathbb{R}\mathrm{P}^N$, this is why in Figure 1.1 there are two regions with different properties. Therefore, over the real numbers we might not have a notion of strong genericity. Instead, we consider the probabilistic picture. The next definition is the probabilistic analogue of Definition 1.3.

Definition 1.6 (Expected Properties)**.** Let S be a semialgebraic set. A measurable property is a measurable function $\beta : S\to\mathbb{C}^m$. If we have a (reasonable) probability distribution on S, we call $\mathbb{E}_{s\in S}\,\beta(s)$ the expected property.

In fact, Definition 1.3 is a special case of Definition 1.6. We will discuss this in Section 1.5 below. First, let us revisit Example 1 from a probabilistic point of view. For this, we have to agree on a reasonable probability distribution. The distribution of our choice is the *Kostlan polynomial ensemble*. Kostlan polynomials have coefficients which are Gaussian random variables.

Definition 1.7 (Gaussian random variables)**.** A random variable $X\in\mathbb{R}$ is Gaussian with mean μ and variance σ^2, if $\mathbb{P}\{X\in A\}=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{x\in A} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x$ for any measurable subset $A\subset\mathbb{R}$. We write $X\sim N(\mu,\sigma^2)$.

For defining the Kostlan ensemble it is helpful to use the following notation:

$$\binom{d}{\alpha}:=\frac{d!}{\alpha_0!\cdots\alpha_n!}.$$

Definition 1.8 (Kostlan polynomials)**.** Let $f=\sum_{|\alpha|=d} f_\alpha\, x^\alpha\in\mathbb{R}[x_0,\ldots,x_n]_{(d)}$ be a homogeneous polnomial of degree d. We call f a Kostlan polynomial, if the coefficients f_α are all independent Gaussian random variables with mean $\mu=0$ and variance $\sigma^2=\binom{d}{\alpha}$.

Kostlan polynomials are also called *Bombieri-Weyl polynomial*, or *Shub-Smale random polynomial.* Note that we defined Kostlan polynomials to be elements in $\mathbb{R}[x_0,\ldots,x_n]_{(d)}$, and not in the projective space $P(\mathbb{R}[x_0,\ldots,x_n]_{(d)})$. The random polynomial $P(f)$, where f is a Kostlan polynomial, will also be important in this book. But it will be called a polynomial from the *uniform distribution*. The reason for this is that $P(f)$ follows the uniform distribution defined by the so-called *Bombieri-Weyl metric*, which is in a sense the natural metric for polynomials as we discuss in Subsection 3.3.2 and Section 2.1.

Kostlan polynomials are reasonable in the following rigorous sense (the fact that the Kostlan probability distribution on $\mathbb{R}[x_0,\ldots,x_n]_{(d)}$ is absolutely continuous with respect to Lebesgue measure is clear).

Lemma 1.9. *Let $O(n+1) \subset \mathbb{R}^{(n+1)\times(n+1)}$ be the orthogonal group and let f be a Kostlan polynomial. For all $O \in O(n+1)$ we have $f \circ O \sim f$.*

We postpone the proof of this lemma until we give a thorough discussion of probability distribution which are invariant under group actions. To measure volumes of zero sets in $\mathbb{R}\mathrm{P}^n$ it is reasonable to take a measure that is invariant under the action of the orthogonal group, too. The *Fubini-Study measure* serves this purpose. In the following, when talking about volumes of sets in projective space, we will always mean volume with respect to the Fubini-Study measure.

Definition 1.10 (Uniform measure)**.** Let $S^n = \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ be the n-dimensional unit sphere, where $\|x\| = \sqrt{\langle x, x\rangle}$ is the norm induced by the euclidean inner product $\langle\cdot,\cdot\rangle$. The uniform measure on $\mathbb{R}\mathrm{P}^n$ is the push-forward measure of the standard measure on S^n under $\pi|_{S^n}$, where π is the projection (1.2). In particular, for any measurable subset $A \subset \mathbb{R}\mathrm{P}^n$ we have $\mathrm{vol}(A) = \frac{1}{2}\mathrm{vol}(\pi|_{S^n}^{-1}(A))$.

To get an intuition for the uniform measure, the reader can think of $\mathbb{R}\mathrm{P}^n$ as the "upper half" of the sphere S^n. Volumes of sets in $\mathbb{R}\mathrm{P}^n$ are then precisely the volumes of the corresponding sets in the upper half. Furthermore, as we will see, $\mathbb{R}\mathrm{P}^n$ is a smooth manifold, and the uniform measure is induced by a volume density coming from a riemannian metric.

Now, we revisit Example 1 over the real numbers.

Example 2. 1. Let $f \in \mathbb{R}[x_0,x_1]_{(d)}$ be a Kostlan polynomial of degree d in 2 variables. Then, the expected number of real zeros of f is $\sqrt{d}$.

2. Let $f \in \mathbb{R}[x_0,x_1,x_2]_{(d)}$ be a Kostlan polynomial of degree d in 3 variables. There exist constants $c, C > 0$ such that the expected value of the zero-th Betti number $b_0(f)$ of $Z(f)$ satisfies $cd \le \mathbb{E}\, b_0(f) \le Cd$.
3. Let $f \in \mathbb{R}[x_0,x_1,x_2,x_3]_{(3)}$ be a Kostlan polynomial of degree 3 in 3 variables. Then, the expected number of real lines on $Z(f)$ is $6\sqrt{2}-3$.

The first example was proven in [8], the second in [9], and the third in [2], but we will also prove them in the remainder of this book. We want to emphasize that the first two of those examples obey a square-root law – the expected value of the real property has the order of the square root of the generic value of the complex property.

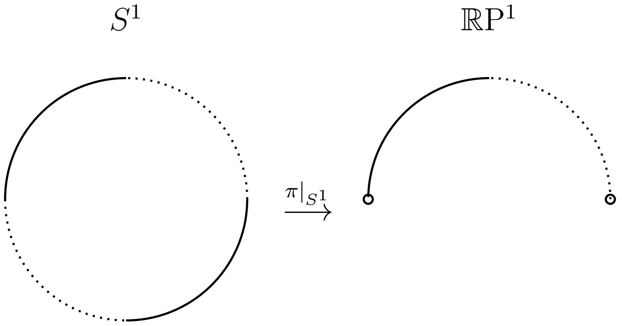

Figure 1.3: A cartoon of the projection $\pi|_{S^1} : S^1 \to \mathbb{R}\mathrm{P}^1$. The volume of the dotted area in $\mathbb{R}\mathrm{P}^1$ is equal to half the volume of the dotted area in S^1. The two antipodal points in the right picture represent the same point in $\mathbb{R}\mathrm{P}^1$.

1.4 The number of zeros of a Kostlan polynomial

We will sketch two proofs of Example 2 1, using the two major tools for approaching these type of problems presented in this book: the *Kac-Rice formula* and the *Integral Geometry formula.* For this, we consider a Kostlan polynomial $f = \sum_{i=0}^{d} c_i x_0^i x_1^{d-i}$ of degree d in two variables.

1.4.1 The Kac-Rice approach

The *Kac-Rice* formula is the main topic of Chapter 2. The full formula is given in Theorem 2.9, but in the case of Kostlan polynomials in two variables it is

$$\mathbb{E}\ \#\{z \in \mathbb{R}\mathrm{P}^1 \mid f(z) = 0\} = \frac{\sqrt{\pi}}{\sqrt{2}}\ \mathbb{E} \left| \frac{\partial f}{\partial x_1}\ \Big|_{(x_0,x_1)=(1,0)} \right| ;$$

see (2.12). This implies, that the expected number of zeros of f is $\sqrt{\frac{\pi}{2}}\ \mathbb{E}\,|c_1|$ with $c_1 \sim N(0, d)$. We get

$$\mathbb{E}\ \#\{z \in \mathbb{R}\mathrm{P}^1 \mid f(z) = 0\} = \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2d^2}}\ \mathrm{d}x = \sqrt{\frac{2}{\pi d}} \int_0^{\infty} x e^{-\frac{x^2}{2d^2}}\ \mathrm{d}x,$$

the last step by symmetry. The change of variables $y = \frac{x}{d}$ yields

$$\mathbb{E}\ \#\{z \in \mathbb{R}\mathrm{P}^1 \mid f(z) = 0\} = \sqrt{\frac{2d}{\pi}} \int_0^{\infty} y e^{-\frac{y^2}{2}}\ \mathrm{d}y = \sqrt{\frac{2d}{\pi}}.$$

This shows, that $\mathbb{E}\,\#\{z \in \mathbb{RP}^1 \mid f(z) = 0\} = \sqrt{d}$.

1.4.2 The Integral Geometry approach

The integral geometry approch takes a different point of view on equations, namely as linear sections of an appropriate *rational normal curve.*

Definition 1.11 (Kostlan Rational normal curve)**.** The map

$$\nu_d : \mathbb{R}\mathrm{P}^1 \to \mathbb{R}\mathrm{P}^n,\ [x_0, x_1] \mapsto \left[x_0^d : \sqrt{d} x_0^{d-1} x_1, \ldots, \binom{d}{k}^{1/2} x_0^{d-k} x_1^k, \ldots, \sqrt{d} x_0 x_1^{d-2}, x_1^d \right]$$

will be called the d-th Kostlan-Veronese embedding of $\mathbb{R}\mathrm{P}^1$ into $\mathbb{R}\mathrm{P}^d$. The image $\nu(\mathbb{R}\mathrm{P}^1)$ is called the d-th Kostlan rational normal curve.

It is straightforward to show that ν_d is injective, so that it is indeed an embedding. Now, zeros of

$$f(x_0, x_1) = \sum_{k=0}^{d} \xi_k \binom{d}{k}^{1/2} x_0^{d-k} x_1^k$$

correspond to points in the intersection of $\nu(\mathbb{R}\mathrm{P}^1)$ with the linear space $L = \{[y_0, \ldots, y_d] \in \mathbb{R}\mathrm{P}^d \mid \xi_0 y_0 + \cdots + \xi_d y_d = 0\}$. If L is random, such that $\xi_k \sim N(0,1)$, the integral geometry formula for the rational normal curve is as follows:

$$\mathbb{E}\ \#(L \cap \nu(\mathbb{R}\mathrm{P}^1)) = \frac{\mathrm{vol}(\nu(\mathbb{R}\mathrm{P}^1))}{\mathrm{vol}(\mathbb{R}\mathrm{P}^1)}. \tag{1.4}$$

An elementary proof of the next theorem is given in [8]. We want to give a more intuitive and shorter proof, but for this we need a bit more machinery. This is why we postpone the proof of the next theorem.

Theorem 1.12 (Volume of the rational normal curve)**.** *We have*

$$\mathrm{vol}(\nu(\mathbb{R}\mathrm{P}^1)) = \sqrt{d}\,\mathrm{vol}(\mathbb{R}\mathrm{P}^1).$$

This theorem, combined with (1.4) shows that the expected number of zeros of a Kostlan polynomial is $\sqrt{d}$. Readers familiar in algebraic geometry can see the formula in (1.4) as a metric analogue of a divisor, especially in the light of formula (1.6) below.

1.5 Generic properties are expected properties

In closing of this introductory chapter we want to explain why generic properties are, in fact, random properties in disguise. This essence of this is a simple observation: suppose $z \in \mathbb{C}\mathrm{P}^N$ is a random variable that is supported on some open subset of $\mathbb{C}\mathrm{P}^N$. In particular, this implies that, if P is a property with discriminant Σ, and if $\Sigma \subsetneq \mathbb{C}\mathrm{P}^N$ is an algebraic variety, then $\mathbb{P}\{z \in \Sigma\} = 0$, and so $\mathbb{P}\{P(z) \text{ has the generic value}\} = 1$. Therefore

$$\mathbb{E}\, P(z) = \text{ generic value of } P(z).$$

It is interesting to approach the problem of computing generic properties from a probablistic point of view. For instance, below we give a proof of a probabilistic Fundamental Theorem of Algebra. This strategy becomes more effective as the counting problem over the complex numbers becomes more complicated. Consider in fact the random polynomial

$$f = \sum_{i=0}^{d} c_i x_0^i x_1^{d-i} \in \mathbb{C}[x_0, x_i]_{(d)}, \tag{1.5}$$

where the real and imaginary parts of the c_i are independent Gaussian random variables such that $\Re(c_i) \sim N(0, \frac{1}{2}\binom{d}{i})$ and $\Im(c_i) \sim N(0, \frac{1}{2}\binom{d}{i})$ (the factor $\frac{1}{2}$ is for normalizing the variance to $\mathbb{E}\,|c_i|^2 = 1$). Such a polynomial is called a *complex Kostlan polynomial.* The distribution we have put on the coefficients is absolutely continuous with respect to Lebesgue measure on the space of coefficients, and in fact the distribution of $P(f)$ is supported on the whole $\mathbb{C}\mathrm{P}^d$. Therefore, we know that with probability one we have that $\#Z^{\mathbb{C}}(f)$ equals some constant (we know this constant is d, but let's pretend for a second that we did not know this).

As in Subsection 1.4.2 each zero of f corresponds to a point in the intersection of the random linear space $L = \{[y] = [y_0, \ldots, y_d] \in \mathbb{C}\mathrm{P}^d \mid \zeta_0 y_0 + \cdots + \zeta_d y_d = 0\}$, where $\{\zeta_k\}$ is a family of i.i.d. standard complex Gaussians, with the complex Kostlan rational normal curve

$$\nu_d(\mathbb{C}\mathrm{P}^1) = \left\{ \left[z_0^d, \sqrt{d} z_0^{d-1} z_1, \ldots, \binom{d}{k}^{1/2} z_0^{d-k} z_1^k, \ldots, \sqrt{d} z_0 z_1^{d-1}, z_1^d \right] \;\middle|\; [z_0, z_1] \in \mathbb{C}\mathrm{P}^1 \right\}.$$

The complex integral geometry formula for the expected value of the number of points in this intersection is $\mathbb{E}\ \#(L \cap \nu(\mathbb{C}\mathrm{P}^1)) = \frac{\mathrm{vol}(\nu(\mathbb{C}\mathrm{P}^1))}{\mathrm{vol}(\mathbb{C}\mathrm{P}^1)}$. We will show below that

$$\mathrm{vol}(\nu(\mathbb{C}\mathrm{P}^1)) = d\, \mathrm{vol}(\mathbb{C}\mathrm{P}^1). \tag{1.6}$$

In fact, what the complex integral geometry formula shows is that the degree of a complex algebraic variety is a multiple of its volume. This observation will be discussed in a later chapter.

2 Counting formulas

In this lecture we address the basic problem of counting the number of solutions of a random equation – many interesting questions from geometry to topology can be reduced to a problem where we have to count points. To be more specificy, suppose we are given a *random* map

$$f : \mathbb{R}^m \to \mathbb{R}^m$$

(the precise meaning of random will be given in Section 2.1) and we are interested, in computing $\mathbb{E}\#\left(\{f = 0\} \cap A\right)$, where $A \subseteq \mathbb{R}^m$ is a measurable subset.

We have met such a random map in the first lecture, namely real Kostlan polynomials: as a random map above, we might take $f : \mathbb{R} \to \mathbb{R}$ to be given by $f(x) = \sum_{i=0}^{d} c_i x^i$, where $c_i \sim N(0, \binom{d}{i})$ (note that, however, this polynomial is not a homogeneous polynomial, but we will move to homogeneous polynomials later in this section). Another random map is given by using *Kac-polynomials.*

Definition 2.1 (Kac polynomials)**.** Let $f = \sum_{i=0}^{d} \xi_i \, x^i$ be a univariate polnomial of degree d. We call f a Kac polynomial, if the coefficients ξ_i are i.i.d. $N(0, 1)$ random variables.

We can also consider the case where the map f is given by a complex Kac polynomial. Using the identification of real vector space $\mathbb{R}^2 \cong \mathbb{C}$, we can take $f : \mathbb{C} \to \mathbb{C}$ defined by $f(z) = \sum_{i=0}^{d} \zeta_i \, z^i$, where $\{\zeta_k\}_{k=0,\dots,d}$ is a family of i.i.d. complex random variables such that $\Re(\zeta_k)$ and $\Im(\zeta_k)$ are independent and $N(0, \frac{1}{2})$.

2.1 What do we mean by a "random map"?

In the sequel, following the spirit of what we have called "reasonable" probability distribution, a random map is an element of some finite-dimensional Gaussian space of smooth functions. The more general case of random sections of vector bundles is discussed in Section 2.5.

In the following, $C^\infty(\mathbb{R}^m, \mathbb{R}^m)$ denotes the space of smooth function $\mathbb{R}^m \to \mathbb{R}^m$.

Definition 2.2 (Random Gaussian maps)**.** Let $\mathcal{F} = \{f_1, \dots, f_\ell\} \subset C^\infty(\mathbb{R}^m, \mathbb{R}^m)$ be finite. The random Gaussian map induced by $\mathcal{F}$ is

$$f(x) = \xi_1 f_1(x) + \cdots + \xi_\ell f_\ell(x),$$

where $\{\xi_k\}_{k=1,\ldots,\ell}$ is a family of i.i.d. $N(0,1)$ random variables.

Let $W := \operatorname{span}(\{f_1,\ldots,f_\ell\})$ be the vector space spanned by $\mathcal{F}$. If the elements $f_1,\ldots,f_\ell$ are linearly independent (i.e., if they form a basis for W), we can define a scalar product $\langle\cdot,\cdot\rangle_{\mathcal{F}}$ on W by declaring $\{f_1,\ldots,f_\ell\}$ to be an *orthonormal* basis. With this, the probability distribution induced by $\mathcal{F}$ on W is the same as defining for a measurable set $A \subseteq W$:

$$\mathbb{P}(A) = \frac{\int_A e^{-\frac{\|w\|^2}{2}}\,\mathrm{d}w}{\int_W e^{-\frac{\|w\|^2}{2}}\,\mathrm{d}w}, \text{ where } \|w\|^2 = \langle w,w\rangle_{\mathcal{F}}. \tag{2.1}$$

For example, in the case of real Kac polynomials we have $\mathcal{F} = \{1, x, \ldots, x^d\}$. For the complex Kac polynomials we can choose $\mathcal{F} = \{a_0,\ldots,a_d,b_0,\ldots,b_d\}$ where $a_k(x,y) = \frac{1}{2}\left(\begin{smallmatrix}\Re(x+iy)^k\\0\end{smallmatrix}\right)$ and $b_k(x,y) = \frac{1}{2}\left(\begin{smallmatrix}0\\\Im(x+iy)^k\end{smallmatrix}\right)$.

2.2 The Kac-Rice formula

We will now establish the framework, in which we can count the expected number of zeros of random maps.

For this, we assume that $f : \mathbb{R}^m \to \mathbb{R}^m$ is a random map induced by a set of smooth functions $\mathcal{F} = \{f_1,\ldots,f_\ell\}$, in the sense of Definition 2.2. Furthermore, we let $Jf(x)$ be the jacobian matrix of f at x. Then, for every fixed $x \in \mathbb{R}^m$ we have the random vector-matrix pair $(f(x), Jf(x)) \in \mathbb{R}^m \times \mathbb{R}^{m\times m}$. To assure that this random point is not supported on some lower dimensional subset of $\mathbb{R}^m\times\mathbb{R}^{m\times m}$, we assume that for every fixed $x \in \mathbb{R}^m$ the random vector $f(x)$ has a nondegenerate distribution; i.e., the that the covariance matrix is positive definite:

$$\mathbb{E}\ f(x)f(x)^T \succ 0.$$

Let us denote by $p : \mathbb{R}^m \times \mathbb{R}^{m\times m} \times \mathbb{R}^m \to \mathbb{R}$ the joint density of the pair $(f(x), Jf(x))$; i.e. p is the function defined by the requirement that for every measurable subset $A \subseteq \mathbb{R}^m \times \mathbb{R}^{m\times m}$ we have for a fixed $x \in \mathbb{R}^m$

$$\mathbb{P}\left((f(x), Jf(x)) \in A\right) = \int_A p(w,v,x)\,\mathrm{d}w\mathrm{d}v.$$

The Kac-Rice formula for $\mathcal{F}$ is given in terms of the *Kac-Rice density.*

Definition 2.3 (The Kac-Rice density)**.** The Kac-Rice density of $\mathcal{F}$ is

$$\rho(x) = \int_{\mathbb{R}^{m\times m}} |\det(v)|\, p(0,v,x)\,\mathrm{d}v.$$

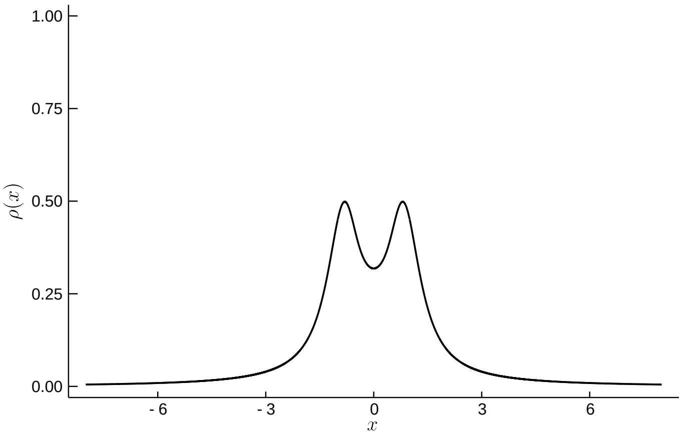

Figure 2.1: The root density $\rho(x)$ of Kac polynomials of degree $d = 4$. The two peaks are at $x = -1$ and $x = 1$. This means that a root of a Kac polynomial is most likely close to -1 or 1.

Now comes the Kac-Rice formula. A proof for the formula can be found, e.g., in [1, Theorem 6.2]. But later in this book we will give a self-contained proof.

Theorem 2.4 (Kac-Rice formula). *Let $f : \mathbb{R}^m \to \mathbb{R}^m$ be the random map induced by $\mathcal{F} \subset C^\infty(\mathbb{R}^m, \mathbb{R}^m)$, and let $A \subseteq \mathbb{R}^m$ be a measurable set. Assume that for every fixed $x \in \mathbb{R}^m$ we have*

1. *$f(x)$ has a nondegenerate distribution; i.e., $\mathbb{E}\, f(x)f(x)^T \succ 0$, and*
2. *the probability of $\det(Jf(x)) = 0$ conditioned on the event $f(x) = 0$ is equal to zero; i.e., $\mathbb{P}\{\det(Jf(x)) = 0 \mid f(x) = 0\} = 0$.*

Then, the expected number of zeroes of f in A is given by the formula:

$$\mathbb{E}\,\#(\{f = 0\} \cap A) = \int_A \rho(x)\,\mathrm{d}x.$$

The theorem shows the role of the Kac-Rice density $\rho(x)$. It is the density of the zeros of f. We thus call $\rho(x)$ also *root density.*

2.2.1 The root density of Kac polynomials

In this subsection we want to illustrate Theorem 2.4 in the case of Kac polynomials. The following theorem appeared in [13].

Theorem 2.5. *The root density of a Kac polynomial is*

$$\rho(x) = \frac{1}{\pi} \frac{\sqrt{1-h(x)^2}}{1-x^2}, \qquad \textit{where} \qquad h(y) = \frac{(d+1)x^d(1-x^2)}{1-x^{2(d+1)}}.$$

As one can see in Figure 2.1 the roots of Kac polynomials tend to accumulate in the interval $[-1,1]$. In particular, for $A = [a,b]$ with either $a,b \gg 0$ or $a,b \ll 0$, the expected number of zeros $\mathbb{E}\,\#(\{f=0\} \cap A)$ is almost zero.

It should be mentioned that one has not been able to derive a closed formula for $\int_{\mathbb{R}} \rho(x)\,\mathrm{d}x$. In Kac's original paper [13] he considered instead the *asymptotic* behavior of this integral. This is a recurring theme in this book: often one can not compute closed expressions of expected properties, but one can estimate the asymptotics when one (or several) parameters go to infinity. In the case of Kac polynomials we have the following.

Theorem 2.6. *The number of zeros of Kac polynomials satisfies for $d \to \infty$:*

$$\mathbb{E}\,\#\{f=0\} \sim \frac{2}{\pi} \log(d).$$

Before we prove the two theorems, we have to revisit *multivariate Gaussians.*

Definition 2.7 (Multivariate Gaussian random variables)**.** Let $X = (X_1, \ldots, X_n)^T$ be a vector of n random variables and $\mu \in \mathbb{R}^n$. We say that X is Gaussian with mean μ and covariance matrix Σ, if Σ is positive semidefinite, and the joint density of X is $(2\pi)^{\frac{-n}{2}}\, e^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)}$. We write $X \sim N(\mu, \Sigma)$.

Multivariate Gaussians behave nicely under linear transformations. For a proof of the next lemma see, e.g., [23, Theorem 1.2.6].

Lemma 2.8. *Let $X = (X_1, \ldots, X_n)$ be Gaussian with mean μ and covariance matrix Σ, and let $A \in \mathbb{R}^{m\times n}$. Then, $AX \sim N(A\mu, A\Sigma A^T)$.*

Now we prove Theorem 2.5 and Theorem 2.6.

Proof of Theorem 2.5 and Theorem 2.6. The following proof appeared in [13]. For a Kac polynomial f we have $f(x) = \sum_{i=0}^d \xi_i x^i$, and $Jf(x) = \sum_{i=0}^d i\xi_i x^{i-1}$. Consider the *Kac Rational normal curve*

$$\vartheta_d(x) = (1, x, x^2, \ldots, x^d)^T.$$

This is similar to the Kostlan Rational normal curve from Definition 1.11, the differences being that the Kac Rational normal curve is in affine space, and the scaling of the monomials. Let us also write $\xi = (\xi_0, \dots, \xi_d)^T$. Then, we have

$$f(x) = \langle \xi, \vartheta_d(x) \rangle \quad \text{and} \quad Jf(x) = \langle \xi, \vartheta_d'(x) \rangle.$$

By Lemma 2.8, $(f(x), Jf(x))$ is a vector of Gaussian random variables with covariance matrix

$$\Sigma = \begin{bmatrix} \langle \vartheta_d(x), \vartheta_d(x) \rangle & \langle \vartheta_d(x), \vartheta_d'(x) \rangle \\ \langle \vartheta_d'(x), \vartheta_d(x) \rangle & \langle \vartheta_d'(x), \vartheta_d'(x) \rangle \end{bmatrix}.$$

The joint density of $(u, v) = (f(x), Jf(x))$ therefore is

$$p(u, v, x) = \frac{1}{2\pi\sqrt{\det \Sigma}}\, e^{-\frac{1}{2}\,(u,v)\Sigma^{-1}(u,v)^T}.$$

This implies $p(0, v, x) = \frac{1}{2\pi\sqrt{\det\Sigma}} e^{-\frac{v^2}{2}\,\alpha}$, where $\alpha = (\det \Sigma)^{-1} \langle \vartheta_d(x), \vartheta_d(x) \rangle$. The corresponding Kac-Rice density is

$$\rho(x) = \int_{\mathbb{R}} |v|\, p(0, v, x)\, \mathrm{d}v = \frac{1}{2\pi\sqrt{\det \Sigma}} \int_{\mathbb{R}} |v|\; e^{-\frac{v^2}{2}\,\alpha}\, \mathrm{d}v = \frac{1}{\pi}\, \frac{\sqrt{\det \Sigma}}{\langle \vartheta_d(x), \vartheta_d(x) \rangle}.$$

Let us simplify this expression a little further. We have

$$\begin{aligned} \frac{\sqrt{\det \Sigma}}{\langle \vartheta_d(x), \vartheta_d(x) \rangle} &= \sqrt{\frac{\langle \vartheta_d(x), \vartheta_d(x) \rangle \langle \vartheta_d'(x), \vartheta_d'(x) \rangle - \langle \vartheta_d'(x), \vartheta_d(x) \rangle^2}{\langle \vartheta_d(x), \vartheta_d(x) \rangle}} \\ &= \sqrt{\frac{\partial^2}{\partial x \partial y}\,\Big|_{x=y}\; \log \langle \vartheta_d(x), \vartheta_d(y) \rangle}. \end{aligned}$$

We have $\langle \vartheta_d(x), \vartheta_d(y) \rangle = \sum_{i=0}^{d} (xy)^i = (1 - (xy)^{d+1})/(1 - xy)$ and so

$$\frac{\partial^2}{\partial x \partial y}\,\Big|_{x=y} = -\frac{(d+1)^2 x^{2d}}{(1 - x^{2(d+1)})^2} + \frac{1}{(1 - x^2)^2}.$$

Thus, we have

$$\rho(x) = \frac{1}{\pi}\, \frac{\sqrt{1 - h(x)^2}}{1 - x^2}, \qquad \text{where} \qquad h(y) = \frac{(d+1) x^d (1 - x^2)}{1 - x^{2(d+1)}}.$$

This finishes the proof of Theorem 2.5.

For proving Theorem 2.5 we observe that $\rho(x)$ is symmetric around 0, so that

$$\mathbb{E}\ \#\{f = 0\} = 2 \int_0^\infty \rho(x)\ \mathrm{d}x = 2 \left(\int_0^1 \rho(x)\ \mathrm{d}x + \int_1^\infty \rho(x)\ \mathrm{d}x \right).$$

In the right interval, we make a change of variables $y = \frac{1}{x}$, which reveals that the two integrals in the sum are equal. Thus, we have

$$\mathbb{E}\,\#\{f = 0\} = 4 \int_0^1 \rho(y)\, \mathrm{d}y.$$

Using the formula for the geometric sum, and the fact that $0 \leq y \leq 1$, we have

$$h(y) = \frac{(d+1)y^d(1-y^2)}{(1-y)(1+y+\cdots+y^{2d+1})} \geq \frac{y^d(1+y)}{2}, \tag{2.2}$$

which implies $1 - h(y)^2 \leq (1 - \frac{1}{2}y^d(1+y))(1 + \frac{1}{2}y^d(1+y)) \leq 2 - y^d(1+y)$. By the mean value theorem, for a fixed y there exists some $y < \theta < 1$ such that

$$\frac{(2 - 1^d(1+1)) - (2 - y^d(1+y))}{1-y} = \frac{y^d(1+y) - 2}{1-y} = -d\theta^{d-1}(1+\theta) - \theta^d.$$

Since $\theta < 1$, this implies $1 - h(y)^2 \leq 2 - y^d(1+y) < (1-y)(2d+1)$. Moreover, for $0 \leq y \leq 1$ we have $1 - y^2 \geq 1 - y$, so that

$$\frac{\sqrt{1-h(y)^2}}{1-y^2} \leq \frac{\sqrt{1-h(y)^2}}{1-y} < \sqrt{\frac{2d+1}{1-y}}. \tag{2.3}$$

On the other hand, (2.2) implies that $h(y) \geq 0$, and so

$$\frac{\sqrt{1-h(y)^2}}{1-y^2} \leq \frac{1}{1-y^2}. \tag{2.4}$$

A combination of (2.3) and (2.4) yields

$$\begin{aligned} \mathbb{E}\,\#\{f = 0\} &\leq \frac{4}{\pi}\left(\int_0^{1-\frac{1}{d}} \frac{1}{1-y^2}\,\mathrm{d}y + \int_{1-\frac{1}{d}}^1 \sqrt{\frac{2d+1}{1-y}}\,\mathrm{d}y\right) \\ &= \frac{4}{\pi}\left(\frac{\log(2-\frac{1}{d})}{2} + \frac{\log(d)}{2} + 2\sqrt{\frac{2d+1}{d}}\right) \\ &\leq \frac{4\log(2)}{\pi} + \frac{2\log(d)}{\pi} + \frac{8\sqrt{3}}{\pi} \\ &\leq \frac{2\log(d)}{\pi} + 6, \end{aligned} \tag{2.5}$$

where in the penultimate step we have used $d \geq 1$. To finish the proof of Theorem 2.6 we also need a lower bound for $\mathbb{E}\,\#\{f = 0\}$. For this we consider a number $0 \leq \delta < 1$. Then, since $h(y) \leq (d+1)y^d$, we have $h(y) \leq (d+1)(1-d^{\delta-1})^d$ for $0 \leq y \leq 1 - d^{\delta-1}$. This implies

$$\begin{aligned}\mathbb{E}\ \#\{f = 0\} &= \frac{4}{\pi}\int_0^1 \rho(y)\ \mathrm{d}y \qquad (2.6)\\ &\geq \frac{4}{\pi}\int_0^{1-d^{\delta-1}} \frac{\sqrt{1-(d+1)^2(1-d^{\delta-1})^{2d}}}{1-y^2}\ \mathrm{d}y\\ &= \frac{2}{\pi}\sqrt{1-(d+1)^2(1-d^{\delta-1})^{2d}}\,\big(\log(2-d^{\delta-1}) + (1-\delta)\log(d)\big)\end{aligned}$$

Let us write $d' = d^{1-\delta}$. Since $1-\delta > 0$ we have

$$(d+1)(1-d^{\delta-1})^d = (d+1)\Big(\Big(1-\frac{1}{d'}\Big)^{d'}\Big)^{d^\delta} \overset{d\to\infty}{\sim} (d+1)e^{-d^\delta} \overset{d\to\infty}{\to} 0,$$

which in combination with (2.6) and (2.5) shows that

$$\mathbb{E}\ \#\{f = 0\} \sim \tfrac{2}{\pi}\log(d) \text{ for } d \to \infty.$$

This finishes the proof of Theorem 2.6. □

2.3 The Kac-Rice formula for random maps on manifolds

The computation of the root density of Kac polynomials and its asymptotics was a tedious task. Furthermore, Kac polynomials are not entirely reasonable in the spirit of Section 1.1. Admittely, the distributuion of Kac polynomials $f(x)$ is invariant under the transformation $x \mapsto -x$. By contrast, as we have learned from Lemma 1.9, Kostlan polynomials are invariant under a much larger group, namely the orthogonal group. For us this means, that Kostlan polynomials are reasonable, because they have no preferred locations for their projective zeros – unlike Kac polynomials whose projective zeros are most likely to be close to $[\pm 1, 0] \in \mathbb{R}\mathrm{P}^1$. However, applying the Kac-Rice formula to Kostlan polynomials directly is difficult. The computations is simplified dramatically, if we use the Kac-Rice formula for random Gaussian maps on the sphere.

Let us consider a Riemannian manifold (M, g) of dimension m, and let us denote by $C^\infty(M, \mathbb{R}^m)$ be the space of smooth functions $M \to \mathbb{R}^m$. Similar to Definition 2.2 we say that a finite collection of functions $\mathcal{F} = \{f_1, \ldots, f_\ell\} \subset C^\infty(\mathbb{R}^m, \mathbb{R}^m)$

induces the random Gaussian map

$$f(x) = \xi_1 f_1(x) + \cdots + \xi_\ell f_\ell(x), \tag{2.7}$$

where $\{\xi_k\}_{k=1,\dots,\ell}$ is a family of i.i.d. $N(0,1)$ random variables.

Clearly, as we are looking at zeroes of a map, their number does not depend on the coordinates we are using and (up to composing with a coordinate chart), we can still work with the case $\mathbb{R}^m \to \mathbb{R}^m$, but with the additional datum of the Riemannian metric g on $\mathbb{R}^m$. We denote the *volume form* of this metric by:

$$\mathrm{vol}_g(x) = \sqrt{g(x)}\, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_m.$$

Let us introduce a way to take derivatives that "keeps track" of the metric. For this, let $x = x(u)$, $u \in U \subset \mathbb{R}^m$, local coordinates for M, and let us write $\widehat{f}(u) := f(x(u))$. Let $\{e_1, \dots, e_m\}$ be a frame field around u which is orthonormal at u with respect to the metric g and define the matrix:

$$\widehat{J}f(x) = J\widehat{f}(u)\, \begin{bmatrix} e_1 & \dots & e_m \end{bmatrix}, \tag{2.8}$$

i.e., the entries of $\widehat{J}f(x)$ are directional derivatives of $\widehat{f}(u)$ with respect to the frame field $\{e_1, \dots, e_m\}$. We denote now by $\widehat{p} : \mathbb{R}^m \times \mathbb{R}^{m\times m} \times \mathbb{R}^m \to \mathbb{R}$ the joint density of the random vector $(f(x), \widehat{J}f(x))$. The analogue of Definition 2.3 for manifolds is

$$\widehat{\rho}(x) = \int_{\mathbb{R}^{m\times m}} |\det(v)|\, \widehat{p}(0, v, x)\, \frac{1}{\sqrt{g(x)}}\, \mathrm{d}v.$$

Here is the corresponding Kac-Rice formula

Theorem 2.9 (Kac-Rice formula for random maps on manifolds)**.** *Let (M, g) be a Riemannian manifold of dimension m and $f : M \to \mathbb{R}^m$ be the random map induced by $\mathcal{F} \subset C^\infty(M, \mathbb{R}^m)$, and let $A \subseteq M$ be a measurable set. Assume that for every fixed $x \in M$ we have*

1. *$f(x)$ has a nondegenerate distribution; i.e., $\mathbb{E}\, f(x) f(x)^T \succ 0$, and*
2. *the probability that $\det(\widehat{J}f(x)) = 0$ conditioned on the event $f(x) = 0$ is equal to zero; i.e., $\mathbb{P}\{\det(\widehat{J}f(x)) = 0 \mid f(x) = 0\} = 0$.*

Then, the expected number of zeroes of f in A is given by the formula:

$$\mathbb{E}\, \#(\{f = 0\} \cap A) = \int_A \widehat{\rho}(x)\, \mathrm{dvol}_g(x). \tag{2.9}$$

Proof. Using a partition of unity (see, e.g., [6, Chapter 0, Section 5]), we can reduce the proof to $A \subset N$, where $N \subset M$ is a local coordinate chart of M. Let us write $x = x(u)$ for the coordinates on N, where $u \in U \subset \mathbb{R}^m$. Let us also denote the random function $\widehat{f}(u) := f(x(u))$ on $\mathbb{R}^m$.

As mentioned before, the number of zeroes of f does not depend on the coordinates we are using. Therefore, by Theorem 2.4 we have

$$\mathbb{E}\#\left(\{f=0\}\cap A\right) = \int_{\substack{u\in U:\\ x(u)\in A}} \rho(u)\,\mathrm{d}u, \tag{2.10}$$

where $\rho(u) = \int_{\mathbb{R}^{m\times m}} |\det(v)|\, p(0,v,u)\,\mathrm{d}v$, and where $p(w,v,u)$ is the joint density of the random pair $(\widehat{f}(u), J\widehat{f}(u))$ at $u=0$. With this notation, and with the notation from (2.8), we observe that $\sqrt{g(x)}\,|\det J\widehat{f}(u)| = |\det \widehat{J}f(x)|$ so that

$$\rho(u) = \sqrt{g(x)}\int_{\mathbb{R}^{m\times m}} |\det(v)|\,\widehat{p}(0,v,u)\,\mathrm{d}v = \sqrt{g(x)}\,\widehat{\rho}(x).$$

Using the change of variables formula, we can therefore rewrite (2.10) as:

$$\mathbb{E}\#\left(\{f=0\}\cap A\right) = \int_A \rho(u)\,\frac{1}{\sqrt{g(x)}}\,\mathrm{dvol}_g(x) = \int_A \widehat{\rho}(x)\,\mathrm{dvol}_g(x).$$

This finishes the proof. □

2.4 How many zeroes of a random polynomial are real?

As an illustrative example, we apply Theorem 2.9 to univariate Kostlan polynomials. Recall from Definition 1.8 that a (univariate) Kostlan polynomial is a random homogeneous polynomial defined by

$$f(x) = \sum_{k=0}^{d} \xi_k \binom{d}{k}^{1/2} x_0^{d-k} x_1^k, \tag{2.11}$$

where $\{\xi_k\}_{k=0,\dots,d}$ is a family of i.i.d. $N(0,1)$ random variables. The goal of this subsection is proving the next Theorem.

Theorem 2.10 (Edelman, Kostlan)**.** *Let f be a Kostlan polynomial of degree d. Then, the expected number of zeros of f in $\mathbb{R}\mathrm{P}^1$ is $\mathbb{E}\,\#\{f=0\} = \sqrt{d}$.*

For pedagogical purposes, we will give two proofs, almost identical, the first one treating $\mathbb{R}\mathrm{P}^1$ as double covered by S^1 and the second one treating polynomials as section of an appropriate line bundle over $\mathbb{R}\mathrm{P}^1$. Going from the first to the second proof we will make a conceptual jump, i.e. we will introduce the notion of a random section of a vector bundle, and we will discuss this new idea better in Section 2.5.

2.4.1 The circle approach

The number of zeros of f in $\mathbb{R}\mathrm{P}^1$ is half the number of zeros of f over the unit circle S^1. Let us denote the latter number by $\#_{S^1}\{f = 0\}$. The polynomial f defines a random map on S^1, induced by the family:

$$\mathcal{F} = \left\{ x_0^d\Big|_{S^1}, \dots, \sqrt{\binom{d}{k}} x_0^{d-k} x_1^k \Big|_{S^1}, \dots, x_1^d\Big|_{S^1} \right\},$$

where for a polynomial $f \in \mathbb{R}[x_0, x_1]_{(d)}$ we denote by $f|_{S^1}$ its restriction to the unit circle $S^1 = \{x_0^2 + x_1^2 = 1\}$. It is straightforward to check that f satisfies the assumptions of Theorem 2.9, so that

$$\mathbb{E}\ \#\{x \in \mathbb{R}\mathrm{P}^1 \mid f(x) = 0\} = \frac{1}{2}\ \mathbb{E}\,\#_{S^1}\{f = 0\} = \frac{1}{2}\int_{S^1} \widehat{\rho}(x)\,\mathrm{dvol}_g(x),$$

where $\mathrm{dvol}_g(x)$ is the volume form of S^1. By Lemma 1.9 we have $f \sim f \circ O$, and so $\widehat{\rho}(Ox) = \widehat{\rho}(x)$. Moreover, $O(2)$ acts transitively on S^1. This implies

$$\mathbb{E}\,\#_{S^1}\{f = 0\} = \int_{S^1} \widehat{\rho}(e_0)\,\mathrm{dvol}_g(x) = 2\pi\,\widehat{\rho}(e_0), \quad \text{where } e_0 = (1,0)^T \in S^1.$$

For computing the root density at e_0 we choose as coordinate chart for S^1 around e_0 the orthogonal complement $e_0^\perp := \{y \in \mathbb{R}^2 \mid \langle e_0, y\rangle = 0\}$. It is spanned by the unit norm vector $e_1 = (0,1)^T$. Moreover, since $e_0^\perp$ is the tangent space of S^1 at e_0, we have for these choice of coordinates $g(e_0) = 1$. We therefore have

$$\widehat{\rho}(e_0) = \int_{\mathbb{R}} |\det(v)|\,\widehat{p}(0, v, e_0)\ \mathrm{d}v,$$

where $\widehat{p}(0, v, e_0)$ is the density of the pair $(f(e_0), \widehat{J}f(e_0))$ at $f(e_0) = 0$. Recall that, by definition, $\widehat{J}(e_0)$ is the directional derivative of $f(x_0, x_1)$ in the direction e_1 at e_0. This implies $f(e_0) = \xi_0$ and $\widehat{J}(e_0) = \frac{\partial f}{\partial x_1}|_{(x_0,x_1)=(1,0)} = \sqrt{d}\,\xi_1$. In particular, $f(e_0)$ and $\widehat{J}(e_0)$ are independent. Therefore, the joint density is $\widehat{p}(0, v, e_0) = \frac{1}{2\pi} e^{-\frac{1}{2}\xi_1^2}$, and we find that

$$\mathbb{E}\ \#\{x \in \mathbb{R}\mathrm{P}^1 \mid f(x) = 0\} = \pi\,\widehat{\rho}(e_0) = \sqrt{\frac{\pi}{2}}\ \mathbb{E}\left|\frac{\partial f}{\partial x_1}\ \Big|_{(x_0,x_1)=(1,0)}\right|. \tag{2.12}$$

We have shown in Section 1.4.1 that this evaluates to $\sqrt{d}$. This finishes the proof. □

To wrap up, let us comment on the differences between this proof and the proof of Theorem 2.6. While the latter proof consisted in bounding the Kac-Rice density from below and from above for obtaining asymptotics of its integral, the proof in this section is based on using a group action to show that the Kac-Rice density is a constant function. This simplifies the integration dramatically, as we are left with integrating a constant.

2.4.2 The line bundle approach

Le us look again at the same problem of computing the expectation of the number of real zeroes of the random polynomial (2.11) on $\mathbb{R}\mathrm{P}^1$, yet from another perspective, i.e. without passing to the circle. A crucial first step here is to switch from the affine picture to the projectve one and to observe that f defines a section σ_f of a real line bundle $\mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d) \to \mathbb{R}\mathrm{P}^1$ (defined just below):

$$\begin{tikzcd} \mathbb{R} \arrow[r, hook] & \mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d) \arrow[d] \\ & \mathbb{R}\mathrm{P}^1 \arrow[u, bend right, "\sigma_f"'] \end{tikzcd} \tag{2.13}$$

The construction of the line bundle and the section is the following. Let us first cover $\mathbb{R}\mathrm{P}^1$ with the two open sets $U_0 = \{[x_0, x_1] \,|\, x_0 \neq 0\}$ and $U_1 = \{[x_0, x_1] \,|\, x_1 \neq 0\}$ and define the vector bundle $\mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d)$ by the cocycle $h_{10} : U_0 \cap U_1 \to \mathrm{GL}(\mathbb{R})$:

$$h_{10}([x_0, x_1]) = \left(\frac{x_0}{x_1}\right)^d. \tag{2.14}$$

In the language of vector bundles, this means that $\mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d)$ is defined to be the quotient topological space:

$$\mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d) = \Big((U_0 \times \mathbb{R}) \bigsqcup (U_1 \times \mathbb{R})\Big) \,/ \sim, \tag{2.15}$$

where the relation "$\sim$" means that the pair $([x_0, x_1], v_0) \in U_0 \times \mathbb{R}$ is identified with the pair $([x_0, x_1], h_{10} v_0) \in U_1 \times \mathbb{R}$. We denote by ψ_0, ψ_1 the trivializations of this vector bundle:

$$\psi_0 : \mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d)|_{U_0} \longrightarrow U_0 \times \mathbb{R} \quad \text{and} \quad \psi_1 : \mathcal{O}_{\mathbb{R}\mathrm{P}^1}(d)|_{U_1} \longrightarrow U_1 \times \mathbb{R}. \tag{2.16}$$

The trivializations satisfy, by definition, the property:

$$\psi_1\psi_0^{-1}([x_0,x_1],v)=([x_0,x_1],h_{10}([x_0,x_1])v). \tag{2.17}$$

In order to construct the claimed section σ_f, we consider first the homogenization of f, which is the polynomial defined by:

$$^h f(x_0,x_1)=\xi_0 {x_0}^d+\cdots+\xi_k\sqrt{\tbinom{d}{k}}x_0^{d-k}x_1^k+\cdots+\xi_d {x_1}^d. \tag{2.18}$$

Notice that, if we restrict $^h f$ to the affine line $\{x_0=1\}\simeq\mathbb{R}$, then we get back our original polynomial on the real line. We construct now the section $\sigma_f:\mathbb{RP}^1\to\mathcal{O}_{\mathbb{RP}^1}(d)$ by defining first the two local sections $\sigma_0:U_0\to\mathcal{O}_{\mathbb{RP}^1}(d)|_{U_0}$ and $\sigma_1:U_1\to\mathcal{O}_{\mathbb{RP}^1}(d)|_{U_1}$

$$\sigma_0([x_0,x_1])=\big([x_0,x_1],{}^h f(1,x_1/x_0)\big)\quad\text{and}\quad\sigma_1([x_0,x_1])=\big([x_0,x_1],{}^h f(x_0/x_1,1)\big). \tag{2.19}$$

Observe now that for $[x_0,x_1]\in U_0\cap U_1$ we have

$$\begin{aligned}\psi_1\psi_0^{-1}\left(\sigma_0([x_0,x_1])\right)&=\big([x_0,x_1],h_{10}([x_0,x_1])^h f(1,x_1/x_0)\big) &(2.20)\\ &=\left([x_0,x_1],\left(\frac{x_0}{x_1}\right)^d {}^h f(1,x_1/x_0)\right) &(2.21)\\ &=\big([x_0,x_1],{}^h f(x_0/x_1,1)\big) &(2.22)\\ &=\sigma_1([x_0,x_1]), &(2.23)\end{aligned}$$

which implies that the two local sections agree on their overlap and therefore define a global section σ_f, through the equations:

$$\psi_0\sigma_f([x_0,x_1])=\sigma_0([x_0,x_1])\quad\text{and}\quad\psi_1\sigma_f([x_0,x_1])=\sigma_1([x_0,x_1]). \tag{2.24}$$

Because f is a random polynomial, σ_f is now a random section of the line bundle $\mathcal{O}_{\mathbb{RP}^1}(d)\to\mathbb{RP}^1$. To put this into the previous language, we have a linear map $f\mapsto\sigma(f)=\sigma_f$ from the space of homogeneous polynomials of degree d to the space of global sections of this line bundle and in this way we have put a gaussian measure on the span of the corresponding sections:

$$\Gamma=\operatorname{span}\left\{\sigma({x_0}^d),\ldots,\sqrt{\tbinom{d}{k}}\sigma({x_0}^{d-k}{x_1}^k),\ldots,\sigma({x_1}^d)\right\}\subset H^0(\mathbb{RP}^1,\mathcal{O}_{\mathbb{RP}^1}(d)). \tag{2.25}$$

Now comes the second observation, which also explains the binomial scaling in front of the monomials. There is a natural action of the group $O(2)$ on $\mathbb{RP}^1$ given by:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot [x_0, x_1] = [x_0\cos\theta + x_1\sin\theta, -x_0\sin\theta + x_1\cos\theta] \tag{2.26}$$

If we now see $\mathbb{RP}^1$ as the quotient of the circle S^1 under the antipodal map and we endow the circle with the standard metric, because the antipodal map is an isometry, this metric descends to a metric on $\mathbb{RP}^1$; we call this the *quotient metric* and denote it by g. It is now easy to see that the above action (2.26) is by isometries with respect to the quotient metric. Moreover this action induces an action on Γ by change of variables and (crucial point!) it also preserves the gaussian measure we have defined through the linear map $\sigma : W \to \Gamma$ (this second observation is less trivial).

We claim now that there is a well defined function $\hat{\rho} : \mathbb{RP}^1 \to \mathbb{R}$ such that for every measurable subset $A \subseteq \mathbb{RP}^1$

$$\mathbb{E}\#\left(\{\sigma_f = 0\} \cap A\right) = \int_{\mathbb{RP}^1} \hat{\rho}([x_0, x_1])\, \mathrm{dvol}_g([x_0, x_1]). \tag{2.27}$$

To this end let us go back to (2.9) and let us apply it to the two random maps[1] $\sigma_0 : U_0 \to \mathbb{R}$ and $\sigma_1 : U_1 \to \mathbb{R}$. Observe that for every Borel set $A \subseteq U_0 \cap U_1$ we have:

$$\begin{aligned} \int_A \hat{\rho}_0([x_0, x_1])\, \mathrm{dvol}_g([x_0, x_1]) &= \mathbb{E}\#\left(\{\sigma_0 = 0\} \cap A\right) && (2.28)\\ &= \mathbb{E}\#\left(\{\sigma_f = 0\} \cap A\right) && (2.29)\\ &= \mathbb{E}\#\left(\{\sigma_1 = 0\} \cap A\right) && (2.30)\\ &= \int_A \hat{\rho}_1([x_0, x_1])\, \mathrm{dvol}_g([x_0, x_1]), && (2.31) \end{aligned}$$

which implies that the two functions $\hat{\rho}_0$ and $\hat{\rho}_1$ agree on their overlap and define the global function $\hat{\rho}$. Finally, because the action of $O(2)$ is transitive and by isometries on $\mathbb{RP}^1$, and since it preserves the probability distribution on Γ, then the distribution of the random variable in (2.8) does not depend on the point $[x_0, x_1] \in \mathbb{RP}^1$ at which we evaluate it and:

$$\hat{\rho}([x_0, x_1]) \equiv \hat{\rho}([1, 0]) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi d}} \int_{\mathbb{R}} |v| e^{-\frac{v^2}{2d}}\, \mathrm{d}v = \frac{\sqrt{d}}{\pi} \tag{2.32}$$

[1] Strictly speaking to the second components of the two random sections...but let's abuse notation silghtly!

For the last identity we have used the fact that under the parametrization $\phi : \mathbb{R} \to U_0$ given by $t \mapsto [1, t]$ the field $\frac{\partial}{\partial t}$ is sent to a field $\frac{\partial \phi}{\partial t}$ which is orthonormal at $[1, 0]$ and we can therefore use the expression already derived in (1.4.1).

If we combine now (2.27) with (2.32) and with the fact that the volume of $\mathbb{R}\mathrm{P}^1$ with respect to the quotient metric is $\pi = \int_{\mathbb{R}\mathrm{P}^1} \mathrm{dvol}_g$, we finally get Theorem 2.10. □

2.5 Random sections of vector bundles

Let us now mimic the construction of the previous Section 2.4 and put the whole set of ideas into a more general framework. Let (M, g) be a Riemannian manifold of dimension m and consider a smooth, rank-m vector bundle $p : E \to M$. A *random section* of this vector bundle is an element of a finite dimensional Gaussian space of smooth sections:

$$\Gamma = \mathrm{span}\left\{\sigma_1, \dots, \sigma_\ell\right\}. \tag{2.33}$$

As we did before, we can assume that $\{\sigma_1, \dots, \sigma_\ell\}$ is a basis of Γ and we can put a scalar product on Γ by declaring this basis to be orthonormal. A *random section* is therefore a section:

$$\sigma = \xi_1\sigma_1 + \cdots + \xi_\ell\sigma_\ell \tag{2.34}$$

with $\{\xi_k\}_{k=1,\dots,\ell}$ a family of i.i.d. standard real gaussian variables (note the analogy with (2.7), which corresponds now to the special case of the trivial bundle $M \times \mathbb{R}$). Using this notation, the gaussian measure on Γ can be defined exactly as in (2.1). Theorem 2.4 admits a simple generalization in this context.

Theorem 2.11 (Kac-Rice formulas for random sections of vector bunldes)**.** *Let (M, g) be a smooth, m-dimensional Riemannian manifold with volume density* vol_g *and let $\sigma \in \Gamma$ be a random gaussian section of the rank-m vector bundle:*

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, hook] & E \arrow[d, "p"', shift right] \\ & M \arrow[u, bend right, "\sigma"'] \end{tikzcd} \tag{2.35}$$

Let us also consider vector bundle trivializations $\{(V_\alpha, \psi_\alpha)\}_{\alpha \in I}$:

$$\psi_\alpha : E|_{V_\alpha} \to V_\alpha \times \mathbb{R}^m \tag{2.36}$$

and denote by $\sigma_\alpha : V_\alpha \to \mathbb{R}^m$ the (vector component of the) section σ in the trivialization (i.e. we write $\psi_\alpha\sigma(x) = (x, \sigma_\alpha(x))$). Assume that

1. *for every $\alpha \in I$ and for every $x \in M$ the vector $(\sigma_\alpha(x), \widehat{J}\sigma_\alpha(x))$ (defined as in Section 2.3) has a nondegenerate distribution;*
2. *the probability that σ has a degenerate zero in A is zero.*

For $x \in V_\alpha$ define:

$$\hat{\rho}(x) = \int_{\mathbb{R}^{m\times m}} |\det v|\, p_\alpha(0, v, x)\, \mathrm{d}v \tag{2.37}$$

Then (2.37) *gives a well defined function $\hat{\rho} : M \to \mathbb{R}$ (i.e.* (2.37) *does not depend on the choice of $\alpha \in I$ such that $x \in V_\alpha$) and:*

$$\mathbb{E}\left(\#\{\sigma = 0\} \cap A\right) = \int_A \hat{\rho}(x)\, \mathrm{dvol}_g(x). \tag{2.38}$$

Proof. For the moment the proof is left to the reader as an "easy" exercise. □

The previous Theorem becomes particularly effective when we have a group G such that:

1. G acts transitively on M by isometries;
2. the action induced on Γ by change of variables preserves the probability measure (2.1).

2.5.1 How many lines are there on a cubic surface?

Here is an example where these approach (invariance under some group action) can be applied. It is a classical result in algebraic geometry that on a generic complex cubic surface there are exactly 27 lines. On the other hand, if the cubic surface is defined by a real polynomial, the number of *real* lines contained on the *real* part of this surface depends on the choice of the polynomial: there could be 27, 15, 7 or 3 real lines on it and it is therfore natural to ask for the expected number of real lines on a random cubic surface. Let us put the problem of counting real lines on hypersurfaces in the above framework. Let $\mathbb{G}(1, 3)$ be the Grassmannian of lines in $\mathbb{RP}^3$ and consider on it the vector bundle[2]

$$\begin{tikzcd}
\mathbb{R}^4 \arrow[r, hook] & \mathrm{sym}^{(3)}(\tau^*) \arrow[d, "p"] \\
 & \mathbb{G}(1,3)
\end{tikzcd} \tag{2.39}$$

Let us now pick a random Kostlan polynomial $f \in \mathbb{R}[x_0, x_1, x_2, x_3]_{(3)}$

$$f(x) = \sum_{|\alpha|=3} \xi_\alpha \cdot \left(\frac{3!}{\alpha_0!\alpha_1!\alpha_2!\alpha_3!} \right)^{1/2} x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_0} x_3^{\alpha_3} \tag{2.40}$$

where $\{\xi_\alpha\}_{|\alpha|=3}$ is a family of i.i.d. standard real gaussian variables. The zero set $\mathscr{Z}(f) \subset \mathbb{RP}^3$ is a random real cubic surface. The random polynomial f defines a random section σ_f of the bundle $\mathrm{sym}^{(3)}(\tau^*)$ by restriction $\sigma_f(\ell) = f|_\ell$ and in particular

$$\{\text{real lines on } \mathscr{Z}(f)\} = \{\text{zeroes of } \sigma_f\}. \tag{2.41}$$

Now we can use Theorem 2.11 for computing the expectation of (2.41). The crucial observation for this computation is that, if we endow $\mathbb{G}(1,3)$ with the Plücker metric, the group $O(4)$ acts on it transitively and by isometries and the induced action by change of variables on the space of sections given by homogeneous polynomials preserves the probability distribution (2.40). The computation of the (constant) quantity $\hat{\rho} \equiv \frac{6\sqrt{2}-3}{2\pi^2}$ from (2.37) in this case it is still not straightforward and, combined with the fact that the Plücker volume og $\mathbb{G}(1,3)$ is $2\pi^2$ this gives (see [2, Theorem 3]):

$$\mathbb{E}\#\{\text{real lines on } \mathscr{Z}(f)\} = 6\sqrt{2} - 3 \tag{2.42}$$

The same strategy can be used to compute the expected (i.e. generic) number of complex lines on the complex cubic $\mathscr{Z}_{\mathbb{C}}(f) \subset \mathbb{CP}^3$, by replacing real with complex Kostlan polynomials, the real Grassmannian with the complex one and the orthogonal group with the unitary one, recovering in this way the number 27 from complex algebraic geometry (see [2, Corollary 8]).

[2] The projective grassmannian $\mathbb{G}(1,3)$ is naturally identified with the Grassmannian $G(2,4)$ of 2-dimensional vector spaces in $\mathbb{R}^4$, and under this identification $\tau = \tau_{2,4} \to G(2,4)$ is the tautological bundle. The fiber over a line $\ell = P(L)$ of τ^* is the set of linear forms on the two-dimensional vector space $L \subset \mathbb{R}^4$, and the fiber over $\mathrm{sym}^{(3)}(\tau^*)$ is the set of cubic forms on L.

3 Topology of random submanifolds

In this lecture we will address the problem of understanding the topology of a random algebraic submanifold of the real projective space $\mathbb{R}\mathrm{P}^n$. Our motivation comes from Hilbert's Sixteenth Problem, which was posed by D. Hilbert at the ICM in Paris in 1900. In its general formulation, this problem asks for the study of the maximal number and the possible arrangements of the components of a *generic* real algebraic hypersurface of degree d in real projective space. This is an extremely complicated problem already for the case of plane curves: the possibilities for the arrangement of the components of such curves grow super-exponentially as the degree goes to infinity.

Notice that the same problem over the complex numbers has a simple solution: the topology of a generic complex hypersurface and the way it is embedded in the complex projective space is determined just by the degree of the defining polynomial (this is a consequence of Thom's Isotopy Lemma).

An interesting approach to the real version of the problem, the one posed by Hilbert, is to look at it from the probabilistic point of view, by replacing the word *generic* with the world *random*. "What is the structure of a random plane curve of degree d? And how is it embedded in the real projective plane?" We already see that this problem has two sides: one is intrinsic (it concerns the topological structure of the hypersurface itself, e.g. its Betti numbers) and the other is extrinsic (it concerns the way the hypersurfaces sits in the projective space). Because both problems become increasingly complicated as the degree grows, deterministically they are both approached with a "prohibitive spirit" (e.g. proving that the Betti numbers of an algebraic set defined by equations of degree d in $\mathbb{R}^n$ cannot be bigger than $d(2d-1)^{n-1}$, or that some specific embedding cannot be realized as the zero set of a polynomial of degree d). We will study from a probabilistic point of view these two sides separately in Section 3.2 and Section 3.3. Before moving to the random questions, in Section 3.1 we will start by discussing some general properties of the topology of algebraic manifolds and classical approximation results.

3.1 Topology of algebraic manifolds

"How large is the class of compact *algebraic* manifolds?" The quick answer to this question is: as large as the class of compact *smooth* manifolds. In fact this was proved by Nash and Tognoli.

Theorem 3.1 (Nash-Tognoli)**.** *Every smooth compact manifold is diffeomorphic to a real algebraic set.*

The compactness condition cannot be dropped: a surface of infinite genus cannot be diffeomorphic to an algebraic set, because algebraic sets have bounded topology, by Theorem 3.6 below. Theorem 3.1 is not easy to prove: in fact Nash proved that every smooth compact manifold is diffeomorphic to a component of a real algebraic set and Tognoli proved that the algebraic set can be chosen to be connected (this step was absolutely not trivial). A nice survey on the structure of the proof can be found in [16]. In this lecture we will prove a simpler theorem, originally due to Seifert. We will need a preliminary result, which is a version of what is often referred to as the Thom's Isotopy Lemma.

Lemma 3.2 (Thom's Isotopy Lemma)**.** *Let $H : D \times I \to \mathbb{R}$ a smooth homotopy and set $h_t = H(\cdot, t) : D \to \mathbb{R}$. Assume that:*

1. $\{H = 0\} \subset \mathrm{int}(D) \times I$
2. *for every $t \in I$ the equation $\{h_t = 0\}$ is regular on D.*

Then there exists a family of diffeomorphisms $\varphi_t : D \to D$ such that:

$$\varphi_t(Z(h_t)) = Z(h_0) \quad \forall t \in I.$$

In particular:

$$(D, Z(h_t)) \sim (D, Z(h_0)) \quad \forall t \in I.$$

Proof. We will realize the family of diffeomorphisms as the flow of a time dependent vector field X_t on D. We look for a flow $\phi : D \times I \to D$ such that

$$\frac{d}{dt}\left[H(\phi(x,t),t)\right] \equiv 0. \tag{3.1}$$

Denoting by $\varphi_t = \phi(\cdot, t)$, the above condition guarantees that $\varphi_t(Z(f)) = Z(h_t)$. Expanding the r.h.s. of (3.1), we see that:

$$\frac{\partial H}{\partial x}(\phi(x,t),t)\frac{\partial \phi}{\partial t}(\phi(x,t)) + \frac{\partial H}{\partial t}(\phi(x,t),t) = d_{\varphi_t(x)}hX_t(x) + \frac{\partial H}{\partial t}(\phi(x,t),t). \tag{3.2}$$

In particular, in order to have (3.1) verified it is enough that X_t solves:

$$d_{\varphi_t(x)}hX_t(x) = -\frac{\partial H}{\partial t}(\phi(x,t),t). \tag{3.3}$$

Condition (2) of the statement ensures that we have that we can always solve equation (3.3) on every small enough neighborhood of a point $(x,t) \in D \times I$ such that $H(x,t) = 0$. We define X_t to be zero away from $\{H(x,t) = 0\}$ and on a neighborhood of this set we glue together the local solutions through a partition of unity. □

From the previous lemma we deduce another useful result with a similar flavor, we call it the Stability Lemma. In order to state it, let us introduce on the space $C^1(D,\mathbb{R})$ the norm:

$$\|f\|_{C^1(D,\mathbb{R})} = \sup_{x\in D}|f(x)| + \sup_{x\in D}\|\nabla f(x)\|. \tag{3.4}$$

With this norm $C^1(D,\mathbb{R})$ is a Banach space.

Lemma 3.3 (Stability Lemma). *Let $\{f = 0\} \subset \mathrm{int}(D(0,R))$ be a smooth and compact hypersurface defined by a regular equation on the disk $D = D(0,R)$. Then there exists $\epsilon > 0$ such that*

$$\|g - f\|_{C^1(D,\mathbb{R})} < \epsilon \implies (D, Z(f)) \sim (D, Z(g)).$$

Proof. Define the homotopy of maps $H(x,t) = (1-t)f(x)+tg(x)$, so that $H(x,0) = f(x)$ and $H(x,1) = g(x)$. Denoting by $h_t = H(\cdot,t) : D \to \mathbb{R}$, we have:

$$\|h_t - h_0\|_{C^1} = t\|f - g\|_{C^1} < \epsilon.$$

Now for $\epsilon > 0$ small enough, zero is a regular value of h_t for all $t \in [0,1]$ and the set $\{H = 0\} \subset \mathrm{int}(D) \times I$, so that the conclusion follows from Lemma 3.2. □

Remark 3.4. With the help of a partition of unity, one can prove similar results for smooth manifolds. Lemma 3.2 in its general formulation takes the following form: if N is a smooth compact manifold and $h_t : N \to \mathbb{R}$ is a smooth homotopy such that $\{h_t = 0\}$ is regular for every $t \in I$, then

$$(N, Z(h_0)) \sim (N, Z(h_t)) \quad \forall t \in I.$$

The proof is essentially the same as the proof we gave above, globalized using partition if unities. A similar statement can be proved for Lemma 3.3. To be more specific, let g be a riemannian metric on N (the introduction of a riemannian metric is not necessary, but it is simplifies the presentation). One can define the gradient of a function $f : N \to \mathbb{R}$ at a point $x \in N$ by the unique vector $\nabla f(x)$ such that:

$$g(v, \nabla f(x)) = d_x f v \quad \forall v \in T_x N.$$

Mimicking (3.4), we can define the C^1 topology on $C^1(N, \mathbb{R})$ by:

$$\|f\|_{C^1} = \sup_{x\in N} |f(x)| + \sup_{x\in N} \|\nabla f(x)\|.$$

If now $M \subset N$ is defined by a regular equation $\{f = 0\}$, then Lemma 3.3 takes the following form: there exists $\epsilon > 0$ such that

$$\|g - f\|_{C^1} < \epsilon \implies (N, Z(f)) \sim (N, Z(g)). \tag{3.5}$$

Theorem 3.5 (Seifert)**.** *Let $M \subset \mathbb{R}^n$ be a compact, smooth hupersurface defined by a regular equation $M = \{f = 0\}$, with $f \in C^\infty(\mathbb{R}^n, \mathbb{R})$. Then there exists a polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ such that:*

$$(\mathbb{R}^n, M) \sim (\mathbb{R}^n, Z(p)),$$

i.e. the two pairs are diffeomorphic. In particular M is diffeomorphic to a real algebraic set.

Proof. Since M is compact, we can assume that, it is contained in the interior of the disk $D = D(0, R)$ for some $R > 0$ and by Lemma 3.2 there exists an $\epsilon > 0$ such that:

$$\|g - f\|_{C^1(D,\mathbb{R})} < \epsilon \implies (D, Z(f)) \sim (D, Z(g)).$$

Pick now $\tilde{p} \in \mathbb{R}[x_1, \dots, x_n]$ such that $\|f - \tilde{p}\|_{C^1} < \frac{\epsilon}{2}$. Such a $\tilde{p}$ exists because polynomials are dense in the space $C^1(D, \mathbb{R})$ with the C^1-topology. Now, using Lemma 3.2 we get:

$$(D, Z(f)) \sim (D, Z(\tilde{p}))$$

however $\tilde{p}$ might have additional zeroes in $\mathbb{R}^n$. To fix this consider the polynomial:

$$p(x) = \tilde{p}(x) + \left(\frac{1}{R^2} \sum_{i=1}^n x_i^2\right)^s$$

depending on the parameter $s \in \mathbb{N}$ (when $s \to \infty$ the degree of such polynomial is $2s$). When s is large enough we have $\|p - \tilde{p}\|_{C^1} < \frac{\epsilon}{2}$, and consequently:

$$(D, Z(p)) \sim (D, Z(\tilde{p})) \sim (D, M).$$

Moreover for large enough s we also have $p|_{D(0,R)^c} > 0$ and in particular p has no extra zeroes other than those in the disk. □

The previous statement says nothing about the degree of the approximating polynomial p: in fact the more complicated is the topology of M, the higher will be the degree of p. For example, because of (3.6) we have at least $\deg(p) \geq b(M)^{\frac{1}{n}}$. Next Theorem is due to Milnor [22] and addresses the question of controlling the topology of a real algebraic hypersurface of degree d (this has to do with the intrinsic side of Hilbert's Sixteenth Problem).

Theorem 3.6 (Milnor)**.** *Let $Z \subset \mathbb{R}^n$ be a compact and smooth hypersurface defined by equations of degree at most d. Then the sum of the Betti numbers of Z is bounded by:*

$$b(Z) \leq d(d-1)^{n-1}. \tag{3.6}$$

Proof. Let $f : Z \to \mathbb{R}$ be the function $f = x_1|_Z$, where $x_1 : \mathbb{R}^n \to \mathbb{R}$ is the first coordinate function. Up to a linear change of coordinates, we can assume that f is a Morse function and consequently:

$$b(Z) \leq \#\{\text{critical points of } f\}.$$

The key point now is that the critical points of f are defined by the algebraic equations:

$$f = \frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0. \tag{3.7}$$

The system of equation (3.7) is nondegenerate in $\mathbb{R}^n$ and, by possibly perturbing its coefficients without changing the number of real solutions, we can assume it is also nondegenerate in $\mathbb{C}^n$ and the conclusion follows now from Bézout's Theorem.

□

Remark 3.7. Milnor actually proved that if $Z \subset \mathbb{R}^n$ is an algebraic set defined by equations of degree at most d (regardless the fact that it is or it is not smooth and or compact) the sum of the Betti numbers of Z is bounded by $b(Z) \leq d(2d-1)^{n-1}$. Using a similar argument, one can also prove a bound for projective algebraic sets $Z \subset \mathbb{R}\mathrm{P}^n$ defined by equations of degree at most d:

$$b(Z) \leq nd(2d-1)^n.$$

There are analogous bounds for the spherical case, the complex projective and the complex affine case, see [22].

3.2 The local structure of a random algebraic hypersurface and its Betti numbers

In this section we address the problem of understanding the Betti numbers of a random algebraic hypersurface of degree d in $\mathbb{R}\mathrm{P}^n$. Of course the result depends on the probability distribution that we pick – and we will work with the Kostlan one, which is seemingly the most reasonable to consider. We will prove two results due to Gayet and Welschinger [9,10]: the first one Theorem 3.9 gives a local probabilistic version of Seifert's theorem and it implies a lower bound on the expectation of the Betti numbers of a random Kostlan hypersurface. The second one Theorem 3.12 gives instead an upper bound on the expectation of Betti numbers. These two bounds have the same order, but not the same leading coefficients and in fact the exact asymptotics for these expected quantities, as the degree goes to infinity, are still not known.

3.2.1 Rescaling limits

For the proof of the two theorems, we will use a result from [20], which exploits the idea of "rescaling limit" and that we discuss now. The idea is to look at the behaviour of a random polynomial $p_d \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ (viewed as a section of an appropriate line bundle on the projective space) near a fixed point $[x] \in \mathbb{R}\mathrm{P}^n$ and then prove that the geometry of this map has a limit when we are sufficiently close to the point, in fact at the scale $O(d^{-1/2})$.

Let us make this idea rigorous. Using the invariance of the Kostlan distribution under the action of the Orthogonal group by change of variables, we can assume that the point is $[x] = [1, \ldots, 0] \in \mathbb{R}\mathrm{P}^n$. Let $D^n \subset \mathbb{R}^n$ be the unit disk and can consider the sequence of maps:

$$\begin{tikzcd}
D^n \arrow[r, "\text{“shrink”}"] \arrow[rr, bend right, "f_d"'] & \{x_0 \neq 0\} \arrow[r, "p_d"] & \mathbb{R} \\
u \arrow[r, mapsto] & {[1, ud^{-1/2}]} \arrow[r, mapsto] & p_d(1, ud^{-1/2}).
\end{tikzcd} \tag{3.8}$$

Because p_d is a random polynomial, then the resulting map $f_d : D^n \to \mathbb{R}$ is a random smooth function, i.e. a random variable with values in $C^\infty(D^n, \mathbb{R})$. Given a function $f \in C^\ell(D, \mathbb{R})$, for $k \leq \ell$ we denote by $f^{(k)}$ the k-th derivative of f, i.e. the symmetric tensor whose entries are all the partial derivatives of f of order k. We assume a fixed norm on the vector space of tensors has been chosen, and on the space of functions $C^\ell(D,\mathbb{R})$ we introduce the norm:

$$\|f\|_{C^\ell(D,\mathbb{R})} = \sum_{k=0}^{\ell} \sup_{x\in D} \left\| f^{(k)}(x) \right\| . \tag{3.9}$$

In order to state the next Theorem consider the space:

$$J^\ell(D,\mathbb{R}) = D \times \bigoplus_{k=0}^{\ell} \mathbb{R}^{N_k}, \quad N_k = \binom{n-1+k}{k}$$

where $\mathbb{R}^{N_k}$ denotes the set of (components of) symmetric tensors of degree k. Given a C^ℓ function $f : D \to \mathbb{R}$, its ℓ-th jet extension is the function $j^\ell f : D \to J^\ell(D,\mathbb{R})$ defined by:

$$j^\ell f(x) = (x, f(x), f^{(1)}(x), \ldots, f^{(\ell)}(x)).$$

In [20] the following Theorem 3.8 is proved.

Theorem 3.8. *Let $U \subset J^\ell(D^n,\mathbb{R})$ be a semialgebraic and open set and consider the set of functions:*

$$\mathcal{U} = \{ f \in C^\infty(D^n,\mathbb{R}) \,|\, j^\ell f(x) \in U \quad \forall x \in D \}.$$

Then there exists a constant $c_A > 0$ such that:

$$\lim_{d\to\infty} \mathbb{P}(f_d \in \mathcal{U}) = c_{\mathcal{U}}.$$

Proof. Rather then giving a proof we will try to convince the reader that the statement is true by looking at a specific example of a polynomial of one variable: this example clarifies why the theorem is true (continue, in order to believe it).

Constructing f_d as in (3.8), we see that:

$$f_d(u) = \sum_{k=0}^{d} \xi_k \binom{d}{k}^{1/2} \left(\frac{u}{d^{1/2}} \right)^k \tag{3.10}$$

$$= \sum_{k=0}^{d} \left(\frac{d!}{k!(d-k)!d^k} \right)^{1/2} u^k \tag{3.11}$$

$$\sim \sum_{k=0}^{d} \xi_k \left(\frac{1}{k!} \right)^{1/2} u^k \quad \text{as } d \to \infty. \tag{3.12}$$

Thus essentially the family of random variables $\{f_d\}$ converges to a limit $f_\infty(u) = \sum_k^\infty \xi_k \left(\frac{1}{k!}\right)^{1/2} u^k$ as $d \to \infty$. Of course one has to be careful with the word "converges", but this convergence is strong enough to ensure that the probabilities of Whitney open sets like $\mathcal{U}$ (whose definition depends only on finitely many derivatives) have a limit. □

3.2.2 A probabilistic Seifert's Theorem

Theorem 3.9. *Let $p_d \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ be a random Kostlan polynomial. Let also $M = \{f = 0\} \subset \mathbb{R}^n$ be a smooth and compact hypersurface defined by a regular equation. Then there exists $c > 0$ such that for $d > 0$ large enough and for every point $y \in \mathbb{R}\mathrm{P}^n$*

$$(\mathbb{R}^n, M) \sim (D(y, d^{-1/2}), Z(p_d) \cap D(y, d^{-1/2})),$$

i.e. the two pairs are diffeomorphic. In particular, with positive probability M is diffeomorphic to a union of components of $Z(p_d)$ (and these components are exactly the components of $Z(p_d)$ contained in a disk of radius $d^{-1/2}$).

Proof. Up to diffeomorphisms of $\mathbb{R}^n$, we can assume that $M = \{f = 0\} \subset D(0, 1)$. Moreover, because of Theorem 3.5, we can also assume f is already a polynomial (of some possibly very large degree). Let now $\epsilon > 0$ be given by Lemma 3.2 and consider the open set $\mathcal{U} \subset C^\infty(D, \mathbb{R})$ given by:

$$\mathcal{U} = \{g \in C^\infty(D, \mathbb{R}) \text{ such that } \|f - g\|_{C^1} < \epsilon\}.$$

Observe that this open set can be described by a condition of the form $j^1 f \in U$, with U the semialgebraic set:

$$U = \{(u, v, w) \in J^1(D, \mathbb{R}) \text{ such that } \|v - g(u)\| < \epsilon \text{ and } \|w - \nabla g(u)\| < \epsilon\}$$

and in particular we can apply Theorem 3.8 and conclude that there exists $c = c_{\mathcal{U}}$ such that:

$$\lim_{d\to\infty} \mathbb{P}(f_d \in \mathcal{U}) = c.$$

The conclusion of the Theorem follows now from Lemma 3.2. □

As a corollary of the previous result, we can deduce a lower bound for the expectation of the Betti numbers of $Z(p_d)$.

Corollary 3.10. *For every $k = 0, \ldots, n-1$ there exists $c_k > 0$ such that:*

$$\mathbb{E} b_k(Z(p_d)) \geq c_k d^{n/2}.$$

Proof. Fix $M \subset \mathbb{R}^n$ compact hypersurface with $b_k(M) > 0$. Put $\Theta(d^{n/2})$ many disjoint balls of radius $d^{-1/2}$ in $\mathbb{R}\mathrm{P}^n$:

$$D(y_1, d^{-1/2}) \sqcup \cdots \sqcup D(y_L, d^{-1/2}) \subset \mathbb{R}\mathrm{P}^n, \quad \text{with } L = \Theta(d^{n/2}).$$

Now we have:

$$b_k(Z(p_d)) \geq \sum_{i=1}^{L} b_k(\text{components of } Z(p_d) \text{ entirely contained in } D(y_i, d^{-1/2}))$$

and consequently, using Theorem 3.9:

$$\mathbb{E}b(k(Z(p_d)) \geq \sum_{i=1}^{L} \mathbb{E}b_k(\text{components of } Z(p_d) \text{ entirely contained in } D(y_i, d^{-1/2})) \tag{3.13}$$

$$\geq L \cdot b_k(M) \cdot c \geq c_k d^{n/2}. \tag{3.14}$$

This concludes the proof. □

Remark 3.11. Here is another example where the “rescaling limit” idea can be used. Let (M, g) be a compact riemannian manifold of dimension m and $\{p_1, \ldots, p_n\}$ be i.i.d. points sampled from the uniform distribution on M. Given $\alpha > 0$ one can consider the random geometric complex:

$$\mathcal{U}_n = \bigcup_{j=1}^{n} D(p_j, \alpha n^{-1/m}).$$

In a similar way to what we have done in the proof Theorem 3.9, one can prove that for every geometric complex $\mathcal{U} \subset \mathbb{R}^m$ there exists $R > 0$ and $c > 0$ such that for any $p \in M$ and for $n > 0$ large enough:

$$(D(p, Rn^{-1/m}), D(p, Rn^{-1/m}) \cap \mathcal{U}_n) \sim (\mathbb{R}^m, \mathcal{U}).$$

Here the local geometry has a limit at the scale $O(n^{-1/m})$.

3.2.3 Random Morse Theory

Given that we have a lower bound of the order $d^{n/2}$ on the expectation of the Betti numbers of $Z(p_d)$, it is natural to ask wether there is an upper bound of the same order. This is the content of next result.

Theorem 3.12. *Let $P \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ be a random Kostlan polynomial and $Z(p_d) \subset \mathbb{R}\mathrm{P}^n$ its zero set. Then $\mathbb{E}b_k(Z(p_d)) \leq O(d^{n/2})$.*

Proof. The proof goes through Morse Theory and uses a variation of the Kac-Rice formula. The main idea is to fix a Morse function $f : \mathbb{R}\mathrm{P}^n \to \mathbb{R}$ and consider its restriction to $Z(p_d)$:

$$g = f|_{Z(p_d)} : Z(p_d) \to \mathbb{R}.$$

With probability one g is a Morse function on the smooth manifold $Z(p_d)$ and the expectation of the number of its critical points can be computed using the Kac-Rice formula. Manipulating the integral (with appropriate change of variables) one gets:

$$\mathbb{E}b_{k(Z(p_d))} \leq \mathbb{E}\#\{\text{critical points of } g\} \leq \int_{\mathbb{R}\mathrm{P}^n} \rho = O(d^{n/2}).$$

□

Combining the bounds from Theorem 3.9 and Theorem 3.12 we immediately get the following result.

Corollary 3.13. *Let $P \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ be a random Kostlan polynomial. Then for every $k = 0, \ldots, n-1$ we have $\mathbb{E}b_k(Z(p_d)) = \Theta(d^{n/2})$.*

Remark 3.14. There are of course other probability distributions on the space of polynomials that are invariant under orthigonal change of coordinates other than the Kostlan distribution; the study of these distributions will be the subject of a separate chapter.

Remark 3.15. So far we have considered the case when n is fixed and $d \to \infty$, but the asymptotic keeping the degree fixed and letting the number of variables go to infinity is also interested. In the case $Z(q_1, \ldots, q_k) \subset \mathbb{R}\mathrm{P}^n$ is the intersection of k independent random Kostlan quadrics one can prove that for every fixed j one has:

$$b_j(Z(q_1, \ldots, q_k)) = 1 \text{ with probability } 1 - O(n^{-a}) \text{ for all } a > 0.$$

Moreover, in the case o the intersection of two quadrics:

$$\mathbb{E}b(Z(q_1, q_2)) = n + \frac{2}{\sqrt{\pi}} n^{1/2} + O(n^c) \quad \forall c \in (0, 1/2).$$

The proofs of both statements go through a random spectral sequence argument, see [19].

3.3 The global structure of a random algebraic hypersurface: approximations by small degree

In this section we discuss the global geometry of a random Kostlan hypersurface. We will adopt the spherical point of view, rather than the projective one, because we want to look at polynomials as functions (the whole discussion could also be done using the language of line bundles, but paying a rather technical price).

Given $p \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ we consider its restriction to the unit sphere S^n and think of it as a function:

$$p : S^n \to \mathbb{R},$$

and we will be interested in $Z_{S^n}(p)$; of course this set double covers $Z(p) \subset \mathbb{R}\mathrm{P}^n$ and one can deduce consequences of the content of this section for the projective case using standard tools from algebraic topology. The main result of this section will be that, with probability that goes to one as $d \to \infty$, the pair $(S^n, Z_{S^n}(p))$ is diffeomorphic to a pair $(S^n, Z_{S^n}(\tilde{p}))$ with $\tilde{p}$ the restriction to the unit sphere of a homogeneous polynomial of degree $O(d^{1/2} \log d)$. Notice that, already this statement, combined with Theorem 3.6 implies that with probability that tends to one as $d \to \infty$ we have $b(Z(p_d)) \leq O(d^{n/2} \log d)$. Before getting to the details, we will need to introduce some additional tools.

3.3.1 Decomposition of polynomials into spherical harmonics

In order to study the global geometry of $p : S^n \to \mathbb{R}$, we will need a different basis for the space of polynomials, more adapted to real studies than the monomial basis. This basis is given by harmonic polynomials.

For every $\ell \in \mathbb{N}$ let $\Delta : \mathbb{R}[x_0, \ldots, x_n]_{(\ell)} \to \mathbb{R}[x_0, \ldots, x_n]_{(\ell-2)}$ be the Laplace operator on $\mathbb{R}^{n+1}$:

$$\Delta = \sum_{j=0}^{n} \frac{\partial^2}{\partial x_j^2}.$$

This is a linear and surjective operator, with kernel the harmonic polynomials:

$$H_{n,\ell} = \{p \in \mathbb{R}[x_0, \ldots, x_n]_{(\ell)} \text{ such that } \Delta p = 0\}.$$

We will need the following theorem, which gives a decomposition of the space of homogeneous polynomials.

Theorem 3.16. *The space of homogeneous polynomials decomposes as:*

$$\mathbb{R}[x_0, \ldots, x_n]_{(d)} = \bigoplus_{0 \leq \ell \leq d,\, \ell \in 2\mathbb{N}} \|x\|^{\ell} \cdot H_{n,\ell}. \tag{3.15}$$

Moreover the decomposition is orthogonal for any scalar product on $\mathbb{R}[x_0, \ldots, x_n]_{(d)}$ *which is invariant under the action of the group* $O(n+1)$ *by change of variables.*

Proof. See [17]. □

When restricting a polynomial to the unit sphere we have $\|x\|^2 = 1$, and (3.15) gives a decomposition:

$$\mathbb{R}[x_0, \ldots, x_n]_{(d)}\big|_{S^n} = \bigoplus_{0 \le \ell \le d,\, \ell \in 2\mathbb{N}} V_{n,\ell},$$

where now $V_{n,\ell} = H_{n,\ell}|_{S^n}$ coincides with the space of spherical harmonics of degree ℓ, see [17]. In particular, given $p \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ we can decompose $p : S^n \to \mathbb{R}$ as:

$$p = \sum_{0 \le \ell \le d,\, \ell \in 2\mathbb{N}} p_\ell, \qquad p_\ell \in V_{n,\ell}.$$

For every $L = 0, \ldots, d$ with the same parity as d, we will also consider the truncation:

$$\pi_L(p) = \sum_{0 \le \ell \le d,\, \ell \in 2\mathbb{N}} p_\ell.$$

Notice that $\pi_L(p)$ is the restriction to the unit sphere of a homogeneous polynomial of degree L.

3.3.2 A quantitative Stability Lemma

Let $\{p = 0\}$ be a regular equation on S^n. Recall from Lemma 3.3 that there exists $\epsilon = \epsilon(p) > 0$ such that for every $g \in C^1(S^n, \mathbb{R})$ with $\|p - g\|_{C^1} < \epsilon$ we have $(S^n, Z(p)) \sim (S^n, Z(g))$. The question that we address in this section is whether it is possible to estimate ϵ in the case p is a polynomial.

On the space $\mathbb{R}[x_0, \ldots, x_n]_{(d)}$ consider the scalar product $\langle \cdot, \cdot \rangle_{BW}$ defined by:

$$\langle x_0^{\alpha_0} \cdots x_n^{\alpha_n}, x_0^{\beta_0} \cdots x_n^{\beta_n} \rangle_{BW} = \delta_{\alpha\beta} \frac{d!}{\alpha_0! \cdots \alpha_n!}.$$

This is called the *Bombieri-Weyl* scalar product. We denote $\Sigma \subset \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ the semialgebraic set:

$$\Sigma = \{p \text{ such that there exists } x \in S^n \text{ with } p(x) = 0 \text{ and } d_x p = 0\}.$$

This *is not* the real part of the complex discriminant, but it coincides with it up to semialgebraic subsets of codimension two. The following theorem from [5] gives a bound of $\epsilon(p)$ in terms of the distance of p from the discriminant Σ.

Theorem 3.17. *Let $p \in \mathbb{R}[x_0, \dots, x_n]_{(d)}$ and assume that $p \notin \Sigma$ (i.e. that the equation $\{p = 0\}$ is regular on S^n). Then for $g \in C^1(S^n, \mathbb{R})$ we have:*

$$\|p - g\|_{C^1} < \frac{\mathrm{dist}_{BW}(p, \Sigma)}{2} \implies (S^n, Z(p)) \sim (S^n, Z(g)). \tag{3.16}$$

Proof. For $t \in [0, 1]$ let us consider now the function $g_t = p + t(g - p)$. Since $\|p - g\|_{C^1} < \frac{\mathrm{dist}_{BW}(p,\Sigma)}{2}$, for all $\theta \in S^n$ we have:

$$|g_t(\theta)| > |p(\theta)| - \frac{\delta}{2}.$$

Moreover, since $d \geq 1$, from $\|g - p\|_{C^1} < \frac{\mathrm{dist}_{BW}(p,\Sigma)}{2}$ we also deduce $\frac{\|g-p\|_{C^1}}{\sqrt{d}} < \frac{\mathrm{dist}_{BW}(p,\Sigma)}{2}$, which in turn implies for every $t \in [0, 1]$ and $\theta \in S^n$:

$$\frac{\|\nabla g_t(\theta)\|}{\sqrt{d}} > \frac{\|\nabla p(\theta)\|}{\sqrt{d}} - \frac{\mathrm{dist}_{BW}(p, \Sigma)}{2}. \tag{3.17}$$

Recall from [24, Theorem 5.1] the following explicit expression[1] for $\mathrm{dist}_{BW}(p, \Sigma)$:

$$\mathrm{dist}_{BW}(p, \Sigma) = \min_{\theta \in S^n} \left(|p(\theta)|^2 + \frac{\|\nabla p(\theta)\|^2}{d} \right)^{1/2}. \tag{3.18}$$

Note that $\left(|p(\theta)|^2 + \frac{\|\nabla p(\theta)\|^2}{d} \right)^{1/2}$ equals the distance in $\mathbb{R}^2$ between the two vectors $v_1(\theta) = (|p(\theta)|, 0)$ and $v_2(\theta) = \left(0, \frac{\|\nabla p(\theta)\|}{\sqrt{d}}\right)$. Observe also that the two vectors $w_1(t, \theta) = (|g_t(\theta)|, 0)$ and $w_2(t, \theta) = \left(0, \frac{\|\nabla g_t(\theta)\|}{\sqrt{d}}\right)$, in virtue of (3.17) and (3.18), satisfy:

$$w_1(t, \theta) \in B_1(\theta) = B_{\mathbb{R}^2}\left(v_1(\theta), \frac{\mathrm{dist}_{BW}(p, \Sigma)}{2}\right)$$

and

$$w_2(t, \theta) \in B_2(\theta) = B_{\mathbb{R}^2}\left(v_2(\theta), \frac{\mathrm{dist}_{BW}(p, \Sigma)}{2}\right). \tag{3.19}$$

[1] This expression can also be derived from [4, Theorem 19.3], where it is proved that the distance from the real discriminant equals the reciprocal of the condition number. We prefer to quote directly [24] because it seems that this nice work of Raffalli has been forgotten from the literature on the subject.

In particular:

$$\left(|g_t(\theta)|^2 + \frac{\|\nabla g_t(\theta)\|^2}{d}\right)^{1/2} = \|w_1(t,\theta) - w_2(t,\theta)\| \tag{3.20}$$

$$> d_{\mathbb{R}^2}\left(B_1(\theta), B_2(\theta)\right) \tag{3.21}$$

$$= \|v_1(\theta) - v_2(\theta)\| - \mathrm{dist}_{BW}(p,\Sigma), \tag{3.22}$$

where the strict inequality comes from the fact that w_1 and w_2 belong to the *interior* of the balls.

Taking the minimum over $\theta \in S^n$ in the above expression gives:

$$\min_{\theta \in S^n} \left(|g_t(\theta)|^2 + \frac{\|\nabla g_t(\theta)\|^2}{d}\right)^{1/2} > 0 \quad \forall t \in [0,1]. \tag{3.23}$$

In particular the equation $\{g_t = 0\}$ on S^n is regular for all $t \in [0,1]$: whenever $g_t(\theta) = 0$, then $\nabla g_t(\theta)$ cannot vanish because of the strict inequality in (3.23). The result follows now from Thom's Isotopy Lemma. □

3.3.3 The small degree approximation theorem

In this section we giva a proof of the following theorem from [5].

Theorem 3.18. *Let $p \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ be a random Kostlan polynomial. Then for all $a > 0$ there exists $b > 0$ such that, denoting by*

$$\tilde{p} = \sum_{0 \le \ell \le b\sqrt{d}\log d,\, \ell \in 2\mathbb{N}} p_\ell,$$

(notice that $\tilde{p}$ is a polynomial of degree $b\sqrt{d}\log d$) then with probability at least $1 - O(d^{-a})$ we have:

$$(S^n, Z(p)) \sim (S^n, Z(\tilde{p})).$$

Proof. The proof consists in applying Theorem 3.17 with the choice of $g = \tilde{p}$ and showing that $\|p - \tilde{p}\|_{C^1} < \frac{\mathrm{dist}_{BW}(p,\Sigma)}{2}$ holds true with probability $1 - O(d^{-a})$, so that (3.16) implies the statement.

We will in fact prove a more general inequality depending on three parameters $t, s > 0$ and L (this is the degree of the approximation) which holds true with some probability. More precisely we will show that there exist constants $c_1(n), c_2(n), c_3(n), c_4(n) > 0$ such that

$$\|p - \pi_L(p)\|_{C^1} \le c_1(n) d^{\frac{1}{2}}\, t\, s\, \mathrm{dist}_{BW}(p, \Sigma),$$

which holds for every $q \geq \frac{n+1}{2}$, $t > 0$ and $s \geq c_4(n)d^{2n}$, with probability

$$\mathbb{P} \geq 1 - \left(c_2(n) \frac{d^{\frac{-3n}{2}+1}(L)^{2q+n-2} e^{-\frac{L^2}{d}}}{t^2} + c_3(n) \frac{d^{2n}}{s} \right).$$

□

Bibliography

[1] Jean-Marc Azais and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[2] Saugata Basu, Antonio Lerario, Erik Lundberg, and Christopher Peterson. Random fields and the enumerative geometry of lines on real and complex hypersurfaces. *Math. Ann. (to appear)*.

[3] Kathleen Booth. An investigation into the real roots of certain polynomials. *Math. Tables and Aids to Computation*, 8:47 pp., 1954.

[4] Peter Bürgisser and Felipe Cucker. *Condition: The geometry of numerical algorithms*, volume 349 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2013.

[5] Daouda Niang Diatta and Antonio Lerario. Low degree approximation of random polynomials, 2018.

[6] Manfredo do Carmo. *Riemannian Geometry*. Birkhäuser, 1993.

[7] Freeman Dyson. Statistical Theory of the Energy Levels of Complex Systems. *J. Math. Phys.*, 3(1):140–156, 1962.

[8] Alan Edelman and Eric Kostlan. How many zeros of a random polynomial are real? *Math. Soc. Mathematical Reviews*, 32:1–37, 05 1995.

[9] Damien Gayet and Jean-Yves Welschinger. Lower estimates for the expected Betti numbers of random real hypersurfaces. *J. Lond. Math. Soc. (2)*, 90(1):105–120, 2014.

[10] Damien Gayet and Jean-Yves Welschinger. Betti numbers of random real hypersurfaces and determinants of random symmetric matrices. *J. Eur. Math. Soc. (JEMS)*, 18(4):733–772, 2016.

[11] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.*, 6:440–449, 1965.

[12] Corey Harris and Yoav Len. Tritangent planes to space sextics: the algebraic and tropical stories. *G.G. Smith and B. Sturmfels (Eds.): Combinatorial Algebraic Geometry*, pages 47–63, 2018.

[13] Mark Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 49:314–320, 1943.

[14] Mark Kac. On the average number of real roots of a random algebraic equation II. *Proc. London Math. Soc.*, 50:390–408, 1949.

[15] Felix Klein. A comparative review of recent researches in geometry. *Bull. Amer. Math. Soc.*, 2(10):215–249, 1893.

[16] János Kollár. Nash's work in algebraic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 54(2):307–324, 2017.

[17] Eric Kostlan. On the expected number of real roots of a system of random polynomial equations. In *Foundations of computational mathematics (Hong Kong, 2000)*, pages 149–188. World Sci. Publ., River Edge, NJ, 2002.

[18] Avinash Kulkarni, Yue Ren, Mahsa Sayyary Namin, and Bernd Sturmfels. Real Space Sextics and their Tritangents. *arXiv:1712.06274*.

[19] Antonio Lerario and Erik Lundberg. Gap probabilities and Betti numbers of a random intersection of quadrics. *Discrete Comput. Geom.*, 55(2):462–496, 2016.

[20] Antonio Lerario and Michele Stecconi. Differential topology of gaussian random fields: applications to random algebraic geometry, 2019.

[21] John Littlewood and Albert Offord. On the number of real roots of a random algebraic equation. *J. London Math. Soc.*, 13:288–295, 1938.

[22] John Milnor. On the Betti numbers of real varieties. *Proc. Amer. Math. Soc.*, 15:275–280, 1964.

[23] Robb J. Muirhead. *Aspects of Multivariate Statistical Theory*, volume 131. John Wiley & Sons, NY, 1982.

[24] Christophe Raffalli. Distance to the discriminant. *arXiv 1404.7253.*

[25] Eugene Wigner. On a Class of Analytic Functions from the Quantum Theory of Collisions. *Annals of Mathematics*, 53(1):36–67, 1951.