# Statistics on Hilbert's Sixteenth Problem

Antonio Lerario
Purdue University

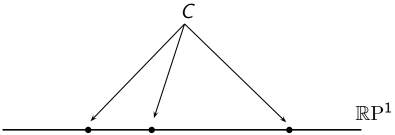
Randomness in Mathematics and Physics
ZIF, 2013

# Hilbert's Sixteenth Problem

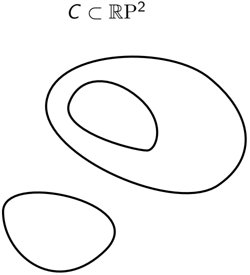
**Question (D. Hilbert, 1900)**

*Investigation of the relative position [...] and the number of connected components of a hypersurface of degree d in* $\mathbb{R}P^n$.

Hypersurface of degree $d$ = zero set $C$ of a homogeneous polynomial $f$ of degree $d$ in $n+1$ variables.




$n = 1$: points on a line




$n = 2$: ovals in the plane

# The definition of a projective hypersurface

$f \in \mathbb{R}[x_0, \ldots, x_n]_{(d)}$ (homogeneous of degree $d$) defines:

$$C = \{[x_0, \ldots, x_n] \in \mathbb{R}\mathrm{P}^n \mid f(x_0, \ldots, x_n) = 0\}$$

$\mathbb{R}\mathrm{P}^n$ is a compactification of $\mathbb{R}^n$

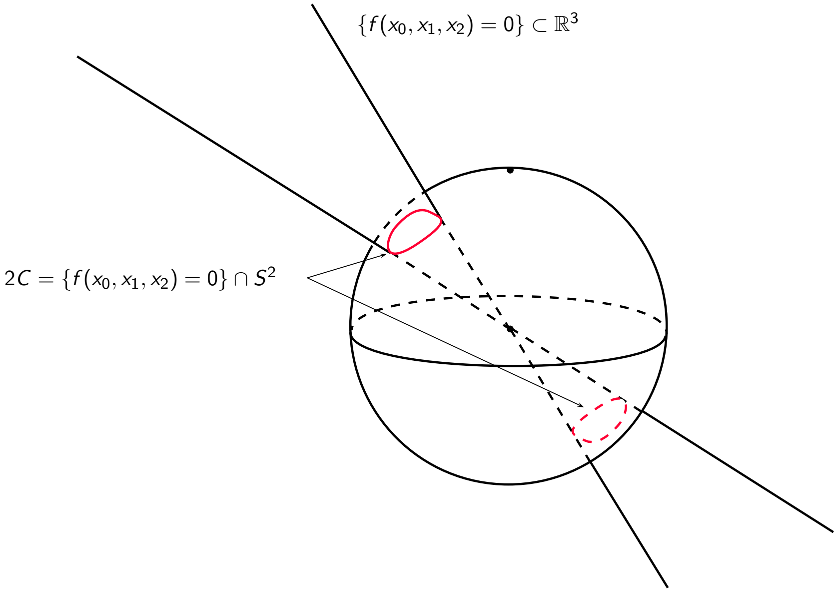
We can see the “affine” part of $C$ by setting:

$$C' = \{f(1, x_1, \ldots, x_n) = 0\} \subset \mathbb{R}^n.$$

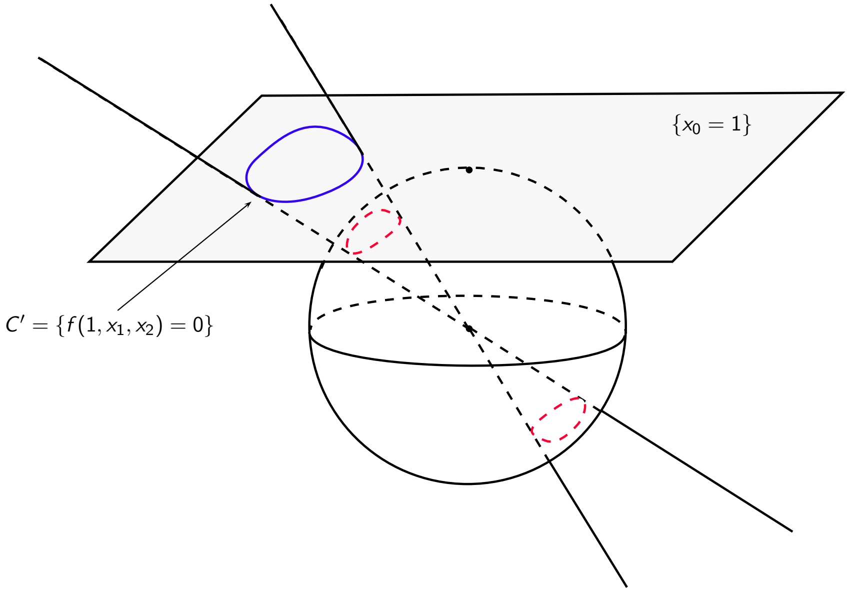
Two copies of $C$ appear on the unit sphere:

$$2C = \{x \in S^n \mid f(x_0, \ldots, x_n) = 0\}.$$

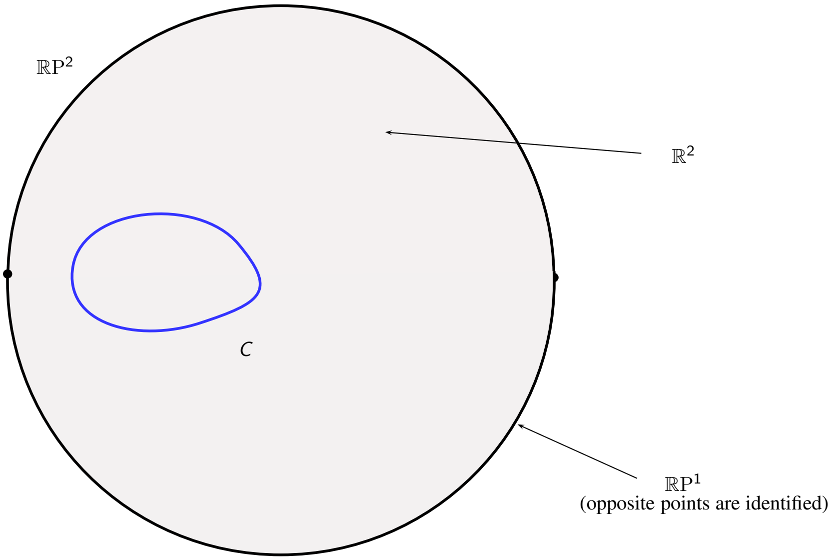
# The double covering of a projective hypersurface



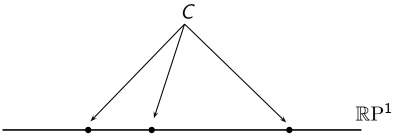
# The affine part of a projective hypersurface



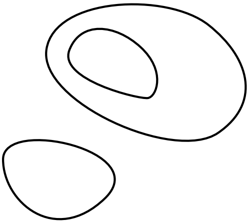
# The projective part



# $n = 1, n = 2$

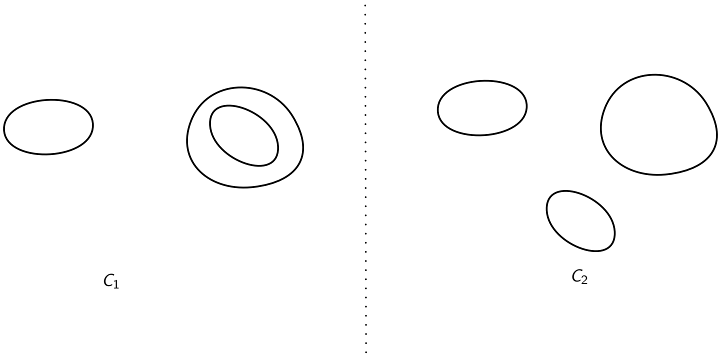



$n = 1$: points on a line




$n = 2$: ovals in the plane

# H16: different arrangements



# Statistical Hilbert's Sixteenth Problem

For $d > 8$ the classification of the possible arrangements of ovals ($n = 2$) is not known; for $n \geq 3$ even the sharp upper bound on the number of components is not known.

**Question (Statistical H16)**

*Investigation of the typical relative position [...] and the typical number of connected components of a hypersurface of degree d in* $\mathbb{RP}^n$.

The main idea is that the topology of "real" objects (as opposed to "complex") should be studied statistically.

# Probability distributions

"typical" $=$ endow $\mathbb{R}[x_0, \ldots, x_n]_{(d)}$ with a probability distribution, and then take expectation;
we make it a Gaussian space by:

$$\mathbb{P}\{f \in A\} = \frac{1}{c} \int_A e^{-\|f\|^2} d\mu, \quad A \subset \mathbb{R}[x_0, \ldots, x_n]_{(d)}.$$

It remains to choose a norm on $\mathbb{R}[x_0, \ldots, x_n]_{(d)}$.

# Interesting choices for random polynomials

Possible interesting choices for the norm of $f = \sum_\alpha f_\alpha x_0^{\alpha_0} \cdots x_n^{\alpha_n}$:

- naive choice (considered by Kac):

$$\|f\|^2 \doteq \sum_\alpha f_\alpha^2$$

- Kostlan distribution (also called Fubini-Study):

$$\|f\|^2 \doteq \sum_\alpha \frac{\alpha_0! \cdots \alpha_n!}{d!} f_\alpha^2$$

- Riemannian geometer's choice (real Fubini-Study):

$$\|f\|^2 \doteq \|f\|^2_{L^2(S^n)} = \int_{S^n} f^2 d\mathrm{Vol}_{S^n}$$

# n=1: random univariate polynomials

$E_{d,1} = \mathbb{E}\{\text{number of zeroes of a random } f \text{ of degree } d\}.$

- naive choice:

$$E_{d,1} \sim \frac{2}{\pi}\log(d) = \Theta(\log(\text{a priori bound}))$$

- Kostlan distribution (Fubini-Study):

$$E_{d,1} = \sqrt{d} = \Theta(\sqrt{\text{a priori bound}})$$

- Riemannian geometer's choice (real Fubini-Study):

$$E_{d,1} = \sqrt{\frac{d(d+2)}{3}} = \Theta(\text{a priori bound})$$

# A deterministic bound for hypersurfaces

$E_{d,n} = \mathbb{E}\{\text{number of components of a random } f \in \mathbb{R}[x_0, \ldots, x_n]\}$.

**Theorem (Milnor, 1964)**

*The a priori bound ("a.p.b." for short) for the number of connected components of a hypersurface of degree $d$ in $\mathbb{RP}^n$ is $O(d^n)$.*

In particular $E_{d,n} \leq O(d^n)$.
We are interested in asymptotic behavior of $E_{d,n}$ for fixed $n$ and $d \to \infty$.

# Random hypersurfaces

**Theorem (P. Sarnak and I. Wigman, 2012, [6])**

*For the naive choice:* $E_{d,2} \leq O(\log(d^2)) = O(\log(a.p.b.))$

**Theorem (D. Gayet and J.-Y. Welschinger, 2013 [1])**

*For the Kostlan distribution (Fubini-Study):*

$$E_{d,n} = \sqrt{d^n} = \Theta(\sqrt{a.p.b.})$$

**Theorem (A. L. and E. Lundberg, 2012 [4])**

*For the Riemannian geometer's choice (real Fubini-Study):*

$$E_{d,n} = O(d^n) = \Theta(a.p.b.)$$

# The large variables limit

Up to now we considered the asymptotic for fixed $n$ and $d \to \infty$. What about the other asymptotic? I.e.

$$\lim_{n\to\infty} E_{d,n} = ? \quad (\text{fixed } d).$$

# Random quadric hypersurfaces, large $n$ limit

Consider the case $d = 2$.

$$q \in \mathbb{R}[x_1, \ldots, x_n]_{(2)} \text{ Kostlan} \quad \Longleftrightarrow \quad Q \in GOE(n)$$

$$\text{where} \quad q(x) = \langle x, Qx \rangle \quad \forall x \in \mathbb{R}^n$$

$$Z(q) \subset \mathbb{R}\mathrm{P}^n, \quad b(Z(q)) = \sum_i b_i(Z(q)).$$

$$b(Z(q)) = \text{ number of “holes” in } Z(q).$$

## Quadric hypersurfaces: expected topology

$$i^+(q) = \text{number of positive eigenvalues of } Q$$
$$\mu(q) = \max\{i^+(q), i^+(-q)\}$$
$$b(Z(q)) = 2(n - \mu(q))$$

**Theorem (A. L., 2012, [3])**

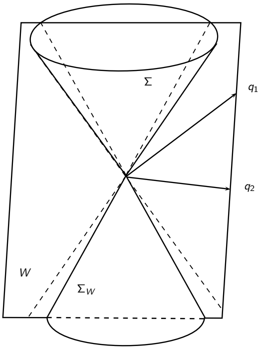
$$\mathbb{E}b(Z(q)) \sim n \quad \textit{as } n \to \infty.$$

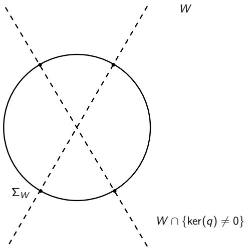
(Essentially concentration near matrices with vanishing signature)

# Intersection of two quadrics: the pencil

$$Z(q_1, q_2) \subset \mathbb{RP}^{n-1}, \quad q_1, q_2 \in \mathbb{R}[x_1, \ldots, x_n]_{(2)}$$

$$W = \mathrm{span}\{q_1, q_2\} \subset \mathbb{R}[x_1, \ldots, x_n]_{(2)}$$
$$\Sigma = \{\det(Q) = 0\}, \quad \Sigma_W = \Sigma \cap W.$$



# Intersection of two quadrics: deterministic formula

$$\mu_W = \max_{q \in W} i^+(q)$$

$$s_W = \{\text{number of singular lines in } W\}$$

**Theorem (A. L. 2011, [2])**

$$b(Z(q_1, q_2)) = 3n - 4\mu_W + s_W + \epsilon_W, \quad \epsilon_W \in \{0, 1\}$$

$\epsilon_W$ is the "linking number" of $W \backslash \{0\}$ with the set of matrices with multiple eigenvalues.

# Random intersection of two independent quadrics

As a corollary:

$$\mathbb{E}b(Z(q_1, q_2)) = 3n - 4\mathbb{E}\mu_W + \mathbb{E}s_W + \mathbb{E}\epsilon_W$$

$$\mathbb{E}\mu_W \sim \frac{n}{2} \quad \text{(concentration near vanishing signature).}$$

$$\mathbb{E}\epsilon_W \to 0 \quad \text{(algebraic topology computations).}$$

What about $\mathbb{E}s_W$?

# Gap probability in GOE

$$\mathbb{E}s_W = \text{average number of singular lines in } W$$

$$\mathbb{E}s_W = \frac{\mathrm{Vol}(\Sigma)}{\mathrm{Vol}(S^{N-2})} \quad \text{(integral geometry formula)}$$

$$\mathrm{Vol}(\Sigma) = \lim_{\epsilon \to 0} \frac{\mathrm{Vol}(\Sigma_\epsilon)}{2\epsilon}.$$

Eckart-Young: $d_F(q, \Sigma) = \sigma(Q)^{-1}$

$$\mathrm{Vol}(\Sigma) = \lim_{\epsilon \to 0} \frac{1 - \mathbb{P}\{\text{no eigenvalues in } (-\epsilon, \epsilon)\}}{2\epsilon}$$

$$\mathrm{Vol}(\Sigma) = \underbrace{O(\sqrt{n})\mathrm{Vol}(S^{N-2})}_{\text{minus derivative at zero of the gap probability}}$$

# Random intersection of two quadrics

Putting al this together we get:

**Theorem (A. L. 2012, [3])**

$$\mathbb{E}Z(q_1, q_2) \sim n + O(\sqrt{n}) + o(n)$$

There are many other interesting questions (expected arrangement, concentration, other topological invariants, random algebraic geometry...)

# Bibliography


- D. Gayet, J-Y. Welschinger, *Betti numbers of random real hypersurfaces and determinants of random symmetric matrices*, arXiv:1207.1579
- A. Lerario, *Convex pencils of real quadratic forms*, Discrete and computational Geometry, Volume 48, Number 4 (2012), 1025-1047
- A. Lerario, *Random matrices and the average topology of the intersection of two quadrics*, Proceedings of the American Mathematical Society, to appear.
- A. Lerario, E. Lundberg, *Statistics on Hilbert's sixteenth problem*, arXiv:1212.3823
- P. Sarnak, *Letter to B. Gross and J. Harris on ovals of random plane curves* publications.ias.edu/sarnak/paper/510 (2011)
- P. Sarnak, I. Wigman, private communication