Random polynomials and their zero sets

Erik Lundberg
(joint work with Antonio Lerario)
elundber@math.purdue.edu
http://www.math.purdue.edu/ elundber

Purdue University
West Lafayette, IN

May 23, 2013

Complex versus real algebraic geometry

Given a polynomial of degree d:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_d x^d.$$

(a) How many complex zeros?

$$= d$$

(b) How many *real* zeros?

$$\leq d$$

-Over the complex numbers: complete rigidity.

-Over the reals: complete freedom (besides the upper bound).

Complex versus real algebraic geometry

Given a homogeneous polynomial f of degree d in $n+1$ variables:

The zero set Z_f is a *hypersurface*.

- What is the total volume of $\mathbb{C}Z_f$ in $\mathbb{C}P^n$?

$$= C_n \cdot d.$$

- What is the total volume of $\mathbb{R}Z_f$ in $\mathbb{R}P^n$?

$$\leq \tilde{C}_n \cdot d.$$

- What is the total Betti number of $\mathbb{C}Z_f$ in $\mathbb{C}P^n$?

$$= \text{polynomial of degree } n \text{ in } d$$

- What is the total Betti number of $\mathbb{R}Z_f$ in $\mathbb{R}P^n$?

$$\leq \text{polynomial of degree } n \text{ in } d$$

Hilbert's sixteenth problem

First part of Hilbert's sixteenth problem: study the number and possible arrangement of the connected components of the zero set of a polynomial.

The (sharp) maximum for curves:

$$\text{(A. Harnack)} \quad N_f \leq \frac{(d-1)(d-2)}{2} + 1$$

As for possible arrangements, "parity" results are known and some bounds such as ($\nu_0 =$ number of empty ovals, $d = 2k$):

$$\text{(V. I. Arnold)} \quad \nu_0 \geq N_f - \frac{(k-1)(k-2)}{2}$$

Especially little is known in more than $2+1$ variables.

How many zeros of a random polynomial are real?

How should we define "random"?

Choose a basis for polynomials of degree at most d. Make a random linear combination with i.i.d. Gaussian coefficients ξ_k.

An obvious choice: monomials as a basis

$$f(x) = \sum_{k=0}^{d} \xi_k x^k$$

Or in projective space $\mathbb{RP}^1$:

$$f(x, y) = \sum_{k=0}^{d} \xi_k x^k y^{d-k}.$$

Random polynomials over $\mathbb{RP}^1$: three Gaussian ensembles

ξ_k - i.i.d. standard normal Gaussians:

$$f = \sum_{k=0}^{d} \xi_k f_k.$$

1. Naive model: $f_k(x, y) = x^k y^{d-k}$
2. Algebraic geometer's model: $f_k(x, y) = \sqrt{\binom{d}{k}} x^k y^{d-k}$
3. Analyst's model: trigonometric polynomials

$$\{f_k(\theta)\} = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(2\theta)}{\sqrt{\pi}}, \frac{\sin(2\theta)}{\sqrt{\pi}}, \ldots, \frac{\cos(d\theta)}{\sqrt{\pi}}, \frac{\sin(d\theta)}{\sqrt{\pi}} \right\}$$

(similar for d odd)

How many zeros of a random polynomial are real?

Three models, three answers:

(i) Naive model: $\mathbb{E}_d = \frac{2}{\pi} \log d + .6257... + \frac{2}{\pi d} + O(1/d^2)$

(ii) Algebraic geometer's model: $\mathbb{E}_d = \sqrt{d}$

(iii) Analyst's model: $\mathbb{E}_d = \sqrt{\frac{d(d+2)}{3}}$

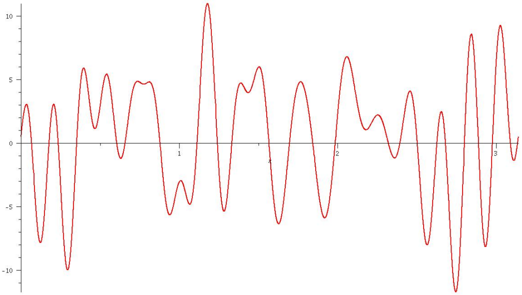


Figure : A random polynomial of degree 40.

Calculating the expectation exactly:

If ξ_k are i.i.d. Gaussians, the joint density function is **radial**:

$$\rho_\xi(a) = \prod_{k=0}^{d} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-a_i^2}{2}\right\} = \frac{\exp\left\{-r^2/2\right\}}{(2\pi)^{(d+1)/2}}$$

The expectation $\mathbb{E}_d$ of number of zeros:

$$\mathbb{E}_d := \frac{1}{(2\pi)^{(d+1)/2}} \int_{\mathbb{R}^{d+1}} |\{f_a(x) = 0\}| \exp\left\{-r^2/2\right\} \cdot dV(a).$$

Main idea for evaluating this: Integral geometry formula.

$$\frac{\int_{S^d} M_{\gamma^\perp}(a) \cdot dV(a)}{Vol(S^d)} = \frac{|\gamma|}{\pi}.$$

Extending "Analyst's model" to more variables

Consider homogeneous polynomials of degree d in $n+1$ variables.

- Restrict to the sphere S^n.

- Spherical harmonics provide an o.n. basis in $L^2(S^n)$.

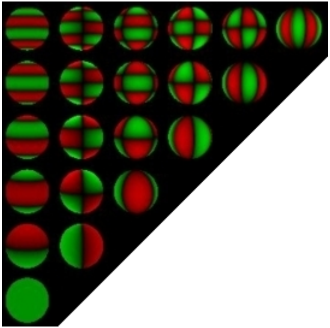
Note: A Fourier polynomial can be extended off the sphere either homogeneously or harmonically.

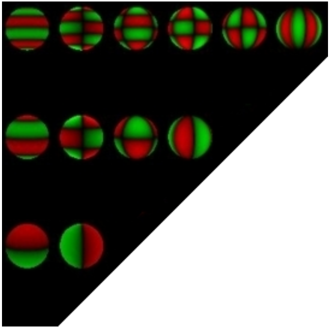
$$P_d = H_d \bigoplus |x|^2 H_{d-2} \bigoplus |x|^4 H_{d-4} \bigoplus \ldots$$

(Chapter 5 of "Harmonic Function Theory" by Axler, Bourdon, Ramey)

Random homogeneous polynomials

Using spherical harmonics as a basis for homogeneous polynomials, build a Gaussian polynomial in $n+1$ variables of degree d.





Spherical harmonics of fixed degree (monochromatic wave)

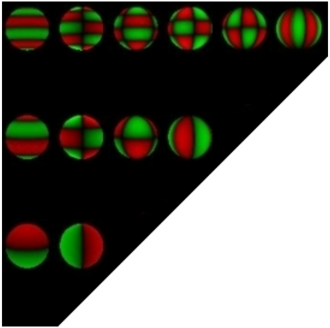
Spherical harmonic = homogeneous harmonic polynomial.

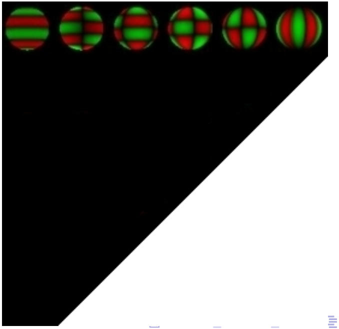
$$\Delta h = 0$$

It is also an eigenfunction of the spherical Laplacian.

$$\Delta_{S^n} h = -\lambda h.$$

Restricting to a single eigenspace H_d: “monochromatic wave”.





Generalizing to zero sets in $\mathbb{RP}^n$: what is the question?

For a hypersurface in $\mathbb{RP}^n$: there are two ways to generalize the question.

Metric question: "What is the expected $n-1$ dimensional volume?"

Topological question: "What is the expected number of connected components?"

(The number of connected components $N_f = b_0(Z_f)$ is the "zeroth Betti number".)

The metric question

Expected length (area, volume): Can be calculated *exactly* using integral geometry.

Obtained by P. Bérard (1984) in the case of random spherical harmonics, P. Burgisser (2007) for random polynomials, and A. Lerario, E.L., (2012) for a band of frequencies.

$$\mathbb{E}\ \mathrm{Vol}(Z_f) = \delta^{1/2} Vol(S^{n-1}),$$

where

$$\delta = \Theta(d^2),$$

can be specified exactly.

The topological question: much more difficult!

How many connected components?

No exact formulas known, and asymptotic answers only recently obtained.

First breakthrough came for the case of spherical harmonics (monochromatic wave) with $n = 2$.

Nodal domains for spherical harmonics

(F. Nazarov, M. Sodin, 2009) Spherical harmonics ($n = 2$):

$$\mathbb{E}_d \geq cd^2, \quad c > 0.$$

Along with a deterministic upper bound of the same order, this implies

$$\mathbb{E}_d = \Theta(d^2).$$

They used a simple and ingenious "barrier" method.

Result had been conjectured in statistical physics by E. Bogomolny and C. Schmit (2001).

Expected topology for homogeneous polynomials

(A. Lerario, E. L., 2012) Analyst's model:

$$\mathbb{E}_d = \Theta(d^n).$$

(D. Gayet, J-Y. Welschinger, 2013) Algebraic Geometer's model:

$$\mathbb{E}_d = \Theta(d^{n/2}).$$

(For $n = 2$, these behaviors were suggested in an inspiring hand-written letter from P. Sarnak to B. Gross and J. Harris, 2011.)

"The random curve is 4% Harnack."

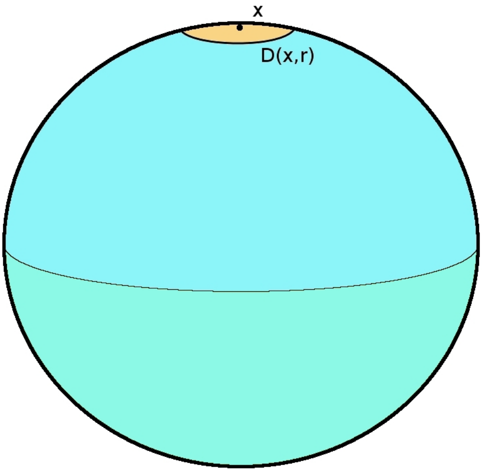
From spherical harmonics to homogeneous polynomials

In order to extend the "barrier methods" from "monochromatic waves" to the full spectrum of frequencies, we used the same basic outline and had to adapt many estimates using special function theory.

The barrier method: Outline

A deterministic upper bound of order d^n is already known from work of J. Milnor. The difficulty is the lower bound.

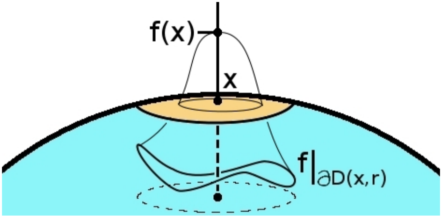
For a point $x \in S^n$ consider a neighborhood $D(x, r) \subset S^n$.



The barrier method: Outline

Define the event:

$$\Omega(x,r) = \{f(x) > 0 \text{ and } f|_{\partial D(x,r)} < 0\}.$$



If $\Omega(x,r)$ occurs then there is a component of Z_f in $D(x,r)$.

Lower bound on expectation

Pack the sphere with disjoint disks $D(x, r)$.

-The radius: take $r = \rho/d$ with ρ a constant chosen later.

-How many such disks? We can find at least $k \cdot d^n$.

The problem is reduced to showing that

$$P(\Omega(x, r)) > c,$$

Indeed, this would immediately imply

$$\mathbb{E}_d > c \cdot k \cdot d^n.$$

Lower bound on probability of $\Omega(x, r)$

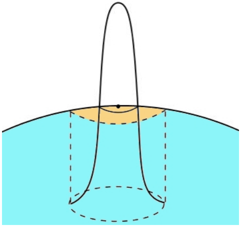
In order to show that

$$P(\Omega(x, r)) > c,$$

we choose a new o.n. basis that includes one element B_x with special behavior:

B_x peaks at x and also takes a large negative value on the boundary $\partial D(x, r)$.

-Namely the order is $\Theta(d^{n/2})$.



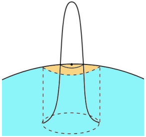
Lower bound on probability of $\Omega(x, r)$

Write the random polynomial in this basis:

$$f = \xi_0 B_x + f^{\perp}.$$

$\Omega(x, r)$ occurs if each of the following occurs:

1) $E_1 : \xi_0 B_x(x) \geq 2C_0 d^{n/2}$
   $\implies \xi_0 B_x|_{\partial D(x,r)} \leq -2C_0 d^{n/2}$
2) $E_2 : |f^{\perp}(x)| \leq C_0 d^{n/2}$
3) $E_3 : \|f^{\perp}|_{\partial D(x,r)}\|_{\infty} \leq C_0 d^{n/2}$



The problem is now reduced to showing that

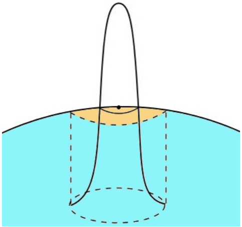
$$P(E_1 \cap E_2 \cap E_3) > c.$$

$$P(E_1 \cap E_2 \cap E_3) = P(E_1) \cdot P(E_2 \cap E_3) \geq P(E_1) \cdot (1 - P(\tilde{E}_2) - P(\tilde{E}_3))$$

Building a better barrier

For spherical harmonics, Nazarov and Sodin used one of the standard basis elements, the normalized "zonal" harmonic.

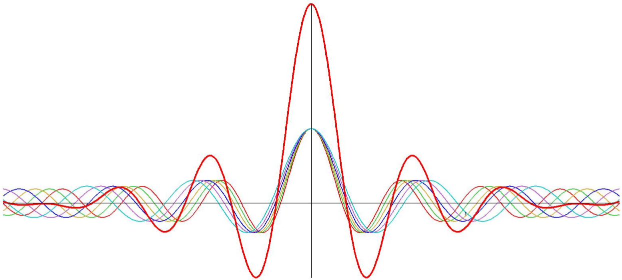
For polynomials, the zonal is insufficient.



A better barrier: Take a window of the zonals of different degrees (near the top degree) and sum them.

Building a better barrier

Addition of zonals:



-increases order of peaking by a factor of d.
-only increases the norm by a factor of $\sqrt{d}$.
-normalized result: magnifies the old barrier by a factor of $\sqrt{d}$.

(Rigorous proof uses Szegö's generalization of the Mehler-Heine asymptotic for Jacobi orthogonal polynomials.)

Controlling the supremum of $|f|$ on $\partial D(x, r)$

Onel step remaining. We want to show that:

$$P(\|f|_{\partial D(x,r)}\|_\infty > C_0 d^{n/2}) < c/C_0.$$

This reduces, by Markov's inequality, to showing:

$$\mathbb{E}\|f|_{\partial D(x,r)}\|_\infty \leq c d^{n/2}.$$

For this, we resorted to explicit construction of hyperspherical harmonics.

Controlling the supremum of $|f|$ on $\partial D(x, r)$

Want to show:

$$\mathbb{E}\|f|_{\partial D(x,r)}\|_\infty \leq cd^{n/2}.$$

Explicit construction of hyperspherical harmonics involves Gegenbauer polynomials, $P_{\ell-m}^{\left(\frac{n-1}{2}+m\right)}(cos\theta)$.

$$f(\theta, \phi) = \sum_{\ell} \sum_{m=0}^{\ell} \sum_{j \in I_m} \xi_\ell^j N_\ell^m (\sin\theta)^m P_{\ell-m}^{\left(\frac{n-1}{2}+m\right)}(\cos\theta) Y_j(\phi),$$

where $\theta \in (0, \pi), \phi \in S^{n-1}$, N_ℓ^m constants, $Y_j(\phi)$ hyperspherical harmonics of one less variable.

A non-technical part of the estimate

After changing order of summation, condensing some coefficients using sum law for Gaussians, and applying asymptotic estimates for small-angle evaluation of Gegenbauer polynomials, the result follows from combining such estimates with an estimate of $\mathbb{E}\|F_m\|_\infty$, which can be deduced from the deterministic estimate:

$$||F_m||_\infty \leq \sqrt{D_m}\|F_m\|_2.$$

where F_m is a degree-m Gaussian random hyperspherical harmonic in S^{n-1} (involving all the basis elements), and D_m is the dimension of the whole space.
(The sup-norm is now over the whole S^{n-1}.)

Fact for hyperspherical harmonics

The D_m-dimensional space H_m of spherical harmonics of degree m has a reproducing kernel $K_x(y)$ with $\|K_x\|_2 = \sqrt{D_m}$.

For any spherical harmonic $F \in H_m$:

$$||F||_\infty \leq \sqrt{D_m}\|F\|_2.$$

Proof.

$$F(x) = \langle F, K_x \rangle.$$

$$|F(x)| \leq \|F\|_2 \|K_x\|_2 = \sqrt{D_m}\|F\|_2.$$

$$||F||_\infty \leq \sqrt{D_m}\|F\|_2.$$

□

Immediate consequences

1. Further addressing Hilbert's Sixteenth Problem: It follows from the proof that for ν_0, the number of empty components:

$$\nu_0 = \Theta(d^n).$$

(not an immediate consequence of V. I. Arnold's bound.)

2. For random surfaces in S^3: For *each* Betti number, b_i:

$$\mathbb{E} b_i = \Theta(d^3).$$

$$\mathbb{E} b_2 = \mathbb{E} b_0 = \Theta(d^3), \quad \text{by Poincare duality, and}$$

$$\mathbb{E} b_1 = \Theta(d^3),$$

using a result of Bürgisser on the Euler characteristic $\mathbb{E}\chi \approx \frac{-d^3}{3\sqrt{3}}$.

Some open questions

Expected arrangement: How much "nesting" occurs on average? A. Lerario and E. L. (2012) proposed a precise version of this question and conjecture a relatively low level of nesting.

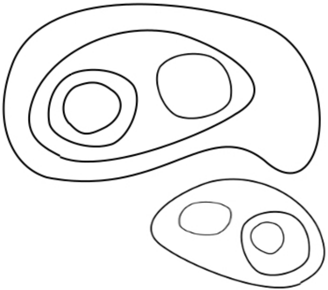
Variance: What is the order of the second moment?

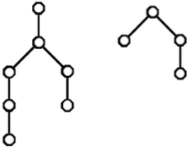
The many-variable limit: Fixing the degree, what happens as the *number of variables* goes to infinity? (We have suspicions of "universality".)

References

- E. Bogomolny, C. Schmit, *Percolation Model for Nodal Domains of Chaotic Wave Functions*. Phys. Rev. Letters, 88 (2002), 114102.
- A. Edelman, E. Kostlan, *How many zeros of a random polynomial are real?*, Bull. Amer. Math. Soc. 32 (1995), 1-37.
- D. Gayet, J-Y. Welschinger, several preprints on ArXiv, 2011-2013
- M. Kac: *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc., 49 (1943), 314-320.
- A. Lerario, E. Lundberg, *Statistics on Hilbert's sixteenth problem*, preprint available at http://arxiv.org/abs/1212.3823 (2012)
- F. Nazarov, M. Sodin, *On the Number of Nodal Domains of Random Spherical Harmonics*, Amer. J. of Math., 131 (2009), 1337-1357.
- P. Sarnak, *Letter to B. Gross and J. Harris*, publications.ias.edu/sarnak
- G. Szegö, *Orthogonal polynomials*, AMS Colloquium Publications, Vol XXIII, 1959.
- ...and more - email me at elundber@math.purdue.edu

Capturing the level of nesting





$$h(C) = 2 \cdot 2 \cdot (2 \cdot 2 + 2) + 2 \cdot (2 + 2 \cdot 2) = 36.$$

Use the tree structure of the components of a curve C to define an "energy" function $h(C)$ that is additive for disjoint unions of trees and multiplicative for appending trees.

Conjecture:

$$\lim_{d \to \infty} \frac{\log \mathbb{E} h(C)}{\log d} = 2$$

A Case Study for the many-variable limit: Quadrics

For a quadric, the zeroth Betti number is either 0 or 1, so it is more interesting to consider the *total Betti number*.

Random quadric ensemble $\to$ random matrix ensemble.

- Kac ensemble $\to$ Wigner matrix ensemble.
- Kostlan ensemble $\to$ GOE.
- Fubini-Study $\to$ matrix ensemble with correlation along diagonal.

The expected total Betti number: Appears to exhibit "universality".

What about higher-degree hypersurfaces?

Monochromatic waves: A surprisingly specific conjecture

Before addressing polynomials (full spectrum of eigenspaces), we consider the restricted case.
(and take n=2)

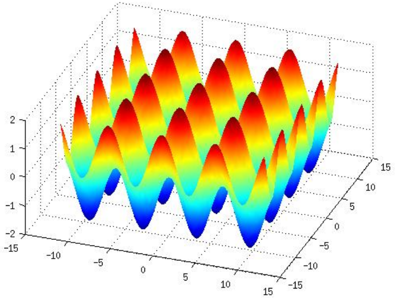
Conjecture (2001, Bogomolny-Schmit): Expected number of nodal domains:

$$\mathbb{E}_d \approx \frac{3\sqrt{3}-5}{2\pi} d^2.$$

(They also conjectured the variance has the same order and gave an explicit constant, and they later conjectured SLE(6) as a scaling limit.)

Bond-percolation model

To give some idea of the heuristics used for the conjecture from physics, instead of the sphere, consider the plane.
Consider the "plain" plane wave:



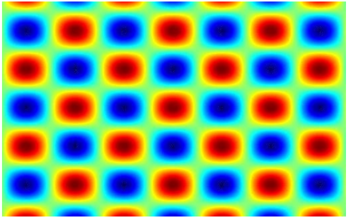
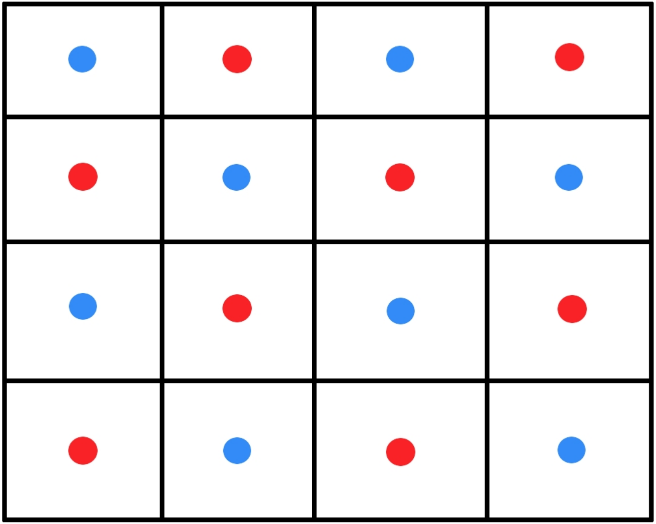


Figure : $\sin(mx) \cdot \sin(my)$

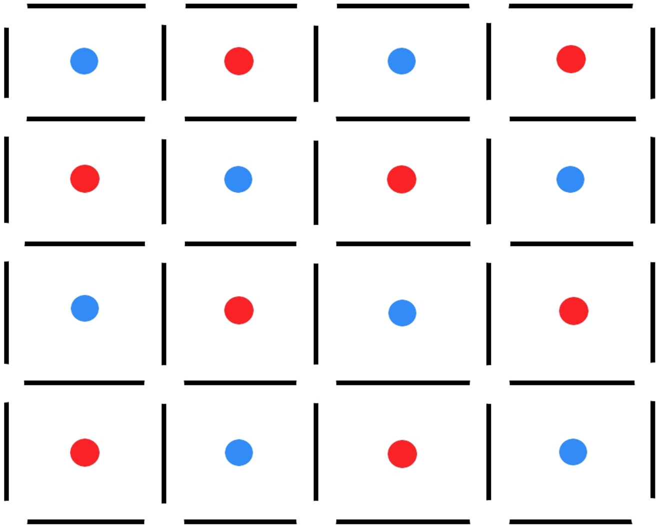
Take a small random perturbation of it.

Resulting nodal domains described by *bond-percolation*.

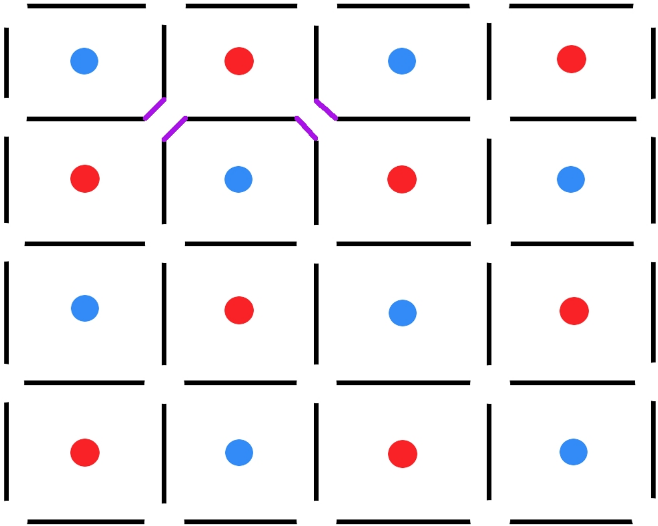
Bond percolation model for nodal domains



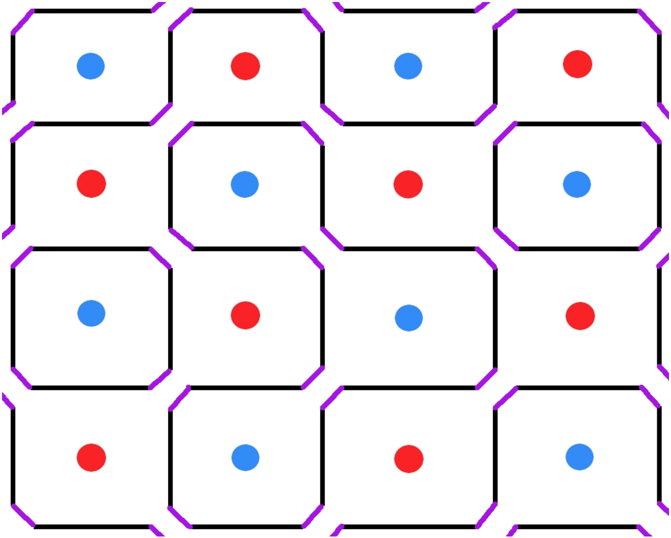
Bond percolation model for nodal domains



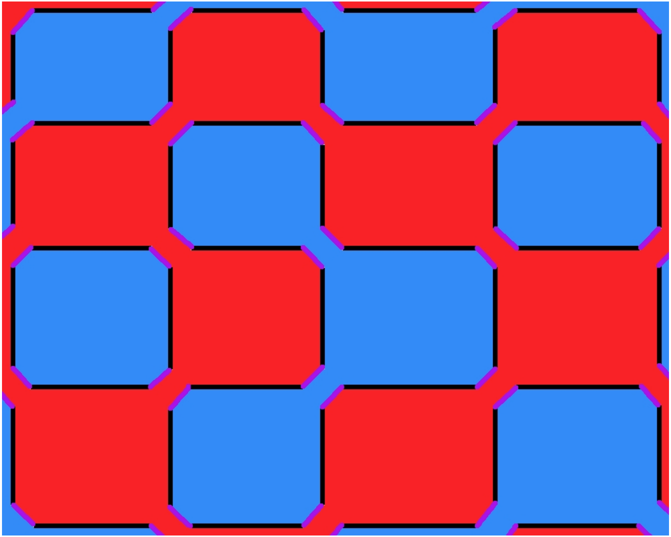
Bond percolation model for nodal domains



Bond percolation model for nodal domains



Bond percolation model for nodal domains



Bond percolation model for nodal domains

Reduces the problem to combinatorics (random graphs), if these leaps of logic are valid.

They obtained the exact constant in the asymptotic (letting the window of the lattice increase) by applying the saddle-point method to a certain generating function.