The geometry of polynomials and the validity of the Cavity Method

Justin Salez (ENS & INRIA)



Joint work with Charles Bordenave (CNRS & Uni. Toulouse) and Marc Lelarge (INRIA & ENS).

In a graph G, a matching is a subset of edges that are pairwise non-adjacent.

In a graph G, a matching is a subset of edges that are pairwise non-adjacent.



In a graph G, a matching is a subset of edges that are pairwise non-adjacent.



The matching number $\nu(G)$ is the maximum possible cardinality of a matching.

In a graph G, a matching is a subset of edges that are pairwise non-adjacent.



The matching number $\nu(G)$ is the maximum possible cardinality of a matching. \triangleright Typical behavior of $\nu(G)$ when *G* is a large random diluted graph ?

 G_n : Erdős-Rényi graph with parameter c/n on n vertices

 G_n : Erdős-Rényi graph with parameter c/n on n vertices

Theorem [Karp & Sipser, '82]

$$\frac{\nu(G_n)}{n} \quad \xrightarrow[n \to \infty]{} 1 - \frac{1}{2} \left(x^* + e^{-cx^*} + cx^* e^{-cx^*} \right),$$

where x^* is the smallest root of $x = e^{-ce^{-cx}}$ in [0, 1].

 G_n : Erdős-Rényi graph with parameter c/n on n vertices

Theorem [Karp & Sipser, '82]

$$\frac{\nu(G_n)}{n} \quad \xrightarrow{P} \quad 1 - \frac{1}{2} \left(x^* + e^{-cx^*} + cx^* e^{-cx^*} \right),$$

where x^* is the smallest root of $x = e^{-ce^{-cx}}$ in [0, 1].

Remark : In the sense of local weak convergence (Aldous & Steele 2003), G_n converges to a Galton-Watson tree with Poisson(c) degree distribution $\phi(x) = e^{-c(1-x)}$.

 G_n : Erdős-Rényi graph with parameter c/n on n vertices

Theorem [Karp & Sipser, '82]

$$\frac{\nu(G_n)}{n} \quad \xrightarrow{P} \quad 1 - \frac{1}{2} \left(x^* + e^{-cx^*} + cx^* e^{-cx^*} \right),$$

where x^* is the smallest root of $x = e^{-ce^{-cx}}$ in [0, 1].

Remark : In the sense of local weak convergence (Aldous & Steele 2003), G_n converges to a Galton-Watson tree with Poisson(c) degree distribution $\phi(x) = e^{-c(1-x)}$.

Theorem For random graphs G_1, G_2, \ldots s.t. $G_n \xrightarrow[n \to \infty]{d} GWT(\phi) \& \phi'(1) < \infty$,

$$\frac{\nu(G_n)}{n} \xrightarrow[n \to \infty]{P} \min_{[0,1]} F,$$

where
$$F = 1 - \frac{1}{2} \left(x \phi'(1-x) + \phi(1-x) + \phi \left(1 - \frac{\phi'(1-x)}{\phi'(1)} \right) \right)$$

PROOF :

PROOF :

The multi-affine polynomial $1 + x_1 + \ldots + x_d$ is non-vanishing whenever all variables lie in the open right half-plane.

A measure $\mu \colon \{0,1\}^E \to \mathbb{R}_+$ over the subsets of a finite ground set E

A measure $\mu \colon \{0,1\}^E \to \mathbb{R}_+$ over the subsets of a finite ground set E

can be represented by a multi-affine polynomial in $\mathbf{x} = (x_e)_{e \in E}$:

$$P(\mathbf{x}) = \sum_{F \subseteq E} \mu(F) \mathbf{x}^F, \text{ with } \mathbf{x}^F = \prod_{e \in F} x_e.$$

A measure $\mu \colon \{0,1\}^E \to \mathbb{R}_+$ over the subsets of a finite ground set E

can be represented by a multi-affine polynomial in $\mathbf{x} = (x_e)_{e \in E}$:

$$P(\mathbf{x}) = \sum_{F \subseteq E} \mu(F) \mathbf{x}^F, \text{ with } \mathbf{x}^F = \prod_{e \in F} x_e.$$

Key idea (Lee-Yang, 1952) : the location of the complex zeros of P is intimately connected to the "combinatorial complexity" of μ .

A measure $\mu \colon \{0,1\}^E \to \mathbb{R}_+$ over the subsets of a finite ground set E

can be represented by a multi-affine polynomial in $\mathbf{x} = (x_e)_{e \in E}$:

$$P(\mathbf{x}) = \sum_{F \subseteq E} \mu(F) \mathbf{x}^F, \text{ with } \mathbf{x}^F = \prod_{e \in F} x_e.$$

Key idea (Lee-Yang, 1952) : the location of the complex zeros of P is intimately connected to the "combinatorial complexity" of μ .

P decomposes as $P = x_e P^{/e} + P^{\setminus e}$, with $P^{/e}, P^{\setminus e}$ multiaffine on $E \setminus e$.

A measure $\mu \colon \{0,1\}^E \to \mathbb{R}_+$ over the subsets of a finite ground set E

can be represented by a multi-affine polynomial in $\mathbf{x} = (x_e)_{e \in E}$:

$$P(\mathbf{x}) = \sum_{F \subseteq E} \mu(F) \mathbf{x}^F$$
, with $\mathbf{x}^F = \prod_{e \in F} x_e$.

Key idea (Lee-Yang, 1952) : the location of the complex zeros of P is intimately connected to the "combinatorial complexity" of μ .

P decomposes as $P=x_eP^{/e}+P^{\backslash e}$, with $P^{/e},P^{\backslash e}$ multiaffine on $E\setminus e.$ The rational function $(P^{/e})/(P^{\backslash e})$ is called the influence of $e\in E$ on μ .

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Goal : compute the weighted number of spanning subgraphs of each given size, i.e.

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}.$$

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Goal : compute the weighted number of spanning subgraphs of each given size, i.e.

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}.$$

Remark : $z \to +\infty \longleftrightarrow$ maximize |F| subject to $\mu_G(F) > 0$.

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Goal : compute the weighted number of spanning subgraphs of each given size, i.e.

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}.$$

Remark : $z \to +\infty \longleftrightarrow$ maximize |F| subject to $\mu_G(F) > 0$. Example $\mu_i(F) = \mathbf{1}_{|F| \le 1}$:

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Goal : compute the weighted number of spanning subgraphs of each given size, i.e.

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}.$$

Remark : $z \to +\infty \longleftrightarrow$ maximize |F| subject to $\mu_G(F) > 0$.

Example $\mu_i(F) = \mathbf{1}_{|F| \le 1}$: μ_G counts matchings, P_G is the matching polynomial.

Given a finite graph G = (V, E), consider a measure $\mu_G \colon \{0, 1\}^E \to \mathbb{R}_+$ that factorizes into a product of local measures around vertices :

$$\mu_G(F) = \prod_{i \in V} \mu_i(F \cap E_i),$$

with $E_i = \{e \in E; e \text{ incident to } i\}$ and $\mu_i \colon \{0, 1\}^{E_i} \to \mathbb{R}_+$.

Goal : compute the weighted number of spanning subgraphs of each given size, i.e.

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}.$$

Remark : $z \to +\infty \longleftrightarrow$ maximize |F| subject to $\mu_G(F) > 0$.

Example $\mu_i(F) = \mathbf{1}_{|F| \le 1}$: μ_G counts matchings, P_G is the matching polynomial. \triangleright even for matchings, computing $P_G(1)$ is known to be # P-complete !

• Partition function :

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$$

• Partition function : P

$$P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$$

• Free entropy :

$$f_G(z) = \frac{1}{|V|} \log P_G(z)$$

- Partition function : $P_G(z) = \sum_{r=r} \mu_G(F) z^{|F|}$ $F \subseteq E$
- Free entropy : $f_G(z) = \frac{1}{|V|} \log P_G(z)$
- Boltzmann distribution with activity z > 0: $\mu_G^z(F) = \frac{\mu_G(F) z^{|F|}}{P_G(z)}$

- Partition function : $P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$
- Free entropy : $f_G(z) = \frac{1}{|V|} \log P_G(z)$
- Boltzmann distribution with activity z > 0:

$$\mu_G^z(F) = \frac{\mu_G(F)z^{|F|}}{P_G(z)}$$

• Internal energy :
$$u_G(z) = \frac{1}{|V|} \sum_{e \in E} \mu_G^z(e \in \mathcal{F})$$

- Partition function : $P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$
- Free entropy : $f_G(z) = \frac{1}{|V|} \log P_G(z)$
- Boltzmann distribution with activity z>0 : μ_C^z

$$\tilde{f}_G(F) = \frac{\mu_G(F)z^{|F|}}{P_G(z)}$$

• Internal energy : $u_G(z) = \frac{1}{|V|} \sum_{e \in E} \mu_G^z(e \in \mathcal{F}) = z f'_G(z)$

- Partition function : $P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$
- Free entropy : $f_G(z) = \frac{1}{|V|} \log P_G(z)$
- Boltzmann distribution with activity z > 0 : $\mu_G^z($

$$(F) = \frac{\mu_G(F)z^{|F|}}{P_G(z)}$$

• Internal energy :
$$u_G(z) = \frac{1}{|V|} \sum_{e \in E} \mu_G^z(e \in \mathcal{F}) = z f'_G(z)$$

▷ Computing the Boltzmann local marginals is generally hard.

- Partition function : $P_G(z) = \sum_{F \subseteq E} \mu_G(F) z^{|F|}$
- Free entropy : $f_G(z) = \frac{1}{|V|} \log P_G(z)$
- $(\mathbf{\Gamma}) \cdot [F]$ Boltzmann distribution with activity z > 0:

$$\mu_G^z(F) = \frac{\mu_G(F)z^{-1}}{P_G(z)}$$

• Internal energy :
$$u_G(z) = \frac{1}{|V|} \sum_{e \in E} \mu_G^z(e \in \mathcal{F}) = z f'_G(z)$$

 \triangleright Computing the Boltzmann local marginals is generally hard.

▷ Cavity Approximation (Mézard & Parisi, 85) : non-rigorous, but really efficient

z : positive parameter (activity)

- z : positive parameter (activity)
- $\mathcal{I}_{\vec{ij}}$: influence of edge ij on the local measure μ_i .

- z : positive parameter (activity)
- $\mathcal{I}_{\vec{ij}}$: influence of edge ij on the local measure μ_i .

 $x_{\vec{ij}}$: positive variables attached to each oriented edge $\vec{ij} \in \vec{E}$.

- z : positive parameter (activity)
- $\mathcal{I}_{\vec{ij}}$: influence of edge ij on the local measure μ_i .

 $x_{\vec{ij}}$: positive variables attached to each oriented edge $\vec{ij} \in \vec{E}$.

1. Find a solution $\{x_{\vec{ij}}; \vec{ij} \in \vec{E}\}$ to the local cavity equations on G:



- z : positive parameter (activity)
- $\mathcal{I}_{\vec{ij}}$: influence of edge ij on the local measure μ_i .

 $x_{\vec{ij}}$: positive variables attached to each oriented edge $\vec{ij} \in \vec{E}$.

1. Find a solution $\{x_{\vec{ij}}; \vec{ij} \in \vec{E}\}$ to the local cavity equations on G:



2. Use the cavity approximation to evaluate the Boltzmann marginals :

$$\mu_{G_n}^z \left(ij \in \mathcal{F} \right) \approx \frac{x_{\vec{ij}} x_{\vec{ji}}}{z + x_{\vec{ij}} x_{\vec{ji}}}.$$

When G is acyclic, the cavity method is valid for any activity z:

When G is acyclic, the cavity method is valid for any activity z:

1. convergence : the cavity equations admit a unique, globally attractive solution

When G is acyclic, the cavity method is valid for any activity z:

- 1. convergence : the cavity equations admit a unique, globally attractive solution
- 2. correctness : the cavity approximation is exact

When G is acyclic, the cavity method is valid for any activity z:

- 1. convergence : the cavity equations admit a unique, globally attractive solution
- 2. correctness : the cavity approximation is exact

In presence of cycles, the heuristic still seems to work well in various cases, leading to

When G is acyclic, the cavity method is valid for any activity z:

- 1. convergence : the cavity equations admit a unique, globally attractive solution
- 2. correctness : the cavity approximation is exact

In presence of cycles, the heuristic still seems to work well in various cases, leading to

 \triangleright powerful predictions e.g. the assignment problem : $(X_{i,j})_{1 \le i,j \le n}$ iid uniform on [0,1],

$$\min_{\pi \in \mathfrak{S}_n} \left(\sum_{i=1}^n X_{i,\pi(i)} \right) \xrightarrow[n \to \infty]{a.s.} \zeta(2) = \frac{\pi^2}{6}.$$

conjectured by Parisi in 1987 and proved by Aldous in 2001.

When G is acyclic, the cavity method is valid for any activity z:

- 1. convergence : the cavity equations admit a unique, globally attractive solution
- 2. correctness : the cavity approximation is exact

In presence of cycles, the heuristic still seems to work well in various cases, leading to

 \triangleright powerful predictions e.g. the assignment problem : $(X_{i,j})_{1 \le i,j \le n}$ iid uniform on [0,1],

$$\min_{\pi \in \mathfrak{S}_n} \left(\sum_{i=1}^n X_{i,\pi(i)} \right) \xrightarrow[n \to \infty]{a.s.} \zeta(2) = \frac{\pi^2}{6}.$$

conjectured by Parisi in 1987 and proved by Aldous in 2001.

efficient decentralized approximation algorithms e.g. belief propagation and its variants (Pearl, Yedida, Mézard, Montanari, Semerjian, Gamarnik, Shah, Weitz, Dembo...)

When G is acyclic, the cavity method is valid for any activity z:

- 1. convergence : the cavity equations admit a unique, globally attractive solution
- 2. correctness : the cavity approximation is exact

In presence of cycles, the heuristic still seems to work well in various cases, leading to

 \triangleright powerful predictions e.g. the assignment problem : $(X_{i,j})_{1 \le i,j \le n}$ iid uniform on [0,1],

$$\min_{\pi \in \mathfrak{S}_n} \left(\sum_{i=1}^n X_{i,\pi(i)} \right) \xrightarrow[n \to \infty]{a.s.} \zeta(2) = \frac{\pi^2}{6}.$$

conjectured by Parisi in 1987 and proved by Aldous in 2001.

efficient decentralized approximation algorithms e.g. belief propagation and its variants
(Pearl, Yedida, Mézard, Montanari, Semerjian, Gamarnik, Shah, Weitz, Dembo...)

But rigorous results remain sparse. Any simple, general conditions for validity ?

A measure μ (over subsets of a finite ground set) is \mathcal{D} -nonvanishing ($\mathcal{D} \subseteq \mathbb{C}$) if its multi-affine generating polynomial is non-vanishing whenever all variables lie in \mathcal{D} .

A measure μ (over subsets of a finite ground set) is \mathcal{D} -nonvanishing ($\mathcal{D} \subseteq \mathbb{C}$) if its multi-affine generating polynomial is non-vanishing whenever all variables lie in \mathcal{D} . Of particular combinatorial interest is the right open half-plane

 $\mathbb{H}_+ := \{ z \in \mathbb{C}; \Re(z) > 0 \}.$

A measure μ (over subsets of a finite ground set) is \mathcal{D} -nonvanishing ($\mathcal{D} \subseteq \mathbb{C}$) if its multi-affine generating polynomial is non-vanishing whenever all variables lie in \mathcal{D} . Of particular combinatorial interest is the right open half-plane

 $\mathbb{H}_+ := \{ z \in \mathbb{C}; \Re(z) > 0 \}.$

 spanning trees (Kirchhoff, 1847), ferromagnetic Ising model (Lee-Yang, 1952), matchings (Heilmann-Lieb, 1972), unbranched subgraphs (Ruelle, 1998), uniform matroïds (Wagner, 2000), determinantal/permanental measures (Choe, Oxley, Sokal, Wagner, 2004). See also the beautiful theory developped by Borcea and Bränden (2008).

A measure μ (over subsets of a finite ground set) is \mathcal{D} -nonvanishing ($\mathcal{D} \subseteq \mathbb{C}$) if its multi-affine generating polynomial is non-vanishing whenever all variables lie in \mathcal{D} . Of particular combinatorial interest is the right open half-plane

 $\mathbb{H}_+ := \{ z \in \mathbb{C}; \Re(z) > 0 \}.$

▷ spanning trees (Kirchhoff, 1847), ferromagnetic Ising model (Lee-Yang, 1952), matchings (Heilmann-Lieb, 1972), unbranched subgraphs (Ruelle, 1998), uniform matroïds (Wagner, 2000), determinantal/permanental measures (Choe, Oxley, Sokal, Wagner, 2004). See also the beautiful theory developped by Borcea and Bränden (2008). Say { μ_i , $i \in V$ } has the uniform half-plane property if there exists an open set $\mathcal{D} \supseteq \overline{\mathbb{H}}_+$ such that every μ_i is \mathcal{D} -non-vanishing.

A measure μ (over subsets of a finite ground set) is \mathcal{D} -nonvanishing ($\mathcal{D} \subseteq \mathbb{C}$) if its multi-affine generating polynomial is non-vanishing whenever all variables lie in \mathcal{D} . Of particular combinatorial interest is the right open half-plane

 $\mathbb{H}_+ := \{ z \in \mathbb{C}; \Re(z) > 0 \}.$

▷ spanning trees (Kirchhoff, 1847), ferromagnetic Ising model (Lee-Yang, 1952), matchings (Heilmann-Lieb, 1972), unbranched subgraphs (Ruelle, 1998), uniform matroïds (Wagner, 2000), determinantal/permanental measures (Choe, Oxley, Sokal, Wagner, 2004). See also the beautiful theory developped by Borcea and Bränden (2008). Say { μ_i , $i \in V$ } has the uniform half-plane property if there exists an open set $\mathcal{D} \supseteq \overline{\mathbb{H}}_+$ such that every μ_i is \mathcal{D} -non-vanishing.

Example : $\mu_i(F) = 1_{\{|F| \le 1\}}$, and more generally $\mu_i(F) = 1_{\{|F| \le r\}}$ for $r \in \mathbb{N}$.

1. Convergence : If the local measures have the uniform half-plane property, then on any graph with bounded degree and for any activity z, the cavity equation admits a unique, globally attractive solution.

1. Convergence : If the local measures have the uniform half-plane property, then on any graph with bounded degree and for any activity z, the cavity equation admits a unique, globally attractive solution.

2. Asymptotical correction: When G is an infinite tree, the cavity approximation can be directly used to construct a law μ_G^z on $\{0,1\}^E$ which turns out to be the weak limit of μ_{Gn}^z along any graph sequence $(G_n)_{n\geq 1}$ converging locally to G.

1. Convergence : If the local measures have the uniform half-plane property, then on any graph with bounded degree and for any activity z, the cavity equation admits a unique, globally attractive solution.

2. Asymptotical correction : When G is an infinite tree, the cavity approximation can be directly used to construct a law μ_G^z on $\{0,1\}^E$ which turns out to be the weak limit of $\mu_{G_n}^z$ along any graph sequence $(G_n)_{n\geq 1}$ converging locally to G.

Corollary : When the convergence $G_n \to G$ holds under uniform choice of the root \circ ,

$$u_{G_n}(z) \to u_G(z) = \frac{1}{2} \mathbb{E}\left[\sum_{i \sim \circ} \mu_G^z \left(i \circ \in \mathcal{F}\right)\right] \text{ and } f_{G_n}(z) \to \int_0^z \frac{u_G(s)}{s} ds.$$

1. contraction at low activity : if z is sufficiently small, then the cavity operator is contracting, hence there exists a unique, exponentially attractive fixed point.

- 1. contraction at low activity : if z is sufficiently small, then the cavity operator is contracting, hence there exists a unique, exponentially attractive fixed point.
- 2. analycity in the activity : the uniform half-plane property guarantees that the cavity operator preserves uniformly bounded analycity in a fixed complex domain containing the positive real line. Hence, the above convergence extends to any z > 0.

- 1. contraction at low activity : if z is sufficiently small, then the cavity operator is contracting, hence there exists a unique, exponentially attractive fixed point.
- 2. analycity in the activity : the uniform half-plane property guarantees that the cavity operator preserves uniformly bounded analycity in a fixed complex domain containing the positive real line. Hence, the above convergence extends to any z > 0.
- 3. local weak convergence : the cavity operator is "local", i.e. continuous with respect to local convergence, so we may pass to the limit in the cavity equations. When the limit is a Galton-Watson tree, the cavity equations may be simplified into a recursive distributional equation, which can sometimes be explicitly solved.

CONCLUSION

Gian-Carlo Rota (1932-1999) :

"The one contribution of mine that I hope will be remembered has consisted in just pointing out that all sorts of problems of combinatorics can be viewed as problems on location of the zeros of certain polynomials and in giving these zeros a combinatorial interpretation."