Coefficients of a chromatic invariant in random graphs

Raúl Penaguião

Research seminar: Discrete mathematics
Valentin Féray
University of Zürich

raul.penaguiao@math.uzh.ch

18th October, 2016

Overview

1. The chromatic symmetric function
   - The ring of symmetric functions
   - Colourings on graphs
2. Tree classification conjecture
   - Classification of caterpillars
3. On random graphs
   - Subgraph containment
   - Matchings

The ring of symmetric functions - Notation

A *weak (strong) composition* is an infinite list of non-negative integers

$$\alpha = (\alpha_1, \cdots),$$

where all but finitely many are zero (finite list of positive).
We say that α is a weak composition of n, or that $\alpha \models n$, if $\sum_{k \geq 1} \alpha_k = n$.
For a given weak composition α, we write

$$x^\alpha = \prod_{k \geq 1} x_k^{\alpha_k}.$$

So for $\alpha = (2, 1, 0, 4, 2, 0, 0, \ldots)$ we have

$$x^\alpha = x_1^2 x_2 x_4^4 x_5^2.$$

The ring of symmetric functions - Notation

An *integer partition* is an infinite multi-set of non-negative integers, where all but finitely many are zero.
We say that λ is a partition of n, or $\lambda \vdash n$, if $\sum_{k\geq 1} \lambda_k = n$.
For a given partition λ, we write

$$x_\lambda = \prod_{k\geq 1} x_{\lambda_k} \,.$$

So, for $\lambda = (4, 2, 2, 1)$ we have

$$x^\alpha = x_1 x_2^2 x_4 \,.$$

The ring of symmetric functions - Notation

We can also write $\lambda = \langle 1^{\lambda^{(1)}} 2^{\lambda^{(2)}} \ldots \rangle$ to represent a partition. In such a way that $\lambda = \langle 1^3 2^2 \rangle$ represents the same partition as $\lambda = (2, 2, 1, 1, 1)$.

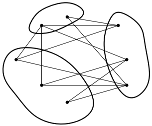

Figure: A set partition π of the vertex set of a graph

By disregarding the order, we obtain a mapping from weak compositions to integer partitions, $\alpha \mapsto \lambda(\alpha)$, called the type of α.
Simply set $\lambda(1, 2, 0, 3, 1, 0, 0, 0, \ldots) =$(3, 2, 1, 1).
For a set partition we write $l(\lambda)$ for the number of non-zero entries of λ.

The ring of symmetric functions

The set Λ_n of homogeneous symmetric function of degree n is the set of formal sums

$$f = \sum_{\alpha \models n} a_\alpha x^\alpha \,,$$

where the coefficients a_α do not depend on the type of α, and there are infinitely many variables $x_1, x_2, \cdots$.
With this, the formal expression is invariant if we change the role of the variables x_i, x_j. The ring of symmetric functions is $\Lambda = \oplus_{n \geq 0} \Lambda_n$.

The ring of symmetric functions - Monomial basis

Naturally, for a partition λ, the symmetric functions

$$m_\lambda = \sum_{\lambda(\alpha)=\lambda} x^\alpha \,,$$

form a basis of Λ, called the *monomial basis*.
Examples of symmetric functions:

$$m_{()} = 1 \,,$$

$$m_{(1)} = x_1 + x_2 + \cdots \,,$$

$$m_{(2,1)} = x_1 x_2^2 + x_2 x_1^2 + x_1 x_3^2 + \cdots \,.$$

The ring of symmetric functions - Power-sum basis

Let $p_n = \sum_{k\geq 1} x_k^n = m_{(n)}$ and for an integer partition $\lambda = (\lambda_1, \dots)$ we define

$$p_\lambda = \prod_{k\geq 1} p_{\lambda_k} .$$

These symmetric functions also form a basis of Λ, and are called the *power-sum basis*.

$$p_{(2,1)} = \left(\sum_i x_i^2\right)\left(\sum_i x_i\right) = \left(\sum_i x_i^3\right) + \left(\sum_{i\neq j} x_i^2 x_j\right)$$

Let $e_n = \sum_{a_1<a_2<\cdots<a_n} x_{a_1}x_{a_2}\cdots x_{a_n} = m_{\underbrace{(1,\dots,1)}_{n \text{ ones}}}$

For an integer partition $\lambda = (\lambda_1, \dots)$ we define

$$e_\lambda = \prod_{k\geq 1} e_{\lambda_k}\,.$$

So, for instance, $e_{(2,1)} = e_2 e_1 = \left(\sum_{i<j} x_i x_j\right)\left(\sum_i x_i\right)$.

$$e_{(2,1)} = \sum_{i<j<k} 3x_i x_j x_k + \sum_{i\neq j} x_i^2 x_j = 3m_{(1,1,1)} + m_{(2,1)}\,.$$

The chromatic symmetric function

A colouring of a graph is a map $k : V(G) \to \mathbb{N}$. A colouring is called *proper* if no edge is monochromatic.
The sum:

$$\chi_G = \sum_k x_{k(v_1)} \cdots x_{k(v_n)} ,$$

where k runs over all proper colourings, is a symmetric function.
We call it the *chromatic symmetric function* of the graph G.
If $G = G_1 \uplus G_2$ then $\chi_G = \chi_{G_1} \chi_{G_2}$.

Expansion of χ_G - the monomial basis

For a graph G on n vertices and a partition $\lambda \vdash n$, let $x_\lambda(G) = x_\lambda$ count the number of stable set partitions of $V(G)$ of type λ.

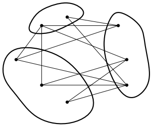

Figure: A set partition π of the vertex set of a graph

Theorem

$$\chi_G = \sum_{\lambda \vdash n} x_\lambda(G)\, aut(\lambda) m_\lambda \,,$$

where $aut(\lambda) = \prod_j \lambda^{(j)}!$.

Monomial basis - simple application

In the complete graph K_n, there is only one stable partition (the finer one). So, we have $\chi_{K_n} = \mathrm{aut}(1, \ldots, 1) m_{(1,\ldots,1)} = n! m_{(1,\ldots,1)}$.

Theorem (Modular relation)

Given a graph G and three edges e_1, e_2, e_3 that form a triangle, we have that:

$$\chi_G + \chi_{G \setminus \{e_1, e_2\}} = \chi_{G \setminus e_1} + \chi_{G \setminus e_2}$$

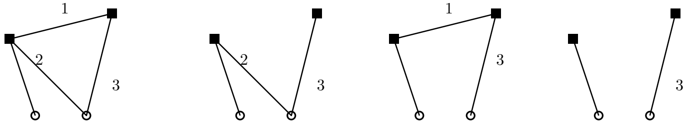

Figure: Relation between set partitions on the different graphs.

This can also be done directly by definition, by counting the contributions of each colouring to each part of the sum.

Expansion of χ_G - the power sum basis

For a graph G and a set of edges $A \subseteq E(G)$, we define $\lambda(A)$ as the partition type of the set partition given by the connected components of the graph $(V(G), A)$.

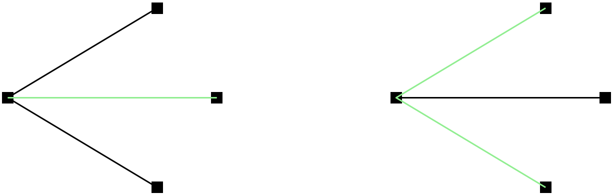

Figure: Sets of edges A and B in black

In this case $\lambda(A) = (3, 1)$ and $\lambda(B) = (2, 1, 1)$.

Expansion of χ_G - the power sum basis

For the power-sum basis we have the following identity:

Theorem

$$\chi_G = \sum_{A \subseteq E(G)} (-1)^{\#A} p_{\lambda(A)} \,.$$

One edge set of type $(1,1,1,1)$, the empty set.
Three edge sets of type $(2,1,1)$.
Three edge sets of type $(3,1)$.
One edge set of type (4).

So $\chi_G = p_{(1,1,1,1)} - 3p_{(2,1,1)} + 3p_{(3,1)} - p_{(4)}$.

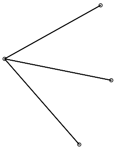

Figure: The claw graph G.

Expansion of χ_G - the power sum basis

This formula becomes quite simple for trees, where the number of edges in A and the number of connected components $l(\lambda(A))$ are directly related.

Theorem

If we set $\theta_\lambda = \#\{A \subseteq E(G) | \lambda(A) = \lambda\}$, and G is a tree, then

$$\chi_G = \sum_{A \subseteq E(G)} (-1)^{\#A} p_{\lambda(A)} = \sum_{k=1}^{n-1} (-1)^{n-k} \sum_{\substack{\lambda \vdash n \\ l(\lambda)=k}} \theta_\lambda p_\lambda \,.$$

Expansion of χ_G - the elementary basis

Over the elementary basis, there is no known combinatorial interpretation. However we can see some examples.

$$\chi_{K_n} = n! m_{(1,\dots,1)} = n! e_n$$

$$\chi_{0_n} = (e_1)^n = e_{(1,\dots,1)}$$

Expansion of χ_G - the elementary basis

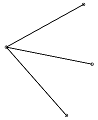

Figure: The claw graph G

$$\chi_G = 24m_{(1,1,1,1)} + 6m_{(2,1,1)} + m_{(3,1)}$$

$$= 6e_{(3,1)} - 12e_{(2,2)} - 5e_{(2,1,1)} + 44e_{(1,1,1,1)}$$

Stanley Conjectures

For a poset P, let $G(P)$ be the *incomparability graph* with vertices $V(G) = P$ and two vertices are connected if they are not comparable in P. A poset P is *(3+1)-free* if it has no embedded (3+1)-poset.

Conjecture

The chromatic symmetric function of the incomparability graph of a (3+1)-free poset is e-positive.

Conjecture

Trees are classifies by the chromatic symmetric function.

Chromatic Symmetric function - a strong invariant

With the chromatic symmetric function we can compute:

- Number of vertices (degree of χ_G);
- Number of edges;
- Number of connected components (from the power sum expansion);
- Degree sequence (Martin, J. L., Morin, M., and Wagner, J. D. 2008);
- Number of subtrees with k vertices and l leafs;
- ...

Classification of subfamilies of trees

To classify a subfamily $\mathcal{F}$ of graphs, we will follow a pattern:

- Identify an invariant Φ that almost classifies the family $\mathcal{F}$, but we can trivially obtain Φ from the chromatic symmetric function.
- Close the remaining gap with the power of chromatic symmetric functions.

Overview of the classification of caterpillars

A hard subfamily of trees: Caterpillars.

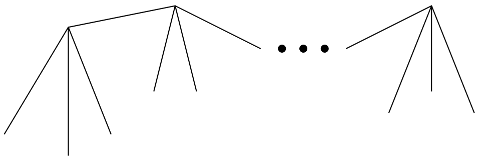

Figure: A generic caterpillar

A caterpillar is a tree G with leafs $L(G)$ in which $T \setminus L(G)$ is a path. A caterpillar is called *proper* if any non-leaf vertex is neighbor of a leaf.

Consider the U-polynomial:

$$U_G = \sum_{A \subseteq E(G)} x_{\lambda(A)} = \sum_{\lambda \vdash n} \theta_\lambda(G) x_\lambda \,.$$

This is trivially as strong as the chromatic symmetric function. Take as well the U^L-polynomial:

$$U^L_G = \sum_{L_e \subseteq A \subseteq E(G)} x_{\lambda(A)} \,.$$

Where L_e is the set of edge-leafs of G. This does not depend on x_1 for proper caterpillars. Indeed, for proper caterpillars we have

$$U_G \big|_{x_1 = 0} = U^L_G \,.$$

Overview of the classification of caterpillars

From a strong composition α we can associate a caterpillar $\Phi(\alpha)$.
Let $[\alpha]$ be the set of strong compositions that generate the same caterpillar (it has at most two elements).

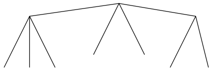

Figure: A caterpillar with main composition $\alpha = (4, 3, 3)$

Take the poset of strong compositions of α, where we have an ordering by refinement, i.e. satisfies

$$(\cdots, \alpha_i, \alpha_{i+1}, \cdots) \leq (\cdots, \alpha_i + \alpha_{i+1}, \cdots)$$

Then, we define,

$$\mathcal{L}_\alpha = \sum_{\beta \geq \alpha} x_\beta \, .$$

Overview of the classification of caterpillars

$$\mathcal{L}_\alpha = \sum_{\beta \geq \alpha} x_\beta \, .$$

$$U_T^L = \sum_{L_e \subseteq A \subseteq E(G)} x_{\lambda(A)} \, .$$

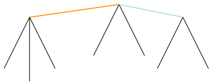

Figure: Four possible sets $L_e \subseteq A \subseteq E(G)$

Sets $L_e \subseteq A \subseteq E(G)$ correspond bijectively to compositions $\beta \geq \alpha$.
The U_T^L-polynomial of the tree $T = \Phi(\alpha)$ is the $\mathcal{L}$-polynomial $\mathcal{L}_\alpha$ in the poset of strong compositions of n.

Overview of the classification of caterpillars

Theorem ($\mathcal{L}$-polynomial in the poset of strong compositions)

If $\beta = \beta_1 \circ \cdots \circ \beta_k$ and $\alpha = \alpha_1 \circ \cdots \circ \alpha_j$ are irreducible factorizations, then $\mathcal{L}_\alpha = \mathcal{L}_\beta \Leftrightarrow k = j$ and $[\alpha_i] = [\beta_i]$.

For that we look to the coefficient $\theta_{(n-\delta,\delta-1,1)}(G)$, which counts the number of leafs of subtrees of size δ that are not leafs of the original tree (this part uses the fact that the caterpillar is proper).

Theorem (José Aliste-Prieto, and José Zamora (2104))

Proper caterpillars are distinguished by the chromatic symmetric function.

Erdös Renyi model

The random graph $G(n, p)$ has n vertices, every edge occurs in $G(n, p)$ independently with probability p.
Hence, the coefficients of $\chi_{G(n,p)}$ are random variables.
In the monomial basis, X_λ represents the number of stable set partitions of type λ, which will be the random variables determining $\chi_{G(n,p)}$.

Example (Behaviour of X_{31^s})

Let X_{v_1,v_2,v_3} be a r.v. that is:

- One if $\{v_1, v_2, v_3\}$ is an independent set.
- Zero otherwise.

Then, from $X_{31^s} = \sum_{v_1,v_2,v_3} X_{v_1,v_2,v_3}$ we have,
$\mathbb{E}[X_{31^s}] = \sum_{v_1,v_2,v_3} \mathbb{E}[X_{v_1,v_2,v_3}] = \binom{n}{3}(1-p)^3$.
So, if $1 - p = o(n^{-1})$ then $\mathbb{E}[X_{31^s}] = o(1)$.

The first moment method

If $\mathbb{E}[X_n] = o(1)$, what can we say about $\mathbb{P}[X_n = 0]$?
We have that $\mathbb{P}[X_n = 0] \geq 1 - \mathbb{E}[X_n] = 1 - o(1)$.
So,

Theorem (First moment method)

If $\mathbb{E}[X_n] = o(1)$ then $X_n = 0$ asymptotically almost surely (a.a.s.), i.e. $\mathbb{P}[X_n = 0] \to 1$.

Since we have $\mathbb{E}[X_{31^s}] = \binom{n}{3}(1-p)^3$, if $p = 1 - o(n^{-1})$ we have that $X_{31^s} = 0$ a.a.s. on $G(n, p)$.

The threshold problem

In the $G(n,p)$ model, we want to find functions $\hat{p}$ that satisfy

- If $p << \hat{p}$ then $\mathbb{P}[G(n,p) \in Q] \to 0$.
- If $p >> \hat{p}$ then $\mathbb{P}[G(n,p) \in Q] \to 1$.

Such functions always exist for *increasing* properties, i.e. if $G_1 \subseteq G_2$ and G_1 has property Q, then G_2 has property Q.
Thresholds are of general interest in the literature of random graphs.

Relation with subgraphs and connected components

Existence of a subgraph of type $\biguplus_i K_{n_i}$ in $G(n, p)$

$$\Leftrightarrow$$

$x_{(n_1, n_2, \ldots, n_k, 1, \ldots, 1)} \neq 0$ in $G(n, p)^c \sim G(n, 1-p)$.

Biggest component has size k in $G(n, p)$

$$\Rightarrow$$

$\theta_{(k+1, \lambda_2, \lambda_3, \cdots)} = 0$ in $G(n, p)$.

Subgraph Containment

Result from the literature: For a graph G let $d(G) = \frac{e(G)}{v(G)}$ be the graph density, and $m(G) = \max_{H \subseteq G}\{d(H)\}$.

Theorem

Let G be a graph, then

- *If $p << n^{-\frac{1}{m(G)}}$ then $\mathbb{P}[G \subseteq G(n,p)] \to 0$.*
- *If $p >> n^{-\frac{1}{m(G)}}$ then $\mathbb{P}[G \subseteq G(n,p)] \to 1$.*

This solves the subgraph problem for threshold problem for $\biguplus_i K_{\mu_i}$
Which relates to the partition $\lambda = \mu 1^s$ where μ is a constant partition and the threshold problem of $X_\lambda = 0$.

Main results in matchings

What about the case $\lambda = 1^s 2^r$ for $r \to \infty$?
From the literature

Theorem (Existence of a perfect matching in $G(n, p)$)

If $p = n^{-1} \log n + \omega(n^{-1})$ then a.a.s. there is a matching that covers all but at most one vertex in $G(n, p)$.

Then we can conclude that $X_{2^{n/2}} \neq 0$ for $p = n^{-1} \log n + \omega(n^{-1})$.
But we want to do better, by trying to evaluate the case $\lambda = 2^r 1^s$ for different growth rates of r and s.

Main results in matchings

Theorem (Janson, Svante, Tomasz Luczak, and Andrzej Rucinski. 2011)

If $p = n^{-1} \log n/2 + \omega(n^{-1})$ *then a.a.s., the set of non-isolated vertices have a matching that covers all but at most one vertex in* $G(n, p)$*.*

This theorem is shown using Hall's theorem to show that any bipartite subgraph with high minimum degree in $G(n, p)$ will have a perfect matching a.a.s.

Detour - 2nd moment method

If X_n is a sequence of random variables with $\mathbb{E}[X_n] \to \infty$, when do we have that $\mathbb{P}[X_n = 0] \to 0$?

Theorem (Chebichev inequality)

For any random variable, we have that

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq a] < a^{-2}\mathrm{Var}[X] \, .$$

So, if $\mathrm{Var}[X_n] = o(\mathbb{E}[X_n]^2)$ then $\mathbb{P}[X_n = 0] \to 0$. This is the *2nd moment method*.

We can even go a little bit further and show that

$$\mathbb{P}[|X - \mathbb{E}[X]| < \epsilon\mathbb{E}[X]] = o(1) \, .$$

Thresholds for big matchings

Case $np = k \log n$, $k \in (0.5, 1)$:
Less isolated vertices $\Rightarrow$ bigger matchings.
We know that $X_\lambda \neq 0$ a.a.s. for some $\lambda = 1^r 2^s$ in $G(n, 1-p)$, a.a.s, where r is roughly the number of isolated vertices.

$$X = \#\{ \text{ Isolated vertices } \}.$$

$$\mathbb{E}[X] = n(1-p)^{n-1} = n \exp(-np)(e^p + o(1)) \sim n^{1-k}.$$

Simple calculation shows that
$\mathrm{Var}[X] = \sum_{v,w} \mathrm{Cov}[I_v, I_w] = \frac{p}{1-p}\mathbb{E}[X]^2 + \Theta(\mathbb{E}[X])$.

Thresholds for big matchings

Recall: $\mathbb{E}[X] \sim n^{1-k}$.

2nd moment method concludes that for any given constant ϵ
$r = (1+\epsilon)n^{1-k} \geq X + 1$ a.a.s.
So for $\lambda = 1^r 2^s$ we have that $X_\lambda \neq 0$ a.a.s. in $G(n, 1-p)$.

References

Janson, S., Luczak, T., & Rucinski, A.
Random Graphs (Vol. 45).
John Wiley & Sons, 2011.

Stanley, Richard P.
A symmetric function generalization of the chromatic polynomial of a graph.
Advances in Mathematics 111.1 (1995): 166-194.

Aliste-Prieto, Jos, and Jos Zamora.
Proper caterpillars are distinguished by their chromatic symmetric function.
Discrete Mathematics 315 (2014): 158-164.