The number of spanning trees in a graph

Данила Черкашин

ИМИ БАН

on joint works with Федор Петров and Павел Прозоров

Национален колоквиум по математика

22 януари 2025

Introduction

Let G be a simple graph. A spanning tree is a connected subgraph of G that has no cycles and contains all vertices of G. We denote by $\tau(G)$ the number of spanning trees in G.

K_n, C_n and P_n denote complete, cycle and path graphs with n vertices, respectively. A complete bipartite graph is denoted by $K_{n,m}$ where n, m are the sizes of the parts.

Enumeration of spanning trees started at 1847 by a celebrated Matrix-tree theorem of Kirchhoff, 1847.

Cayley in 1889 showed that for a complete graph K_n one may found much more than just a number of spanning trees.

Then a lot of nice formulas for special graph classes were obtained by Moon, Kelmans, Pak, Postnikov, Ehrenborg and others.

Kirchhoff Matrix-tree theorem

Theorem (Kirchhoff, 1847)

The number of spanning trees $\tau(G)$ is determined by Laplacian matrix cofactors (first minors):

$$\tau(G) = \mathrm{Det}(L^*(G)),$$

where $L^*(G)$ is the reduced Laplacian of G.

The proof is an application of the Binet–Cauchy identity to

$$B(G)B(G)^T = L(G),$$

where B is the incidence matrix of G.

Thus $\tau(G) = \frac{\lambda_1 \lambda_2 \cdots \lambda_{n-1}}{n}$, where $\{\lambda_i\}$ is the spectrum of L.

This also works for weights on edges!

Cayley theorem

Polynomial P_G (vertex degree enumerator) encodes spanning trees:

$$P_G(x_1, \ldots, x_n) = \sum_{T \in S(G)} \prod_{v \in V} x_v^{\deg_T(v)-1},$$

where $S(G)$ is the set of spanning trees of G and $\deg_T(v)$ is the degree of v in a tree T.

Theorem (Cayley, 1889)

$P_{K_n} = (x_1 + x_2 + \cdots + x_n)^{n-2}.$

A combinatorial proof uses Prüfer codes.

Consider the variable x_e for every edge e. Define the *edge spanning enumerator* of a graph G as

$$Q_G\left(\{x_e\}_{e \in E}\right) = \sum_{T \in S(G)} \prod_{e \in E(T)} x_e.$$

Why polynomials?

They gives more structure.

Magic example: estimate a variance of a random variable $\deg_T(v)$ in a randomly (uniformly) chosen spanning tree $T \in S(G)$.

By a magical reason (we discuss it later)

$$q_v(x) := \sum_{T \in S(G)} x^{\deg_T(v)} = x \prod_{i=1}^{\deg_G(v)-1} (x + t_i)$$

for $t_i \geq 0$.

Then $\deg_T(v) - 1$ is a sum of independent Bernoulli coins with parameters $\frac{1}{t_i+1}$. This immediately gives the upper bound $\frac{1}{4}(\deg(v) - 1)$.

This is sharp for $K_{2,n}$.

Stable and real stable polynomials

A polynomial $P(x_1, x_2, \ldots, x_n)$ with complex coefficients is called *stable*, if $P(z_1, z_2, \ldots, z_n) \neq 0$ whenever $z_1, z_2, \ldots, z_n$ all have positive imaginary parts.

If P is stable and has real coefficients, then P is *real stable*.

Some examples are:

$$a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n, \quad a_0 \in \mathbb{R}, \quad a_1, \ldots, a_n \geqslant 0$$

$$c\prod_{i=1}^{n}(x - \alpha_i), \quad c, \alpha_i \in \mathbb{R}$$

$$p_1p_2 \ldots p_n, \quad \text{if } p_i \text{ are real stable}$$

Real stable polynomials: basic properties

Let $P(x_1, x_2, \ldots, x_n)$ be a real stable polynomial. Then the following polynomials are also real stable or identically zero:

- $x_1^{d_1} P\left(-\frac{1}{x_1}, x_2, \ldots, x_n\right)$, where d_1 is the degree of P with respect to the variable x_1;
- $\frac{\partial P}{\partial x_1}(x_1, x_2, \ldots, x_n)$;
- $Q(x_1, x_2, \ldots, x_{n-1}) := P(x_1, x_2, \ldots, x_{n-1}, a)$ for any real a;
- $Q = \lim P_k$ if $P_1, P_2, \ldots$ are real stable of bounded degree and coefficient-wise converge to Q;
- $\mathrm{Det}(B_0 + x_1 B_1 + x_2 B_2 + \ldots + x_k B_k)$ if B_0 is Hermitian, $B_1, \ldots, B_k$ are Hermitian non-negative definite.

Vertex spanning tree enumerator

Now it is clear that $Q(G)$ is always real stable which implies the magic property.

But what about vertex enumerator?

Recall that $P_G(x_1, x_2, \dots, x_n) = \sum_{T \in S(G)} \prod_{v \in V} x_v^{\deg_T(v)-1}$ and $P_{K_n} = (x_1 + \dots + x_n)^{n-2}$.

Another well-known result is $P_{K_{n,m}} = \left(\sum_{i=1}^{n} x_i\right)^{m-1} \left(\sum_{i=1}^{m} y_i\right)^{n-1}$.

The vertex spanning enumerator is not always real stable! Indeed,
$P_{C_5}(x_1, x_2, x_3, x_4, x_5) = x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_5 + x_4x_5x_1 + x_5x_1x_2$,
$P_{C_5}(1, x_2, -1, x_4, x_5) = x_2(x_5 - x_4 - 1)$.

QUESTION. For which graphs G is the polynomial P_G is real stable?

The answer

Theorem (C.–Petrov–Prozorov, 2023)

The statements (i)–(iii) are equivalent.

(i) *A graph G is distance-hereditary.*

(ii) *The polynomial P_G factors into linear terms.*

(iii) *The polynomial P_G is real stable.*

We focus on a constructive proof of the implication from (i) to (ii).

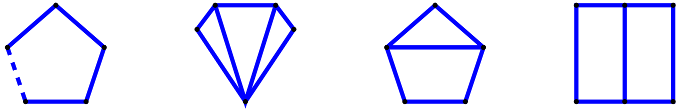
This theorem recovers a lot results by Bogdanowicz, Ehrenborg and van Willigenburg, Gao and Liu, Klee and Stamps and many others.

Distance-hereditary graphs

A *distance-hereditary graph* is a connected graph in which any connected induced subgraph preserves distances.

(i) These are graphs in which any induced path is a shortest path.

(ii) These are graphs that do not contain the following induced subgraphs: a cycle of length ≥ 5, a gem, a house or a domino.



(iii) These are graphs that can be constructed from a single vertex through a sequence of the following three operations:

- Adding a new pendant vertex.
- Replacing any vertex with a pair of vertices, each having the same neighbors as the removed vertex.
- Replacing any vertex with a pair of vertices, each having the same neighbors as the removed vertex, including the other vertex in the pair.

Main tool

Theorem (C.–Prozorov, 2024+)

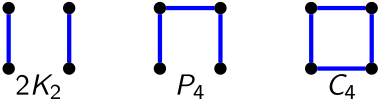
Let a vertex v_1 be marked in G_1 and a vertex v_2 marked in G_2. A graph H is the disjoint union of G_1 and G_2, with removed vertices v_1 and v_2, and added all edges between $N_1(v_1)$ and $N_2(v_2)$. Then

$$P_H = P_{G_1}\left(x_1, \ldots, x_{n-1}, \sum_{i=1}^{m-1} y_i\right) \cdot P_{G_2}\left(y_1, \ldots, y_{m-1}, \sum_{i=1}^{n-1} x_i\right).$$

We use the latter definition of distance-hereditary graphs.
Adding a pendant vertex to v multiplies P_G by the variable x_v (or put $G_1 = G$, $v_1 = v$, $G_2 = (\{1,2,3\}, \{\{1,2\},\{1,3\}\})$, $v_2 = 2$).
For duplicating vertex v in a graph G with an edge substitute $G_1 = G$, $v_1 = v$, $G_2 = K_3$ (and any v_2).
Duplicating vertex v in a graph G without an edge corresponds to $G_1 = G$, $v_1 = v$, $G_2 = (\{1,2,3\}, \{\{1,2\},\{1,3\}\})$, $v_2 = 1$.
In these cases, the second factor is obviously linear.

Subclasses of distance-hereditary graphs for which the results were known

Cographs: no induced P_4 subgraph.
Ferrers–Young are the graphs derived from Ferrers–Young diagrams (bipartite graphs without induced $2K_2$).
Threshold graphs have no P_4, C_4 and $2K_2$ as induced subgraphs.



Klee and Stamps noted that for arbitrary vectors a and b holds

$$\mathrm{Det}\left(L(G) + a \cdot b^T\right) = \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right)\tau(G),$$

where $L(G)$ is the Laplacian matrix of G. Klee–Stamps methods give an explicit formula for a certain subclass of threshold graphs.

Application to forest counting

The *extension* of a graph G is a graph $\widetilde{G}$ obtained by adding a vertex v_0 to G, connected to all vertices. It is well-known that $P_{\widetilde{G}}(x_0, x_1, x_2, \ldots, x_n)$ enumerates the spanning forests of G.

These polynomials are related as follows

$$P_G(x_1, x_2, \ldots, x_n) \cdot (x_1 + x_2 + \cdots + x_n) = P_{\widetilde{G}}(0, x_1, x_2, \ldots, x_n).$$

Theorem (C.–Prozorov, 2024+)

The extended enumerator of a graph G factors into linear terms if and only if G is a cograph.

An open Question

Ehrenborg's conjecture says that for a bipartite graph G with parts V_1 and V_2 holds

$$\tau(G) \leq \frac{\prod_{v \in V(G)} d_G(v)}{|V_1| \cdot |V_2|}.$$

The equality is achieved for Ferrers–Young graphs. This conjecture can be strengthened, with equality still being achieved by substituting any non-negative weights on the vertices of a Ferrers–Young graph.

Theorem (C.–Prozorov, 2024+)

Ehrenborg's conjecture is equivalent to the fact that for every substitution $x_1, \ldots, x_n \geq 0$, the inequality

$$P_G(x_1, \ldots, x_n) \leq \frac{\prod_{v \in V(G)} \sum_{u \in N_G(v)} x_u}{\left(\sum_{v \in V_1} x_v\right) \cdot \left(\sum_{v \in V_2} x_v\right)}$$

holds.

Thanks

THANK YOU VERY MUCH
FOR THE INVITATION AND
THE ATTENTION