Combinatorial proof of the Graham–Pollak Determinant formula for the distance matrix of a tree*

Emmanuel Briand, Luis Esquivias Quintero, Adrián Lillo Pinto, Mercedes Rosas
Universidad de Sevilla

Álvaro Gutiérrez Cáceres
Universität Bonn

May 30, 2023

Abstract

Graham and Pollak proved in 1971 that the determinant of the distance matrix of any tree with n vertices is $(-1)^{m-1}(n-1)2^{n-2}$. We provide a combinatorial interpretation for this formula, and, at the same time, the first combinatorial proof for this formula. Our approach is based on the Gessel–Viennot–Lindström Lemma. Our methods suits as well for many parametric deformations of the Graham-Pollak Formula. The special case when the tree is a path graph allows to recover results on statistics on derangements.

1 Introduction

Graham and Pollak [7] proved in 1971 the following surprising theorem: the determinant of the distance matrix of any tree with n vertices is independent on its structure, and always equal to

$$(-1)^{n-1}(n-1)2^{n-2}.$$

This result is not difficult to prove using row and column operations, and induction. Yet, such a formula strongly suggests that this determinant *counts* something. But what? We answer this question (raised for instance in [9, 12]), and also provide a *combinatorial proof* of the formula.

Evaluating the Graham-Pollak determinant is easily seen to be a *signed enumeration problem*, i.e. it consists in simplifying a sum:

$$\sum_{x\in X} \varepsilon(x)$$

where x are the elements of some set X and ε is a sign function (taking values $+1$ and -1). Here the set X is the set of all pairs (σ, f) where σ is a permutation of the vertices of the tree; and f is a function assigning to each vertex v an edge $f(v)$ in the unique path from v to $\sigma(v)$. We call these pairs the *compatible pairs* for the tree T.

A standard strategy to deal with signed enumeration problems consists in exhibiting a *sign–reversing involution* on X, i.e. an involution τ of X such that (i) if x is not a fixed point of τ then

*Partially supported by Project PID2020-117843GB-I00.

its image has opposite sign and (ii) all fixed points x have the same sign. Then most terms in the signed sum cancel, and the sum simplifies into the cardinality of the set fixed points, up to the sign.

A further refinement of the sign–reversing involution strategy, based on the *Gessel-Viennot-Lindström Lemma* [6] (see also the exposition [1, Ch. 32 (Lattice paths and determinants)]), consists in interpreting each element of X as a *n–path* (a family of n–paths from n fixed sources to n fixed sinks) in a acyclic digraph G. Any such n–path matches the sources with the sinks and defines therefore a permutation of $\{1, 2 \ldots, n\}$, whose sign is the sign of the n–path. Then the signed sum of these n-paths is the number of intersection free n–paths. Indeed, there is a signed reversing involution that kills the n-paths that have an intersection.

2 Main results and sketch of the proofs

Let T be a tree with n vertices. Let X be the set of its compatible pairs (σ, f).

1. Define a *diagram on T* as a multiset D of n oriented edges of the tree T. To each compatible pair (σ, f) we associate the following diagram D: the multiset of the edges $f(v)$, where each $f(v)$ is oriented in the same direction as the path from v to $\sigma(v)$. This splits X into a disjoint union of classes $X(D)$, and therefore
$$\sum_{x \in X} \varepsilon(x) = \sum_{D} \sum_{x \in X(D)} \varepsilon(x)$$

2. Say that a diagram is *standard* if for each edge $x - y$ of the tree T, there is one oriented edge ($x \to y$ or $x \leftarrow y$) in D, except for one, for which both oriented edges ($x \to y$ *and* $x \leftarrow y$) appear. Note that there are $(n-1)2^{n-2}$ standard diagrams. Indeed, the tree T has $n-1$ edges; there are $n-1$ choices for the edge with double orientation, and 2^{n-2} choices of orientations for the remaining $n-2$ edges.

3. For D non-standard, we exhibit a sign-reversing involution on $X(D)$ with no fixed point. Its existence shows that:
$$\sum_{x \in X(D)} \varepsilon(x) = 0$$

4. For D standard, we build an acyclic digraph $G(D)$ with n sources and n sinks, such that the elements of $X(D)$ are in a sign-preserving bijection with the n–paths in $G(D)$. We show that there is only one intersection-free n–path in $G(D)$, and the corresponding permutation is a full cycle. Therefore,
$$\sum_{x \in X(D)} \varepsilon(x) = (-1)^{n-1}.$$

5. This finishes the combinatorial proof and the combinatorial interpretation of the formula: $(n-1)2^{n-2}$ counts the standard diagrams and $(-1)^{n-1}$ is the sign of a full cyclic permutation of $\{1, 2, \ldots, n\}$.

3 Parametric deformations

Several deformations of the Graham-Pollak Formula, obtained by introducing parameters, have been obtained [3–5, 10, 11].

For instance, when the distances are replaced with their q-analogues (each distance d is replaced with the corresponding "q–integer" $1+q+\cdots+q^{d-1}$, where q is a variable), the determinant of the distance matrix of any tree becomes, as shown in [10, corollary 2.3],

$$(-1)^{n-1}(n-1)(q+1)^{n-2}. \tag{1}$$

Another example comes from attaching to each edge e a variable x_e. In the determinant of the distance matrix, replace the distance between vertices v and w with $\sum x_e$, where the sum is over all edges in the path from v to w. Then the determinant of the distance matrix becomes, as shown in [3, Corollary 2.5],

$$(-1)^{n-1}2^{n-2}\sum_e x_e \prod_e x_e. \tag{2}$$

Our approach fits like a glove for the study of these deformations. Indeed, the Gessel–Viennot–Lindström Lemma applies to *weighted* acyclic digraphs, and serves to count *weighted* signed sums. We recover many of the deformations of the Graham-Pollak Formula found in the literature (and all of [3–5, 10, 11]), simply by equipping the graphs $G(D)$ with suitable weights.

4 Derangements

We apply the same strategy to problems of signed enumeration of *derangements* (permutations with no fixed point) . In [8] it is proved that:

$$\sum_\sigma \varepsilon(\sigma) q^{\mathrm{exc}(\sigma)} = (-1)^{n-1}(q+q^2+\cdots+q^{n-1}), \tag{3}$$

where the sum is over all derangements σ of $\{1,2,\ldots,n\}$, $exc(\sigma)$ is the number of excedances of σ (the number of indices i such that $\sigma(i)>i$) and $\varepsilon(\sigma)$ is the sign of the permutation σ.

The identity (3) was refined recently in [2, Formula (3)] as:

$$\sum_\sigma \varepsilon(\sigma)\left(\prod_{j\in \mathrm{RLM}(\sigma)} x_j\right)\left(\prod_{j\in \mathrm{EXC}(\sigma)} y_j\right) = (-1)^{n-1}\sum_{j=1}^{n-1} y_1\cdots y_j x_{j+1}\cdots x_n, \tag{4}$$

where $\mathrm{RLM}(\sigma)$ is the set of right-to-left minima of σ (the set of all indices i such that $\sigma(i)<\sigma(j)$ for all $j>i$) and $\mathrm{EXC}(\sigma)$ is the set of excedances of σ (the set of all indices i such that $\sigma(i)>i$).

Again, we turn the sign enumeration problems (raised by the left–hand side of the formulas) into problems of non-intersecting paths in weighted acyclic digraphs with n sources and n sinks. We find $(n-1)$ non-intersecting n–paths, that correspond to $n-1$ terms of the right–hand sides of (3) and (4). The proof is remarkably simple and visually appealing.

5 Further remarks

We observe that many proofs with sign–reversing involutions involve arbitrary choices (in the definition of the involution). Indeed, often the sign–reversing involution is not unique. This sometimes makes the description of the involution, and the subsequent proofs, cumbersome. Turning the problem into a non-intersecting n–path problem avoids this difficulty, by changing the problem of constructing an involution (one among many), into a problem of finding a set of non-intersecting n–paths (a well–defined set). This provides clearer proofs.

References

[1] Martin Aigner and Günter M. Ziegler, *Proofs from The Book*, sixth ed., Springer, Berlin, 2018.

[2] Per Alexandersson and Frether Getachew Kedebe, *An involution on derangements preserving excedances and right-to-left minima*, 2021, Arxiv 2105.08455 [math.CO].

[3] R. Bapat, S.J. Kirkland, and M. Neumann, *On distance matrices and laplacians*, Linear Algebra and its Applications **401** (2005), 193–209, Special Issue in honor of Graciano de Oliveira.

[4] R. B. Bapat, A. K. Lal, and Sukanta Pati, *A q-analogue of the distance matrix of a tree*, Linear Algebra Appl. **416** (2006), no. 2-3, 799–814.

[5] Projesh Nath Choudhury and Apoorva Khare, *Distance matrices of a tree: two more invariants, and in a unified framework*, 2019, Arxiv 1903.11566v4 [math.CO].

[6] Ira Gessel and Gérard Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. in Math. **58** (1985), no. 3, 300–321.

[7] R. L. Graham and H. O. Pollak, *On the addressing problem for loop switching*, Bell System Tech. J. **50** (1971), 2495–2519.

[8] Roberto Mantaci and Fanja Rakotondrajao, *Exceedingly deranging!*, Adv. in Appl. Math. **30** (2003), no. 1-2, 177–188, Proceedings of FPSAC 2001.

[9] Valerie Tillia, *Determinants of distance matrices of trees*, Master's thesis, Portland State University, 2010.

[10] Weigen Yan and Yeong-Nan Yeh, *The determinants of q-distance matrices of trees and two quantiles relating to permutations*, Adv. in Appl. Math. **39** (2007), no. 3, 311–321.

[11] Hui Zhou and Qi Ding, *The distance matrix of a tree with weights on its arcs*, Linear Algebra and its Applications **511** (2016), 365–377.

[12] Leqi Zhou, *Determinants of distance matrices of trees*, 2020, Euler circle papers, at `http://simonrs.com/eulercircle/pftb2020/leqi-determinants.pdf`.