On the Maximum Number of Spanning Trees in a Planar Graph With a Fixed Number of Edges: A Linear-Algebraic Connection

Alan Bu* Yuchong Pan†

November 14, 2023

Abstract

There has been a long series of works concerning the maximum number of spanning trees in a graph with different constraints on the graph and dependencies on graph parameters. We study the maximum number of spanning trees in a planar graph with a fixed number m of edges. In this paper, we prove that this quantity is at most $(\sqrt[3]{7})^m \simeq 1.913^m$, where the trivial upper bound is 2^m. Our proof is based on a novel linear-algebraic reduction to a determinant maximization problem concerning a special class of matrices. A nontrivial upper bound on the latter problem is obtained through an elementary linear-algebraic argument using the pigeonhole principle.

We conjecture that the maximum number of spanning trees in a planar graph with m edges is equal to the maximum determinant of an $m \times m$ matrix $[M|N]$ obtained by concatenating two matrices M and N, where each row of M and N has at most one 1 and at most one -1 with all other entries 0. We present partial results towards this conjecture. In particular, we study a quantity which we call the "excess" of a graph. We prove that the excess of any planar graph is 0, while the excess of any nonplanar graph is at least 18. This illustrates a dichotomy between planar and nonplanar graphs in terms of the excess, offering a new characterization of planarity from perspectives of linear algebra and spanning trees, which might be of independent interest. As a first step towards the conjecture, we prove that subdivisions of $K_{3,3}$ and K_5 underperform the best planar graph with the same number of edges in some linear-algebraic sense.

1 Introduction

Counting spanning trees in a graph is a central problem in the field of counting and sampling. The number of spanning trees has various connections and applications to other fields such as statistical physics, telecommunication networks and geometry. In statistical physics, [Big99] showed that the number of spanning trees in a graph counts stable and recurrent chip configurations in chip-firing games on the graph, which can be reformulated into Abelian sandpile models. It is also closely related to the partition functions of many models of ferromagnetism, including ice-type models, the Potts model and the Ising model [Wu77, Bax73, Vis17]. In the study of network reliability, it is well known that the reliability measure of a network can be approximated using the number of spanning trees [Kel67], while the exact computation is NP-hard [PB83]. When restricted to planar graphs, the quantity was used by [BS10] to study realization spaces by bounding the size of the grid embedding of a 3-dimensional polytope. Upper bounds on the number of spanning trees often help us understand respective models in these applications and corresponding algorithmic questions.

*Phillips Exeter Academy, Exeter, New Hampshire, U.S.A., abu@exeter.edu.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A., yuchong@mit.edu.

Connections between matrix representations of a graph and intrinsic properties of the graph have long been a source of study. One of the earliest and most celebrated results in counting spanning trees is Kirchhoff's matrix-tree theorem, which says that the number of spanning trees in an undirected graph is equal to the determinant of any cofactor of the graph Laplacian [Kir47]. Kirchhoff's theorem was later generalized by Tutte to count the number of arborescences in a directed graph [Tut48]. The whole field of spectral graph theory studies relations between the eigenvalues of corresponding matrices and graph properties such as bipartiteness. Furthermore, in algorithmic counting and sampling, a common technique for lower or upper bounding the mixing time of a Markov chain defined on some graph is to establish respective bounds on the spectral gap [LP17]. These examples show that matrices are highly effective tools for analyzing graphs.

In this paper, we introduce a novel linear-algebraic method for obtaining an upper bound on the number of spanning trees in a planar graph with m edges. This question was asked by user Adam Lowrance on the Mathematics Stack Exchange online forum in 2018, and user JimT commented in 2021 that "this is a very, very tough question" and conjectured that this quantity is upper bounded by c^m for some constant $c < 2$, where 2^m is the trivial upper bound. In a follow-up comment, user JimT claimed that they proved this conjecture with $c \simeq 1.9328$, but we were not able to find their proof online or in the literature.[1] Our work improves this bound to $(\sqrt[3]{7})^m \simeq 1.913^m$ using a linear-algebraic reduction to a determinant maximization problem concerning a special class of matrices. An upper bound on the latter problem is obtained through an elementary linear-algebraic argument using the pigeonhole principle. In particular, given a planar graph, we provide two different constructions of a square matrix $[M|N]$, whose determinant counts spanning trees, obtained by concatenating two matrices M and N, where each row of M and N has at most one 1 and at most one -1 with all other entries 0.

This problem was originally motivated from a question in linear algebra and optimization: how far can the concatenation $[M|N]$ of two incidence matrices M and N be from total unimodularity, i.e., how large can the determinant of a square submatrix of such a concatenation be? Recall that an incidence matrix is totally unimodular. It turns out that our linear-algebraic connection also offers a graph-theoretic approach to derive exponential lower bounds on this quantity, i.e., by constructing planar graphs with many spanning trees.

1.1 Our Contributions

Our work initiates the study of the maximum number of spanning trees in a planar graph with a fixed number of edges. To state our contributions, we first introduce several notations. Throughout this paper, we allow multiple edges between two vertices. Given a graph G, we denote by $\tau(G)$ the number of spanning trees in G. We say that a matrix is a *bi-incidence matrix* if it is the concatenation $[M|N]$ of two matrices M and N, where each row of M and N has at most one 1 and at most one -1 with all other entries 0. We call M and N the *left* and *right sides* of $[M|N]$, respectively. For all $m \in \mathbb{N}$, we denote by τ_m the maximum number of spanning trees in a planar graph with m edges, and by Δ_m the maximum determinant of an $m \times m$ bi-incidence matrix. Our main theorem is the following:

Theorem 1. *For sufficiently large $m \in \mathbb{N}$,*

$$1.791^m \leq \exp\left(\frac{2C}{\pi} \cdot \left(m - \Theta\left(\sqrt{m}\right)\right)\right) \leq \tau_m \leq \Delta_m \leq \left(\sqrt[3]{7}\right)^m \simeq 1.913^m,$$

[1] See `https://math.stackexchange.com/questions/2832917`.

where

$$C = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2} \simeq 0.916$$

is Catalan's constant. In particular, the third and fourth inequalities hold for all $m \in \mathbb{N}$*.*

Along the way, we develop several techniques that might be of independent interest for future research. First, we give two alternative constructions of an $m \times m$ bi-incidence matrix M from a connected planar graph G with m edges such that $|\det(M)| = \tau(G)$. One uses duality in planar graphs (Theorem 22), while the other uses the notion of the cycle space of a graph and Mac Lane's planarity criterion (Theorem 24). We note that these two alternative constructions are fundamentally different, as the former can be generalized to a construction whose determinant has absolute value equal to $\tau_0(G)$, i.e., the number of maximal acyclic subgraphs when G is not connected, while the latter gives a construction with determinant 0 which is equal to $\tau(G)$ in the disconnected case. The two constructions both imply the third inequality in Theorem 1, i.e., $\tau_m \leq \Delta_m$ for all $m \in \mathbb{N}$.

A second technique that is very powerful for the analysis of bi-incidence matrices is what we call the merge-cut lemma, which might be of independent interest. It is motivated by the following question. In both of the two constructions above, the left side of the constructed $m \times m$ bi-incidence matrix is a *truncated incidence matrix* of G, i.e., the matrix obtained by removing an arbitrary column from an incidence matrix of G. Fixing *any* graph G with n vertices and m edges, how large can the determinant of an $m \times m$ bi-incidence matrix be if we set the left side to be a *truncated incidence matrix* of G? We denote by $\max\det(G)$ the maximum of such a determinant. If $n - 1 > m$, then we define $\max\det(G) := 0$.

Given a graph $G = (V, E)$ and $e \in E$, we denote by G/e the graph obtained by contracting e in G, and by $G \setminus e$ the graph obtained by deleting e from G. The merge-cut lemma states the following:

Lemma 2 (merge-cut lemma)**.** *For any graph* $G = (V, E)$ *and for any* $e \in E$*, we have that* $\max\det(G) \leq \max\det(G/e) + \max\det(G \setminus e)$*.*

An inductive application of the merge-cut lemma shows that $\max\det(G)$ cannot exceed the number of spanning trees in G for any graph G. Therefore, the two above constructions for planar graphs are the best possible.

Theorem 3. *For any graph G, we have* $\max\det(G) \leq \tau(G)$*.*

Theorem 3 motivates us to study the relation between planarity and the difference between $\tau(G)$ and $\max\det(G)$. We call this difference the *excess* of G and denote it by $\varepsilon(G)$. Note that the two aforementioned constructions show that the excess of any planar graph is equal to 0. On the other hand, the merge-cut lemma allows us to derive a nontrivial constant lower bound on the excess of a nonplanar graph. This lower bound demonstrates a dichotomy between planar and nonplanar graphs in terms of their excesses, offering a new characterization of planarity from perspectives of linear algebra and spanning trees. It allows one to give a numerically verifiable certificate for planarity.

Theorem 4. *Let G be a graph. Then*

$$\varepsilon(G) \begin{cases} = 0, & \text{if } G \text{ is planar,} \\ \geq 18, & \text{otherwise.} \end{cases}$$

In other words, G is planar if and only if $\varepsilon(G) = 0$*.*

We remark that the lower bound of 18 on $\varepsilon(G)$ when G is nonplanar is sharp; it is achieved by $K_{3,3}$. Theorem 4 shows that it is not possible to construct an $m \times m$ bi-incidence matrix for nonplanar graphs analogous to the two constructions for planar graphs. In addition, we prove the following theorem, showing that planar graphs outperform subdivisions of $K_{3,3}$ and K_5 with the same number of edges, in terms of the $\max\det(\cdot)$ function.

Theorem 5. *If G is a subdivision of $K_{3,3}$ or K_5 with m edges, then* $\max\det(G) \leq \tau_m$.

The proof of Theorem 5 in the case of K_5 follows from the following two useful observations. First, by Wagner's theorem, modifying one edge in K_5 to coincide with another edge in K_5 results in a planar graph with the same number of edges. Second, by certain operations on the matrix attaining $\max\det(G)$, one can show that there must exist paths in the original graph created from subdividing edges in K_5 such that changing their endpoints to coincide with two vertices of the original K_5 does not decrease the determinant of the matrix. We call this proof technique the "edge relocation" method. It might be the case that the edge relocation method has further applications in extending Theorem 5 to arbitrary nonplanar graphs, or even in problems involving determinants and other graph properties.

Theorem 5, together with computations of τ_m and Δ_m for small values of m, motivates us to conjecture the following. Indeed, if one were able to generalize Theorem 5 to *any* nonplanar graph, then this conjecture would follow.

Conjecture 6. *For all $m \in \mathbb{N}$, we have $\tau_m = \Delta_m$.*

Finally, we remark that the fourth inequality in Theorem 1, i.e., $\Delta_m \leq (\sqrt[3]{7})^m$ for all $m \in \mathbb{N}$, is proven by an elementary linear-algebraic argument using the pigeonhole principle. Moreover, the second inequality in Theorem 1, i.e., $\tau_m \geq \exp(\frac{2C}{\pi} \cdot (m - \Theta(\sqrt{m})))$, follows directly from a result in [Tap21].

1.2 Related Works

In the literature, there has been a long series of works concerning the maximum number of spanning trees in a graph with different constraints on the graph and dependencies on graph parameters. Cayley's classical formula states that the number of spanning trees in the complete graph with n vertices is precisely n^{n-2}. In fact, this formula was originally discovered by Borchardt in 1860, and Cayley extended this formula in several directions by considering the degrees of the vertices in 1889. Kirchhoff's matrix-tree theorem also allows one to compute the number of spanning trees in terms of the eigenvalues of the Laplacian of the graph. Kelmans studied operations on graphs that increase the number of spanning trees, and later gave an upper bound on the maximum number of spanning trees in a graph with n vertices and m edges, namely $(2m/(n-1))^{n-1}/n$ [Kel76b, Kel76a, Kel96]. In addition, there have been several independent efforts to characterize graphs achieving the maximum number of spanning trees with different ranges of parameters n and m [KC74, Shi74, Che81, BLS91, Kel96, PBS98]. Among these results, [Che81] proved that a regular complete multipartite graph has the maximum number of spanning trees over all simple graphs with the same numbers of vertices and edges. Fixing the number n of vertices, the number m of edges and the maximum degree d, [Das07] showed that the number of spanning trees is at most $((2m-d-1)/(n-2))^{n-2}$ and that both star graphs and complete graphs achieve this bound. Moreover, [FXD$^+$16] and [LZD21] considered the problem of the maximum number of spanning trees with a fixed number of vertices and a fixed matching number.

For planar graphs, the maximum number of spanning trees was first studied by statistical physicists, who computed the asymptotic behavior of this quantity on various 2-dimensional (and higher dimensional) lattices, such as the square lattice, the triangular lattice and the honeycomb lattice [Wu77, SW00, CS06, TW11, ZLWZ11]. Extending the result on the square lattice, [Tap21] studied the number of spanning trees on general grid graphs (that are not necessarily rectangular), and proved lower and upper bounds in terms of the "area" of the graph. With a flavor similar to our results, [BS10] proved that a planar graph on n vertices has at most $O(5.2852^n)$ spanning trees, with several other upper bounds given certain constraints on connectivity, triangle-freeness and quadrilateral-freeness. As an application, they used their results to bound necessary grid sizes for embedding 3-dimensional polytopes. To the best of our knowledge, there have been no previous results on the maximum number of spanning trees in a planar graph with a fixed number of edges.

More broadly, there has been interest in the maximum numbers of various combinatorial objects. This includes stable matchings [Thu02, KOGW18], minimal dominating sets [FGPS08, CLL15], Latin transversals [Tar15, GL16] and non-crossing subgraphs [GNT00, HSS$^+$12, HdM15]. The results in this paper add one more color to this vibrant palette.

As mentioned in the prior subsection, we give a novel characterization of planar graphs in terms of linear algebra and spanning trees. In the literature, there has been a multitude of works studying various characterizations of planar graphs. Perhaps the most well known ones are Kuratowski's and Wagner's theorems, which essentially state that a graph is planar if and only if it does not have $K_{3,3}$ or K_5 as a minor [Kur30, Wag37]. The de Fraysseix–Rosenstiehl planarity criterion is based on properties of depth-first search trees [dFR85]. Colin de Verdière gave a characterization based on the maximum multiplicity of the second eigenvalue of certain Schrödinger operators defined by the graph [dV90]. Moreover, The Hanani-Tutte theorem gives a characterization using the number of crossings of independent edges in graph drawings [Cho34, Tut70].

Similar to ours, there are also algebraic characterizations of planar graphs. Whitney's planarity criterion is a matroid-theoretic one which states that a graph is planar if and only if its graphic matroid is cographic [Whi31]. Mac Lane's planarity criterion provides a characterization using the notion of the cycle space of a graph [ML36]. Schnyder's theorem characterizes planar graphs in terms of the order dimension of their incidence posets [Sch89]. However, the flavors of these planarity characterizations are quite distinct from ours, as our conceptually simple, linear-algebraic characterization allows one to give a certificate for a planar graph that can be easily verified by computing the determinant of a sparse matrix.

Planarity testing algorithms are closely related to characterizations of planar graphs, and can be categorized based on their methods. Among many known algorithms, we mention the first linear-time planarity testing algorithm by [HT74] using the path addition method. The vertex addition method was first used by [Lem67] and later improved by [BL76] who developed the PQ tree data structure. Recently, [BM06] gave an $O(n)$ time algorithm using edge additions, and this is one of the state-of-the-art algorithms for planarity testing, while the other is based on the de Fraysseix–Rosenstiehl planarity criterion [dFR85].

1.3 Future Directions

The main question that remains is Conjecture 6. We remark several potential approaches for proving Conjecture 6. First, as mentioned above, it might be possible to generalize Theorem 5 to show that any arbitrary nonplanar graph underperforms the best planar graph with the same number of edges in terms of the $\max\det(\cdot)$ function. The edge relocation method used in proving Theorem 5 might help with the generalization. Second, we note that the application of the merge-cut lemma in the proof of Theorem 4 is extremely loose. Many nonplanar graphs contain many copies of $K_{3,3}$ and K_5 as minors. This observation could potentially be used to strengthen Theorem 4, e.g., in terms of the crossing number. Third, there might exist an algorithmic proof of Conjecture

6. We imagine that it might be possible to devise a procedure which transforms a square matrix $[M|N]$ to some other square matrix $[M|N']$, where M and N are truncated incidence matrices of nonplanar graphs G and H, respectively, and where N' is a truncated incidence matrix of some graph H' which has fewer spanning trees than H while preserving the determinant. Repeatedly applying this procedure would imply Conjecture 6. Fourth, one might be able to prove the equality of the two quantities by closing the gap between the lower and upper bounds in Theorem 1.

In addition to Conjecture 6, our work raises two algorithmic questions. First, it is hopeful that the constructions from Theorems 22 and 24 inspire further research on exact or approximate algorithms for counting spanning trees in a planar graph that run faster than $O(n^{1.5})$, which is the running time of the algorithm by [LRT79] using the planar separator theorem. Furthermore, we hope that our linear-algebraic characterization of planar graphs stimulates further investigation into efficient algorithms for deciding and testing planarity of a graph.

Finally, we remark that, according to Theorem 4, the excess can be viewed as a measure of nonplanarity of a graph. There are many other measures of nonplanarity that have been extensively studied, such as the crossing number, the genus and the thickness. It would be interesting to compare the excess and these measures, and investigate their relations.

1.4 Organization of the Paper

In Section 2, we introduce notations and conventions that we use throughout the paper. We also review prior results from graph theory and linear algebra that we use in our proofs.

In Section 3, we prove the merge-cut lemma, with results regarding operations that preserve the bi-incidence property of a matrix and its determinant (when the matrix is square), which are also used in Sections 5 and 6.

In Section 4, we provide two constructions of an $m \times m$ bi-incidence matrix whose determinant has absolute value $\tau(G)$ given a planar graph G with n vertices and m edges. The first construction requires the graph to be connected, and the second merely requires $m - n + 1 \geq 0$. These two constructions each imply that $\tau_m \leq \Delta_m$ for all $m \in \mathbb{N}$. In addition, using the merge-cut lemma, we prove that $\max\det(G) \leq \tau(G)$ for any graph G. A direct proof of this inequality using the Cauchy-Binet formula is given in Appendix A.

In Section 5, we use the inequality $\tau_m \leq \Delta_m$ for all $m \in \mathbb{N}$ to derive an upper bound on τ_m. We also show that a lower bound on τ_m follows directly from a result in [Tap21]. An alternative proof of this lower bound using standard tools from spectral graph theory is given in Appendix B.

In Section 6, we extend our methods in Section 4 to study the excess of nonplanar graphs. In particular, we use the merge-cut lemma to prove a nontrivial constant lower bound on the excess of a nonplanar graph. Furthermore, we prove that matrices corresponding to subdivisions of $K_{3,3}$ and K_5 with m edges have determinants at most τ_m.

2 Preliminaries

Given an $m \times n$ matrix M, we denote by $M[i,j]$ the (i,j)-minor of M for all $i \in [m]$ and $j \in [n]$, and by $M_{S,T}$ the submatrix of M formed by rows with indices in S and columns with indices in T for all $S \subseteq [m]$ and $T \subseteq [n]$. Given an $m \times n$ matrix $M = (a_{i,j})$ and an $m \times k$ matrix $N = (b_{i,j})$, we define their *concatenation* to be an $m \times (n+k)$ matrix, which we denote by $[M|N] = (c_{i,j})$, where

$$c_{i,j} := \begin{cases} a_{i,j}, & \text{if } j \leq n, \\ b_{i,j-n}, & \text{otherwise.} \end{cases}$$

For all $m \in \mathbb{N}$ and for all $k \in \{0, \dots, m\}$, we denote by $\binom{[m]}{k}$ the collection of subsets of $[m]$ with cardinality k.

Given a graph G, we denote by G/e the graph obtained by contracting e in G, and by $G \setminus e$ the graph obtained by deleting e from G. Throughout this paper, we allow multiple edges between two vertices in a graph.

We define *subdivision* to be a graph operation that creates a new vertex w, takes an edge uv and replaces it with edges uw and vw. We say that a graph obtained from repeated subdivision operations on a graph G is a *subdivision* of G. We call the vertices created in subdivision operations *internal* vertices.

2.1 Incidence Matrices

Let $D = (V, A)$ be a directed graph. We define its *incidence matrix* to be an $A \times V$ matrix denoted by $\iota_D = (a_{e,v})$, where

$$a_{e,v} := \begin{cases} 1, & \text{if } e \text{ enters } v, \\ -1, & \text{if } e \text{ leaves } v, \\ 0, & \text{otherwise.} \end{cases}$$

We say that a matrix is an *incidence matrix* if it is the incidence matrix of some directed graph. Furthermore, we define a *truncated incidence matrix* of D to be a matrix obtained by removing a column from ι_D, and which we denote by $\tilde{\iota}_D$. We define an *incidence matrix* (respectively, a *truncated incidence matrix*) of an undirected graph to be an incidence matrix (respectively, a truncated incidence matrix) of an orientation of the graph. We say that a matrix is an *incidence submatrix* if each row has at most one 1 and at most one -1, with all other entries 0. The following two observations are easy to see:

Proposition 7. *For each incidence submatrix M, there exists a unique directed graph D such that $\tilde{\iota}_D$ is equal to M up to all-zero rows.*

Proposition 8. *For any directed graph D, we have* $\operatorname{rank}(\iota_D) = \operatorname{rank}(\tilde{\iota}_D)$.

Proposition 7 can be shown by appending to M a vector such that the sum of the columns of the resulting matrix is the all-zero vector.

We say that a matrix is a *bi-incidence matrix* if it is the concatenation $[M|N]$ of two incidence submatrices M and N, which we call the *left* and *right sides* of $[M|N]$, respectively. For all $m \in \mathbb{N}$, we denote by Δ_m the maximum determinant of an $m \times m$ bi-incidence matrix. Since flipping the signs of the entries in a row changes the sign of the determinant, the minimum determinant of an $m \times m$ bi-incidence matrix is equal to $-\Delta_m$. It follows that $\Delta_m \geq 0$.

Given a graph G, we denote by $\max\det(G)$ the maximum determinant of a concatenation $[\tilde{\iota}_D\,|M]$ over all incidence submatrices M such that $[\tilde{\iota}_D\,|M]$ is square, where D is a fixed orientation of G. Note that the specific choice of the orientation is irrelevant since we can flip the signs of entries in some of the rows, and that the specific choice of the truncated vertex in $\tilde{\iota}_D$ is also irrelevant by elementary column operations. Hence, we can fix an arbitrary orientation of G with an arbitrary truncated vertex in $\tilde{\iota}_D$.

2.2 Spanning Trees and Kirchhoff's Matrix-Tree Theorem

Given a graph $G = (V, E)$, we denote by $\tau(G)$ the number of spanning trees in G, and by $\tau_0(G)$ the number of maximal acyclic subgraphs in G. Then $\tau(G) = \tau_0(G)$ if G is connected, and $\tau(G) = 0$ if

G is not connected. Moreover, we denote by L_G the *Laplacian* of G, which is defined to be a $V \times V$ matrix $(\ell_{u,v})$, where

$$\ell_{u,v} = \begin{cases} \deg_G(v), & \text{if } u = v, \\ -1, & \text{if } uv \in E, \\ 0, & \text{otherwise.} \end{cases}$$

Given a graph G, we define the *excess* of a graph to be $\varepsilon(G) := \tau(G) - \max\det(G)$. Theorem 3 shows that $\varepsilon(G) \geq 0$ for any graph G, justifying the name "excess." For all $m \in \mathbb{N}$, we denote by τ_m the maximum number of spanning trees in a planar graph with m edges.

Kirchhoff's celebrated matrix-tree theorem offers an algebraic approach to compute the number of spanning trees in a graph:

Theorem 9 (Kirchhoff's matrix-tree theorem, 1847, [Kir47])**.** *For any graph $G = (V, E)$, we have $\tau(G) = \det(L_G[v, v])$ for any $v \in V$.*

One way to prove Kirchhoff's matrix-tree theorem is to use the following deletion-contraction relation for the number of spanning trees:

Proposition 10. *For any graph $G = (V, E)$ and for any $e \in E$, we have $\tau(G) = \tau(G/e) + \tau(G \setminus e)$.*

A second way to prove Kirchhoff's matrix-tree theorem is to use the Cauchy-Binet formula:

Theorem 11 (Cauchy-Binet formula)**.** *Let M and N be an $m \times n$ matrix and an $n \times m$ matrix, respectively. Then*

$$\det(MN) = \sum_{S \in \binom{[n]}{m}} \det\left(M_{[m],S}\right) \det\left(N_{S,[m]}\right).$$

A direct consequence of Kirchhoff's matrix-tree theorem is the following formula for computing the number of spanning trees in a graph based on the eigenvalues of the Laplacian:

Corollary 12 (folklore)**.** *Let G be a graph. Let $0 = \lambda_1 \leq \ldots \leq \lambda_n$ be the eigenvalues of L_G. Then*

$$\tau(G) = \frac{1}{n} \prod_{j=2}^{n} \lambda_j.$$

2.3 Planar Graphs

Given a planar directed graph D, we construct its *directed planar dual* D^* as follows. The vertices of D^* are the faces of a planar embedding of the underlying undirected graph of D (including the outer face). For each arc e in D, we introduce a new arc in D^* connecting the two vertices in D^* corresponding to the two faces in D that meet at e, whose orientation is obtained by "rotating" the orientation of e by 90° counterclockwise. We define the *planar dual* of an undirected graph to be the underlying undirected graph of the directed planar dual of an arbitrary orientation of the graph. Euler's polyhedral formula implies that the planar dual of a graph with n vertices and m edges has exactly $m - n + 2$ vertices.

Planar graphs are very well studied. Wagner's theorem is a characterization of planar graphs in terms of forbidden minors.

Theorem 13 (Wagner's theorem, 1937, [Wag37])**.** *A graph is planar if and only it does not contain $K_{3,3}$ or K_5 as a minor.*

Another widely used characterization of planar graphs is given by Mac Lane in terms of their cycle spaces. Given a graph $G = (V, E)$, we define its *cycle space*, denoted by $\mathcal{C}(G)$, to be the vector space over GF(2) generated by the characteristic vectors of cycles in G, and a 2-*basis* of G is a basis of $\mathcal{C}(G)$ such that, for all $e \in E$, at most two vectors in the basis have nonzero components corresponding to e. The dimension of the cycle space of a graph with n vertices, m edges and κ connected components is $m - n + \kappa$. Mac Lane's planarity criterion states the following:

Theorem 14 (Mac Lane's planarity criterion, 1936, [ML36])**.** *A graph is planar if and only if its cycle space has a* 2*-basis.*

To prove the existence of a 2-basis, one can simply take the collection of boundaries of the bounded faces of any planar embedding of the given graph. For necessity, Lefschetz [Lef65] gave a proof with a slightly different formulation, which implies Mac Lane's criterion by leaving any one of the cycles out:

Theorem 15 (Lefschetz's formulation of Mac Lane's planarity criterion, 1965, [Lef65])**.** *A graph is planar if and only if it has a set of cycles (that are not necessarily simple) covering each edge exactly twice, such that the only nontrivial relation among these cycles in $\mathcal{C}(G)$ is that their sum be zero.*

The following folklore theorem relates the numbers of spanning trees in a connected planar graph and its dual:

Theorem 16 (folklore)**.** *Let G be a connected planar graph and G^* its dual. Then $\tau(G) = \tau(G^*)$.*

3 Merge-Cut Lemma

In this section, we prove the merge-cut lemma (Lemma 2). To do so, we first prove results on operations that preserve the bi-incidence property of a matrix and its determinant (when the matrix is square). These operations are also used in subsequent sections to derive an upper bound on Δ_m and to study the excess of a nonplanar graph. We start with a list of operations that preserve the bi-incidence property of a matrix.

Lemma 17. *Let M be a bi-incidence matrix. The following operations on M result in a bi-incidence matrix:*

(i) deleting a column from M;

(ii) deleting a row from M;

*(iii) (*combination*) replacing two columns from the same side of M by their sum;*

(iv) swapping two columns from the same side of M;

(v) swapping two rows of M;

(vi) swapping the left and right sides of M;

*(vii) (*realignment*) replacing a column by -1 times the sum of all columns from the its side of M.*

In particular, the last four operations do not change the dimensions of M and preserve the absolute value of the determinant of M when M is square.

Proof. It suffices to prove for combination and realignment, as all other operations satisfy the desired properties by definition.

For combination, since each row of each side of M has at most one 1 and at most one -1, with all other entries 0, it follows that the operation does not yield elements outside $\{-1, 0, 1\}$, and that each 1 (respectively, -1) in the sum comes from a 1 (respectively, -1) in a replaced column.

For realignment, note that the operation can be interpreted as a series of elementary column operations (i.e., column additions and a scalar multiplication by -1), each of which preserve the absolute value of the determinant. To prove the bi-incidence property, we have the following three cases. Fix a component of the new column. If the component is 0, then we are done. If the component is 1, then no entry in the corresponding row of the same side is 1, and we are done. The third case is symmetric. $\square$

Next, we provide an operation that removes a row with all entries equal to 0 on one side of a square bi-incidence matrix, while preserving the bi-incidence property of the matrix and the absolute value of its determinant:

Lemma 18. *Let $[M|N]$ be a square bi-incidence matrix with left and right sides M and N, such that N has at least one column. Suppose that there exists a row r of $[M|N]$ that has all entries equal to 0 when restricted to the left side.*

- *If row r has all entries equal to 0 when restricted to the right side, then removing row r and any column on the right side of $[M|N]$ results in a square bi-incidence matrix P with $\det([M|N]) = 0 \leq |\det(P)|$.*
- *If row r has exactly one nonzero entry in column c, then removing row r and column c from $[M|N]$ results in a square bi-incidence matrix P with $|\det([M|N])| = |\det(P)|$.*
- *If row r has exactly two nonzero entries in columns c_1 and c_2, respectively, then removing row r and replacing columns c_1 and c_2 by $c_1 + c_2$ (i.e. combining c_1 and c_2) in $[M|N]$ results in a square bi-incidence matrix P with $|\det([M|N])| = |\det(P)|$.*

Proof. The first case is trivial. The second case follows from the expansion of the determinant along row r. For the third case, since the (r, c_1)-entry and the (r, c_2)-entry are exactly one 1 and one -1, adding column c_1 to column c_2 in $[M|N]$ results in a square matrix P_0 that has exactly one nonzero entry in row r, such that $|\det([M|N])| = |\det(P_0)|$. Now the lemma follows from the expansion of the determinant along row r. $\square$

Repeatedly applying Lemma 18 allows one to remove all rows with every entry equal to 0 on the same side of a square bi-incidence matrix, while preserving the bi-incidence property of the matrix and the absolute value of its determinant:

Corollary 19. *Let $[M|N]$ be a square bi-incidence matrix with left and right sides M and N. Suppose that there exists a set R of rows of $[M|N]$ that have all entries equal to 0 when restricted to the left side. Let M_0 be the matrix obtained by removing rows in R from M. Then either $\det([M|N]) = 0$, or there exists a matrix N', obtained by removing rows in R from N and a sequence of $|R|$ removal and combination operations on columns of N, such that $|\det([M|N])| \leq |\det([M_0|N'])|$.*

Proof. Suppose that $[M|N]$ is an $m \times m$ matrix and that N has ℓ columns. If $|R| \leq \ell$, then repeatedly applying Lemma 18 for $|R|$ times proves the lemma. Otherwise, repeatedly applying Lemma 18 for ℓ times results in a square bi-incidence matrix P with an all-zero row such that $|\det([M|N])| \leq |\det(P)| = 0$, so $\det([M|N]) = 0$. $\square$

By swapping the left and right parts in the the concatenation, which does not change the absolute value of the determinant, we obtain the following analogous corollary:

Corollary 20. *Let $[M|N]$ be a square bi-incidence matrix with left and right sides M and N. Suppose that there exists a set R of rows of $[M|N]$ that have all entries equal to 0 when restricted to the right side. Let N_0 be the matrix obtained by removing rows in R from N. Then either $\det([M|N]) = 0$, or there exists a matrix M', obtained by removing rows in R from M and a sequence of $|R|$ removal and combination operations on columns of M, such that $|\det([M|N])| \leq |\det([M'|N_0])|$.*

Now, we are ready to prove the merge-cut lemma:

Proof of Lemma 2. Let $G = (V, E)$ be a graph with m edges. Let $e \in E$. Fix an orientation D of G. Let $P = [\tilde{\iota}_D\,|M]$ attain $\max\det(G)$, i.e., $\max\det(G) = \det(P)$. Let $r = (r_1, \ldots, r_m) \in \mathbb{R}^m$ be the row of P corresponding to e. Define $r^L = (r_1^L, \ldots, r_m^L), r^R = (r_1^R, \ldots, r_m^R) \in \mathbb{R}^m$ by

$$r_j^L = \begin{cases} r_j, & \text{if } j \leq n-1, \\ 0, & \text{otherwise,} \end{cases}$$

$$r_j^R = \begin{cases} 0, & \text{if } j \leq n-1, \\ r_j, & \text{otherwise.} \end{cases}$$

Then $r = r^L + r^R$. Let P^L and P^R be the matrices obtained by replacing row r with r^L and with r^R, respectively. By multilinearity of determinants, $\det(P) = \det(P^L) + \det(P^R)$.

First, we show that $\det(P^R) \leq \max\det(G \setminus e)$. Without loss of generality, we assume that $\det(P^R) \neq 0$. Let L_0 be the matrix obtained by removing row r from $\tilde{\iota}_D$. By Corollary 19, there exists a matrix M', obtained by removing row r from M followed by combining two columns or removing one column, such that $|\det(P^R)| \leq |\det([L_0|M'])|$. It is easy to see that L_0 is a truncated incidence matrix of $G \setminus e$. Hence, $\det(P^R) \leq |\det(P^R)| \leq |\det([L_0|M'])| \leq \max\det(G \setminus e)$.

Second, we show that $\det(P^L) \leq \max\det(G/e)$. Without loss of generality, we assume that $\det(P^L) \neq 0$. Let M_0 be the matrix obtained by removing row r from M. By Corollary 20, there exists a matrix L', obtained by removing row r from $\tilde{\iota}_D$ followed by combining two columns or removing one column, such that $|\det(P^L)| \leq |\det([L'|M_0])|$. We have the following two cases:

Case 1: One column is removed from $\tilde{\iota}_D$ to obtain L'. Then the two endpoints of e correspond to the truncated column and the removed column, respectively. It is easy to see that L' is a truncated incidence matrix of G/e. Hence, $\det(P^L) \leq |\det(P^L)| \leq |\det([L'|M_0])| \leq \max\det(G/e)$.

Case 2: Two columns are combined in $\tilde{\iota}_D$ to obtain L'. Then the two endpoints of e correspond to the two combined columns, respectively. It is easy to see that L' is a truncated incidence matrix of G/e. Hence, $\det(P^L) \leq |\det(P^L)| \leq |\det([L'|M_0])| \leq \max\det(G/e)$.

This completes the proof. □

Applying the deletion-contraction relation for the number of spanning trees (Proposition 10), we obtain the following alternative form of the merge-cut lemma in terms of the excess of a graph:

Corollary 21. *For any graph $G = (V, E)$ and for any $e \in E$, we have $\varepsilon(G) \geq \varepsilon(G/e) + \varepsilon(G \setminus e)$.*

4 Two Proofs of $\tau_m \leq \Delta_m$

In this section, we prove that the maximum number of spanning trees in a planar graph with m edges is always upper bounded by the maximum determinant of an $m \times m$ bi-incidence matrix, i.e., $\tau_m \leq \Delta_m$ for all $m \in \mathbb{N}$. We give two proofs of this inequality. The first proof (Theorem 22) makes use of duality in planar graphs, while the second proof (Theorem 24) uses the notion of the cycle space of a graph and Mac Lane's planarity criterion. Indeed, Theorem 22 can be adapted to a construction whose determinant has absolute value $\tau_0(G)$, i.e., the number of maximal acyclic subgraphs in G, for *any* planar graph that is not necessarily connected (whose proof we omit for the sake of conciseness). In contrast, Theorem 24 gives a construction whose determinant has absolute value $\tau(G)$ even in the disconnected case. Since $\tau(G) \neq \tau_0(G)$ if G is not connected, these two proofs are fundamentally different. We hope that the use of the cycle space inspires future work.

4.1 Proof Using Planar Duality

We start with the following theorem which uses duality in planar graphs:

Theorem 22. *Let G be a connected planar graph. Let D be an orientation of G. Let D^* be the directed planar dual of D. Suppose that for each i, the i^{th} rows of $\tilde{\iota}_D$ and $\tilde{\iota}_{D^*}$, respectively, correspond to the same arc in D (and its dual arc in D^*). Then*

$$\det \left[\begin{array}{c|c} \tilde{\iota}_D & \tilde{\iota}_{D^*} \end{array} \right] \in \{\tau(G), -\tau(G)\}.$$

Proof. By Euler's polyhedral formula and by the assumption that G is connected, $[\tilde{\iota}_D \,|\, \tilde{\iota}_{D^*}]$ is an $m \times m$ matrix, where m is the number of edges in G. Since $\iota_D^\mathsf{T} \iota_D = L_G$, it follows that $\det(\tilde{\iota}_D^\mathsf{T}\, \tilde{\iota}_D)$ is a principal first minor of L_G, which is equal to $\tau(G)$ by Kirchhoff's matrix-tree theorem. Similarly, $\det(\tilde{\iota}_{D^*}^\mathsf{T}\, \tilde{\iota}_{D^*}) = \tau(G^*) = \tau(G)$ by Theorem 16, where G^* is the planar dual of G.

We show that $\tilde{\iota}_D^\mathsf{T}\, \tilde{\iota}_{D^*} = 0$. Let c be a column vector of $\tilde{\iota}_D$ and c' a column vector of $\tilde{\iota}_{D^*}$. Let v be the vertex in G corresponding to c and f the face of G corresponding to c'. If there is no edge e of G that is incident to both v and f, then $c^\mathsf{T} c' = 0$. Otherwise, since edges in G incident to f form a cycle, it follows that there are exactly two edges that are incident to both v and f. A casework on the orientations of these two edges together with the definition of the directed planar dual implies that $c^\mathsf{T} c' = 0$. This proves that $\tilde{\iota}_D^\mathsf{T}\, \tilde{\iota}_{D^*} = 0$.

Now, we have

$$\begin{aligned}\left(\det \left[\begin{array}{c|c} \tilde{\iota}_D & \tilde{\iota}_{D^*} \end{array} \right]\right)^2 &= \det\left(\left[\begin{array}{c|c} \tilde{\iota}_D & \tilde{\iota}_{D^*} \end{array} \right]^\mathsf{T} \left[\begin{array}{c|c} \tilde{\iota}_D & \tilde{\iota}_{D^*} \end{array} \right]\right) = \det \begin{bmatrix} \tilde{\iota}_D^\mathsf{T}\, \tilde{\iota}_D & 0 \\ 0 & \tilde{\iota}_{D^*}^\mathsf{T}\, \tilde{\iota}_{D^*} \end{bmatrix} \\ &= \det\left(\tilde{\iota}_D^\mathsf{T}\, \tilde{\iota}_D\right) \cdot \det\left(\tilde{\iota}_{D^*}^\mathsf{T}\, \tilde{\iota}_{D^*}\right) = \tau(G)^2.\end{aligned}$$

This completes the proof. $\square$

Note that flipping the signs of all entries in a row changes the sign of the determinant. Hence, Theorem 22 implies that $\tau_m \leq \Delta_m$ for all $m \in \mathbb{N}$, as $[\tilde{\iota}_D \,|\, \tilde{\iota}_{D^*}]$ is an $m \times m$ bi-incidence matrix for any graph G with m edges with orientation D, where D^* is the directed planar dual of D.

Corollary 23. *For all $m \in \mathbb{N}$, we have $\tau_m \leq \Delta_m$.*

4.2 Proof Using the Cycle Space and Mac Lane's Planarity Criterion

Alternatively, using the notion of the cycle space and Mac Lane's planarity criterion, we give another construction of a square bi-incidence matrix, whose determinant has absolute value $\tau(G)$, with left side being a truncated incidence matrix of a planar graph G and with right side being some "incidence-like" matrix M from its 2-basis.

Theorem 24. *Let $G = (V, E)$ be a planar graph, and suppose that $n := |V|$ and $m := |E|$ satisfy $m - n + 1 \geq 0$. Let D be an orientation of G. Let $\mathcal{C} = \{C_1, \ldots, C_{m-n+1}\} \subseteq 2^E$ be a 2-basis of the cycle space of G. For each $i \in [m - n + 1]$, fix a direction of flow along C_i. Let M be an $m \times (m - n + 1)$ matrix where* *(i) for $i \in [m]$, the i^{th} row corresponds to the same arc as the i^{th} row of $\tilde{\iota}_D$;*

(ii) for $i \in [m]$ and $j \in [m-n+1]$, the (i,j)-entry of M is 1 if the arc corresponding to the i^{th} row points along the fixed direction of flow along C_j, -1 if it points against this fixed direction, and 0 otherwise.

Then

$$\det \left[\begin{array}{c|c} \tilde{\iota}_D & M \end{array} \right] \in \{\tau(G), -\tau(G)\}.$$

We remark that Theorem 24 applies to *any* graph (that is not necessarily connected), as long as $m - n + 1 \geq 0$. This weakens the assumption of connectedness in Theorem 22. If we take the 2-basis to be the collection of boundaries of the bounded faces of a planar embedding of G and fix all directions of flow to be counterclockwise, then M is indeed an incidence submatrix. Therefore, Theorem 24 again implies Corollary 23.

To prove Theorem 24, we first prove the following lemmata about truncated incidence matrices.

Lemma 25. *Let $G = (V, E)$ be a graph with κ connected components. Let D be an orientation of G. Then $\operatorname{rank}(\tilde{\iota}_D) = |V| - \kappa$.*

Proof. First, suppose that G is connected. Since $\tilde{\iota}_D$ has exactly $|V|-1$ columns, $\operatorname{rank}(\tilde{\iota}_D) \leq |V|-1$. Let $u = (u_j : j \in V) \in \ker(\iota_D)$. Then $u_i - u_j = 0$ for all $ij \in E$. Since G is connected, $u = \alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}$. Hence, $\dim(\ker(\iota_D)) \leq 1$, so $\operatorname{rank}(\tilde{\iota}_D) = \operatorname{rank}(\iota_D) \geq |V| - 1$.

Now, let $D_1, \dots, D_\kappa$ be the connected components of D (in the undirected sense). Rearranging columns and rows gives that $\iota_D = \operatorname{diag}(\iota_{D_1}, \dots, \iota_{D_\kappa})$. Hence,

$$\operatorname{rank}(\tilde{\iota}_D) = \operatorname{rank}(\iota_D) = \sum_{j=1}^{\kappa} \operatorname{rank}\left(\iota_{D_j}\right) = \sum_{j=1}^{\kappa} (|V(D_j)| - 1) = \sum_{j=1}^{\kappa} |V(D_j)| - \kappa = |V| - \kappa.$$

This completes the proof. □

Lemma 26. *Let $G = (V, E)$ be a graph. Let D be an orientation of G. Then $\tau_0(G)$ is equal to the number of maximal subsets of linearly independent row vectors of $\tilde{\iota}_D$.*

Proof. Let $F \subseteq E$. It suffices to show that F is acyclic if and only if the row vectors corresponding to F are linearly independent. For necessity, let $F' \subseteq F$ be a cycle. Fix a direction along F'. For each $e \in F'$, set $\alpha_e := 1$ if the orientation of e in D is the same as the fixed direction of F', and $\alpha_e := -1$ otherwise. Then $\sum_{e \in F'} \alpha_e r_e = \mathbf{0}$, where r_e denotes the row vector corresponding to e.

For sufficiency, suppose that (V, F) is acyclic with κ connected components. Let M be the submatrix of $\tilde{\iota}_D$ formed by the row vectors corresponding to F. By Lemma 25, $|F| = |V| - \kappa = \operatorname{rank}(M)$. Hence, the row vectors corresponding to F are linearly independent. □

Lemma 27. *Let $G = (V, E)$ be a planar graph. Let $\mathcal{C} = \{C_1, \dots, C_{m-n+1}\} \subseteq 2^E$ be a 2-basis of the cycle space of G, where $n = |V|$ and $m = |E|$. Let M be an $m \times (m-n+1)$ matrix constructed as in Theorem 24. Let $F \subseteq E$. Then F corresponds to a maximal subset of linearly independent row vectors of M if and only if $|F| = m - n + 1$ and e belongs to a cycle in $(V, E \setminus (F \setminus \{e\}))$ for all $e \in F$.*

Proof. Without loss of generality, we assume that F does not contain an edge whose corresponding row vector in M is an all-zero vector. Let N be the matrix obtained by removing all-zero rows from M. Let D' be the unique directed graph such that $\tilde{\iota}_{D'} = N$. By Lefschetz's formulation of Mac Lane's planarity criterion, D' is connected. By Lemma 25, $\operatorname{rank}(N) = m - n + 1$. Let N' be the matrix formed by the row vectors of N corresponding to F.

For necessity, note that $|F| = \mathrm{rank}(N) = m - n + 1$. Since N' is a square incidence submatrix whose rows are linearly independent, it follows that $\det(N') \in \{1, -1\}$ by total unimodularity of incidence matrices. Let $e \in F$. By Cramer's rule and by the fact that $\det(N') \in \{1, -1\}$, χ_e is a linear combination of column vectors of N' with integer coefficients. This implies that e belongs to a cycle in $(V, E \setminus (F \setminus \{e\}))$.

For sufficiency, since e belongs to a cycle in $(V, E \setminus (F \setminus \{e\}))$ for all $e \in F$, it follows that χ_e is in the column space of N' for all $e \in F$. Since $|F| = m - n + 1$, it follows that N' is a square matrix with $\det(N') \neq 0$. This completes the proof. $\square$

Now, we are ready to prove Theorem 24.

Proof of Theorem 24. If $m - n + 1 = 0$, then $\tau(G) = 1$ and $(\det[\tilde{\iota}_D \,|M])^2 = (\det \tilde{\iota}_D)^2 = \det(\tilde{\iota}_D^{\mathsf{T}} \tilde{\iota}_D)$ is the determinant of a principal first minor of L_G, which is equal to $\tau(G) = 1$, so $\det[\tilde{\iota}_D \,|M] \in \{1, -1\} = \{\tau(G), -\tau(G)\}$. Hence, without loss of generality, we assume that $m - n + 1 \geq 1$.

First, we show that $\tilde{\iota}_D^{\mathsf{T}} M = 0$. Let c be a column vector of $\tilde{\iota}_D$ corresponding to vertex v in G. Let c' be a column vector of M corresponding to cycle C. Then $c^{\mathsf{T}} c'$ is equal to the amount of accumulation of the fixed flow of cycle C at vertex v, which is always equal to 0 for a cycle. This proves that $\tilde{\iota}_D^{\mathsf{T}} M = 0$.

Next, we show that $\det(M^{\mathsf{T}} M) = \det(\tilde{\iota}_D^{\mathsf{T}} \tilde{\iota}_D) = \tau(G)$. Let N be the matrix obtained by removing all-zero rows from M. Then $M^{\mathsf{T}} M = N^{\mathsf{T}} N$. Let D' be the unique directed graph such that $\tilde{\iota}_{D'} = N$. By Lefschetz's formulation of Mac Lane's planarity criterion, D' is connected. By Kirchhoff's matrix-tree theorem, $\det(N^{\mathsf{T}} N) = \tau(G') = \tau_0(G')$, where G' the underlying undirected graph of D'. By Lemma 26, $\tau_0(G')$ is equal to the number of maximal subsets of linearly independent row vectors of N. Hence, it suffices to exhibit a bijection between maximal subsets of linearly independent row vectors of N and spanning trees in G.

We define such a bijection as follows. For each $F \subseteq E$ corresponding to a subset of row vectors of N, we map F to $E \setminus F$. It suffices to show that F corresponds to a maximal subset of linearly independent row vectors of N if and only if $E \setminus F$ forms a spanning tree in G. For necessity, suppose that $E \setminus F$ does not form a spanning tree in G. Since D' is connected, Lemma 25 implies that $|F| = \mathrm{rank}(N) = m - n + 1$, so $|E \setminus F| = m - (m - n + 1) = n - 1$. Hence, $(V, E \setminus F)$ has at least two connected components. Let $e \in F$ be such that its two endpoints are in different connected components of $(V, E \setminus F)$. Then e does not belong to a cycle in $(V, (E \setminus F) \cup \{e\}) = (V, E \setminus (F \setminus \{e\}))$, a contradiction to Lemma 27. For sufficiency, suppose that $E \setminus F$ forms a spanning tree in G. Then $|E \setminus F| = n - 1$, which implies that $|F| = m - (n - 1) = m - n + 1$. For each $e = uv \in F$, there exists a path from u to v in $(V, E \setminus F)$, and adding e to this path forms a cycle containing e in $(V, (E \setminus F) \cup \{e\}) = (V, E \setminus (F \setminus \{e\}))$. This proves that the map is bijective by Lemma 27.

Now, we have

$$\left(\det \left[\ \tilde{\iota}_D \mid M\ \right]\right)^2 = \det\left(\left[\ \tilde{\iota}_D \mid M\ \right]^{\mathsf{T}} \left[\ \tilde{\iota}_D \mid M\ \right]\right) = \det \begin{bmatrix} \tilde{\iota}_D^{\mathsf{T}} \tilde{\iota}_D & 0 \\ 0 & M^{\mathsf{T}} M \end{bmatrix}$$
$$= \det\left(\tilde{\iota}_D^{\mathsf{T}} \tilde{\iota}_D\right) \cdot \det\left(M^{\mathsf{T}} M\right) = \tau(G)^2.$$

This completes the proof. $\square$

4.3 Proof of Theorem 3

Theorems 22 and 24 give two ways to construct a square bi-incidence matrix from a connected planar graph, whose determinant is equal to the number of spanning trees in the graph up to the sign. This naturally leads to the following question: Is it possible to concatenate an incidence submatrix to a fixed truncated incidence matrix of a planar graph G that attains a larger determinant than that from Theorem 22? Recall we defined in Section 2 that $\max\det(G)$ denotes the maximum determinant of a square matrix of the form $[\tilde{\imath}_D\,|M]$ over all incidence submatrices M such that $[\tilde{\imath}_D\,|M]$ is a square matrix, where D is a fixed orientation of G. Using the merge-cut lemma, we prove Theorem 3, i.e., $\max\det(G) \leq \tau(G)$ for any graph G.

Proof of Theorem 3. We proceed by induction on the number of edges in G. The base case is trivial. For the induction step,

$$\max\det(G) \leq \max\det(G/e) + \max\det(G \setminus e) \leq \tau(G/e) + \tau(G \setminus e) = \tau(G),$$

where the first inequality follows from the merge-cut lemma, the second inequality follows from the inductive hypothesis as both G/e and $G \setminus e$ have one edge fewer than G, and the equality follows from the deletion-contraction relation for the number of spanning trees (Proposition 10). This completes the proof. □

A direct, alternative proof of Theorem 3 using the Cauchy-Binet formula is given in Appendix A. It turns out that the constructions from Theorems 22 and 24 give the best possible determinant.

Corollary 28. *For any planar graph G, we have* $\max\det(G) = \tau(G)$.

5 Asymptotic Bounds

This section completes the proof of Theorem 1 by establishing the second and fourth inequalities in the statement of Theorem 1, i.e., $\Delta_m \leq (\sqrt[3]{7})^m$ for all $m \in \mathbb{N}$ and $\tau_m \geq \exp(\frac{2C}{\pi} \cdot (m - \Theta(\sqrt{m})))$, where $C = \sum_{k=0}^{\infty}(-1)^k/(2k+1)^2 \simeq 0.916$ is Catalan's constant. The proof of the former follows from an elementary linear-algebraic argument using the pigeonhole principle, and the proof of the latter is a direct consequence of a result in [Tap21].

5.1 Upper Bounds

We start with a trivial upper bound on Δ_m:

Proposition 29. *For all $m \in \mathbb{N}$, we have $\Delta_m \leq 2^m$.*

Proof. Let $P = [M|N]$ be an $m \times m$ bi-incidence matrix that attains Δ_m, i.e., $\Delta_m = \det(P)$. Since the identity matrix is a bi-incidence matrix, it follows that $\det(P) = \Delta_m > 0$. By Corollary 19, $\det(P) \leq |\det([M_0|N'])|$, where M_0 is the matrix obtained by removing every all-zero row from M, and where N' is some matrix such that $[M_0|N']$ is a square bi-incidence matrix. By Proposition 7, there exists a directed graph D such that $\tilde{\imath}_D = M_0$. Let G be the underlying undirected graph of D. Then G has at most m edges. By Theorem 3, $\max\det(G) \leq \tau(G)$. Let n be the number of vertices in G. Since each spanning tree in G, if any, has exactly $n-1$ edges, a crude upper bound on $\tau(G)$ can be obtained by the number of subsets of E with cardinality $n-1$. Hence, $\tau(G) \leq \binom{m}{n-1} \leq \sum_{k=0}^{m}\binom{m}{k} = 2^m$, completing the proof. □

The analysis of this trivial upper bound is very crude and does not use sparsity of bi-incidence matrices. In what follows, we exploit this sparsity and obtain an upper bound on Δ_m that is exponentially stronger than the trivial upper bound.

Theorem 30. *For all $m \in \mathbb{N}$, we have $\Delta_m \leq (\sqrt[3]{7})^m \simeq 1.913^m$.*

Proof. We proceed by strong induction on m. The base cases $m = 1, 2, 3$ are easy to check. For the induction step, assume that $\Delta_j \le (\sqrt[3]{7})^j$ for all $j \in [m-1]$. Let $P = [M|N] = (a_{i,j})$ be an $m \times m$ bi-incidence matrix that attains Δ_m, i.e., $\Delta_m = \det(P)$. Let c and d be the sums of the columns in M and in N, respectively. Let $M' := [-c|M]$ and $N' := [-d|N]$. Then $P' := [M'|N']$ is a bi-incidence matrix with at most $4m$ nonzero entries and $m+2$ columns. By the pigeonhole principle, there exists a column c^* of P' with at most $\lfloor 4m/(m+2) \rfloor = 3$ nonzero entries. By realignment and by possibly interchanging M and N, we assume without loss of generality that c^* is a column in M.

Without loss of generality, we assume that the nonzero entries of c^* lie in the first three rows of P. Let k and ℓ be the numbers of columns in M and in N, respectively. For all $i \in [3]$, let r_i be the i^{th} row of P, and define $r_i^0 = ((r_i^0)_1, \ldots, (r_i^0)_m), r_i^1 = ((r_i^1)_1, \ldots, (r_i^1)_m) \in \mathbb{R}^m$ by

$$\left(r_i^0\right)_j := \begin{cases} a_{i,j}, & \text{if } j \le k, \\ 0, & \text{otherwise,} \end{cases}$$

$$\left(r_i^1\right)_j := \begin{cases} 0, & \text{if } j \le k, \\ a_{i,j}, & \text{otherwise.} \end{cases}$$

Then $r_i = r_i^0 + r_i^1$ for all $i \in [m]$. Let $P_{\alpha,\beta,\gamma}$ be the matrix formed by rows $r_1^\alpha, r_2^\beta, r_3^\gamma, r_4, \ldots, r_m$. By multilinearity of determinants,

$$\det(P) = \sum_{\alpha,\beta,\gamma \in \{0,1\}} \det\left(P_{\alpha,\beta,\gamma}\right).$$

If $\alpha = \beta = \gamma = 1$, then column c^* of $P_{\alpha,\beta,\gamma}$ is all-zero, so $\det(P_{\alpha,\beta,\gamma}) = 0$. If $(\alpha, \beta, \gamma) \in \{0,1\}^3 \setminus \{(1,1,1)\}$, we successively apply Corollaries 19 or 20 to the first three rows of $P_{\alpha,\beta,\gamma}$ to obtain an $(m-3) \times (m-3)$ bi-incidence matrix, which has determinant at most $(\sqrt[3]{7})^{m-3}$ by the inductive hypothesis. Hence,

$$\det([M|N]) = \sum_{(\alpha,\beta,\gamma) \in \{0,1\}^3 \setminus \{(1,1,1)\}} \det\left(P_{\alpha,\beta,\gamma}\right) \le 7 \cdot \left(\sqrt[3]{7}\right)^{m-3} = \left(\sqrt[3]{7}\right)^m.$$

This completes the proof. $\square$

5.2 Lower Bound

In this subsection, we prove that $\tau_m \ge \exp(\frac{2C}{\pi} \cdot (m - \Theta(\sqrt{m})))$ by constructing a planar graph with m edges that achieves this lower bound.

We use a result of [Tap21] on the number of spanning trees in a grid graph. For simplicity, we state the result in the special case of a $k \times \ell$ grid graph for $k, \ell \in \mathbb{N}$ to avoid introducing unnecessary definitions. By a $k \times \ell$ *grid graph*, we mean the graph $P_k \times P_\ell$, where we denote by P_j the path graph with j vertices for all $j \in \mathbb{N}$, and by $\times$ the Cartesian product of two graphs.

Theorem 31 (Tapp, 2021, [Tap21])**.** *For all $k, \ell \in \mathbb{N}$, the number of spanning trees in the $k \times \ell$ grid graph is at least* $\exp(\frac{4C}{\pi} \cdot (k-1)(\ell-1))$.

Now, we are ready to prove our lower bound:

Theorem 32. $\tau_m \ge \exp(\frac{2C}{\pi} \cdot (m - \Theta(\sqrt{m})))$.

Proof. Let $m \in \mathbb{N}, m \geq 4$. Let ℓ be the largest positive integer such that $2\ell^2 + 2\ell \leq m$, i.e., $\ell = \lfloor \sqrt{m/2 + 1/4} - 1/2 \rfloor$. Let G be the $(\ell + 1) \times (\ell + 1)$ grid graph. By Theorem 31,

$$\begin{aligned}
\tau_m \geq \tau(G) &\geq \exp\left(\frac{4C}{\pi} \cdot \ell^2\right) = \exp\left(\frac{4C}{\pi} \cdot \left\lfloor \sqrt{\frac{m}{2} + \frac{1}{4}} - \frac{1}{2} \right\rfloor^2\right) \geq \exp\left(\frac{4C}{\pi} \cdot \left(\sqrt{\frac{m}{2} + \frac{1}{4}} - \frac{3}{2}\right)^2\right) \\
&= \exp\left(\frac{4C}{\pi} \cdot \left(\frac{m}{2} + \frac{1}{4} + \frac{9}{4} - 3\sqrt{\frac{m}{2} + \frac{1}{4}}\right)\right) = \exp\left(\frac{4C}{\pi} \cdot \left(\frac{m}{2} - \Theta\left(\sqrt{m}\right)\right)\right) \\
&= \exp\left(\frac{2C}{\pi} \cdot \left(m - \Theta\left(\sqrt{m}\right)\right)\right).
\end{aligned}$$

This completes the proof. □

Alternatively, using standard tools from spectral graph theory, one can prove the following proposition, which also implies that $\tau_m \geq 1.791^m$ for sufficiently large $m \in \mathbb{N}$, albeit with a slightly weaker statement compared to Theorem 32. For completeness, we include its proof in Appendix B.

Proposition 33. $\lim_{n\to\infty} \ln(\tau(P_n \times P_n))/|E(P_n \times P_n)| = 2C/\pi$.

6 Excess of a Nonplanar Graph

In this section, we present some partial results towards Conjecture 6. In particular, we extend our linear-algebraic methods from Section 4 to study the $\max\det(\cdot)$ function and the excess of a nonplanar graph. First, the merge-cut lemma gives us a powerful tool for analyzing graph excess that allows us to show as a direct corollary that $\varepsilon(G) \geq 18$ for any nonplanar graph G.

Theorem 34. *For any nonplanar graph G, we have $\varepsilon(G) \geq 18$.*

Proof. It can be computed that $\varepsilon(K_{3,3}) = 18$ and that $\varepsilon(K_5) = 25$. (Note that it is not feasible to use the naïve brute-force algorithm to compute the $\max\det(\cdot)$ function in reasonable time. Our proof is computer-assisted and requires several propositions to reduce the running time. We defer the details to the full version of this paper.) By Theorem 3, $\varepsilon(G) \geq 0$ for any graph G. Hence, Corollary 21 implies that

$$\varepsilon(G) \geq \min\{\varepsilon(G/e), \varepsilon(G \setminus e)\}. \tag{1}$$

By Wagner's theorem, one can obtain either $K_{3,3}$ or K_5 by a sequence of edge contractions and deletions. Hence, applying (1) inductively gives that $\varepsilon(G) \geq \min\{\varepsilon(K_{3,3}), \varepsilon(K_5)\} = 18$. This completes the proof. □

Recall that Corollary 28 says that $\varepsilon(G) = 0$ for any planar graph G. Combining Corollary 28 and Theorem 34 illustrates a dichotomy between planar and nonplanar graphs in terms of the excess, offering a new characterization of planarity in terms of linear algebra and spanning trees. This proves Theorem 4.

We remark that the application of the merge-cut lemma in the proof of Theorem 34 is extremely loose in the sense that it only uses *one* minor of the nonplanar graph that is $K_{3,3}$ or K_5. Many nonplanar graphs, however, have many copies of $K_{3,3}$ or K_5 as minors. It is hopeful that this observation can be exploited to strengthen Theorem 34, e.g., in terms of the crossing number.

Furthermore, note that $\max\det(K_{3,3}) = 63$, $\tau(K_{3,3}) = 81$, $\max\det(K_5) = 100$ and $\tau(K_5) = 125$. The values of $\max\det(\cdot)$ and $\tau(\cdot)$ have a mysteriously large greatest common divisor in these two cases. Formalizing this observation for a general nonplanar graph would be interesting on its own and might illustrate deeper connections between determinants and spanning trees.

Theorem 34 also shows that one cannot find a construction for nonplanar graphs that is similar to the ones given in Theorems 22 and 24. However, it does not rule out the possibility that a nonplanar graph has a large number of spanning trees with a positive but small excess, resulting in a larger determinant than the maximum determinant from planar graphs. We proceed to rule out this possibility for subdivisions of $K_{3,3}$ and K_5. Together, the following lemmata prove Theorem 5.

Lemma 35. *If G is a subdivision of $K_{3,3}$ with m edges, then* $\max\det(G) \leq \tau_m$.

Proof. Let D be an orientation of G such that any internal vertex of G has one incoming edge and one outgoing edge. Then each column in $\tilde{\iota}_D$ corresponding to an internal vertex has exactly one 1 and one -1, with all other entries 0. Let $P = [\tilde{\iota}_D \,|M]$ attain $\max\det(G)$, i.e., $\max\det(G) = \det(P)$.

Fix an edge e of $K_{3,3}$. Let $V_e \subseteq V(G)$ be the set of internal vertices created from subdividing e. Let $S_e \subseteq E(G)$ be the set of edges created from subdividing e. We successively perform the following procedure until V_e is empty:

- pick an internal vertex $v \in V_e$, whose corresponding column has exactly two nonzero entries 1 and -1 in rows corresponding to two incident edges e_1 and e_2, respectively;
- adding row e_1 to row e_2 in P, so the (e_2, v)-entry becomes 0;
- now column v has exactly one nonzero entry 1, so we expand the determinant of P along that column, i.e., we remove column v and row e_1;
- remove v from V_e, and replace P with the submatrix from the determinant expansion.

This results in a 9×9 matrix $P_0 = [L_0|M_0]$ (that is not necessarily a bi-incidence matrix) such that L_0 is a truncated incidence matrix of $K_{3,3}$ and that $\det(P_0) = \det(P)$. By our choice of D, rows in P corresponding to edges in S_e are replaced with a single row corresponding to e, whose restriction to the right side is the sum of the rows in M corresponding to edges in S_e. This new row can be interpreted as a convex combination after scaling. By multilinearity of determinants, convexity and the maximality of M, we can assume without loss of generality that all rows in M corresponding to S_e are identical.

We apply this argument to every edge in $K_{3,3}$. Then M has at most 9 distinct rows, plus all-zero rows. Without loss of generality, we assume that $\det(P) > 0$. Let M' be the matrix obtained by removing every all-zero row from M. By Corollary 20, there exists a matrix L' such that $P' := [L'|M']$ is a square bi-incidence matrix such that $\det(P') \geq \det(P)$. By Proposition 7, there exists a directed graph D' such that $\tilde{\iota}_{D'} = M'$. Let G' be the underlying undirected graph of D'. Then G' has $m + 2 - (m - 3) = 5$ vertices and at most m edges, with at most 9 *distinct* edges. By Wagner's theorem, G' is planar. Hence,

$$\max\det(G) = \det(P) \leq \det\left(\left[L' \mid \tilde{\iota}_{D'}\right]\right) = \left|\det\left(\left[\tilde{\iota}_{D'} \mid L'\right]\right)\right| \leq \max\det\left(G'\right) \leq \tau\left(G'\right) \leq \tau_m.$$

This completes the proof. □

Lemma 36. *If G is a subdivision of K_5 with m edges, then* $\max\det(G) \leq \tau_m$.

Proof. Let us apply the same argument as in the proof of Lemma 35. We obtain a square bi-incidence matrix $P' = [L'|M']$ that attains $\max\det(G)$, as well as a graph G' with $m+2-(m-5) = 7$ vertices and m edges, among which there are at most 10 *distinct* edges, such that $M' = \tilde{\iota}_{D'}$ for some orientation D' of G'. We also obtain a 10×10 bi-incidence matrix $P_0 = [L_0|M_0]$ with $\det(P_0) = \det(P')$ such that L_0 is a truncated incidence matrix of K_5 (we omit the details for conciseness).

If G' is planar, then we are done. If G' is disconnected, then $\max\det(G) \le \tau(G') = 0 \le \tau_m$ and we are done.

Now, suppose that G' is nonplanar and connected. Let H be the underlying simple graph of G', which has 7 vertices and at most 10 edges. Since the sum of the degree of vertices in H is at most $2 \cdot 10 = 20$, there exists a vertex v of H with degree at most $\lfloor 20/7 \rfloor = 2$. Since G' is connected, so is H, implying that $\deg_H(v) \in \{1, 2\}$. We have the following two cases:

Case 1: $\deg_H(v) = 1$. Then column v in P_0 has exactly one nonzero entry. Expanding the determinant of P_0 along column v results in a 9×9 submatrix $P_1 = [L_1|M_1]$, where L_1 is a truncated incidence matrix of $K_5 \setminus e$, i.e., the graph obtained by deleting an arbitrary edge from K_5. Since row e is eliminated in this expansion, the endpoints of e do not affect the determinant, so we modify e to coincide with another edge e' in K_5, resulting in a planar graph K_5' and a new 10×10 matrix P_0' in place of P_0 with the same determinant. Moreover, we modify the endpoints of the path of edges created by subdividing e to coincide with those of edge e', obtaining a planar subdivision G'' of K_5' with m edges, while preserving the determinant. Hence, $\max\det(G) \le \max\det(G'') \le \tau(G'') \le \tau_m$.

Case 2: $\deg_H(v) = 2$. Then column v in P_0 has exactly two nonzero entries with values $s, t \in \mathbb{Z} \setminus \{0\}$, respectively. Let $r_1, r_2, \ldots, r_{10}$ be the rows of P_0. Let e_1 and e_2 be edges of K_5 whose rows in P_0 correspond to the two nonzero entries in column v, respectively. Without loss of generality, we assume that r_1 and r_2 correspond to e_1 and e_2, respectively. Then

$$\begin{aligned}\det(P_0) &= \det[r_1, r_2, r_3, \ldots, r_{10}] \\ &= -st \cdot \det\left[s^{-1}r_1, -t^{-1}r_2, r_3, \ldots, r_{10}\right] \\ &= -st \cdot \det\left[s^{-1}r_1, s^{-1}r_1 - t^{-1}r_2, r_3, \ldots, r_{10}\right], \end{aligned} \tag{2}$$

where we denote by $[u_1, \ldots, u_m]$ the matrix formed by rows $u_1, \ldots, u_m$. Let r_1' and r_2' be rows r_1 and r_2 restricted to the left side of P_0. Fix all entries of P_0 except the entries in r_1' and r_2'. By multilinearity of determinants and by expanding 2 along column v, it follows that $\det(P_0)$ is a linear function in $s^{-1}r_1' - t^{-1}r_2'$. This quantity can be viewed as a convex combination of r_1' and r_2' after scaling, so the maximum determinant of P_0, over all choices of r_1' and r_2' with at most one 1 and at most one -1, is achieved when $r_1' = r_2'$ or $r_1' = -r_2'$. In particular, this maximum occurs when r_1' and r_2' both correspond to the same edge e' of K_5.

Hence, we can change both e_1 and e_2 to coincide with edge e' in K_5 (with the orientation possibly reversed), resulting in a planar graph K_5' and a new 10×10 matrix P_0' in place of P_0 that attains the same determinant. Moreover, we modify the endpoints of the paths from subdividing e_1 and e_2 to coincide with the endpoints of the path from subdividing e' (with the orientation possibly reversed), obtaining from G a planar subdivision G'' of K_5' with m edges, in a manner that does not decrease the determinant of P'. Therefore, $\max\det(G) \le \max\det(G'') \le \tau(G'') \le \tau_m$. This completes the proof. □

Theorem 5 rules out the possibility for subdivisions of $K_{3,3}$ and K_5 to have larger $\max\det(\cdot)$ values than all planar graphs with the same number of edges. It would be interesting to generalize Theorem 5 to show that any arbitrary nonplanar graph underperforms the best planar graph with the same number of edges in this sense. By Corollary 19, such a generalization would prove Conjecture 6.

Intuitively, the proof of Lemma 36 rests on the following two observations. First, by Wagner's theorem, modifying one edge in K_5 to coincide with another edge in K_5 results in a planar graph with the same number of edges. Second, by certain operations on the matrix attaining $\max\det(G)$, one can show that there must exist paths in the original graph created from subdividing edges in

K_5 such that changing their endpoints to coincide with two vertices of the original K_5 does not decrease the determinant of the matrix. We call this proof technique the "edge relocation" method. A natural question is whether the edge relocation method can be applied to derive further results, e.g., to generalize Theorem 5 to any arbitrary nonplanar graph as mentioned in the prior paragraph.

Acknowledgments. This work was partially done while A.B. was participating in the Research Science Institute (RSI) program at MIT under the mentorship of Y.P. A.B. would like to thank RSI, the Center for Excellence in Education, MIT and all of the sponsors of RSI for making the mentorship possible. A.B. would also like to thank Tanya Khovanova for her invaluable advice, and thank Peter Gaydarov and Allen Lin for their helpful suggestions. A.B. and Y.P. would like to thank Michel Goemans for proposing the question of how far the concatenation of two incidence matrices is from total unimodularity, and for his insightful discussions.

References

[Bax73] R. J. Baxter. Potts model at the critical temperature. *Journal of Physics C: Solid State Physics*, 6(23):L445, 1973.

[Big99] N. L. Biggs. Chip-firing and the critical group of a graph. *Journal of Algebraic Combinatorics*, 9:25–45, 1999.

[BL76] K. S. Booth and G. S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using pq-tree algorithms. *Journal of computer and system sciences*, 13(3):335–379, 1976.

[BLS91] F. T. Boesch, X. Li, and C. Suffel. On the existence of uniformly optimally reliable networks. *Networks*, 21(2):181–194, 1991.

[BM06] J. M. Boyer and W. J. Myrvold. Simplified $O(n)$ planarity by edge addition. *Graph Algorithms and Applications*, 5:241, 2006.

[BS10] K. Buchin and A. Schulz. On the number of spanning trees a planar graph can have. In *Algorithms–ESA 2010: 18th Annual European Symposium, Liverpool, UK, September 6-8, 2010. Proceedings, Part I 18*, pages 110–121. Springer, 2010.

[Che81] C.-S. Cheng. Maximizing the total number of spanning trees in a graph: two related problems in graph theory and optimum design theory. *Journal of Combinatorial Theory, Series B*, 31(2):240–248, 1981.

[Cho34] C. Chojnacki. Über wesentlich unplättbare kurven im dreidimensionalen raume. *Fundamenta Mathematicae*, 23:135–142, 1934.

[Chu97] F. R. K. Chung. *Spectral graph theory.* American Mathematical Society, 1997.

[CLL15] J.-F. Couturier, R. Letourneur, and M. Liedloff. On the number of minimal dominating sets on some graph classes. *Theoretical Computer Science*, 562:634–642, 2015.

[CS06] S.-C. Chang and R. Shrock. Some exact results for spanning trees on lattices. *Journal of Physics A: Mathematical and General*, 39(20):5653, 2006.

[Das07] K. C. Das. A sharp upper bound for the number of spanning trees of a graph. *Graphs and Combinatorics*, 23(6):625–632, 2007.

[dFR85] H. de Fraysseix and P. Rosenstiehl. A characterization of planar graphs by Trémaux orders. *Combinatorica*, 5:127–135, 1985.

[dV90] Y. C. de Verdiere. On a new graph invariant and a planarity criterion. *Journal of Combinatorial Theory, Series B*, 50(1):11–21, 1990.

[FGPS08] F. V. Fomin, F. Grandoni, A. V. Pyatkin, and A. A. Stepanov. Combinatorial bounds via measure and conquer: bounding minimal dominating sets and applications. *ACM Transactions on Algorithms (TALG)*, 5(1):1–17, 2008.

[FXD+16] L. Feng, K. Xu, K. C. Das, A. Ilić, and G. Yu. The number of spanning trees of a graph with given matching number. *International Journal of Computer Mathematics*, 93(6):837–843, 2016.

[GL16] R. Glebov and Z. Luria. On the maximum number of Latin transversals. *Journal of Combinatorial Theory, Series A*, 141:136–146, 2016.

[GNT00] A. Garcıa, M. Noy, and J. Tejel. Lower bounds on the number of crossing-free subgraphs of K_N. *Computational Geometry*, 16(4):211–221, 2000.

[HdM15] C. Huemer and A. de Mier. Lower bounds on the maximum number of non-crossing acyclic graphs. *European Journal of Combinatorics*, 48:48–62, 2015.

[HSS+12] M. Hoffmann, A. Schulz, M. Sharir, A. Sheffer, C. D. Tóth, and E. Welzl. Counting plane graphs: flippability and its applications. In *Thirty Essays on Geometric Graph Theory*, pages 303–325. Springer, 2012.

[HT74] J. Hopcroft and R. Tarjan. Efficient planarity testing. *Journal of the ACM (JACM)*, 21(4):549–568, 1974.

[KC74] A. K. Kelmans and V. M. Chelnokov. A certain polynomial of a graph and graphs with an extremal number of trees. *Journal of Combinatorial Theory, Series B*, 16(3):197–214, 1974.

[Kel67] A. K. Kelmans. Connectivity of probabilistic networks. *Automation and Remote Control*, 3:98–116, 1967.

[Kel76a] A. K. Kelmans. Comparison of graphs by their number of spanning trees. *Discrete Mathematics*, 16(3):241–261, 1976.

[Kel76b] A. K. Kelmans. Operations on graphs that increase the number of their spanning trees. *Issledovanie po Discretnoy Optimizacii, Nauka, Moscow*, pages 406–424, 1976.

[Kel96] A. K. Kelmans. On graphs with the maximum number of spanning trees. *Random Structures & Algorithms*, 9(1-2):177–192, 1996.

[Kir47] G. Kirchhoff. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik*, 148(12):497–508, 1847.

[KOGW18] A. R. Karlin, S. Oveis Gharan, and R. Weber. A simply exponential upper bound on the maximum number of stable matchings. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 920–925, 2018.

[Kur30] C. Kuratowski. Sur le probleme des courbes gauches en topologie. *Fundamenta mathematicae*, 15(1):271–283, 1930.

[Lef65] S. Lefschetz. Planar graphs and related topics. *Proceedings of the National Academy of Sciences*, 54(6):1763–1765, 1965.

[Lem67] A. Lempel. An algorithm for planarity testing of graphs. In *Theory of Graphs: International Symposium.*, pages 215–232. Gorden and Breach, 1967.

[LP17] D. A. Levin and Y. Peres. *Markov chains and mixing times.* American Mathematical Society, 2017.

[LRT79] R. J. Lipton, D. J. Rose, and R. E. Tarjan. Generalized nested dissection. *SIAM journal on numerical analysis*, 16(2):346–358, 1979.

[LZD21] M. Liu, G. Zhang, and K. C. Das. The maximum number of spanning trees of a graph with given matching number. *Bulletin of the Malaysian Mathematical Sciences Society*, 44(6):3725–3732, 2021.

[ML36] S. Mac Lane. *A combinatorial condition for planar graphs.* Seminarium Matematik, 1936.

[PB83] J. S. Provan and M. O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM Journal on Computing*, 12(4):777–788, 1983.

[PBS98] L. Petingi, F. Boesch, and C. Suffel. On the characterization of graphs with maximum number of spanning trees. *Discrete mathematics*, 179(1-3):155–166, 1998.

[Sch89] W. Schnyder. Planar graphs and poset dimension. *Order*, 5:323–343, 1989.

[Shi74] D. R. Shier. Maximizing the number of spanning trees in a graph with n nodes and m edges. *Journal Research National Bureau of Standards, Section B*, 78(193-196):3, 1974.

[SW00] R. Shrock and F. Y. Wu. Spanning trees on graphs and lattices in d dimensions. *Journal of Physics A: Mathematical and General*, 33(21):3881, 2000.

[Tap21] K. Tapp. Spanning tree bounds for grid graphs. *arXiv preprint arXiv:2109.05987*, 2021.

[Tar15] A. A. Taranenko. Multidimensional permanents and an upper bound on the number of transversals in Latin squares. *Journal of Combinatorial Designs*, 23(7):305–320, 2015.

[Thu02] E. G. Thurber. Concerning the maximum number of stable matchings in the stable marriage problem. *Discrete Mathematics*, 248(1-3):195–219, 2002.

[Tut48] W. T. Tutte. The dissection of equilateral triangles into equilateral triangles. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 44, pages 463–482. Cambridge University Press, 1948.

[Tut70] W. T. Tutte. Toward a theory of crossing numbers. *Journal of Combinatorial Theory*, 8(1):45–53, 1970.

[TW11] E. Teufl and S. Wagner. Resistance scaling and the number of spanning trees in self-similar lattices. *Journal of Statistical Physics*, 142(4):879–897, 2011.

[Vis17] G. M. Viswanathan. Correspondence between spanning trees and the Ising model on a square lattice. *Physical Review E*, 95(6):062138, 2017.

[Wag37] K. Wagner. Über eine eigenschaft der ebenen komplexe. *Mathematische Annalen*, 114(1):570–590, 1937.

[Whi31] H. Whitney. Non-separable and planar graphs. *Proceedings of the National Academy of Sciences*, 17(2):125–127, 1931.

[Wu77] F. Y. Wu. Number of spanning trees on a lattice. *Journal of Physics A: Mathematical and General*, 10(6):L113, 1977.

[ZLWZ11] Z. Zhang, H. Liu, B. Wu, and T. Zou. Spanning trees in a fractal scale-free lattice. *Physical Review E*, 83(1):016116, 2011.

A Alternative Proof of Theorem 3

Proof of Theorem 3. Let D be an orientation of G. Let M be an incidence submatrix such that $A = (a_{i,j}) := [\tilde{\iota}_D \,|M]$ is a square matrix. Let $m := |E|$. Without loss of generality, we assume that $E = [m]$. For each $T \in \binom{[m]}{n-1}$, we denote by $\sigma(T, [m] \setminus T)$ the unique permutation in S_m such that $T = \{\sigma(1), \ldots, \sigma(n-1)\}$, that $\sigma(1) < \ldots < \sigma(n-1)$ and that $\sigma(n) < \ldots < \sigma(m)$. For each $T \in \binom{[m]}{n-1}$, $i \in [n-1]$ and $j \in [m-n+1]$, we denote by $T(i)$ the i^{th} smallest element in T, and by $\overline{T}(j)$ the j^{th} smallest element in $[m] \setminus T$. Then

$$\begin{aligned}
\det(A) &= \det\left(A^{\mathsf{T}}\right)\\
&= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} a_{\sigma(i),i}\\
&= \sum_{T \in \binom{[m]}{n-1}} \sum_{\sigma_1 \in S_T} \sum_{\sigma_2 \in S_{[m]\setminus T}} \operatorname{sgn}\left(\sigma_1 \circ \sigma_2 \circ \sigma(T,[m]\setminus T)\right) \prod_{i=1}^{n-1} a_{\sigma_1(T(i)),i} \prod_{i=n}^{m} a_{\sigma_2\left(\overline{T}(i-n+1)\right),i}\\
&= \sum_{T \in \binom{[m]}{n-1}} \operatorname{sgn}(\sigma(T,[m]\setminus T)) \left(\sum_{\sigma_1 \in S_T} \operatorname{sgn}(\sigma_1) \prod_{i=1}^{n-1} a_{\sigma_1(T(i)),i}\right) \left(\sum_{\sigma_2 \in S_{[m]\setminus T}} \operatorname{sgn}(\sigma_2) \prod_{i=n}^{m} a_{\sigma_2\left(\overline{T}(i-n+1)\right),i}\right)\\
&= \sum_{T \in \binom{[m]}{n-1}} \operatorname{sgn}(\sigma(T,[m]\setminus T)) \det\left(A_{T,[n-1]}\right) \det\left(A_{[m]\setminus T,\{n,\ldots,m\}}\right)\\
&\le \sum_{T \in \binom{[m]}{n-1}} \det\left(A_{T,[n-1]}\right)^2,
\end{aligned}$$

where the inequality above follows from the facts that $\operatorname{sgn}(\sigma) \in \{-1,0,1\}$ for all $\sigma \in S_m$ and that $\det(A_{T,[n-1]}), \det(A_{[m]\setminus T,\{n,\ldots,m\}}) \in \{-1,0,1\}$ for all $T \in \binom{[m]}{n-1}$ because of total unimodularity of incidence matrices.

Since $\iota_D^{\mathsf{T}}\iota_D = L_G$, it follows that $\det(\tilde{\iota}_D^{\mathsf{T}}\,\tilde{\iota}_D)$ is a principal first minor of L_G, which is equal to $\tau(G)$ by Kirchhoff's matrix-tree theorem. By the Cauchy-Binet formula,

$$\det\left(\tilde{\iota}_D^{\mathsf{T}}\,\tilde{\iota}_D\right) = \sum_{T\in\binom{[m]}{n-1}} \det\left(\left(\tilde{\iota}_D^{\mathsf{T}}\right)_{[n-1],T}\right)\det\left(\left(\tilde{\iota}_D\right)_{T,[n-1]}\right) = \sum_{T\in\binom{[m]}{n-1}} \det\left(A_{T,[n-1]}\right)^2 \geq \det(A).$$

This completes the proof. □

B Proof of Proposition 33

The following results are standard in spectral graph theory, whose proofs we omit for conciseness. We refer the interested reader to any standard textbook in spectral graph theory, e.g., [Chu97].

Lemma 37. *The eigenvalues of L_{P_n} are $2 - 2\cos(\pi j/n)$ for $j = 0, \ldots, n-1$.*

Lemma 38. *Let G and H be undirected graphs. If L_G has eigenvalues $\lambda_1, \ldots, \lambda_n$ and L_H has eigenvalues $\mu_1, \ldots, \mu_m$, then the eigenvalues of $L_{G\times H}$ are $\lambda_j + \mu_k$ for $j \in [n]$ and $k \in [m]$.*

Corollary 39. *The eigenvalues of $L_{P_n\times P_n}$ are $4-2\cos(\pi j/n)-2\cos(\pi k/n)$ for $j, k \in \{0, \ldots, n-1\}$.*

Now, we are ready to prove Proposition 33.

Proof of Proposition 33. By Corollaries 12 and 39,

$$\tau\left(P_n \times P_n\right) = \frac{1}{n^2}\prod_{(j,k)\in\{0,\ldots,n-1\}^2\setminus\{(0,0)\}}\left(4 - 2\cos\left(\frac{\pi j}{n}\right) - 2\cos\left(\frac{\pi k}{n}\right)\right).$$

Since $|E(P_n \times P_n)| = 2n^2 + O(n)$,

$$\begin{aligned}
&\lim_{n\to\infty}\frac{\ln\tau\left(P_n\times P_n\right)}{\left|E\left(P_n\times P_n\right)\right|}\\
&= \lim_{n\to\infty}\frac{1}{2n^2}\ln\left(\frac{1}{n^2}\prod_{(j,k)\in\{0,\ldots,n-1\}^2\setminus\{(0,0)\}}\left(4 - 2\cos\left(\frac{\pi j}{n}\right) - 2\cos\left(\frac{\pi k}{n}\right)\right)\right)\\
&= \lim_{n\to\infty}\frac{1}{2n^2}\left(-2\ln n + \sum_{(j,k)\in\{0,\ldots,n-1\}^2\setminus\{(0,0)\}}\ln\left(4 - 2\cos\left(\frac{\pi j}{n}\right) - 2\cos\left(\frac{\pi k}{n}\right)\right)\right)\\
&= \lim_{n\to\infty}\frac{1}{2n^2}\sum_{(j,k)\in\{0,\ldots,n-1\}^2\setminus\{(0,0)\}}\ln\left(4 - 2\cos\left(\frac{\pi j}{n}\right) - 2\cos\left(\frac{\pi k}{n}\right)\right)\\
&= \frac{1}{2}\int_0^1\int_0^1 \ln(4 - 2\cos(\pi x) - 2\cos(\pi y))\,dx\,dy\\
&= \frac{2C}{\pi},
\end{aligned}$$

where the last equality can be obtained by standard calculus techniques, completing the proof. □