Disjoint Paths and Matrix Determinants: A Survey on Lindström-Gessel-Viennot Theorem

Johan Cristobal

Abstract

In this survey, we state the Lindström-Gessel-Viennot theorem similar to its statement in Combinatorial Mathematics by Douglas B. West with some adjustment [Wes20]. The theorem connects the number of disjoint paths between the sets of vertices in a digraph, along with permutations on the pairing of vertices, with the determinant of a special matrix associated with the set of vertices. We prove the statement using the concepts introduced. Lastly, we explore consequences of this theorem.

1 Introduction

In 1973, Lindström published a paper giving a generalization of a previous result about digraphs and the induced matroids in these graphs [Lin73]. In doing so, Lindström set the ground for the main theorem of this survey when he built a special matrix with entries corresponding to a 'weight' of a collection of disjoint paths in a digraph [Lin73]. However, it was Gessel and Viennot who built on this ground work to obtain the identity in the main theorem [GV85]. Note that Lindström used lattice paths in his discussion, however we will show the result for finite acyclic digraphs.

We will begin by introducing background concepts necessary to understand the statement of the theorem as well as the proof.

2 Background

Definition 2.1 Given a finite set of numbers $[n] = \{1, ..., n\}$, a *permutation* on $[n]$ is a bijective function $\sigma : [n] \to [n]$.

Example 2.2 An example of a permutation σ on $[6]$ has the mappings: $1 \mapsto 3 \mapsto 5 \mapsto 1$, $2 \mapsto 4 \mapsto 2$, and 6.

Definition 2.3 An *involution* ϕ is a permutation such ϕ^2 is the identity mapping on $[n]$.

The special properties of permutations that are at the crux of this theorem are *inversions* and the *sign of a permutation*.

Definition 2.4 An *inversion* of a permutation σ on $[n]$ is a pair (i, j) in $[n] \times [n]$ where $i < j$ and $\sigma(i) > \sigma(j)$.

Definition 2.5 The *sign of a permutation* is $(\text{sign } \sigma) = (-1)^{N(\sigma)}$ where $N(\sigma)$ is the number of inversions in σ.

Lemma 2.6 Let σ be a permutation of $[n]$. If σ' is constructed by swapping the mappings of i and j, denoted as $\sigma' = (i\ j)\sigma$, then $(\text{sign } \sigma') = -1 \cdot (\text{sign } \sigma)$.

Proof. Let σ be a permutation of $[n]$ and $i < j$ be elements of $[n]$. Suppose the elements $\sigma(i)$ and $\sigma(j)$ are adjacent, meaning there does not exist $\sigma(k)$ such that $\sigma(k)$ is in between $\sigma(i)$ and $\sigma(j)$. If $\sigma(i) < \sigma(j)$, then swapping them would add one inversion. If $\sigma(i) > \sigma(j)$, then swapping them would remove an inversion. In either case, we have changed the number of inversions by 1. Thus, $(\text{sign } \sigma') = -1 \cdot (\text{sign } \sigma)$.

Suppose $\sigma(i)$ and $\sigma(j)$ are not adjacent. Therefore, there exists $S \subseteq [n]$ with $|S| = m \geq 1$ such that $\sigma(s)$ lies in between $\sigma(i)$ and $\sigma(j)$ for all $s \in S$. If $\sigma(i) < \sigma(j)$, we can perform m adjacent swaps to have $\sigma(i)$ be adjacent to $\sigma(j)$. We do so by swapping $\sigma(i)$ with the $\sigma(s)$ it is adjacent to that it has not already swapped with. After m swaps, we have $\sigma(i) < \sigma(j)$ are adjacent, so by performing one more swap, we reach $\sigma(i) > \sigma(j)$. Afterwards, we will perform m adjacent swaps with $\sigma(j)$ to the $\sigma(s)$ that it has not already swapped with. The end result is that $\sigma(j)$ is now in the original position of $\sigma(i)$ before the adjacent swappings. Further, the other positions $\sigma(s)$ are in their original positions before any swappings. The total number of adjacent swaps we performed is $2m + 1$, which is an odd number for any integer $m \geq 1$. By the previous case, this results in $(\text{sign } \sigma)$ changing an odd number of times. Thus, $(\text{sign } \sigma') = -1 \cdot (\text{sign } \sigma)$. $\square$

Definition 2.7 A graph G is a pair (V, E) where V is a set of elements called vertices and E is a set of pairs of vertices called edges. More precisely, $E \subseteq V \times V = \{(u, v) : u, v \in V\}$. If an edge exists and connects a pair of vertices, we call these vertices *adjacent*. We only consider graphs with finite number of vertices and edges connect two distinct vertices.

Definition 2.8 A *walk* in $G = (V, E)$ is a list of elements of V, $v_0 v_1 \ldots v_k$, such that $v_{i-1} v_i$ is an edge in E for all $i = 1, 2, \ldots, k$. A walk is said to be *closed* if the starting vertex v_0 is equal to the ending vertex v_k. A *path* in G is walk such that no vertex is repeated. If we allow for only the starting vertex and ending vertex of a path to be the same, we call that a *cycle*. A graph with no cycles is called *acyclic*. We say two paths are *distinct* if they do not share any internal vertices, but we allow for their starting vertex or ending vertex to be the same.

We can extrapolate that for any two vertices u, v in a cycle, there are two paths that can come about from the cycle. We do so by partitioning the cycle into two paths one that starts at u and ends with v and the other starting at v and going to u using the other part of the cycle. Further, the converse is true. For vertices u, v, if there are two distinct paths that starts at u and ends at v, we say that u and v are in a cycle. This idea is further explored and generalized in Menger's Theorem [Die17].

Therefore, in this discussion, we will define more structure to graphs that allow for paths to intersect but cycles do not form.

Definition 2.9 A directed graph '*digraph*' imposes an order on the pair of vertices for each edge, denoted as $G = (V, \vec{E})$. More precisely, if the edge $uv = (u, v)$ and $vu = (v, u)$ are in $\vec{E}$, then they must be distinct elements of $\vec{E}$.

The definitions of paths and acyclic are the same for digraphs as is in the general graph. However with the imposition of order on the edges, if there are two paths that start at u and end at v, we do not necessarily have a cycle. A cycle can only arise if in addition to a path from u to v, there is a path v to u.

Definition 2.10 Given two subsets of a vertex set $X = \{x(1), ..., x(n)\}, Y = \{y(1), ..., y(n)\} \subseteq V(G)$, an X, Y*-path system* $\mathcal{P}$ consists of a permutation $\sigma_{\mathcal{P}}$ of $[n]$ and paths $P_1, ..., P_n$ such that P_i is a path from $x(i)$ to $y(\sigma_{\mathcal{P}}(i))$. An X, Y-path system is called *disjoint* if the paths $P_1, ..., P_n$ are pairwise non-intersecting.

Definition 2.11 The *weight of an edge* e, denoted by $w(e)$, is a non-negative real number. The *weight of a path* Q, denoted by $w(Q)$, is the product of the weights of the edges in the path.

The *weight of a path system* $\mathcal{P}$, denoted as $W(\mathcal{P})$, is the product over all paths in the path system:

$$W(\mathcal{P}) = \prod_{Q \in \mathcal{P}} w(Q) = \prod_{Q \in \mathcal{P}} \prod_{e \in Q} w(e).$$

Definition 2.12 Given two subsets of a vertex set $X = \{x(1), ..., x(n)\}, Y = \{y(1), ..., y(n)\} \subseteq V(G)$, let $P(x(i), y(j))$ be the set of all paths that start at $x(i)$ and end at $y(j)$. An X, Y*-path matrix* has the sum of the weights of all paths from $x(i)$ to $y(j)$ for its (i, j)-th entry

$$a_{ij} = \sum_{Q \in P(x(i), y(j))} w(Q).$$

Definition 2.13 Given a $n \times n$ path matrix $A = [a_{ij}]$, the *determinant* of A is given by the sum over all permutations σ of $[n]$,

$$\det(A) = \sum_{\sigma} (\text{sign } \sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

3 Lindström-Gessel-Viennot Theorem

Theorem 1 (Lindström-Gessel-Viennot Theorem) [Wes20] Let X and Y be sets of n vertices in a finite acyclic digraph G with edges weighted by w. If A is the X, Y-path matrix, and $\mathbf{P}$ is the set of disjoint X, Y-path systems weighted by W as above, then

$$\sum_{\mathcal{P} \in \mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}) = \det(A).$$

Proof. We will first consider the set $\mathbf{T}$ of all X, Y-path systems, partitioning $\mathbf{T}$ into the set of disjoint X, Y-path systems $\mathbf{P}$ and the set of intersecting X, Y-path systems $\mathbf{Q}$. By defining a special signed involution $\varphi : \mathbf{Q} \to \mathbf{Q}$, we can establish that the net contribution over $\mathbf{Q}$ is zero. Thus,

$$\sum_{\mathcal{P}\in\mathbf{T}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}) = \sum_{\mathcal{P}\in\mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}).$$

Lastly, we will interpret the determinant of A combinatorially and relate it to the signed sum over $\mathbf{T}$ to conclude the theorem.

Let $X = \{x(1), \ldots, x(n)\}$ and $Y = \{y(1), \ldots, y(n)\}$ be n-sets of vertices in a finite acyclic digraph. Keep in mind that for every element $\mathcal{T}$ of $\mathbf{T}$, there are n paths in $\mathcal{T}$ that start from a unique $x(i)$ and end at a unique $y(\sigma_{\mathcal{P}}(i))$.

Let $\mathbf{P}$ be the set of disjoint X, Y-path systems and $\mathbf{Q}$ be the set of intersecting X, Y-path systems. For each element $Q = \{P_1, \ldots, P_n\}$ of $\mathbf{Q}$ where P_i is a path from $x(i)$ to $y(\sigma(i))$, there are at least two paths that intersect. Let k be the smallest integer such that P_k intersects another path in Q. Then, let l be the minimum index distinct from k such that P_k intersects P_l. Note that $k < l$. We impose the minimality in order to obtain a well-defined involution.

Let $P_k := x(k)u_2 \cdots u_{i-1}y(\sigma_{\mathcal{P}}(k))$ and $P_l := x(l)v_2 \cdots v_{j-1}y(\sigma_{\mathcal{P}}(l))$. Our goal is to swap the endings of P_k and P_l after their final intersection. Since Q is a path-system, the final intersection of P_k and P_l cannot be at their end point in Y. Although it should be noted that, it is possible for at most one of them to end at the intersection, therefore swapping the endings of P_k and P_l would result in swapping which path ends at the intersection.

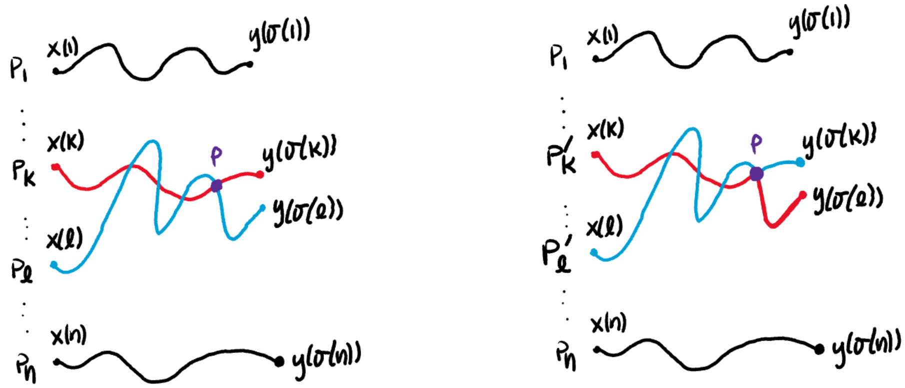

Figure 1: (left) The original configuration of Q.
(right) The configuration of Q' after swapping the endings of P_k and P_l

Suppose the final intersection of P_k and P_l is the vertex p, then define P'_k and P'_l as $P'_k = x(k)u_2 \cdots p \cdots v_{j-1}y(\sigma_{\mathcal{P}}(l))$ and $P'_l = x(l)v_2 \cdots p \cdots u_{i-1}y(\sigma_{\mathcal{P}}(k))$. We can visualize the change in Figure 1.

Note, the collection of paths $Q' = \{P_1, \dots, P'_k, P_{k+1}, \dots, P_{l-1}, P'_l, \dots, P_n\}$ is still a path system, and since the swapping did not eliminate the intersection, $Q' \in \mathbf{Q}$.

Define the function $\varphi : \mathbf{Q} \to \mathbf{Q}$ to do this ending swapping process given a path system Q. Roughly summarizing above, φ will find the smallest index k such that $P_k \in Q \in \mathbf{Q}$ intersects another path, then find the smallest index l such that P_k intersects P_l. Then, if p is the final intersection of P_k and P_l, φ will swap the part of P_k and P_l after p.

If we apply φ to Q', the smallest index it finds will be k for path P'_k. Otherwise, there exists an index $m < k$ such that P_m intersects another path. Since P_m was not chosen and altered by φ on Q, P_m is the same path in both Q and Q'. This contradicts the minimality of k since we specifically chose it to be the smallest index of an intersecting path in Q. Then, φ will find that the smallest index that intersects P'_k is l. Otherwise, there exists an index $k < \mu < l$ such that P'_k intersects with P_μ in Q'. This contradicts the minimality of l for Q, since it would imply that P_k intersects P_μ in Q.

Since the interesting pair of paths that φ chooses in Q' is P'_k and P'_l, we know they intersect at p. If $m \neq p$ is the final intersection of P'_k and P'_l in Q', then this implies that P_k and P_l had another intersection at m in Q, which contradicts the final property of p. So, applying φ to Q' will swap the parts of P'_k and P'_l after the intersection p. However, swapping these endings will return us to P_k and P_l. Thus, $\varphi(Q') = Q$ and so $\varphi^2(Q) = Q$. We can extend φ to all the elements of $\mathbf{T}$ by defining $\varphi(P) = P$ if $P \in \mathbf{P}$. Therefore, $\varphi^2(T) = T$ for any T in $\mathbf{T}$ and so φ is an involution.

We investigate the contribution of P'_k and P'_l of Q' to the weight of the path system. Since the associated permutation $\sigma_{Q'}$ of Q' only differs from σ_Q of Q at two places $\sigma_Q(k)$ and $\sigma_Q(l)$ by a swap, by Lemma 2.6, we have that

$$(\operatorname{sign} \sigma_{Q'}) = -1 \cdot (\operatorname{sign} \sigma_Q).$$

For a path $P_i = v_1 \dots v_k$ and a vertex v in P_i, let $\operatorname{pre}(P_i, v)$ be the path within P_i that starts at v_1 and ends at v, and let $\operatorname{post}(P_i, v)$ be the path within P_i that starts at v and ends at v_k. Further, by the swapping process,

$$\begin{aligned} \operatorname{pre}(P'_k, p) &= \operatorname{pre}(P_k, p) \quad , \quad \operatorname{pre}(P'_l, p) = \operatorname{pre}(P_l, p) \quad , \\ \operatorname{post}(P'_k, p) &= \operatorname{post}(P_l, p) \quad , \quad \operatorname{post}(P'_l, p) = \operatorname{post}(P_k, p). \end{aligned}$$

With this notation, we can use the commutativity of the definition of weights to obtain the following identity

$$\begin{aligned} w(P'_k)w(P'_l) = \prod_{e \in P'_k} w(e) \prod_{e \in P'_l} w(e) &= \prod_{e \in \operatorname{pre}(P'_k, p)} w(e) \prod_{e \in \operatorname{post}(P'_k, p)} w(e) \prod_{e \in \operatorname{pre}(P'_l, p)} w(e) \prod_{e \in \operatorname{post}(P'_l, p)} w(e) \\ &= \prod_{e \in \operatorname{pre}(P_k, p)} w(e) \prod_{e \in \operatorname{post}(P_k, p)} w(e) \prod_{e \in \operatorname{pre}(P_l, p)} w(e) \prod_{e \in \operatorname{post}(P_l, p)} w(e) \\ &= \prod_{e \in P_k} w(e) \prod_{e \in P_l} w(e) = w(P_k)w(P_l). \end{aligned}$$

Note that $Q' \setminus \{P'_k, P'_l\} = Q \setminus \{P_k, P_l\}$. Therefore, we have

$$\begin{aligned} W(Q') = \prod_{P_i \in Q'} w(P_i) &= \left[\prod_{P_i \in Q' \setminus \{P'_k, P'_l\}} w(P_i) \right] w(P'_k) w(P'_l) \\ &= \left[\prod_{P_i \in Q \setminus \{P_k, P_l\}} w(P_i) \right] w(P_k) w(P_l) = \prod_{P_i \in Q} w(P_i) = W(Q). \end{aligned}$$

We can separate $\mathbf{Q}$ into two disjoint sets A and B using the involution φ. For element Q of $\mathbf{Q}$, put Q in A (if it is not already in B), and put $Q' = \varphi(Q)$ in B. Then, with $(\text{sign } \sigma_{Q'}) = -(\text{sign } \sigma_Q)$ and $W(Q') = W(Q)$, we have that the net contribution of the elements of $\mathbf{Q}$ is zero.

$$\begin{aligned} \sum_{Q \in \mathbf{Q}} (\text{sign } \sigma_Q) W(Q) &= \sum_{Q \in A} (\text{sign } \sigma_Q) W(Q) + \sum_{Q' \in B} (\text{sign } \sigma_{Q'}) W(Q') \\ &= \sum_{Q \in A} (\text{sign } \sigma_Q) W(Q) - \sum_{Q \in A} (\text{sign } \sigma_Q) W(Q) = 0. \end{aligned}$$

Therefore,

$$\sum_{\mathcal{P} \in \mathbf{T}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}) = \sum_{\mathcal{P} \in \mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}).$$

For each permutation σ, there are path systems $\mathcal{P}$ in $\mathbf{T}$ such that $\sigma_{\mathcal{P}} = \sigma$. Recall that for each $i = 1, \ldots, n$, $a_{i,\sigma(i)} = \sum_{Q \in P(x(i), y(\sigma(i)))} w(Q)$, the sum over $P(x(i), y(\sigma(i)))$, the set of paths Q that start at $x(i)$ and ends at $y(\sigma(i))$. The total weight of such path system $\mathcal{P}$ is

$$\prod_{i=1}^{n} a_{i,\sigma(i)} = \prod_{i=1}^{n} \sum_{Q \in P(x(i), y(\sigma(i)))} w(Q) = \sum_{\substack{\mathcal{P} \in \mathbf{T} \\ \text{assoc.} \\ \text{to } \sigma}} W(\mathcal{P}).$$

The final equality arises since the product of the sums of weights of path systems Q over $P(x(i), y(\sigma(i)))$ can be grouped up into path systems $\mathcal{P}$ in $\mathbf{T}$ such that every $\mathcal{P}$ is associated with σ.

Lastly, we can partition $\mathbf{T}$ into its association to each permutation σ. Thus, by Definition 2.13 and the net zero contribution over $\mathbf{Q}$, we have that the determinant of A is

$$\begin{aligned} \det A = \sum_{\sigma} (\text{sign } \sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} &= \sum_{\sigma} (\text{sign } \sigma) \sum_{\substack{\mathcal{P} \in \mathbf{T} \\ \text{assoc.} \\ \text{to } \sigma}} W(\mathcal{P}) \\ &= \sum_{\mathcal{P} \in \mathbf{T}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}) \\ &= \sum_{\mathcal{P} \in \mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}). \end{aligned}$$

Thus, we have shown the theorem. $\square$

4 Applications

Nice consequences of the Lindström-Gessel-Viennot theorem gives classical determinant properties their own combinatorial interpretation [SAL03]. In this section, we will show how Lindström-Gessel-Viennot is used to show that the determinant of a transpose is equal to the determinant of the original matrix and we also prove the Cauchy-Binet formula.

4.1 Determinant of Transpose Matrix

Definition 4.1 Given n, m positive integers, let A be an $n \times m$ matrix given by $[a_{ij}]$ with $1 \leq i \leq n$ and $1 \leq j \leq m$. The transpose A^T is the $m \times n$ matrix obtained by taking the rows of A and making them columns of A^T. Thus, A^T is given by $[a_{ji}]$.

Definition 4.2 For a permutation π of $[n]$, a *cycle* of π is a subset S of $[n]$ such that for any integer $k \geq 1$ and for all $s \in S$, we have $\pi^k(s) \in S$. A *transposition* is a cycle of size 2.

Lemma 4.3 The sign of the composition of permutations is the product of the signs of permutation, in other words, for any permutations π_1 and π_2 of $[n]$, $(\text{sign } (\pi_1 \circ \pi_2)) = (\text{sign } \pi_1)(\text{sign } \pi_2)$.

Proof. Every permutation π of $[n]$ has a unique decomposition and factorization into transpositions [Cla71]. In [Jac09], there is an equivalent definition of the sign of a permutation in terms of the number of transpositions in the decomposition and factorization. We have

$$(\text{sign } \pi) = (-1)^m$$

where m is the number of transpositions in the decomposition and factorization of π.

Let π_1 and π_2 be permutations of $[n]$. Suppose π_1 and π_2 has m_1 and m_2 number of transpositions in their decomposition and factorization, respectively. From [Cla71], we know that the decomposition and factorization of their composition $\pi_1 \circ \pi_2$ has $m_1 + m_2$ transpositions.

Therefore, we have

$$(\text{sign } (\pi_1 \circ \pi_2)) = (-1)^{m_1+m_2} = (-1)^{m_1}(-1)^{m_2} = (\text{sign } \pi_1)(\text{sign } \pi_2).$$

Thus, proving the theorem.

□

Lemma 4.4 Given any permutation σ, we have $(\text{sign } \sigma) = (\text{sign } \sigma^{-1})$.

Proof. We have $\sigma \circ \sigma^{-1}$ is the identity permutation. The identity permutation has zero inversions, thus $(\text{sign } (\sigma \circ \sigma^{-1})) = (\text{sign } 1) = (-1)^0 = 1$.

By Lemma 4.3, $(\text{sign } \sigma)(\text{sign } \sigma^{-1}) = 1$. Therefore, for the product of two terms to be positive, they must be the same sign. □

Proposition 4.5 Let A be a $n \times n$ matrix. $\det(A^T) = \det A$.

Proof. We view A to be a X, Y-path matrix of a finite acyclic digraph G, with $|X| = |Y| = n$. The interpretation of the transpose becomes viewing the direction of all the edges of G in reverse. Therefore, A^T could be seen as a Y, X-path matrix.

By Lindström-Gessel-Viennot theorem, for any given permutation σ of $[n]$, we have a collection of disjoint X, Y-path systems $\mathbf{P}_\sigma$ associated to σ in the calculation of $\det A$. Therefore, for each $\mathcal{P}$ in $\mathbf{P}_\sigma$ 'reverse' the directions of each edge to obtain a Y, X-path system $\mathcal{R}$.

Note that the permutation associated with $\mathcal{R}$ is σ^{-1}. By reversing the direction of edges, we obtain a permutation σ^{-1} of $[n]$ and a collection of disjoint Y, X-path systems $\mathbf{R}_{\sigma^{-1}}$ associated to σ^{-1}.

Since we obtain $\mathcal{R}$ from $\mathcal{P}$, there is a one-to-one correspondence between $\mathbf{P}_\sigma$ and $\mathbf{R}_{\sigma^{-1}}$. Suppose $\mathcal{P} = \{P_1, \ldots, P_n\}$, then $\mathcal{R} = \{\tilde{P}_1, \ldots, \tilde{P}_n\}$ where, for each $i = 1, \ldots, n$, $\tilde{P}_i$ is the path P_i with all directions reversed and the weights are preserved. Therefore,

$$W(\mathcal{P}) = \prod_{i=1}^{n} w(P_i) = \prod_{i=1}^{n} w(\tilde{P}_i) = W(\mathcal{R})$$

and so, this correspondence preserves the path system weights.

Further the summation over all permutations σ of $[n]$ will coincide with the summation over all permutations σ^{-1} of $[n]$. Thus, by Lemma 4.4 and Lindström-Gessel-Viennot theorem,

$$\det(A^T) = \sum_{\sigma^{-1}} (\text{sign } \sigma^{-1}) \sum_{\mathcal{R} \in \mathbf{R}_{\sigma^{-1}}} W(\mathcal{R}) = \sum_{\sigma} (\text{sign } \sigma) \sum_{\mathcal{P} \in \mathbf{P}_\sigma} W(\mathcal{P}) = \det A$$

proving the proposition. □

4.2 Cauchy-Binet Formula

Typical linear algebra text usually contains the following fact: Let A and B be $n \times n$ matrices. Then, $\det(AB) = \det(A)\det(B)$ [SAL03]. The restriction of A and B being the same size $n \times n$ can be alleviated using the Lindström-Gessel-Viennot theorem.

The following proposition gives an identity for the determinant of a product of two matrices that are not the same size (given that the product is well-defined and the product matrix is a square matrix).

Proposition 4.6 (Cauchy-Binet Formula) [Wes20] Let n, p be positive integers with $n \leq p$. If $A = [a_{i,j}]$ is a $n \times p$ matrix and $B = [b_{j,k}]$ is a $p \times n$ matrix, then

$$\det(AB) = \sum_{S \subseteq \mathbf{S}} (\det(A_S))(\det(B_S))$$

where $\mathbf{S}$ is the family of all n-subsets of $[p]$, A_S consists of the columns of A indexed by S, and B_S consists of the rows of B indexed by S.

Proof. Let $X = \{x(1), \ldots, x(n)\}$, $Y = \{y(1), \ldots, y(n)\}$, and $Z = \{z(1), \ldots, z(p)\}$. Form a digraph on the vertex set $X \cup Y \cup Z$. Let with edge sets be

$$A = \{x(i)z(j) : i = 1, \ldots, n\ ,\ j = 1, \ldots, p\} \text{ and}$$

$$B = \{z(j)y(k) : j = 1, \ldots, p\ ,\ k = 1, \ldots, n\}.$$

We assign edge weights: $w(x(i)z(j)) = a_{i,j}$ and $w(z(j)y(k)) = b_{j,k}$.

Thus, the (i,k)-th entry of $C = AB$ is the dot product of the i-th row of A and the k-th column of B, which gives $c_{i,k} = \sum_{j=1}^{n} w(x(i)z(j))w(z(j)y(k))$. Therefore, C is an X, Y-path matrix since $w(x(i)z(j))w(z(j)y(k))$ is the weight of the paths connecting $x(i)$ to $y(k)$. Let $\mathbf{P}$ be the set of all disjoint X, Y-path systems, and by Lindström-Gessel-Viennot theorem,

$$\det(AB) = \sum_{\mathcal{P} \in \mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}).$$

Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ in $\mathbf{P}$. For each i, P_i chooses one element of Z for its path from $x(i)$ to $y(\sigma(i))$. It has to pass exactly one element of Z by design of the digraph, since all edges from X connects to an element of Z and does not return. Similarly, all edges from Z connects to an element of Y and does not return. Let S be the indices of elements of Z that was used by the paths in P. Define σ' as the permutation on $[n]$ such that for each i in $[n]$, $z(\sigma(i))$ is in the path P_i.

We can deconstruct every path system $\mathcal{P}$ in $\mathbf{P}$ as two disjoint path systems $\mathcal{A}_S$ (going from X to Z') and $\mathcal{B}_S$ (going from Z' to Y), where Z' are the elements of Z indexed by S.

More precisely, we can split P_i in $\mathcal{P}$ to be a path A_i connecting $x(i)$ to $z(\sigma'(i))$ and a path B_i connecting $z(\sigma'(i))$ to $y(\sigma(i))$. Thus,

$$\begin{aligned} w(P_i) = \prod_{e \in P_i} w(e) &= \prod_{\substack{e \in P_i \\ \text{ending at} \\ z(\sigma'(i))}} w(e) \prod_{\substack{e \in P_i \\ \text{starting at} \\ z(\sigma'(i))}} w(e) \\ &= \prod_{e \in A_i} w(e) \prod_{e \in B_i} w(e) = w(A_i)w(B_i). \end{aligned}$$

For each S, define the sets

$$\mathbf{M}_S = \{\mathcal{A}_S \mid \mathcal{A}_S \text{ disjoint } X, S\text{-path system}\} \text{ and}$$

$$\mathbf{N}_S = \{\mathcal{B}_S \mid \mathcal{B}_S \text{ disjoint } S, Y\text{-path system}\}.$$

For path systems $\mathcal{P}$ in $\mathbf{P}$, we have disjoint path systems $\mathcal{A}_S$ in $\mathbf{M}_S$ and $\mathcal{B}_S$ in $\mathbf{N}_S$ such that

$$W(\mathcal{P}) = \prod_{P_i \in \mathcal{P}} w(P_i) = \prod_{i=1}^{n} w(A_i)w(B_i) = \prod_{A_i \in \mathcal{A}_S} w(A_i) \prod_{B_i \in \mathcal{B}_S} w(B_i) = W(\mathcal{A}_S)W(\mathcal{B}_S).$$

Consider the matrices A_S consisting of the columns of A indexed by S and B_S consisting of the rows of B indexed by S. Since A_S and B_S are square matrices, by Lindström-Gessel-Viennot theorem,

$$\det(A_S) = \sum_{\mathcal{A}_S \in \mathbf{M}_S} (\text{sign } \sigma_{\mathcal{A}_S}) W(\mathcal{A}_S) \qquad \det(B_S) = \sum_{\mathcal{B}_S \in \mathbf{N}_S} (\text{sign } \sigma_{\mathcal{B}_S}) W(\mathcal{B}_S)$$

where $\sigma_{\mathcal{A}_S}$ is a bijective map $[n] \to S$ and $\sigma_{\mathcal{B}_S}$ is a bijective map $S \to [n]$.

For any pair $\mathcal{A}_S$ in $\mathbf{M}_S$ and $\mathcal{B}_S$ in $\mathbf{N}_S$, we can concatenate every path in $\mathcal{A}_S$ with exactly one path in $\mathcal{B}_S$, since for every $i = 1, \ldots, n$, we have a path from $x(i)$ to $z(\sigma_{\mathcal{A}_S}(i))$ in $\mathcal{A}_S$ and a path from $z(\sigma_{\mathcal{A}_S}(i))$ to $y(\sigma_{\mathcal{B}_S}(i))$ in $\mathcal{B}_S$.

We know these paths are disjoint since $\mathcal{A}_S$ and $\mathcal{B}_S$ are each made up of n disjoint paths, so their concatenation only overlaps at the elements of Z. Further, we know two paths in $\mathcal{A}_S$ do not end at the same element of Z and two paths in $\mathcal{B}_S$ do not start at the same element of Z. Therefore, the concatenation of $\mathcal{A}_S$ and $\mathcal{B}_S$ is a disjoint X, Y-path system, and thus, there is a $\mathcal{P}^*$ in $\mathbf{P}$ such that $\mathcal{A}_S \cup \mathcal{B}_S = \mathcal{P}^*$.

Then, by construction, $\sigma_{\mathcal{A}_S} \circ \sigma_{\mathcal{B}_S} = \sigma_{\mathcal{P}^*}$ is a permutation of $[n]$ to $[n]$ for any $\mathcal{A}_S$ in $\mathbf{M}_S$ and any $\mathcal{B}_S$ in $\mathbf{N}_S$. By Lemma 4.3, $(\text{sign } \sigma_{\mathcal{P}^*}) = (\text{sign } (\sigma_{\mathcal{A}_S} \circ \sigma_{\mathcal{B}_S})) = (\text{sign } \sigma_{\mathcal{A}_S})(\text{sign } \sigma_{\mathcal{B}_S})$.

Therefore, by summing S over all n-subsets of $[p]$, we have

$$\begin{aligned}
\sum_{S \subseteq \mathbf{S}} \det(A_S) \det(B_S) &= \sum_{S \subseteq \mathbf{S}} \sum_{\mathcal{A}_S \in \mathbf{M}_S} (\text{sign } \sigma_{\mathcal{A}_S}) W(\mathcal{A}_S) \sum_{\mathcal{B}_S \in \mathbf{N}_S} (\text{sign } \sigma_{\mathcal{B}_S}) W(\mathcal{B}_S) \\
&= \sum_{S \subseteq \mathbf{S}} \sum_{\mathcal{A}_S \in \mathbf{M}_S} \sum_{\mathcal{B}_S \in \mathbf{N}_S} (\text{sign } \sigma_{\mathcal{A}_S})(\text{sign } \sigma_{\mathcal{B}_S}) W(\mathcal{A}_S) W(\mathcal{B}_S).
\end{aligned}$$

By the above, if we fix $\mathcal{A}_S$ in $\mathbf{M}_S$ and sum over $\mathcal{B}_S$ in $\mathbf{N}_S$, then we will obtain a collection of X, Y-path systems $\{\mathcal{P}^*(\mathcal{A}_S)\}$ associated with the collection of compositions of permutations $\sigma_{\mathcal{A}_S} \circ \sigma_{\mathcal{B}_S} = \sigma_{\mathcal{P}^*(\mathcal{A}_S)}$ where $\sigma_{\mathcal{A}_S}$ is fixed and $\sigma_{\mathcal{B}_S}$ varies over $\mathbf{N}_S$.

Then, collecting these X, Y-path systems over $\mathcal{A}_S$ in $\mathbf{M}_S$ will yield every possible disjoint X, Y-path system that uses the elements of Z indexed by S. Lastly, summing over all possible n-sets of $[p]$ will yield every disjoint X, Y-path system. Therefore, by Lemma 4.3 and the above, we have

$$\begin{aligned}
\sum_{S \subseteq \mathbf{S}} \det(A_S) \det(B_S) &= \sum_{S \subseteq \mathbf{S}} \sum_{\mathcal{A}_S \in \mathbf{M}_S} (\text{sign } \sigma_{\mathcal{P}^*(\mathcal{A}_S)}) W(\mathcal{P}^*(\mathcal{A}_S)) \\
&= \sum_{\mathcal{P} \in \mathbf{P}} (\text{sign } \sigma_{\mathcal{P}}) W(\mathcal{P}) \\
&= \det(AB).
\end{aligned}$$

Thus, we have the Cauchy-Binet formula. □

References

[Cla71] Allan Clark. *Elements of Abstract Algebra*, chapter on The symmetric groups. Dover, 1971.

[Die17] Reinhard Diestel. *Graph Theory*, chapter on Menger's Theorem. Springer, 2017.

[GV85] Ira M. Gessel and Xavier G. Viennot. Binomial determinants, paths, and hook length formulae. *Advances in Mathematics*, 1985.

[Jac09] Nathan Jacobson. *Basic Algebra I*, chapter on Cycle decomposition of permutations. Dover, 2009.

[Lin73] Bernt Lindström. On the vector representation of induced matroids. *Bulletin of the London Mathematical Society*, 1973.

[SAL03] Friedberg, Stephen, Insel, Arnold, and Spence, Lawrence. *Linear Algebra*, chapter Determinants of Order n & Properties of Determinants. Pearson, 2003.

[Wes20] Douglas B. West. *Combinatorial Mathematics*, chapter on Signed Involutions & Determinants and Path Systems, pages 163–171. Cambridge University Press, 2020.