ON THE COMBINATORICS OF DETERMINANTS

SAI CHINTAGUNTA

1. The Lindström-Gessel-Viennot Lemma

The Lindström-Gessel-Viennot Lemma (LGV Lemma) relates the number of non-intersecting path systems in an acyclic directed graph to the determinant of a matrix. The lemma was first introduced by Bernt Lindström in 1972, but it was popularized by Ira Gessel and Gerard Viennot in 1985. Before we can introduce the lemma, we need to talk about some important preliminaries. First, we will talk about permutations.

Definition 1.1. A *permutation* is a bijective function σ where $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. The permutation for which $\sigma(i) = i$ for all $i \in \{1, \ldots, n\}$ is called the *identity* permutation.

In practice, a permutation is often referred to as a reordering of a set $\{1, \ldots, n\}$ where $1, 2, 3, \ldots, n$ is assumed to be the initial order.

Example. A simple example of a permutation can be seen in Figure 1. This reorders $1, 2, 3$ to $2, 1, 3$.

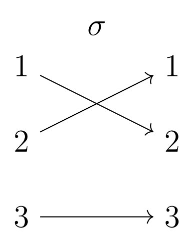

Figure 1. A typical permutation.

We can now introduce a special case of a permutation.

Definition 1.2. We call a permutation a *transposition* if there exist distinct elements i and j such that $i < j$, $\sigma(i) = j$, $\sigma(j) = i$, and every other element is fixed under the permutation. We can denote it by π_{ij} [1].

A transposition can be thought of as swapping two elements and keeping the rest the same.

Example. The permutation illustrated by Figure 1 is a transposition because you are swapping 1 and 2 while sending the 3 to itself. Thus, it can be denoted as π_{12}.

We can now make the following observation.

Proposition 1.3. *Every permutation can be expressed as a composition of transpositions.*

[1] The reason we have $i < j$ is so we can't have π_{ij} and π_{ji} for the sake of notation.

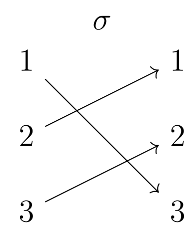

Figure 2. The permutation $\sigma = \pi_{12} \circ \pi_{13}$.

Example. The permutation in Figure 2 reorders $1, 2, 3$ to $2, 3, 1$. So, expressing it as a composition of transpositions, we can write it as $\pi_{12} \circ \pi_{13}$ (the order in which we apply the transpositions is right to left as in function composition).

You can also note here that applying these transpositions in reverse to the permutation will give you the identity permutation. So for Figure 2 we have $\pi_{13} \circ \pi_{12} \circ \sigma = \mathrm{id}$. In general, it is true that if you express a given permutation as a composition of transpositions, reversing the order of transpositions and applying it to the permutation will give you the identity.

A natural question arises: given a reordering of a set, can you say anything about the number of transpositions needed to express the permutation? In fact, there is a neat answer to this question!

Lemma 1.4. *The number of transpositions needed to express a given permutation is always odd or always even.*

The proof of this lemma requires machinery outside of the scope of this paper. However, if you are interested in the proof it can be accessed here: [reference]. With Lemma 1.4 the following definition is quite natural:

Definition 1.5. We call a permutation *even* if it is composed by an even number of transpositions, and *odd* if it is composed by an odd number of transpositions.

Lemma 1.4 and Definition 1.5 allow us to introduce a new notion: the sign of a permutation.

Definition 1.6. Given a permutation σ, we define $\mathrm{sign}(\sigma)$ as

$$\mathrm{sign}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Example. Since permutation in Figure 2 is $\pi_{12} \circ \pi_{13}$, we have that its sign is 1.

A natural question would be how is this related to determinants? So far, we have been talking about mathematical tools used in the world of combinatorics. So, the connection to linear algebra is not immediately obvious. The connection comes in the form of the Leibniz formula for determinants.

Lemma 1.7 (Leibniz Expansion)**.** *The determinant of an $n \times n$ matrix A is given by:*

$$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma)\ a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

where S_n is the set of all permutations on $\{1, \ldots, n\}$.

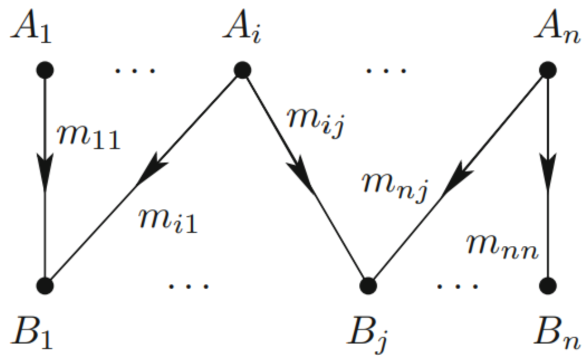

Figure 3. The sets $\mathcal{A}$ and $\mathcal{B}$ along with the edges between them and their respective weights. Taken from [reference].

The proof of this lemma is quite long will and thus be omitted from this paper. However, it can be found in [reference].

Now recall that the LGV-Lemma relates the determinant of a matrix to a graph. Even though the LGV-Lemma applies to the more general acyclic directed graphs, to build intuition, we explore a special case: directed bipartite graphs.

Definition 1.8. A *bipartite* graph is a graph whose vertices can be split into two disjoint sets, U and V such that every edge in the graph connects a vertex in U with a vertex in V.

Notice we want to relate our graph to the determinant of a matrix, and determinants are only defined on matrices of sizes $n \times n$. Thus, we examine a bipartite graph without multi edges, whose disjoint sets of vertices are both of size n. We call these sets $\mathcal{A}$ and $\mathcal{B}$ where $\mathcal{A} = \{A_1, \dots, A_n\}$ and $\mathcal{B} = \{B_1, \dots, B_n\}$. Additionally, we can assign a weight to each edge in the graph.

Now we can very naturally define a matrix. We define our *path matrix* M as the $n \times n$ matrix with m_{ij} equal to the weight of the edge between A_i and B_j [2]. The graph with the weights is depicted in Figure 3.

Now take a look at the formula for the determinant of M:

$$\det M = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\ m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}.$$

How do we interpret the right hand side? First, we need to understand what the permutation σ is actually doing. Every σ, corresponds to the set of edges

$$A_1 \to B_{\sigma(1)}, \dots, A_n \to B_{\sigma(n)}.$$

This gives us a *path system* (We define this later in Definition **??**) . We can denote it by $\mathcal{P}_\sigma$. Now we can define the weight of a path system $\mathcal{P}_\sigma$ as

$$w(\mathcal{P}_\sigma) = w(A_1 \to B_{\sigma(1)}) \cdots w(A_n \to B_{\sigma(n)}).$$

[2]We assume that there exists an edge between every two vertices A_i and B_j. In the case where there does not exist an edge, we can simply set its weight equal to 0

So, we notice immediately that that $w(\mathcal{P}_\sigma) = m_{1\sigma(1)} \cdots m_{n\sigma(n)}$. Thus, we are able to rewrite our formula for the determinant of M as

$$\det M = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) w(\mathcal{P}_\sigma).$$

This is a nice way to think of the determinant. But we can actually generalize this way of thinking to acyclic directed graphs. This is what the LGV-Lemma is all about! Before we can introduce the formal statement of the lemma we need to talk about some of the tools that will be used in its proof.

The graph in the LGV-Lemma is denoted $G = (V, E)$. G is finite, acyclic, and directed.

Definition 1.9. A graph is *acyclic* if it contains no cycles.

The graph being acyclic prevents there from being infinite paths between two vertices[3]. We can also give every edge e in the graph a weight. We can call it $w(e) \in \mathbb{R}$.

Additionally, we are able to split G into two sets of vertices $\mathcal{A} = \{A_1, \dots, A_n\}$ and $\mathcal{B} = \{B_1, \dots, B_n\}$ where they are both of size n.

In the bipartite case we simply took the weight of the edge. However, in the general case we consider paths with more than just one edge. Thus, we come up with the following definition.

Definition 1.10. Suppose we have a path P between vertices $A \in \mathcal{A}$ and $B \in \mathcal{B}$, denoted $P : A \to B$, then we define its weight as

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

Now, just as in the bipartite case, we need to define a path matrix. We can define the path matrix M as having

$$m_{ij} = \sum_{P: A_i \to B_j} w(P).$$

If there do not exist any paths between A_i and B_j we let $m_{ij} = 0$. In other words, m_{ij} is the sum of the weights of all paths between A_i and B_j. Now we can return to the idea of a path system.

Definition 1.11. A *path system* from $\mathcal{A}$ to $\mathcal{B}$ consists of a permutation σ together with n paths from $P_i : A_i \to B_{\sigma(1)}$, for $i = 1, \dots, n$

Example. Figure 4 is a graph with $\mathcal{A} = \{A_1, A_2\}$ and $\mathcal{B} = \{B_1, B_2\}$.The path system contains P_1 (the blue path) and P_2 (the red path). The permutation associated with the path system is π_{12}.

Just as we defined a weight for a path, it makes sense to define weights for a path system.

Definition 1.12. We define the weight of a path system as the product of its path weights. That is,

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

[3]if you have a cycle, you can keep going in loops i.e. infinite paths.

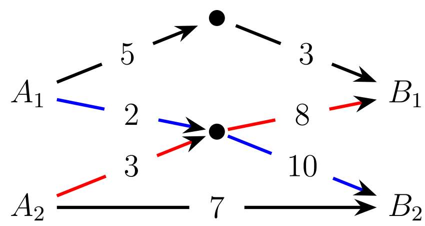

Figure 4. A path system.

where the P_i are the paths associated with the path system.

Since a path system $\mathcal{P}$ is associated with a permutation σ, it makes sense to give it a notion of sign.

Definition 1.13. We define $\operatorname{sign}(\mathcal{P})$ as being equal to $\operatorname{sign}(\sigma)$ where σ is the permutation associated with $\mathcal{P}$.

Now we give a word to describe path systems whose paths are non-intersecting.

Definition 1.14. We say a path system $\mathcal{P} = (P_1, \ldots, P_n)$ is vertex disjoint if the paths of $\mathcal{P}$ are pairwise disjoint.

We now have everything to state the Lindström-Gessel-Veinnot Lemma.

Lemma 1.15 (Lindström-Gessel-Viennot)**.** *Let $G = (V, E)$ be a finite weighted acyclic directed graph, $\mathcal{A} = \{A_1, \ldots, A_n\}$ and $\mathcal{B} = \{B_1, \ldots, B_n\}$ two sets of n vertices, and M the path matrix from $\mathcal{A}$ to $\mathcal{B}$. Then*

$$\det(M) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

Proof. We know from the Leibnitz expansion that

$$\det M = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\ m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}.$$

However, recalling that $m_{i\sigma(i)}$ is the sum of all the paths between A_i and $B_{\sigma(i)}$ we can write

$$m_{i\sigma(i)} = \sum_{P_i : A_i \to B_{\sigma(i)}} w(P_i).$$

Substituting into $m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}$ we get the following:

$$m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)} = \left(\sum_{P_1 : A_1 \to B_{\sigma(1)}} w(P_1) \right) \left(\sum_{P_2 : A_2 \to B_{\sigma(2)}} w(P_2) \right) \cdots \left(\sum_{P_n : A_n \to B_{\sigma(n)}} w(P_n) \right).$$

Now imagine expanding the product of sums. What you will end up with is the sum of all possible products of weights where the i_{th} term (from the left) in each product is the weight of a path from $A_i \to B_{\sigma(i)}$. Recalling Definition 1.12, we can notice that the sum of these products is actually the sum of the weights of all path systems associated with σ. By

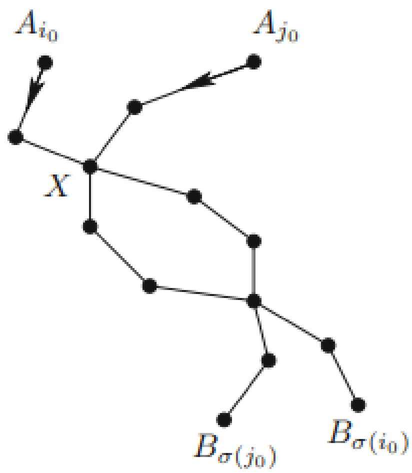

Figure 5. The first intersection of two paths in $\mathcal{P}$.

Definition 1.13 the sign of a path system associated with a permutation σ is $\text{sign}(\sigma)$. Thus, it follows that

$$\text{sign}(\sigma)\ m_{1\sigma(1)}m_{2\sigma(2)}\cdots m_{n\sigma(n)} = \sum_{\substack{\mathcal{P}\text{ associated}\\ \text{with }\sigma}} \text{sign}(\mathcal{P})\ w(\mathcal{P}).$$

This immediately gives us

$$\det(M) = \sum_{\mathcal{P}} \text{sign}(\mathcal{P})\ w(\mathcal{P}).$$

We now need to somehow make the jump to the statement of the LGV-Lemma. We will do this by showing that

$$\sum_{\mathcal{P}\in N} \text{sign}(\mathcal{P})\ w(\mathcal{P}) = 0,$$

where N is the set of all intersecting path systems. We accomplish this by making a very elegant bijective argument.

First take a $\mathcal{P} \in N$ with $P_i : A_i \to B_{\sigma(i)}$, where σ is the permutation associated with $\mathcal{P}$. We know by definition that some two paths in $\mathcal{P}$ will intersect. Now we define i_0 as the minimal index such that $P_{i_0} \in \mathcal{P}$ shares some vertex with another path. We can let the first common such vertex be called X. We let j_0 be minimal index ($j_0 > i_0$) such that P_{j_0} has the vertex X in common with P_{i_0}. This intersection is depicted in Figure 5.

Now we will define a function $\gamma : N \to N$. We say that $\gamma(\mathcal{P}) = (P_1', \ldots, P_n')$ where we have the following:

- $P_k' = P_k$ if $k \neq i_0, j_o$.

- P_{i_0}' goes from A_{i_0} to X and then continues along the path of P_{j_0} until it reaches $B_{\sigma(j_0)}$. Similarly, P_{j_0}' goes from A_{j_0} to X at which point it goes along P_{i_0} until it reaches $B_{\sigma(i_0)}$.

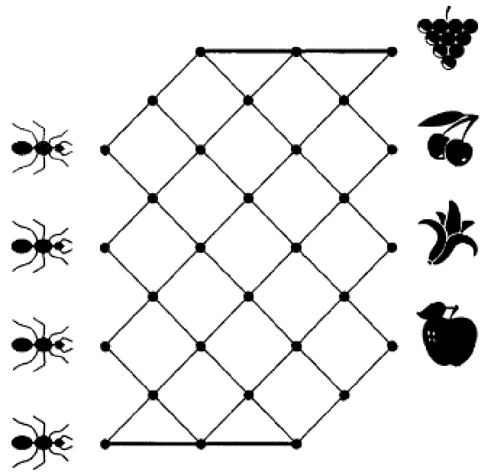

Figure 6

We will clearly have $\gamma(\gamma(\mathcal{P})) = \mathcal{P}$ since all γ does is flip the paths. So, applying it twice will flip it back to its initial state. We also know that $w(\gamma(\mathcal{P})) = w(\mathcal{P})$ since both path systems have exactly the same edges. Finally, we have that if the permutation associated with $\gamma(\mathcal{P})$ is σ', then $\text{sign}(\sigma') = -\text{sign}(\sigma)$. This is because we have $\pi_{i_0 j_o} \circ \sigma = \sigma'$.

Thus, we are able to pair up all $\mathcal{P} \in N$ with another path system with opposite sign and same weight. Therefore, $\sum_{\mathcal{P} \in N} \text{sign}(\mathcal{P})\, w(\mathcal{P}) = 0$. It immediately follows that

$$\det(M) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \text{sign}(\mathcal{P})\, w(\mathcal{P}),$$

proving the lemma.

■

2. Applications Of The Lindström-Gessel-Viennot Lemma

How do we use the LGV-Lemma in practice? This lemma is incredibly adaptable since you can either solve a determinant to get an answer to a combinatorics problem, or you can analyze a combinatorics problem to solve for a determinant. We can first look at solving a combinatorics problem through solving a determinant.

We begin with the problem of the determined ants which was originally stated in [reference].

Question 2.1. *If an ant can only move to the right along the grid, how many ways are there for the ants in Figure 6 to reach different pieces of food without their paths crossing each other?*

The problem is just asking to count the number of non-intersecting path systems. But is the LGV-Lemma even applicable? One thing that we can immediately notice is that even though the graph in Figure 6 doesn't have any arrows to indicate direction, it implicitly has

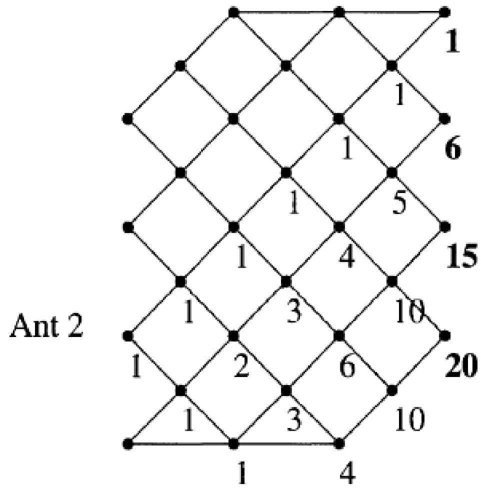

Figure 7. Recursively counting paths for Ant 2.

direction because of the condition that the ants can only move to the right. This condition also forces the graph to be acyclic since in order for a cycle to exist an ant would have to move left at least once. We also have a very natural choice of points for our two sets of vertices. The 4 origin points (the ants), and the 4 destinations (the food). Thus, the LGV-Lemma is applicable! We will first compute the path matrix M.

Label the ants from bottom to top as Ant 1 to Ant 4, and the destinations (apple, banana, etc.) as Destination 1 to Destination 4. We will also set the weight of each edge to be 1. This makes sense because then each path will have weight 1 which forces all path systems to also have weight 1. This is ideal since each path system will only contribute a 1 or -1 to the final sum. With this weight condition the m_{ij} entry of the path matrix becomes the number of paths from Ant i to Destination j. How do we count this? It's actually pretty simple using recursion as shown in Figure 7. Using recursion we can see that $m_{21} = 20$, $m_{22} = 15$, $m_{23} = 6$, and $m_{24} = 1$. Doing the same thing for each of the ants, we end up with the path matrix

$$M = \begin{bmatrix} 14 & 6 & 1 & 0 \\ 20 & 15 & 6 & 1 \\ 15 & 20 & 15 & 6 \\ 6 & 20 & 15 & 14 \end{bmatrix}.$$

A very natural question is why would the LGV-Lemma actually work here? The statement would says that

$$\det(M) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

What is stopping some path systems from contributing a -1 and erasing other path systems from the sum. It all comes from the key observation that the only way for a non-intersecting path system to exist in Figure 6 is if Ant i goes to Destination i.[4] In fact, this property is so special we give a name.

[4]This is pretty clear just by looking at it, but if you are confused trying some examples should help.

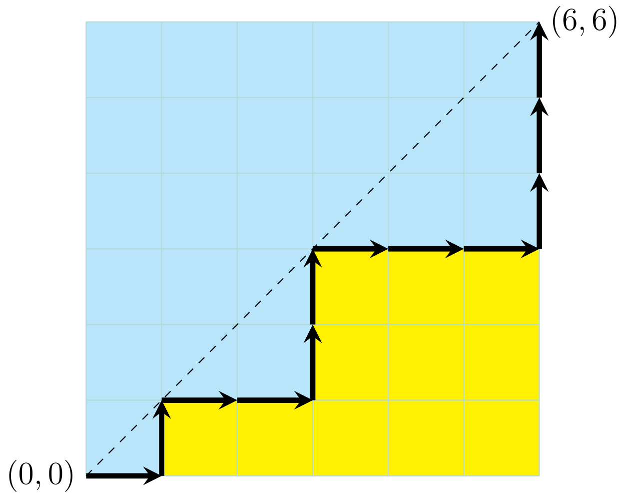

Figure 8. A route from $(0,0)$ to $(6,6)$.

Definition 2.2. graph $G = (V, E)$ that satisfies the conditions of the LGV-Lemma is called *nonpermutable* if all vertex-disjoint path systems are associated with the identity permutation.

Having a nonpermutable graph gives us a crucial piece of information as all vertex-disjoint path systems are now forced to have their sign equal to 1. This leads to a neat result.

Corollary 2.3. *If a graph $G = (V, E)$ satisfies the conditions for the LGV-Lemma, is nonpermutable, and all its edges have weight* 1*, then*

$$\det(M) = \text{number of vertex-disjoint path systems associated with id}\ .$$

Proof. Since G is nonpermutable, we know that all its vertex-disjoint path systems have sign equal to 1. Additionally, since the weight of every edge is equal to 1 we get by Definition 1.10 that the weight of every path is 1. It immediately follows from Definition 1.12 that the weight of all the path systems is 1. Thus, we get

$$\det(M) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} 1$$

which is just the number of vertex-disjoint path systems. ■

It follows directly from Corollary 2.3 that the number of non-intersecting path systems in Figure 6 is just $\det(M) = 889$.

We will now introduce two problems in which we can use a combinatorial way of thinking to prove identities for determinants.

We begin with introducing the Catalan Number. Say we have a standard coordinate grid. Our origin point will be $(0,0)$ and our destination will be (n,n). We can consider paths along the grid from $(0,0)$ to (n,n). Specifically, paths that only go up or right.

It is fairly easy to count paths that go up and right without any other restrictions. This is because in order to get from, $(0,0)$ to (n,n), you need to make n ups and n rights.All of the satisfying paths will just be some reordering of these ups and rights. Thus, the total number of these paths is just $\binom{2n}{n}$.

Example. In Figure 8 the number of paths from $(0, 0)$ to $(6, 6)$ that only go up and right along the grid will be $\binom{12}{6} = 924$.

Now we can explore another special type of path. We call this path a route.

Definition 2.4. We call a path a *route* if it goes up and right along the grid, but every coordinate (x, y) on the path satisfies $y \leq x$.

In other words, a route should stay below the line $y = x$.

Example. The path depicted in Figure 8 is a route.

The problem of counting routes is significantly more difficult then just counting up and right paths. However, we are able to accomplish this using the Catalan Numbers.

Lemma 2.5. *the number of routes from* $(0, 0)$ *to* (n, n) *is the* n_{th} *Catalan Number where*

$$C_n = \frac{1}{n+1}\binom{2n}{n},$$

and $n \geq 0$.

One way you can show this is induction. However, there are many nice visual proofs that can be found on YouTube. This is one of them [reference].

Now that we had a brief introduction to Catalan Numbers, we can start talking about an interesting determinant identity.

Claim 2.6. *If we have*

$$S = \begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_n \\ C_1 & C_2 & C_3 & \cdots & C_{n+1} \\ C_2 & C_3 & C_4 & \cdots & C_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_n & C_{n+1} & C_{n+2} & \cdots & C_{2n} \end{bmatrix},$$

then $\det(A) = 1$.

Proving this identity with a Leibniz Expansion seems pretty difficult. This is especially true since you don't know how big n is. But this problem is solved simply using the LGV-Lemma! To do this you can construct an acyclic directed graph G that satisfies the LGV lemma, and whose path matrix is of the form of matrix S. We will show this for the case $n = 3$ but the proof for all n is the same.

Proof. We construct the graph G in Figure 9. We say that the horizontal edges in the graph have direction going to the right, and all vertical edges have direction going upwards. We also assume that each edge has weight 1. We make it so that our first set of vertices is $\mathcal{O} = \{O_0, O_1, O_2, O_3\}$ (origins) and our second set $\mathcal{D} = \{D_0, D_1, D_2, D_3\}$ (destinations). Clearly with these conditions, G satisfies the assumptions for the LGV-Lemma.

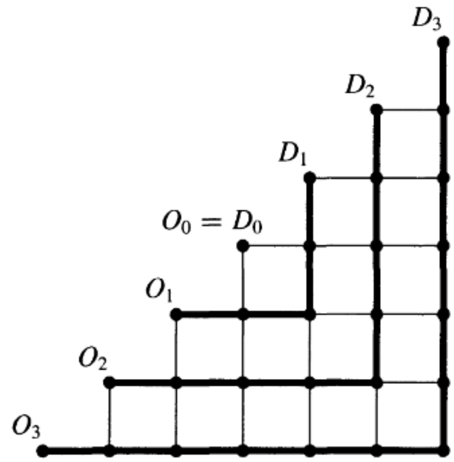

Figure 9. The graph G for $n = 3$. Take from [reference].

We want to determine the path matrix for G. The first thing that we can note is that since all edges have weight 1, every path also has weight 1. Thus, if our path matrix M has

$$m_{ij} = \sum_{P:O_i \to D_j} w(P),$$

then m_{ij} is simply the number of paths from O_i to D_j. Is this simple to compute? Yes! It's fairly easy to see that because of the direction of the edges, all paths between origins and destinations must be routes. Thus, we can use Catalan Numbers! Noting that the number of paths between O_i and D_j is C_{i+j}, we get the path matrix

$$M = \begin{bmatrix} C_0 & C_1 & C_2 & C_3 \\ C_1 & C_2 & C_3 & C_4 \\ C_2 & C_3 & C_4 & C_5 \\ C_3 & C_4 & C_5 & C_6 \end{bmatrix},$$

where M is clearly of the form of S but for $n = 3$. Now we would like to show that the determinant of M is 1. But how do we go about this? From the LGV-Lemma we have that

$$\det(M) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

So, we want to find the vertex disjoint path systems. The key observation is that this graph is nonpermutable. This is easy to see just by looking at G. This allows us to use Corollary 2.3 to say that

$$\det(M) = \text{number of vertex-disjoint path systems associated with id }.$$

In fact, we can notice that there one such 1 path system. This path system is the one depicted in Figure 9. Therefore, $\det(M) = 1$. We can do the same thing for all n by simply constructing a large enough graph of the same type. Doing this will prove Claim 2.6.

■

We now explore a theorem Binet and Cauchy. This theorem is yet another striking example of the flexibility of the LGV-Lemma in determinant problems.

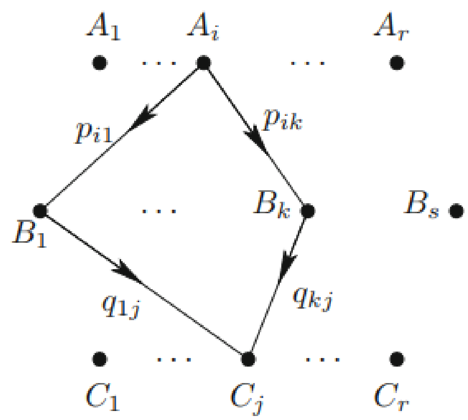

Figure 10. The bipartite graph for Binet-Cauchy. Taken from [reference].

Claim 2.7 (Binet-Cauchy)**.** *If P is a $r \times s$ matrix and Q an $s \times r$ matrix, $r \leq s$, then*

$$\det(PQ) = \sum_{\mathcal{Z}} (\det P_{\mathcal{Z}})(\det Q_{\mathcal{Z}}),$$

where $P_{\mathcal{Z}}$ is the $r \times r$ submatrix of P with column-set $\mathcal{Z}$, and $Q_{\mathcal{Z}}$ the $r \times r$ submatrix with the corresponding rows $\mathcal{Z}$.

Proof. As before, we want to construct a graph whose path matrix is equal to PQ. We construct the graph G so that it is split into 3 disjoint sets of vertices $\mathcal{A} = \{A_1, \ldots, A_r\}$, $\mathcal{B} = \{B_1, \ldots, B_s\}$, and $\mathcal{C} = \{C_1, \ldots, C_r\}$. There will exist no edges between vertices within the same set. We put an edge between every $A \in \mathcal{A}$ and $B \in \mathcal{B}$ with direction going towards B. Similarly, we put an edge between every $B \in \mathcal{B}$ and $C \in \mathcal{C}$ with direction going towards C. Finally, we say that the edge between A_i and B_j has weight p_{ij}, where p_{ij} is $(i,j)_{th}$ entry of P, and the edge between B_i and C_j has weight q_{ij}, where q_{ij} is the $(i,j)_{th}$ entry of Q. This graph is depicted in Figure 10

We must first check if the graph satisfies the conditions of the LGV-Lemma. It is clearly finite and directed. It is also acyclic since all edges are directed “downward”. Thus, there couldn’t possibly exist a cycle since it would need to have an “upward” directed edge. We can also set our two sets of vertices as $\mathcal{A}$ and $\mathcal{C}$. Therefore, G satisfies all the conditions to use the LGV-Lemma!

Now we must construct our path matrix. We know that m_{ij} will be the sum of all the weights of the paths from A_i to C_j. We can notice that to go from A_i to C_j we just need to pick a vertex B_k to go through. The weight of this path will be $p_{ik}q_{kj}$. Since B_k can be any vertex in $\mathcal{B}$, we know that

$$m_{ij} = \sum_{1 \leq k \leq s} p_{ik} q_{kj}.$$

Now that we know what our path matrix looks like, we make the following observation by looking at PQ. Writing out the product in terms of our matrices we have

$$PQ = \begin{pmatrix} p_{11} & \cdots & p_{1s} \\ \vdots & \ddots & \vdots \\ p_{r1} & \cdots & p_{rs} \end{pmatrix} \begin{pmatrix} q_{11} & \cdots & q_{1r} \\ \vdots & \ddots & \vdots \\ q_{s1} & \cdots & q_{sr} \end{pmatrix}.$$

In the matrix multiplication we have that the $(i, j)_{th}$ term of PQ will be the dot product of the i_{th} row of with P with the j_{th} column of Q. This gives us that

$$(i, j)_{th} \text{ entry of } PQ = \sum_{1 \leq k \leq s} p_{ik} q_{kj}.$$

It is now clear why the LGV-Lemma is useful. PQ and the path matrix are the same! So, we can get information about the determinant of PQ by looking at the vertex-disjoint path systems of G. But how do we relate

$$\det(PQ) = \sum_{\substack{\mathcal{P} \text{ vertex-disjoint} \\ \text{path system}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

to Claim 2.7? Well we can first note that every vertex-disjoint path system needs to pass through r of the vertices in $\mathcal{B}$. Therefore, every single path system can be sorted into groups based on which which r vertices in $\mathcal{B}$ they pass through (There are $\binom{s}{r}$ subsets of $\mathcal{B}$ with r vertices). Label the r sized subsets of $\mathcal{B}$ as $\mathcal{Z}_1, \ldots, \mathcal{Z}_{\binom{s}{r}}$. Since we can group the vertex disjoint path systems, we can rewrite our above equation as the following.

$$\det(PQ) = \sum_{\substack{\mathcal{P} \text{ vdps} \\ \text{correspond to } \mathcal{Z}_1}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}) \;+\; \cdots \;+ \sum_{\substack{\mathcal{P} \text{ vdps} \\ \text{correspond to } \mathcal{Z}_{\binom{s}{r}}}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

Our goal is to get each of these sums in a form that is close to that Claim 2.7.

Without loss of generality assume that $\mathcal{Z}_1 = \{B_1, B_2, \ldots, B_r\}$. We will attempt to rewrite the term

$$\sum_{\substack{\mathcal{P} \text{ vdps} \\ \text{correspond to } \mathcal{Z}_1}} \operatorname{sign}(\mathcal{P})\, w(\mathcal{P}).$$

We want to tackle this by breaking down our path systems into simpler parts. We first make the observation that each of the vertex-disjoint path systems that pass through the set $\mathcal{Z}_1$ can actually be broken down into two different vertex disjoint path systems. We have one vertex-disjoint path system from $\mathcal{A}$ to $\mathcal{Z}_1$, and another from $\mathcal{Z}_1$ to $\mathcal{C}$. In fact, we can say more. Let the set of all vertex-disjoint path systems between $\mathcal{A}$ and $\mathcal{Z}_1$ be X, and the set of all vertex-disjoint path systems between $\mathcal{Z}_1$ and $\mathcal{C}$ be Y. Then we know that the set of all vertex-disjoint path systems between $\mathcal{A}$ and $\mathcal{C}$ that pass through the vertices in $\mathcal{Z}_1$ are in one to one correspondence with the cartesian product $X \times Y$. this is clear because say that $\mathcal{P}$ from $\mathcal{A}$ to $\mathcal{C}$ can be broken into $\mathcal{P}_1 \in X$ and $\mathcal{P}_2 \in Y$. Then we can simply pair $\mathcal{P}$ with $(\mathcal{P}_1, \mathcal{P}_2)$.

How can we use this information to solve our problem? We can make a very clever observation. Take the $r \times r$ submatrix of P that has its r rows and its first r columns[5]. We will call this matrix $P_{\mathcal{Z}_1}$. We also take the $r \times r$ submatrix of Q with the first r rows and its r columns. This matrix will be denoted $Q_{\mathcal{Z}_1}$. Then we have that

$$\sum_{\substack{\mathcal{P} \text{ vdps} \\ \text{correspond to } \mathcal{Z}_1}} \text{sign}(\mathcal{P})\ w(\mathcal{P}) = \det(P_{\mathcal{Z}_1}) \det(Q_{\mathcal{Z}_1}).$$

Where does this come from? From the LGV-Lemma we have that

$$\det(P_{\mathcal{Z}_1}) = \sum_{\mathcal{P}_1 \in X} \text{sign}(\mathcal{P}_1)\ w(\mathcal{P}_1),$$

and

$$\det(Q_{\mathcal{Z}_1}) = \sum_{\mathcal{P}_2 \in Y} \text{sign}(\mathcal{P}_2)\ w(\mathcal{P}_2).$$

Therefore, we have that

$$\det(P_{\mathcal{Z}_1}) \det(Q_{\mathcal{Z}_1}) = \left(\sum_{\mathcal{P}_1 \in X} \text{sign}(\mathcal{P}_1)\ w(\mathcal{P}_1) \right) \left(\sum_{\mathcal{P}_2 \in Y} \text{sign}(\mathcal{P}_2)\ w(\mathcal{P}_2) \right).$$

Let us denote the set of all vertex-disjoint path systems from $\mathcal{A}$ to $\mathcal{Z}_1$ as X, and the set of all vertex-disjoint path systems from $\mathcal{Z}_1$ to $\mathcal{C}$ as Y. Then expanding the product on the right hand side we get

$$\sum_{\mathcal{P}_1 \in X \text{ and } \mathcal{P}_2 \in Y} \text{sign}(\mathcal{P}_1) w(\mathcal{P}_1) \text{sign}(\mathcal{P}_2) w(\mathcal{P}_2).$$

Now recall that the set of vertex-disjoint path systems between $\mathcal{A}$ and $\mathcal{C}$ that pass through the vertices in $\mathcal{Z}_1$ are in one to one correspondence with the cartesian product $X \times Y$. Take the path system $\mathcal{P}$ between $\mathcal{A}$ and $\mathcal{C}$, passing through $\mathcal{Z}_1$, and its components $\mathcal{P}_1 \in X$ and $\mathcal{P}_2 \in Y$. Then we can notice that

$$w(\mathcal{P}) = w(\mathcal{P}_1) w(\mathcal{P}_2).$$

Additionally, say that σ is the permutation corresponding with $\mathcal{P}$, σ_1 the permutation corresponding with $\mathcal{P}_1$, and σ_2 the permutation corresponding with $\mathcal{P}_2$. Then we will have that

$$\sigma = \sigma_2 \circ \sigma_1.$$

Now we introduce the following lemma:

Lemma 2.8. *If $\sigma_1, \sigma_2 \in S_n$, then*

$$\text{sign}(\sigma_2 \circ \sigma_1) = \text{sign}(\sigma_2)\text{sign}(\sigma_1)$$

A proof can be found here [reference]. Applying this we see that sign(σ)=sign(σ_2)sign(σ_1). Therefore,

$$\text{sign}(\mathcal{P}) = \text{sign}(\mathcal{P}_1)\text{sign}(\mathcal{P}_2).$$

This gives us

$$\sum_{\substack{\mathcal{P} \text{ vdps} \\ \text{correspond to } \mathcal{Z}_1}} \text{sign}(\mathcal{P})\ w(\mathcal{P}) = \det(P_{\mathcal{Z}_1}) \det(Q_{\mathcal{Z}_1}).$$

[5]We do this since the numbers of the columns correspond to the subscripts of the vertices in $\mathcal{Z}_1$.

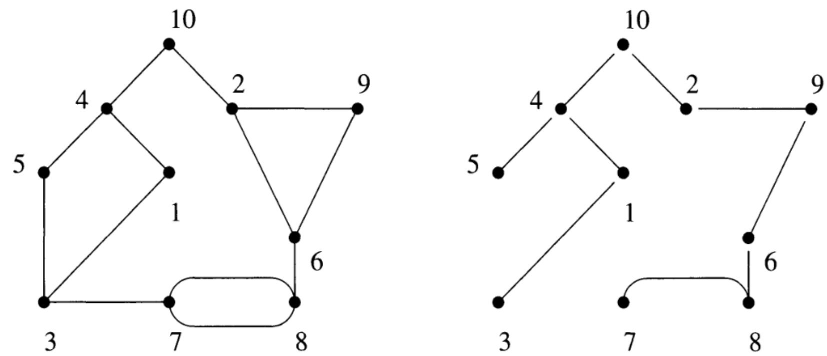

Figure 11. A connected graph G (left) and its spanning tree (right). Taken from [reference].

We showed this for $\mathcal{Z}_1 = \{B_1, B_2, \ldots, B_r\}$. But we can apply this same logic to any r sized subset of $\mathcal{Z} \in \mathcal{B}$ by only changing the part in the proof where we construct the submatrix. In the general case we pick the columns that correspond to the subscripts of the vertices in $\mathcal{Z}$ for our submatrix $P_{\mathcal{Z}}$, and the rows that correspond to the same subscripts for the submatrix $Q_{\mathcal{Z}}$. The Binet-Cauchy Theorem easily follows. ■

3. The Matrix-Tree Theorem

The Matrix-Tree Theorem was discovered by a man named Gustav Kirchhoff in the 1880s. His theorem gives the number of spanning trees of a graph G in terms of the determinant of a matrix derived from G. Before we can introduce the theorem, we need to talk about some definitions. We first talk about the notion of a connected graph.

Definition 3.1. An undirected graph $G = (V, E)$ is *connected* if for every $u, v \in V$ there exists a path between u and v.

Example. Both graphs depicted in Figure 11 are connected graphs.

We also define what it means for a graph to be loopless.

Definition 3.2. We say that a graph is *loopless* if there doesn't exist an edge going from a vertex to itself.

Example. Both graphs depicted in Figure 11 are loopless.

Now we define what a spanning tree is.

Definition 3.3. A *spanning tree* of a graph G is a connected acyclic subgraph of G with the same vertex set.

Example. The graph to the right in Figure 11 is a spanning tree of the graph to the left.

Now just as in the LGV-Lemma, we want to construct a matrix that counts what we want in the graph. In this case, spanning trees. How do we do this here? We do this by first defining something called the adjacency matrix.

Definition 3.4. We define the *adjacency matrix* A of a graph G as the matrix with its a_{ij} entry equal to the number of edges between vertex i and j.

$$D - A = \begin{bmatrix} 2 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & -1 \\ -1 & 0 & 3 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 3 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 3 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 3 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -2 & 3 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 2 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

Figure 12. The matrix $D - A$ for the graph in Figure 11.

We also define a diagonal matrix.

Definition 3.5. We define the *diagonal matrix* D of a graph G as the matrix with d_{ii} equal to the degree of the vertex i. All entries not along the diagonal are equal to 0.

The matrix we work with in the Matrix-Tree Theorem is actually the matrix $D - A$.

Example. The matrix in Figure 12 is $D - A$ for the graph depicted in Figure 11.

Is the determinant of this matrix the number of spanning trees in G? No! Notice that the sum of the entries in each row of $D - A$ is actually 0. This is because in row i, the i_{th} column will just the degree of vertex i. But each of the other columns will have the negative of the number of edges from i to another vertex. Therefore, adding all the entries in the row together must give us 0. But this means that the columns of $D-A$ add to 0 making them linearly dependent and the matrix non-invertible. This forces the determinant of $D-A$ to be 0.

The Matrix-Tree Theorem instead looks at the submatrix that comes from deleting any row of and column of $D - A$. The determinant of this matrix will give us the number of spanning trees of G.

Theorem 3.6. *Let G be a loopless undirected graph with n vertices, adjacency matrix A, and diagonal degree matrix D. If B_{rs} signifies the $(n-1) \times (n-1)$ matrix obtained by deleting from $D - A$ its r_{th} row and s_{th} column, then the number of spanning trees of G is equal to $(-1)^{r+s} \det(B_{rs})$ for any choice of r and s.*

In this paper we will only show the proof of the case when $r = s = n$. However, a proof for the general case can be found here [reference]. The proof will be of the same flavor as our previous proof for the LGV-Lemma. We will show that the acyclic graphs are counted in the determinant of B_{nn}, but then show that the cyclic graphs that are counted will some how cancel each other out and leave us with only the acyclic ones.

Proof. We first note that for any spanning tree of G, there is only one way to orient all the edges in such a way that they point at the vertex n [6]. We can also see that the trees must have $n-1$ edges and the outdegree of every nonroot vertex is 1. Essentially what this graph

[6] This is easy to see just by thinking about the vertices in layers of connections. Combining with the fact that the spanning tree is acyclic gives us this result.

is is a function from the set $\{1, \ldots, n-1\}$ to $\{1, \ldots, n\}$. The function maps the vertex $i \in \{1, \ldots, n-1\}$ to the vertex $j \in \{1, \ldots, n\}$ that it is oriented towards.

Example. The rooted spanning tree in Figure 13 has $f(1) = f(5) = 4$.

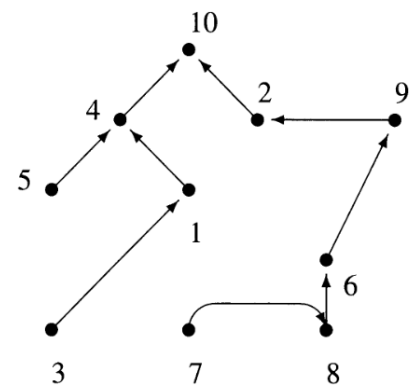

Figure 13. The spanning tree of Figure 11 rooted at $n = 10$.

In general, we call any spanning tree in which every nonroot vertex has outdegree 1 a *functional digraph*. Let the set of all functional digraphs in G be denoted by $\mathcal{F}$. We can now note that all vertices in a functional digraph either lead to n or are contained in a cycle that doesn't contain n. The acyclic functional digraphs are spanning trees of G.

Now we consider the set of signed functional digraphs $\mathcal{S}$. This is the set of all $F \in \mathcal{F}$ but we now give every cycle in F a sign (positive of negative). Note that a functional digraph with k cycles will have 2^k copies in $\mathcal{S}$.

Example. The graph depicted in Figure 14 is a typical example of a signed functional digraph. We can note that the functional digraph in this figure actually has 4 signed copies including this one.

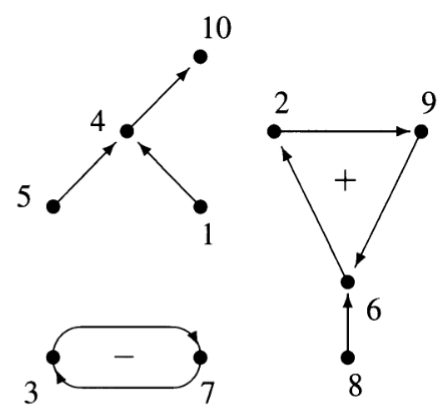

Figure 14. A typical example of a signed functional digraph.

$$
B = \begin{bmatrix}
1+\mathbf{1} & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1+\mathbf{1}+1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\
-1 & 0 & 1+1+1 & 0 & -1 & 0 & \mathbf{-1} & 0 & 0 \\
-1 & 0 & 0 & 1+1+\mathbf{1} & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & -1 & 1+\mathbf{1} & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & \mathbf{1}+1+1 & 0 & -1 & -1 \\
0 & 0 & \mathbf{-1} & 0 & 0 & 0 & 1+1+1 & -1-1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & -1-1 & \mathbf{1}+1+1 & 0 \\
0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 1+\mathbf{1}
\end{bmatrix}
$$

Figure 15. The B matrix for the graph in Figure 11.

Now we define the sign of a graph $S \in \mathcal{S}$ to be the product of the signs of its cycles. If S is acyclic, then it is a spanning tree and its sign is positive. If S is cyclic (contains at least one cycle) then we define its conjugate $\overline{S}$ as the graph with the same functional digraph but with the sign of its first cycle reversed. We define the first cycle of S as the cycle which contains the small vertex.

Example. The first cycle of the graph depicted in Figure 14 is the cycle containing the vertices $2, 9,$ and 6 .

We can notice that $\text{sign}(\overline{S}) = -\text{sign}(S)$ and that $\overline{\overline{S}} = S$. This should look similar as it is the same strategy we used in the proof of the LGV-Lemma except we used the signs of path systems. In fact, because of this condition we know that there is a one-to-one correspondence between the negative signed functional digraphs and positive signed cyclic functional digraphs.

Now we want to somehow relate B (the submatrix gained from deleting the 10_{th} row and 10_{th} column of $D - A$) to this idea. Suppose in G that there is a directed edge from $i \neq n$ to $j \neq n$. Then the directed edge is represented twice in the matrix B. Once on the diagonal positively at b_{ii} and once negatively at b_{ij}. A directed edge from i to n is represented positively only once. Therefore, we can see that every signed function digraph $S \in \mathcal{S}$ is associated with $n - 1$ 1s or -1s in the matrix B.

Example. The bolded 1s in the matrix in Figure 15 represent the signed functional digraph depicted in Figure 14.

So essentially we are trying to count combinations of $n - 1$ 1s and -1s in the matrix B. This motivates us to bring back our Leibniz expansion:

$$\det B = \sum_{\sigma \in S_9} \text{sign}(\sigma)\ b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{9\sigma(9)}.$$

In fact, what we can observe is that non-zero terms in this summation will be in one-to-one correspondence with the set $\mathcal{S}$. This is because every $S \in \mathcal{S}$ can be represented in the determinant by its combination of 1s and -1s, and every combination of 1s and -1s in determinant represents a signed functional digraph [7].

[7] Trying examples with the matrix B in Figure 15 will help you see why this is true.

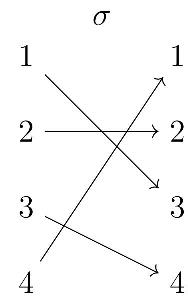

Figure 16. A cyclic permutation.

Now we can see that if a signed functional digraph S has m negative edges, then its contribution to $\det(B)$ will be $X_S = (-1)^m \mathrm{sign}(\pi_S)$ where π_s is the permutation associated with S. All that is left to show is that $X_S = \mathrm{sign}(S)$. Then we will have that

$$\det(B) = \# \text{ of spanning trees.}$$

To do this we will introduce the notion of a cyclic permutation.

Definition 3.7. A permutation σ from $\{1,\ldots,n\} \to \{1,\ldots,n\}$ is called *cyclic* if there exists $a_1 < a_2 < \cdots < a_i \in \{1,\ldots,n\}$ such that

$$\sigma(a_j) = a_{j+1 (\text{mod } i)},$$

and the rest are fixed under the permutation.

The reason why the permutation is called cyclic is clear since the permutation cycles the a_is.

Example. The permutation in Figure 16 is cyclic since $f(1) = 3,\ f(3) = 4,\ f(4) = 1$, and 2 remains fixed.

Remark 3.8. We should note here that the sign of a cyclic permutation where k is the number of elements in the cycle is $(-1)^{k-1}$.

This is incredibly important since if $C_1, C_2, \ldots, C_k$ are the negative cycles that are in S, then $\pi_S = \pi_1 \circ \pi_2 \circ \cdots \circ \pi_k$ where π_i is the natural cyclic permutation associated with C_i. Thus, we are able to apply Lemma 2.8 to say that

$$\mathrm{sign}(\pi_S) = \mathrm{sign}(\pi_1 \circ \pi_2 \circ \cdots \circ \pi_k). = \mathrm{sign}(\pi_1)\cdots\mathrm{sign}(\pi_k).$$

Now say we have that m_i is the number of edges in the cycle C_i and m is the total number of negative edges. Then we have that $m_1 + \cdots + m_k = m$. Thus, we get

$$\begin{aligned} X_S &= (-1)^m \,\mathrm{sign}(\pi_S) = (-1)^m \,\mathrm{sign}(\pi_1 \cdots \pi_k) \\ &= (-1)^{m_1+\cdots+m_k}\, \mathrm{sign}(\pi_1)\cdots\mathrm{sign}(\pi_k) \\ &= (-1)^{m_1+\cdots+m_k}\, (-1)^{m_1-1}\cdots(-1)^{m_k-1} \\ &= (-1)^k = \mathrm{sign}(S). \end{aligned}$$

This proves Theorem 3.6 for the case $r = s = n$. ■

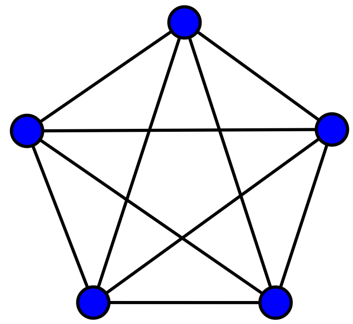

Figure 17. The complete graph K_5.

4. An Application of the Matrix-Tree Theorem

We will first talk about the definition of a complete graph.

Definition 4.1. We say that a loopless graph $G = (V, E)$ is *complete* if for every $u, v \in V$ there exists exactly one edge between u and v. It is denoted K_n where $n = |V|$.

Example. The graph depicted in Figure 17 is the complete graph K_5.

Remark 4.2. It is important to note here that complete graph K_n has $\binom{n}{2}$ edges in total, and every vertex has degree $n - 1$.

So, given a complete graph K_n can we find the number of spanning trees? Yes! We make the following claim:

Claim 4.3. *The complete graph K_n has n^{n-2} spanning trees.*

Proof. We will first construct our adjacency matrix A. This will just be $n \times n$ matrix with 0s along the diagonal and 1s everywhere else. This is clear because by definition every pair of distinct vertices in a complete graph has exactly 1 edge between them. The diagonal matrix D will just be the $n \times n$ matrix with $n - 1$s along the diagonal and 0s everywhere else. This is because every vertex has degree $n - 1$. Therefore $D - A$ is the matrix with $n - 1$s along the diagonal and -1s everywhere else.

Now we can recall that the Matrix-Tree theorem actually requires us to create a submatrix from deleting a row and a column. We will just call B the matrix in which we delete the last row and last column of $D - A$. Since $r = s = n$, the determinant of B is the number of spanning trees of K_n. Now notice that we can actually write B is a clever way. Observe that

$$B = nI - J,$$

where I is the $n - 1 \times n - 1$ identity matrix and J is the $n - 1 \times n - 1$ matrix filled with 1s. What is so special about this form? It allows us to easily see the eigenvalues! We can see that because J has all ones, all the columns are the same. Thus, it has rank 1. Therefore, we can apply Rank-Nullity.

Theorem 4.4 (Rank-Nullity)**.** *For any matrix M we have that the*

$$\text{rank}(M) + \text{null}(M) = \# \text{ of columns in } M.$$

This theorem is really better understood when you think of matrices as linear transformations. The rank of a matrix is the dimension of its image, and the nullity is the dimension of its kernel. A proof can be found here: [reference]. By applying Rank-Nullity to J, we are able to see that the dimension of the kernel of J is $n-2$. Since 0 is an eigenvalue of J with geometric multiplicity (gemu) $n-2$, we get that n is an eigenvalue of B with geometric multiplicity $n-2$. This is because $Bv = nv \Leftrightarrow Jv = 0$ for $v \in \mathbb{R}^{n-1}$. Additionally, we see that if multiply the vector $u = [1, 1 \ldots, 1]$, where u has $n-1$ $1s$, with B we will get u back. Therefore, 1 is also an eigenvalue with geometric multiplicity at least 1. However, we must have that

$$\mathrm{gemu}(n) + \mathrm{gemu}(1) \leq n - 1.$$

Since gemu(n)= $n-2$, we must have that gemu(1)=1. This lets us know that n and 1 are the only eigenvalues of B. Why are the eigenvalues imporant? There is actually a theorem lemma relating the determinant of a matrix B with its eigenvalues.

Lemma 4.5. *Given a matrix M with eigenvalues $\lambda_1, \ldots, \lambda_i$ we have*

$$\det(M) = \prod_{j=1}^{i} \lambda_j^{\mathrm{almu}(\lambda_j)},$$

where almu *is the algebraic multiplicity.*

This is pretty easy to prove by simply looking at the characteristic polynomial of M and using Vieta's Formulas. A full proof is found here: [reference]. But how is the almu related to the gemu? It is through the following lemma:

Lemma 4.6. *Given a matrix M with eigenvalue λ we have*

$$\mathrm{gemu}(\lambda) \leq \mathrm{almu}(\lambda).$$

Combining this statement with the fact the sum of the algebraic multiplicities over all eigenvalues of B is $n-1$, we have almu(n)= $n-2$ and almu(1)=1. Thus, applying Lemma **??** we get $\det(B) = n^{n-2}$. Therefore, K_n has n^{n-2} spanning trees. ■