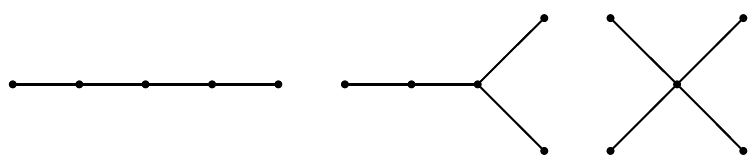
THE MATRIX-TREE THEOREM AND RELATED RESULTS

9 The Matrix-Tree Theorem.

The Matrix-Tree Theorem is a formula for the number of spanning trees of a graph in terms of the determinant of a certain matrix. We begin with the necessary graph-theoretical background. Let G be a finite graph, allowing multiple edges but not loops. (Loops could be allowed, but they turn out to be completely irrelevant.) We say that G is *connected* if there exists a walk between any two vertices of G. A *cycle* is a closed walk with no repeated vertices or edges, except for the the first and last vertex. A *tree* is a connected graph with no cycles. In particular, a tree cannot have multiple edges, since a double edge is equivalent to a cycle of length two. The three nonisomorphic trees with five vertices are given by:



A basic theorem of graph theory (whose easy proof we leave as an exercise) is the following.

9.1 Proposition. *Let G be a graph with p vertices. The following conditions are equivalent.*

(a) *G is a tree.*

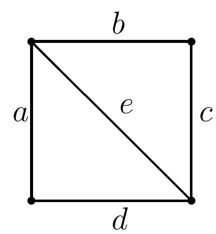
(b) *G is connected and has $p-1$ edges.*

(c) *G is has no cycles and has $p-1$ edges.*

(d) *There is a unique path (= walk with no repeated vertices) between any two vertices.*

A *spanning subgraph* of a graph G is a graph H with the same vertex set as G, and such that every edge of H is an edge of G. If G has q edges, then the number of spanning subgraphs of G is equal to 2^q, since we can choose any subset of the edges of G to be the set of edges of H. (Note that multiple edges between the same two vertices are regarded as *distinguishable*, in accordance with the definition of a graph in Section 1.) A spanning subgraph which is a tree is called a *spanning tree*. Clearly G has a spanning tree if and only if it is connected [why?]. An important invariant of a graph G is its number of spanning trees, called the *complexity* of G and denoted $\kappa(G)$.

9.2 Example. Let G be the graph illustrated below, with edges a, b, c, d, e.



Then G has eight spanning trees, namely, *abc*, *abd*, *acd*, *bcd*, *abe*, *ace*, *bde*, and *cde* (where, e.g., *abc* denotes the spanning subgraph with edge set $\{a, b, c\}$).

9.3 Example. Let $G = K_5$, the complete graph on five vertices. A simple counting argument shows that K_5 has 60 spanning trees isomorphic to the first tree in the above illustration of all nonisomorphic trees with five vertices, 60 isomorphic to the second tree, and 5 isomorphic to the third tree. Hence $\kappa(K_5) = 125$. It is even easier to verify that $\kappa(K_1) = 1$, $\kappa(K_2) = 1$, $\kappa(K_3) = 3$, and $\kappa(K_4) = 16$. Can the reader make a conjecture about the value of $\kappa(K_p)$ for any $p \geq 1$?

Our object is to obtain a "determinantal formula" for $\kappa(G)$. For this we need an important result from matrix theory which is often omitted from a beginning linear algebra course. (Later (Theorem 10.4) we will prove a more general determinantal formula without the use of the Binet-Cauchy theorem. However, the use of the Binet-Cauchy theorem does afford some additional algebraic insight.) This result, known as the Binet-Cauchy theorem (or sometimes as the Cauchy-Binet theorem), is a generalization of the familiar fact that if A and B are $n \times n$ matrices, then $\det(AB) = \det(A)\det(B)$ (where det denotes determinant). We want to extend this formula to the case where A and B are rectangular matrices whose product is a square matrix (so that $\det(AB)$ is defined). In other words, A will be an $m \times n$ matrix and B an $n \times m$ matrix, for some $m, n \geq 1$.

We will use the following notation involving submatrices. Suppose $A = (a_{ij})$ is an $m \times n$ matrix, with $1 \leq i \leq m$, $1 \leq j \leq n$, and $m \leq n$. Given an m-element subset S of $\{1, 2, \ldots, n\}$, let $A[S]$ denote the $m \times m$ submatrix of A obtained by taking the columns indexed by the elements of S. In other words, if the elements of S are given by $j_1 < j_2 < \cdots < j_m$, then $A[S] = (a_{i,j_k})$, where $1 \leq i \leq m$ and $1 \leq k \leq m$. For instance, if

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$$

and $S = \{2, 3, 5\}$, then

$$A[S] = \begin{bmatrix} 2 & 3 & 5 \\ 7 & 8 & 10 \\ 12 & 13 & 15 \end{bmatrix}.$$

Similarly, let $B = (b_{ij})$ be an $n \times m$ matrix with $1 \leq i \leq n$, $1 \leq j \leq m$ and $m \leq n$. Let S be an m-element subset of $\{1, 2, \ldots, n\}$ as above. Then $B[S]$ denotes the $m \times m$ matrix obtained by taking the *rows* of B indexed by S. Note that $A^t[S] = A[S]^t$, where ^{t} denotes transpose.

9.4 Theorem. (the Binet-Cauchy Theorem) *Let $A = (a_{ij})$ be an $m \times n$ matrix, with $1 \leq i \leq m$ and $1 \leq j \leq n$. Let $B = (b_{ij})$ be an $n \times m$ matrix with $1 \leq i \leq n$ and $1 \leq j \leq m$. (Thus AB is an $m \times m$ matrix.) If $m > n$, then* $\det(AB) = 0$. *If $m \leq n$, then*

$$\det(AB) = \sum_S (\det A[S])(\det B[S]),$$

where S ranges over all m-element subsets of $\{1, 2, \ldots, n\}$.

Before proceeding to the proof, let us give an example. We write $|a_{ij}|$ for the determinant of the matrix (a_{ij}). Suppose

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}, \quad B = \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \\ c_3 & d_3 \end{bmatrix}.$$

Then

$$\det(AB) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_3 & d_3 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix}.$$

Proof of Theorem 9.4 (sketch). First suppose $m > n$. Since from linear algebra we know that $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$ and that the rank of an $m \times n$ matrix cannot exceed n (or m), we have that $\operatorname{rank}(AB) \leq n < m$. But AB is an $m \times m$ matrix, so $\det(AB) = 0$, as claimed.

Now assume $m \leq n$. We use notation such as M_{rs} to denote an $r \times s$ matrix M. It is an immediate consequence of the definition of matrix multiplication (which the reader should check) that

$$\begin{bmatrix} R_{mm} & S_{mn} \\ T_{nm} & U_{nn} \end{bmatrix} \begin{bmatrix} V_{mn} & W_{mm} \\ X_{nn} & Y_{nm} \end{bmatrix} = \begin{bmatrix} RV + SX & RW + SY \\ TV + UX & TW + UY \end{bmatrix}. \qquad (59)$$

In other words, we can multiply "block" matrices of suitable dimensions as if their entries were numbers. Note that the entries of the right-hand side of (59) all have well-defined dimensions (sizes), e.g., $RV + SX$ is an $m \times n$ matrix since both RV and SX are $m \times n$ matrices.

Now in equation (59) let $R = I_m$ (the $m \times m$ identity matrix), $S = A$, $T = O_{nm}$ (the $n \times m$ matrix of 0's), $U = I_n$, $V = A$, $W = O_{mm}$, $X = -I_n$, and $Y = B$. We get

$$\begin{bmatrix} I_m & A \\ O_{nm} & I_n \end{bmatrix} \begin{bmatrix} A & O_{mm} \\ -I_n & B \end{bmatrix} = \begin{bmatrix} O_{mn} & AB \\ -I_n & B \end{bmatrix}. \qquad (60)$$

Take the determinant of both sides of (60). The first matrix on the left-hand side is an upper triangular matrix with 1's on the main diagonal. Hence its determinant is one. Since the determinant of a product of square matrices is the product of the determinants of the factors, we get

$$\left| \begin{array}{cc} A & O_{mm} \\ -I_n & B \end{array} \right| = \left| \begin{array}{cc} O_{mn} & AB \\ -I_n & B \end{array} \right|. \tag{61}$$

It is easy to see [why?] that the determinant on the right-hand side of (61) is equal to $\pm\det(AB)$. So consider the left-hand side. A nonzero term in the expansion of the determinant on the left-hand side is obtained by taking the product (with a certain sign) of $m+n$ nonzero entries, no two in the same row and column (so one in each row and each column). In particular, we must choose m entries from the last m columns. These entries belong to m of the bottom n rows [why?], say rows $m+s_1, m+s_2, \ldots, m+s_m$. Let $S = \{s_1, s_2, \ldots, s_m\} \subseteq \{1, 2, \ldots, n\}$. We must choose $n-m$ further entries from the last n rows, and we have no choice but to choose the -1's in those rows $m+i$ for which $i \notin S$. Thus every term in the expansion of the left-hand side of (61) uses exactly $n-m$ of the -1's in the bottom left block $-I_n$.

What is the contribution to the expansion of the left-hand side of (61) from those terms which use exactly the -1's from rows $m+i$ where $i \notin S$? We obtain this contribution by deleting all rows and columns to which these -1's belong (in other words, delete row $m+i$ and column i whenever $i \in \{1, 2, \ldots, n\} - S$), taking the determinant of the $2m \times 2m$ matrix M_S that remains, and multiplying by an appropriate sign [why?]. But the matrix M_S is in block-diagonal form, with the first block just the matrix $A[S]$ and the second block just $B[S]$. Hence $\det M_S = (\det A[S])(\det B[S])$ [why?]. Taking all possible subsets S gives

$$\det AB = \sum_{\substack{S \subseteq \{1,2,\ldots,n\} \\ |S|=m}} \pm(\det A[S])(\det B[S]).$$

It is straightforward but somewhat tedious to verify that all the signs are $+$; we omit the details. This completes the proof. □

In Section 1 we defined the adjacency matrix $\boldsymbol{A}(G)$ of a graph G with vertex set $V = \{v_1, \ldots, v_p\}$ and edge set $E = \{e_1, \ldots, e_q\}$. We now define two related matrices. Assume for simplicity that G has no loops. (This assumption is harmless since loops have no effect on $\kappa(G)$.)

9.5 Definition. Let G be as above. Give G an *orientation* $\mathfrak{o}$, i.e, for every edge e with vertices u, v, choose one of the ordered pairs (u, v) or (v, u). (If we choose (u, v), say, then we think of putting an arrow on e pointing from u to v; and we say that e is directed from u to v, that u is the *initial vertex* and v the *final vertex* of e, etc.)

(a) The *incidence matrix* $\boldsymbol{M}(G)$ of G (with respect to the orientation $\mathfrak{o}$) is the $p \times q$ matrix whose (i, j)-entry $\boldsymbol{M}_{ij}$ is given by

$$\boldsymbol{M}_{ij} = \begin{cases} 1, & \text{if the edge } e_j \text{ has initial vertex } v_i \\ -1, & \text{if the edge } e_j \text{ has final vertex } v_i \\ 0, & \text{otherwise.} \end{cases}$$

(b) The *laplacian matrix* $\boldsymbol{L}(G)$ of G is the $p \times p$ matrix whose (i, j)-entry $\boldsymbol{L}_{ij}$ is given by

$$\boldsymbol{L}_{ij} = \begin{cases} -m_{ij}, & \text{if } i \neq j \text{ and there are } m_{ij} \text{ edges between } v_i \text{ and } v_j \\ \deg(v_i), & \text{if } i = j, \end{cases}$$

where $\deg(v_i)$ is the number of edges incident to v_i. (Thus $\boldsymbol{L}(G)$ is symmetric and does not depend on the orientation $\mathfrak{o}$.)

Note that every column of $\boldsymbol{M}(G)$ contains one 1, one -1, and $q - 2$ 0's; and hence the sum of the entries in each column is 0. Thus all the rows sum to the 0 vector, a linear dependence relation which shows that $\operatorname{rank}(\boldsymbol{M}(G)) < p$. Two further properties of $\boldsymbol{M}(G)$ and $\boldsymbol{L}(G)$ are given by the following lemma.

9.6 Lemma. (a) *We have* $\boldsymbol{M}\boldsymbol{M}^t = \boldsymbol{L}$.

(b) *If G is regular of degree d, then $\boldsymbol{L}(G) = dI - \mathbf{A}(G)$, where $\boldsymbol{A}(G)$ denotes the adjacency matrix of G. Hence if G (or $\boldsymbol{A}(G)$) has eigenvalues $\lambda_1, \ldots, \lambda_p$, then $\boldsymbol{L}(G)$ has eigenvalues $d - \lambda_1, \ldots, d - \lambda_p$.*

Proof. (a) This is immediate from the definition of matrix multiplication. Specifically, for $v_i, v_j \in V(G)$ we have

$$(\boldsymbol{M}\boldsymbol{M}^t)_{ij} = \sum_{e_k \in E(G)} \boldsymbol{M}_{ik}\boldsymbol{M}_{jk}.$$

If $i \neq j$, then in order for $\boldsymbol{M}_{ik}\mathbf{M}_{jk} \neq 0$, we must have that the edge e_k connects the vertices v_i and v_j. If this is the case, then one of $\boldsymbol{M}_{ik}$ and $\boldsymbol{M}_{jk}$ will be 1 and the other -1 [why?], so their product is always -1. Hence $(\boldsymbol{M}\boldsymbol{M}^t)_{ij} = -m_{ij}$, as claimed.

There remains the case $i = j$. Then $\boldsymbol{M}_{ik}\boldsymbol{M}_{ik}$ will be 1 if e_k is an edge with v_i as one of its vertices and will be 0 otherwise [why?]. So now we get $(\boldsymbol{M}\boldsymbol{M}^t)_{ii} = \deg(v_i)$, as claimed. This proves (a).

(b) Clear by (a), since the diagonal elements of $\boldsymbol{M}\boldsymbol{M}^t$ are all equal to d.
□

Now assume that G is connected, and let $\boldsymbol{M}_0(G)$ be $\mathbf{M}(G)$ with its last row removed. Thus $\boldsymbol{M}_0(G)$ has $p-1$ rows and q columns. Note that the number of rows is equal to the number of edges in a spanning tree of G. We call $\boldsymbol{M}_0(G)$ the *reduced incidence matrix* of G. The next result tells us the determinants (up to sign) of all $(p-1)\times(p-1)$ submatrices N of $\boldsymbol{M}_0$. Such submatrices are obtained by choosing a set S of $p-1$ edges of G, and taking all columns of $\boldsymbol{M}_0$ indexed by the edges in S. Thus this submatrix is just $\boldsymbol{M}_0[S]$.

9.7 Lemma. *Let S be a set of $p-1$ edges of G. If S does not form the set of edges of a spanning tree, then* $\det \boldsymbol{M}_0[S] = 0$. *If, on the other hand, S is the set of edges of a spanning tree of G, then* $\det \boldsymbol{M}_0[S] = \pm 1$.

Proof. If S is not the set of edges of a spanning tree, then some subset R of S forms the edges of a cycle C in G. Suppose that the cycle C defined by R has edges $f_1, \ldots, f_s$ in that order. Multiply the column of $\boldsymbol{M}_0[S]$ indexed by f_j by 1 if in going around C we traverse f_i in the direction of its arrow; otherwise multiply the column by -1. Then add these modified columns. It is easy to see (check a few small examples to convince yourself) that we get the 0 column. Hence the columns of $\boldsymbol{M}_0[S]$ are linearly dependent, so $\det \boldsymbol{M}_0[S] = 0$, as claimed.

Now suppose that S is the set of edges of a spanning tree T. Let e be an edge of T which is connected to v_p (the vertex which indexed the bottom row of $\boldsymbol{M}$, i.e., the row removed to get $\boldsymbol{M}_0$). The column of $\boldsymbol{M}_0[S]$ indexed by e contains exactly one nonzero entry [why?], which is ± 1. Remove from $\boldsymbol{M}_0[S]$

the row and column containing the nonzero entry of column e, obtaining a $(p-2)\times(p-2)$ matrix $\boldsymbol{M}_0'$. Note that $\det(\boldsymbol{M}_0[S]) = \pm\det(\boldsymbol{M}_0')$ [why?]. Let T' be the tree obtained from T by contracting the edge e to a single vertex (so that v_p and the remaining vertex of e are merged into a single vertex u). Then $\boldsymbol{M}_0'$ is just the matrix obtained from the incidence matrix $\boldsymbol{M}(T')$ by removing the row indexed by u [why?]. Hence by induction on the number p of vertices (the case $p=1$ being trivial), we have $\det(\boldsymbol{M}_0') = \pm 1$. Thus $\det(\boldsymbol{M}_0[S]) = \pm 1$, and the proof follows. $\square$

Note. An alternative way of seeing that $\det(\boldsymbol{M}_0 S) = \pm 1$ when S is the set of of edges of a spanning tree T is as follows. Let $u_1, u_2, \dots, u_{p-1}$ be an ordering of the vertices $v_1, \dots, v_{p-1}$ such that u_i is an endpoint of the tree obtained from T by removing vertices $u_1, \dots, u_{i-1}$. (It is easy to see that such an ordering is possible.) Permute the rows of $\boldsymbol{M}_0[S]$ so that the ith row is indexed by u_i. Then we obtain a lower triangular matrix with ± 1's on the main diagonal, so the determinant is ± 1.

We have now assembled all the ingredients for the main result of this section (due originally to Borchardt). Recall that $\kappa(G)$ denotes the number of spanning trees of G.

9.8 Theorem. (the Matrix-Tree Theorem) *Let G be a finite connected graph without loops, with laplacian matrix $\boldsymbol{L} = \boldsymbol{L}(G)$. Let $\boldsymbol{L}_0$ denote $\boldsymbol{L}$ with the last row and column removed (or with the ith row and column removed for any i). Then*

$$\det(\boldsymbol{L}_0) = \kappa(G).$$

Proof. Since $\boldsymbol{L} = \boldsymbol{M}\boldsymbol{M}^t$ (Lemma 9.6(a)), it follows immediately that $\boldsymbol{L}_0 = \boldsymbol{M}_0\boldsymbol{M}_0^t$. Hence by the Binet-Cauchy theorem (Theorem 9.4), we have

$$\det(\boldsymbol{L}_0) = \sum_S (\det \boldsymbol{M}_0[S])(\det \boldsymbol{M}_0^t[S]), \tag{62}$$

where S ranges over all $(p-1)$-element subsets of $\{1, 2\dots, q\}$ (or equivalently, over all $(p-1)$-element subsets of the set of edges of G). Since in general $A^t[S] = A[S]^t$, equation (62) becomes

$$\det(\boldsymbol{L}_0) = \sum_S (\det \boldsymbol{M}_0[S])^2. \tag{63}$$

According to Lemma 9.7, $\det(\boldsymbol{M}_0[S])$ is ± 1 if S forms the set of edges of a spanning tree of G, and is 0 otherwise. Therefore the term indexed by S in the sum on the right-hand side of (63) is 1 if S forms the set of edges of a spanning tree of G, and is 0 otherwise. Hence the sum is equal to $\kappa(G)$, as desired. □

The operation of removing a row and column from $\boldsymbol{L}(G)$ may seem somewhat contrived. We would prefer a description of $\kappa(G)$ directly in terms of $\boldsymbol{L}(G)$. Such a description will follow from the next lemma.

9.9 Lemma. *Let M be a $p \times p$ matrix (with entries in a field) such that the sum of the entries in every row and column is 0. Let M_0 be the matrix obtained from M by removing the last row and last column (or more generally, any row and any column). Then the coefficient of x in the characteristic polynomial* $\det(M - xI)$ *of M is equal to $-p \cdot \det(M_0)$. (Moreover, the constant term of* $\det(M - xI)$ *is 0.)*

Proof. The constant term of $\det(M - xI)$ is $\det(M)$, which is 0 since the rows of M sum to 0.

For simplicity we prove the rest of the lemma only for removing the last row and column, though the proof works just as well for any row and column. Add all the rows of $M - xI$ except the last row to the last row. This doesn't effect the determinant, and will change the entries of the last row all to $-x$ (since the rows of M sum to 0). Factor out $-x$ from the last row, yielding a matrix $N(x)$ satisfying $\det(M - xI) = -x \det(N(x))$. Hence the coefficient of x in $\det(M - xI)$ is given by $-\det(N(0))$. Now add all the columns of $N(0)$ except the last column to the last column. This does not effect $\det(N(0))$. Because the columns of M sum to 0, the last column of $N(0)$ becomes the column vector $[0, 0, \ldots, 0, p]^t$. Expanding the determinant by the last column shows that $\det(N(0)) = p \cdot \det(M_0)$, and the proof follows. □

9.10 Corollary. (a) *Let G be a connected (loopless) graph with p vertices. Suppose that the eigenvalues of $\boldsymbol{L}(G)$ are $\mu_1, \ldots, \mu_{p-1}, \mu_p$, with $\mu_p = 0$. Then*

$$\kappa(G) = \frac{1}{p}\mu_1\mu_2\cdots\mu_{p-1}.$$

(b) *Suppose that G is also regular of degree d, and that the eigenvalues of*

$\boldsymbol{A}(G)$ *are* $\lambda_1, \ldots, \lambda_{p-1}, \lambda_p$, *with* $\lambda_p = d$. *Then*

$$\kappa(G) = \frac{1}{p}(d - \lambda_1)(d - \lambda_2) \cdots (d - \lambda_{p-1}).$$

Proof. (a) We have

$$\begin{aligned} \det(\boldsymbol{L} - xI) &= (\mu_1 - x) \cdots (\mu_{p-1} - x)(\mu_p - x) \\ &= -(\mu_1 - x)(\mu_2 - x) \cdots (\mu_{p-1} - x)x. \end{aligned}$$

Hence the coefficient of x is $-\mu_1\mu_2 \cdots \mu_{p-1}$. By Lemma 9.9, we get $-\mu_1\mu_2 \cdots \mu_{p-1} = p \cdot \det(\boldsymbol{L}_0)$. By Theorem 9.8 we have $\det(\boldsymbol{L}_0) = \kappa(G)$, and the proof follows.

(b) Immediate from (a) and Lemma 9.6(b). □

Let us look at a couple of examples of the use of the Matrix-Tree Theorem.

9.11 Example. Let $G = K_p$, the complete graph on p vertices. Now K_p is regular of degree $d = p - 1$, and by Proposition 1.5 its eigenvalues are -1 ($p - 1$ times) and $p - 1 = d$. Hence from Corollary 8.10(b) there follows

$$\kappa(K_p) = \frac{1}{p}((p - 1) - (-1))^{p-1} = p^{p-2}.$$

This surprising result is often attributed to Cayley, who stated it without proof in 1889 (and even cited Borchardt explicitly). However, it was in fact stated by Sylvester in 1857, while a proof was published by Borchardt in 1860. It is clear that Cayley and Sylvester could have produced a proof if asked to do so. There are many other proofs known, including elegant combinatorial arguments due to Prüfer, Joyal, and others.

9.12 Example. Let $G = C_n$, the n-cube discussed in Section 2. Now C_n is regular of degree n, and by Corollary 2.5 its eigenvalues are $n - 2i$ with multiplicity $\binom{n}{i}$ for $0 \leq i \leq n$. Hence from Corollary 8.10(b) there follows the amazing result

$$\begin{aligned} \kappa(C_n) &= \frac{1}{2^n} \prod_{i=1}^{n} (2i)^{\binom{n}{i}} \\ &= 2^{2^n - n - 1} \prod_{i=1}^{n} i^{\binom{n}{i}}. \end{aligned}$$

To my knowledge a direct combinatorial proof is not known.

10 Eulerian digraphs and oriented trees.

A famous problem which goes back to Euler asks for what graphs G is there a closed walk which uses every edge exactly once. (There is also a version for non-closed walks.) Such a walk is called an *Eulerian tour* (also known as an *Eulerian cycle*). A graph which has an Eulerian tour is called an *Eulerian graph*. Euler's famous theorem (the first real theorem of graph theory) states that G is Eulerian if and only if it is connected and every vertex has even degree. Here we will be concerned with the analogous theorem for directed graphs. We want to know not just whether an Eulerian tour exists, but how many there are. We will prove an elegant determinantal formula for this number closely related to the Matrix-Tree Theorem. For the case of undirected graphs no analogous formula is known, explaining why we consider only the directed case.

A (finite) *directed graph* or *digraph* D consists of a *vertex set* $V = \{v_1, \ldots, v_p\}$ and edge set $E = \{e_1, \ldots, e_q\}$, together with a function $\varphi : E \to V \times V$ (the set of ordered pairs (u, v) of elements of V). If $\varphi(e) = (u, v)$, then we think of e as an arrow from u to v. We then call u the *initial vertex* and v the *final vertex* of e. (These concepts arose in the definition of an orientation in Definition 8.5.) A *tour* in D is a sequence $e_1, e_2, \ldots, e_r$ of *distinct* edges such that the final vertex of e_i is the initial vertex of e_{i+1} for all $1 \leq i \leq r - 1$, and the final vertex of e_r is the initial vertex of e_1. A tour is *Eulerian* if every edge of D occurs at least once (and hence exactly once). A digraph which has no isolated vertices and contains an Eulerian tour is called an *Eulerian digraph*. Clearly an Eulerian digraph is connected. The *outdegree* of a vertex v, denoted $\text{outdeg}(v)$, is the number of edges of G with initial vertex v. Similarly the *indegree* of v, denoted $\text{indeg}(v)$, is the number of edges of D with final vertex v. A loop (edge of the form (v, v)) contributes one to both the indegree and outdegree. A digraph is *balanced* if $\text{indeg}(v) = \text{outdeg}(v)$ for all vertices v.

10.1 Theorem. *A digraph D is Eulerian if and only if it is connected and balanced.*

Proof. Assume D is Eulerian, and let $e_1, \ldots, e_q$ be an Eulerian tour. As we move along the tour, whenever we enter a vertex v we must exit it, except at the very end we enter the final vertex v of e_q without exiting it. However, at the beginning we exited v without having entered it. Hence every vertex is entered as often as it is exited and so must have the same outdegree as indegree. Therefore D is balanced, and as noted above D is clearly connected.

Now assume that D is balanced and connected. We may assume that D has at least one edge. We first claim that for any edge e of D, D has a tour for which $e = e_1$. If e_1 is a loop we are done. Otherwise we have entered the vertex $\text{fin}(e_1)$ for the first time, so since D is balanced there is some exit edge e_2. Either $\text{fin}(e_2) = \text{init}(e_1)$ and we are done, or else we have entered the vertex $\text{fin}(e_2)$ once more than we have exited it. Since D is balanced there is new edge e_3 with $\text{fin}(e_2) = \text{init}(e_3)$. Continuing in this way, either we complete a tour or else we have entered the current vertex once more than we have exited it, in which case we can exit along a new edge. Since D has finitely many edges, eventually we must complete a tour. Thus D does have a tour which uses e_1.

Now let $e_1, \ldots, e_r$ be a tour C of maximum length. We must show that $r = q$, the number of edges of D. Assume to the contrary that $r < q$. Since in moving along C every vertex is entered as often as it is exited (with $\text{init}(e_1)$ exited at the beginning and entered at the end), when we remove the edges of C from D we obtain a digraph H which is still balanced, though it need not be connected. However, since D is connected, at least one connected component H_1 of H contains at least one edge and has a vertex v in common with C [why?]. Since H_1 is balanced, there is an edge e of H_1 with initial vertex v. The argument of the previous paragraph shows that H_1 has a tour C' of positive length beginning with the edge e. But then when moving along C, when we reach v we can take the "detour" C' before continuing with C. This gives a tour of length longer than r, a contradiction. Hence $r = q$, and the theorem is proved. □

Our primary goal is to count the number of Eulerian tours of a connected balanced digraph. A key concept in doing so is that of an oriented tree. An *oriented tree* with root v is a (finite) digraph T with v as one of its vertices, such that there is a unique directed path from any vertex u to v. In other words, there is a unique sequence of edges $e_1, \ldots, e_r$ such that (a) $\text{init}(e_1) = u$, (b) $\text{fin}(e_r) = v$, and (c) $\text{fin}(e_i) = \text{init}(e_{i+1})$ for $1 \le i \le r - 1$.

It's easy to see that this means that the underlying undirected graph (i.e., "erase" all the arrows from the edges of T) is a tree, and that all arrows in T "point toward" v. There is a surprising connection between Eulerian tours and oriented trees, given by the next result (due to de Bruijn and van Aardenne-Ehrenfest). This result is sometimes called the BEST Theorem, after de **B**ruijn, van Aardenne-**E**hrenfest, **S**mith, and **T**utte. However, Smith and Tutte were not involved in the original discovery.

10.2 Theorem. *Let D be a connected balanced digraph with vertex set V. Fix an edge e of D, and let $v = \mathrm{init}(e)$. Let $\tau(D, v)$ denote the number of oriented (spanning) subtrees of D with root v, and let $\epsilon(D, e)$ denote the number of Eulerian tours of D starting with the edge e. Then*

$$\epsilon(D, e) = \tau(D, v) \prod_{u \in V} (\mathrm{outdeg}(u) - 1)!. \tag{64}$$

Proof. Let $e = e_1, e_2, \ldots, e_q$ be an Eulerian tour E in D. For each vertex $u \neq v$, let $e(u)$ be the "last exit" from u in the tour, i.e., let $e(u) = e_j$ where $\mathrm{init}(e(u)) = u$ and $\mathrm{init}(e_k) \neq u$ for any $k > j$.

Claim #1. The vertices of D, together with the edges $e(u)$ for all vertices $u \neq v$, form an oriented subtree of D with root v.

Proof of Claim #1. This is a straightforward verification. Let T be the spanning subgraph of D with edges $e(u)$, $u \neq v$. Thus if $|V| = p$, then T has p vertices and $p - 1$ edges [why?]. There are three items to check to insure that T is an oriented tree with root v:

(a) T does not have two edges f and f' satisfying $\mathrm{init}(f) = \mathrm{init}(f')$. This is clear since both f and f' can't be last exits from the same vertex.

(b) T does not have an edge f with $\mathrm{init}(f) = v$. This is clear since by definition the edges of T consist only of last exits from vertices other than v, so no edge of T can exit from v.

(c) T does not have a (directed) cycle C. For suppose C were such a cycle. Let f be that edge of C which occurs after all the other edges of C in the Eulerian tour E. Let f' be the edge of C satisfying $\text{fin}(f) = \text{init}(f')$ ($= u$, say). We can't have $u = v$ by (b). Thus when we enter u *via* f, we must exit u. We can't exit u *via* f' since f occurs after f' in E. Hence f' is not the last exit from u, contradicting the definition of T.

It's easy to see that conditions (a)–(c) imply that T is an oriented tree with root v, proving the claim.

Claim #2. We claim that the following converse to Claim #1 is true. Given a connected balanced digraph D and a vertex v, let T be an oriented (spanning) subtree of D with root v. Then we can construct an Eulerian tour E as follows. Choose an edge e_1 with $\text{init}(e_1) = v$. Then continue to choose any edge possible to continue the tour, except we never choose an edge f of E unless we have to, i.e., unless it's the only remaining edge exiting the vertex at which we stand. Then we never get stuck until all edges are used, so we have constructed an Eulerian tour E. Moreover, the set of last exits of E from vertices $u \neq v$ of D coincides with the set of edges of the oriented tree T.

Proof of Claim #2. Since D is balanced, the only way to get stuck is to end up at v with no further exits available, but with an edge still unused. Suppose this is the case. At least one unused edge must be a last exit edge, i.e., an edge of T [why?]. Let u be a vertex of T closest to v in T such that the unique edge f of T with $\text{init}(f) = u$ is not in the tour. Let $y = \text{fin}(f)$. Suppose $y \neq v$. Since we enter y as often as we leave it, we don't use the last exit from y. Thus $y = v$. But then we can leave v, a contradiction. This proves Claim #2.

We have shown that every Eulerian tour E beginning with the edge e has associated with it a "last exit" oriented subtree $T = T(E)$ with root $v = \text{init}(e)$. Conversely, given an oriented subtree T with root v, we can obtain all Eulerian tours E beginning with e and satisfying $T = T(E)$ by choosing for each vertex $u \neq v$ the order in which the edges from u, except the edge of T, appear in E; as well as choosing the order in which all the edges from v except for e appear in E. Thus for each vertex u we have $(\text{outdeg}(u) - 1)!$ choices, so for each T we have $\prod_u (\text{outdeg}(u) - 1)!$ choices. Since there are $\tau(G, v)$ choices for T, the proof is complete. □

10.3 Corollary. *Let D be a connected balanced digraph, and let v be a vertex of D. Then the number $\tau(D, v)$ of oriented subtrees with root v is independent of v.*

Proof. Let e be an edge with initial vertex v. By equation (64), we need to show that the number $\epsilon(G, e)$ of Eulerian tours beginning with e is independent of e. But $e_1 e_2 \cdots e_q$ is an Eulerian tour if and only if $e_i e_{i+1} \cdots e_q e_1 e_2 \cdots e_{i-1}$ is also an Eulerian tour, and the proof follows [why?]. $\square$

What we obviously need to do next is find a formula for $\tau(G, v)$. Such a formula is due to W. Tutte in 1948. This result is very similar to the Matrix-Tree Theorem, and indeed we will show (Example 10.6) that the Matrix-Tree Theorem is a simple corollary to Theorem 10.4.

10.4 Theorem. *Let D be a loopless connected digraph with vertex set $V = \{v_1, \ldots, v_p\}$. Let $\boldsymbol{L}(D)$ be the $p \times p$ matrix defined by*

$$\boldsymbol{L}_{ij} = \begin{cases} -m_{ij}, & \text{if } i \neq j \text{ and there are } m_{ij} \text{ edges with} \\ & \quad \text{initial vertex } v_i \text{ and final vertex } v_j \\ \operatorname{outdeg}(v_i), & \text{if } i = j. \end{cases}$$

(Thus $\boldsymbol{L}$ is the directed analogue of the laplacian matrix of an undirected graph.) Let $\boldsymbol{L}_0$ denote $\boldsymbol{L}$ with the last row and column deleted. Then

$$\det \boldsymbol{L}_0 = \tau(D, v_p). \tag{65}$$

NOTE. If we remove the ith row and column from $\boldsymbol{L}$ instead of the last row and column, then equation (65) still holds with v_p replaced with v_i.

Proof (sketch). Induction on q, the number of edges of D. The fewest number of edges which D can have is $p-1$ (since D is connected). Suppose then that D has $p-1$ edges, so that as an undirected graph D is a tree. If D is not an oriented tree with root v_p, then some vertex $v_i \neq v_p$ of D has outdegree 0 [why?]. Then $\boldsymbol{L}_0$ has a zero row, so $\det \boldsymbol{L}_0 = 0 = \tau(D, v_p)$. If on the other hand D is an oriented tree with root v_p, then an argument like that used to prove Lemma 9.7 (in the case when S is the set of edges of a spanning tree) shows that $\det \boldsymbol{L}_0 = 1 = \tau(D, v_p)$.

Now assume that D has $q > p - 1$ edges, and assume the theorem for digraphs with at most $q - 1$ edges. We may assume that no edge f of D has initial vertex v, since such an edge belongs to no oriented tree with root v and also makes no contribution to $\boldsymbol{L}_0$. It then follows, since D has at least p edges, that there exists a vertex $u \neq v$ of D of outdegree at least two. Let e be an edge with $\mathrm{init}(e) = u$. Let D_1 be D with the edge e removed. Let D_2 be D with all edges e' removed such that $\mathrm{init}(e) = \mathrm{init}(e')$ and $e' \neq e$. (Note that D_2 is strictly smaller than D since $\mathrm{outdeg}(u) \geq 2$.) By induction, we have $\det \boldsymbol{L}_0(D_1) = \tau(D_1, v_p)$ and $\det \boldsymbol{L}_0(D_2) = \tau(D_2, v_p)$. Clearly $\tau(D, v_p) = \tau(D_1, v_p) + \tau(D_2, v_p)$, since in an oriented tree T with root v_p, there is exactly one edge whose initial vertex coincides with that of e. On the other hand, it follows immediately from the multilinearity of the determinant [why?] that

$$\det \boldsymbol{L}_0(D) = \det \boldsymbol{L}_0(D_1) + \det \boldsymbol{L}_0(D_2).$$

From this the proof follows by induction. □

10.5 Corollary. *Let D be a connected balanced digraph with vertex set $V = \{v_1, \ldots, v_p\}$. Let e be an edge of D. Then the number $\epsilon(D, e)$ of Eulerian tours of D with first edge e is given by*

$$\epsilon(D, e) = (\det \boldsymbol{L}_0(D)) \prod_{u \in V} (\mathrm{outdeg}(u) - 1)!.$$

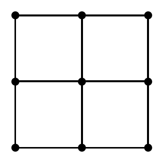
Equivalently (using Lemma 9.9), if $\boldsymbol{L}(D)$ has eigenvalues $\mu_1, \ldots, \mu_p$ with $\mu_p = 0$, then

$$\epsilon(D, e) = \frac{1}{p} \mu_1 \cdots \mu_{p-1} \prod_{u \in V} (\mathrm{outdeg}(u) - 1)!.$$

Proof. Combine Theorems 10.2 and 10.4. □

10.6 Example. (the Matrix-Tree Theorem revisited) Let G be a connected loopless undirected graph. Let $\hat{G}$ be the digraph obtained from G by replacing each edge $e = uv$ of G with a pair of directed edges $u \to v$ and $v \to u$. Clearly $\hat{G}$ is balanced and connected. Choose a vertex v of G. There is an obvious one-to-one correspondence between spanning trees T of G and oriented spanning trees $\hat{T}$ of $\hat{G}$ with root v, namely, direct each edge of T toward v. Moreover, $\mathbf{L}(G) = \boldsymbol{L}(\hat{G})$ [why?]. Hence the Matrix-Tree Theorem is an immediate consequence of the Theorem 10.4.

10.7 Example. (the efficient mail carrier) A mail carrier[2] has an itinerary of city blocks to which he (or she) must deliver mail. He wants to accomplish this by walking along each block twice, once in each direction, thus passing along houses on each side of the street. The blocks form the edges of a graph G, whose vertices are the intersections. The mail carrier wants simply to walk along an Eulerian tour in the digraph $\hat{G}$ of the previous example. Making the plausible assumption that the graph is connected, not only does an Eulerian tour always exist, but we can tell the mail carrier how many there are. Thus he will know how many different routes he can take to avoid boredom. For instance, suppose G is the 3×3 grid illustrated below.



This graph has 128 spanning trees. Hence the number of mail carrier routes beginning with a fixed edge (in a given direction) is $128 \cdot 1!^4 2!^4 3! = 12288$. The total number of routes is thus 12288 times twice the number of edges [why?], viz., $12288 \times 24 = 294912$. Assuming the mail carrier delivered mail 250 days a year, it would be 1179 years before he would have to repeat a route!

10.8 Example. (binary de Bruijn sequences) A *binary sequence* is just a sequence of 0's and 1's. A *binary de Bruijn sequence* of degree n is a binary sequence $A = a_1 a_2 \cdots a_{2^n}$ such that every binary sequence $b_1 \cdots b_n$ of length n occurs exactly once as a "circular factor" of A, i.e., as a sequence $a_i a_{i+1} \cdots a_{i+n-1}$, where the subscripts are taken modulo n if necessary. For instance, some circular factors of the sequence *abcdefg* are *a*, *bcde*, *fgab*, and *defga*. Note that there are exactly 2^n binary sequences of length n, so the only possible length of a binary de Bruijn sequence of degree n is 2^n [why?]. Clearly any cyclic shift $a_i a_{i+1} \cdots a_{2^n} a_1 a_2 \cdots a_{i-1}$ of a binary de Bruijn

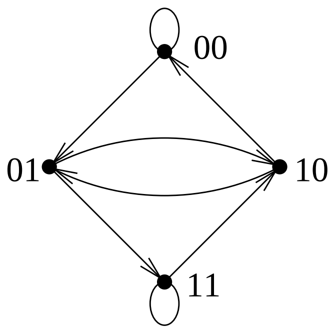
[2] postperson?

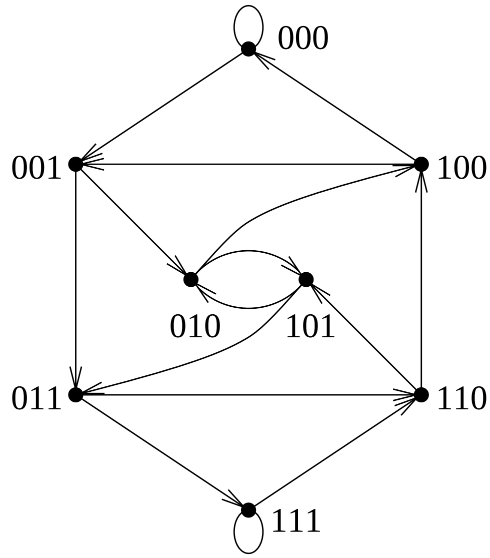
sequence $a_1a_2\cdots a_{2^n}$ is also a binary de Bruijn sequence, and we call two such sequences *equivalent*. This relation of equivalence is obviously an equivalence relation, and every equivalence class contains exactly one sequence beginning with n 0's [why?]. Up to equivalence, there is one binary de Bruijn sequence of degree two, namely, 0011. It's easy to check that there are two inequivalent binary de Bruijn sequences of degree three, namely, 00010111 and 00011101. However, it's not clear at this point whether binary de Bruijn sequences exist for all n. By a clever application of Theorems 10.2 and 10.4, we will not only show that such sequences exist for all positive integers n, but we will also count the number of them. It turns out that there are *lots* of them. For instance, the number of inequivalent binary de Bruijn sequences of degree eight is equal to

$$1329227995784915872903807060280344576.$$

The reader with some extra time on his or her hands is invited to write down these sequences. De Bruijn sequences are named after Nicolaas Govert de Bruijn, who published his work on this subject in 1946. However, it was discovered in 1975 that de Bruijn sequences had been earlier created and enumerated by C. Flye Sainte-Marie in 1894. De Bruijn sequences have a number of interesting applications to the design of switching networks and related topics.

Our method of enumerating binary de Bruijn sequence will be to set up a correspondence between them and Eulerian tours in a certain directed graph D_n, the *de Bruijn graph* of degree n. The graph D_n has 2^{n-1} vertices, which we will take to consist of the 2^{n-1} binary sequences of length $n-1$. A pair $(a_1a_2\cdots a_{n-1}, b_1b_2\cdots b_{n-1})$ of vertices forms an edge of D_n if and only if $a_2a_3\cdots a_{n-1} = b_1b_2\cdots b_{n-2}$, i.e., e is an edge if the last $n-2$ terms of $\text{init}(e)$ agree with the first $n-2$ terms of $\text{fin}(e)$. Thus every vertex has indegree two and outdegree two [why?], so D_n is balanced. The number of edges of D_n is 2^n. Moreover, it's easy to see that D_n is connected (see Lemma 10.9). The graphs D_3 and D_4 look as follows:





Suppose that $E = e_1 e_2 \cdots e_{2^n}$ is an Eulerian tour in D_n. If $\text{fin}(e_i)$ is the binary sequence $a_{i1} a_{i2} \cdots a_{i,n-1}$, then replace e_i in E by the last bit $a_{i,n-1}$. It is easy to see that the resulting sequence $\beta(E) = a_{1,n-1} a_{2,n-1} \cdots a_{2^n,n-1}$ is a binary de Bruijn sequence, and conversely every binary de Bruijn sequence arises in this way. In particular, since D_n is balanced and connected there exists at least one binary de Bruijn sequence. In order to count the total number of such sequences, we need to compute $\det \boldsymbol{L}(D_n)$. One way to do this is by a clever but messy sequence of elementary row and column operations which transforms the determinant into triangular form. We will give instead an elegant computation of the eigenvalues of $\boldsymbol{L}(D_n)$ based on the following simple lemma.

10.9 Lemma. *Let u and v be any two vertices of D_n. Then there is a unique (directed) walk from u to v of length $n-1$.*

Proof. Suppose $u = a_1 a_2 \cdots a_{n-1}$ and $v = b_1 b_2 \cdots b_{n-1}$. Then the unique path of length $n-1$ from u to v has vertices

$$a_1 a_2 \cdots a_{n-1}, a_2 a_3 \cdots a_{n-1} b_1, a_3 a_4 \cdots a_{n-1} b_1 b_2, \ldots, a_{n-1} b_1 \cdots b_{n-2}, b_1 b_2 \cdots b_{n-1}. \square$$

10.10 Theorem. *The eigenvalues of $\boldsymbol{L}(D_n)$ are 0 (with multiplicity one) and 2 (with multiplicity $2^{n-1} - 1$).*

Proof. Let $\boldsymbol{A}(D_n)$ denote the directed adjacency matrix of D_n, i.e., the rows and columns are indexed by the vertices, with

$$\boldsymbol{A}_{uv} = \begin{cases} 1, & \text{if } (u,v) \text{ is an edge} \\ 0, & \text{otherwise.} \end{cases}$$

Now Lemma 10.9 is equivalent to the assertion that $\boldsymbol{A}^{n-1} = J$, the $2^{n-1} \times 2^{n-1}$ matrix of all 1's [why?]. If the eigenvalues of $\boldsymbol{A}$ are $\lambda_1, \ldots \lambda_{2^{n-1}}$, then the eigenvalues of $J = A^{n-1}$ are $\lambda_1^{n-1}, \ldots, \lambda_{2^{n-1}}^{n-1}$. By Lemma 1.4, the eigenvalues of J are 2^{n-1} (once) and 0 ($2^{n-1} - 1$ times). Hence the eigenvalues of $\boldsymbol{A}$ are 2ζ (once, where ζ is an $(n-1)$-st root of unity to be determined), and 0 ($2^{n-1} - 1$ times). Since the trace of $\boldsymbol{A}$ is 2, it follows that $\zeta = 1$, and we have found all the eigenvalues of A.

Now $\boldsymbol{L}(D_n) = 2I - \boldsymbol{A}(D_n)$ [why?]. Hence the eigenvalues of $\boldsymbol{L}$ are $2 - \lambda_1, \ldots, 2 - \lambda_{2^{n-1}}$, and the proof follows from the above determination of $\lambda_1, \ldots, \lambda_{2^{n-1}}$. □

10.11 Corollary. *The number $B_0(n)$ of binary de Bruijn sequences of degree n beginning with n 0's is equal to $2^{2^{n-1}-n}$. The total number $B(n)$ of binary de Bruijn sequences of degree n is equal to $2^{2^{n-1}}$.*

Proof. By the above discussion, $B_0(n)$ is the number of Eulerian tours in D_n whose first edge the loop at vertex $00\cdots 0$. Moreover, the outdegree of every vertex of D_n is two. Hence by Corollary 10.5 and Theorem 10.10 we have

$$B_0(n) = \frac{1}{2^{n-1}} 2^{2^{n-1}-1} = 2^{2^{n-1}-n}.$$

Finally, $B(n)$ is obtained from $B_0(n)$ by multiplying by the number 2^n of edges, and the proof follows. □

Note that the total number of binary sequences of length 2^n is $N = 2^{2^n}$. By the previous corollary, the number of these which are de Bruijn sequences is just $\sqrt{N}$. This suggests the following unsolved problem. Let $\mathcal{A}_n$ be the set of all binary sequences of length 2^n. Let $\mathcal{B}_n$ be the set of binary de Bruijn sequences of degree n. Find an explicit bijection $\varphi : \mathcal{B}_n \times \mathcal{B}_n \to \mathcal{A}_n$, thereby giving a combinatorial proof of Corollary 10.11.